

Morphologie / Morphology

HSK 17.2



# **Handbücher zur Sprach- und Kommunikations- wissenschaft**

Handbooks of Linguistics  
and Communication Science

Manuels de linguistique et  
des sciences de communication

Mitbegründet von Gerold Ungeheuer (†)  
Mitherausgegeben 1985–2001 von Hugo Steger

Herausgegeben von / Edited by / Edités par  
Herbert Ernst Wiegand

Band 17.2

Walter de Gruyter · Berlin · New York

# Morphologie Morphology

Ein internationales Handbuch zur Flexion und  
Wortbildung  
An International Handbook on Inflection and  
Word-Formation

Herausgegeben von / Edited by  
Geert Booij · Christian Lehmann · Joachim Mugdan ·  
Stavros Skopeteas  
in collaboration with Wolfgang Kesselheim

2. Halbband / Volume 2

Walter de Gruyter · Berlin · New York

⊗ Gedruckt auf säurefreiem Papier, das die  
US-ANSI-Norm über Haltbarkeit erfüllt.

ISBN 3-11-017278-X

*Bibliografische Information der Deutschen Bibliothek*

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <<http://dnb.ddb.de>> abrufbar.

© Copyright 2004 by Walter de Gruyter GmbH & Co. KG, 10785 Berlin.  
Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Printed in Germany

Satz: META Systems GmbH, Wustermark

Druck: Tutte Druckerei GmbH, Salzweg

Einbandgestaltung und Schutzumschlag: Rudolf Hübler, Berlin

# Inhalt / Contents

## 2. Teilband / Volume 2

XIII.	Semantische Kategorien und Operationen in der Morphologie I: Entitätsbegriffe Semantic categories and operations in morphology I: Entity concepts	
94.	J. Lachlan Mackenzie, Entity concepts . . . . .	973
95.	Theo A. J. M. Janssen, Deixis and reference . . . . .	983
96.	Jeffrey Heath, Person . . . . .	998
97.	Colette Grinevald, Classifiers . . . . .	1016
98.	Alexandra Y. Aikhenvald, Gender and noun class . . . . .	1031
99.	Peter Bakema & Dirk Geeraerts, Diminution and augmentation .	1045
100.	José Luis Iturrioz-Leza & Stavros Skopeteas, Numerus . . . . .	1053
101.	Maria Koptjevskaja-Tamm, Mass and collection . . . . .	1067
102.	Barry Blake, Case . . . . .	1073
103.	Fritz Serzisko, Possession . . . . .	1091
XIV.	Semantische Kategorien und Operationen in der Morphologie II: Sachverhalts-, Eigenschafts- und verwandte Begriffe Semantic categories and operations in morphology II: State-of-affairs, property and related concepts	
104.	Kees Hengeveld, State-of-affairs concepts . . . . .	1104
105.	Sandra A. Thompson, Property concepts . . . . .	1111
106.	Charles J. Fillmore, Circumstance concepts . . . . .	1117
107.	Martin Haspelmath & Thomas Müller-Bardey, Valency change .	1130
108.	Masayoshi Shibatani, Voice . . . . .	1145
109.	Ronny Boogaart, Aspect and Aktionsart . . . . .	1165
110.	Östen Dahl, Tense . . . . .	1180
111.	Kees Hengeveld, Illocution, mood, and modality . . . . .	1190
112.	Ger P. Reesink, Interclausal relations . . . . .	1202
113.	D. N. S. Bhat, Negation . . . . .	1207
114.	Pierluigi Cuzzolin & Christian Lehmann, Comparison and gradation . . . . .	1212

<b>XV.</b>	<b>Morphologische Typologie und Universalien</b> <b>Morphological typology and universals</b>	
115.	Theodora Bynon, Approaches to morphological typology . . . . .	1221
116.	John Haiman, Types of morphological structure . . . . .	1231
117.	Sebastian Kempgen & Werner Lehfeldt, Quantitative Typologie . .	1235
118.	Johannes Helmbrecht, Cross-linguistic generalizations and their explanation . . . . .	1247
<b>XVI.</b>	<b>Systeme morphologischer Struktur: Sprachskizzen</b> <b>Systems of morphological structure: Illustrative sketches</b>	
119.	Andrew Spencer, English (Indo-European: Germanic) . . . . .	1255
120.	Elke Ronneberger-Sibold, Deutsch (Indogermanisch: Germanisch) . . . . .	1267
121.	† Danielle Corbin, Français (Indo-européen: Roman) . . . . .	1285
122.	Helmut Jachnow, Russisch (Indogermanisch: Slawisch) . . . . .	1300
123.	Helena Kurzová, Altgriechisch (Indogermanisch) . . . . .	1310
124.	Fred Karlsson, Finnish (Finnno-Ugric) . . . . .	1328
125.	Yishai Tobin, Hebrew (Semitic) . . . . .	1343
126.	Armin Bassarak & Gerd Jendraschek, Türkisch (Turk) . . . . .	1358
127.	† Helma van den Berg, Hunzib (North-East Caucasian) . . . . .	1367
128.	Margarita Wall & Igor Kanakin, Ketisch (Jenisseysisch) (Beispielanalyse: Heinrich Werner) . . . . .	1376
129.	Michael Fortescue, West Greenlandic (Eskimo) . . . . .	1389
130.	Michael Fortescue, Koyukon (Athapaskan) . . . . .	1399
131.	Sandra Clarke & Marguerite MacKenzie, Montagnais/ Innu-aimun (Algonquian) . . . . .	1411
132.	Maura Velázquez-Castillo, Guaraní (Tupi-Guarani) . . . . .	1421
133.	Michel Launey, Nahuatl (Uto-Aztecan) . . . . .	1433
134.	Willem F. H. Adelaar, Quechua (Quechua) . . . . .	1453
135.	Doris L. Payne, Yagua (Peba-Yaguan) . . . . .	1464
136.	Nikolaus P. Himmelmann, Tagalog (Austronesian) . . . . .	1473
137.	Peter Austin, Diyari (Pama-Nyungan) . . . . .	1490
138.	Lourens de Vries, Wambon (Awyu) . . . . .	1501
139.	Gerrit Dimmendaal & Manuela Noske, Turkana (Nilo-Saharan) . . . . .	1507
140.	Brigitte Reineke, Twi (Kwa) . . . . .	1517
141.	Alexandre Kimenyi, Kinyarwanda (Bantu) . . . . .	1527
142.	Valerij S. Panfilov, Vietnamesisch (Viet-Muong) (Übersetzung: Holger Baumann & Martin Haspelmath) . . . . .	1545
143.	Renate Fischer & Regina Leven, Deutsche Gebärdensprache . .	1554
144.	Detlev Blanke, Plansprachen . . . . .	1563
<b>XVII.</b>	<b>Morphologischer Wandel I: Theoretische Probleme</b> <b>Morphological change I: Fundamental issues</b>	
145.	Geert Koefoed & Jaap van Marle, Fundamental concepts . . . . .	1574
146.	Scott DeLancey, Grammaticalization: from syntax to morphology . . . . .	1590

147.	† Wolfgang U. Wurzel, Morphologisierung: von der Phonologie zur Morphologie . . . . .	1600
148.	Jeroen van de Weijer, Analogical change . . . . .	1611
149.	Alwin Fill, Remotivation and reinterpretation . . . . .	1615
150.	Hans Sauer, Lexicalization and demotivation . . . . .	1625
151.	Suzanne Romaine, Change in productivity . . . . .	1636
152.	Norbert Boretzky, Morphologische Entlehnung und Lehnübersetzung . . . . .	1644
153.	Pieter Muysken, Pidginization, creolization, and language death . . . . .	1653
154.	Brian D. Joseph, Morphological reconstruction . . . . .	1661
<b>XVIII.</b>	<b>Morphologischer Wandel II: Fallstudien</b> <b>Morphological change II: Case studies</b>	
155.	Klaus Faiß, From Old English to Modern English . . . . .	1668
156.	Hans-Joachim Solms, Vom Althochdeutschen zum Neuhochdeutschen . . . . .	1680
157.	Franz Rainer, From Latin to French . . . . .	1698
158.	George Cardona, From Vedic to modern Indic languages . . . . .	1712
159.	Edwin G. Pulleyblank, From Archaic Chinese to Mandarin . . . . .	1730
160.	Kees Versteegh, From Classical Arabic to the modern Arabic vernaculars . . . . .	1740
161.	Peter Mühlhäusler, Tok Pisin . . . . .	1755
<b>XIX.</b>	<b>Psycholinguistische Perspektiven</b> <b>Psycholinguistic perspectives</b>	
162.	Hartmut Günther, Mentale Repräsentation morphologischer Strukturen . . . . .	1766
163.	Jean Aitchison, Speech production and perception . . . . .	1778
164.	† Victoria A. Fromkin, Speech errors . . . . .	1788
165.	Eve V. Clark & Ruth A. Berman, First language acquisition . . . . .	1795
166.	Peter Jordens, Second language acquisition . . . . .	1806
167.	Chris Schaner-Wolles & Wolfgang U. Dressler, Sprachstörungen	1816
<b>XX.</b>	<b>Morphologie in der Praxis</b> <b>Morphology in practice</b>	
168.	Ursula Wiesemann, Field work . . . . .	1829
169.	Christian Lehmann, Interlinear morphemic glossing . . . . .	1834
170.	Christian Lehmann & Elena Maslova, Grammaticography . . . . .	1857
171.	Howard Jackson, Lexicography . . . . .	1882
172.	Walter Daelemans, Computational linguistics . . . . .	1893
<b>XXI.</b>	<b>Morphologie und Nachbardisziplinen</b> <b>Morphology and related fields</b>	
173.	Rüdiger Harnisch & Damaris Nübling, Namenkunde . . . . .	1901
174.	Elmar Seebold, Etymologie . . . . .	1910

175.	Hartmut Günther, Schriftsysteme . . . . .	1915
176.	Jeanette M. Pugh & Juan C. Sager, Terminology in special languages . . . . .	1924
177.	Reiner Schmidt, Sprachunterricht . . . . .	1929

## Verzeichnisse und Sprachenkarten

### Indices and language maps

Verzeichnis der Abkürzungen / List of abbreviations . . . . .	1937
Sprachenkarten / Language maps . . . . .	1941
Namenregister / Index of names . . . . .	1949
Sprachenregister / Index of languages . . . . .	1972
Sachregister / Index of subjects . . . . .	1985

## 1. Teilband / Volume 1

Vorwort . . . . .	XIII
Preface . . . . .	XXI

I.	Morphologie als Disziplin Morphology as a discipline	
1.	Wolfgang U. Wurzel, Der Gegenstand der Morphologie . . . . .	1
2.	Paul Salmon, The term <i>morphology</i> . . . . .	15
3.	Vladimir A. Plungian, Die Stellung der Morphologie im Sprachsystem (Übersetzung: Robert Hammel) . . . . .	22
II.	Geschichte der morphologischen Forschung I: von der Antike bis zum 19. Jahrhundert History of morphological research I: From antiquity to the 19 <sup>th</sup> century	
4.	Jeremy Black, The Ancient Near East . . . . .	35
5.	George Cardona, Old Indic grammar . . . . .	41
6.	Robert H. Robins, Classical Antiquity . . . . .	52
7.	Jonathan Owens, Traditional Arabic grammar . . . . .	67
8.	Vivien Law, The Middle Ages . . . . .	76
9.	Von der Renaissance bis ca. 1800 . . . . .	90
10.	Sabine Ziegler, Das 19. Jahrhundert . . . . .	91
III.	Geschichte der morphologischen Forschung II: Forschungstraditionen im 20. Jahrhundert History of morphological research II: Research traditions in the 20 <sup>th</sup> century	
11.	Clemens Knobloch, Schulgrammatik als Modell linguistischer Beschreibung . . . . .	104
12.	Henry M. Hoenigswald, Historical-comparative grammar . . . . .	117

13.	Überblick: die europäische Tradition . . . . .	124
14.	Sebastian Kempgen, Osteuropa . . . . .	125
15.	Françoise Kerleroux, France and Switzerland (Translation: Rachel Pankhurst) . . . . .	138
16.	Even Hovdhaugen, Scandinavia . . . . .	145
17.	Francis Katamba, Britain . . . . .	149
18.	Wolfgang Motsch, Germany . . . . .	157
19.	Henk Schultink, The Netherlands (Translation: Geert Booij) . . .	162
20.	John Fought, American Structuralism . . . . .	170
21.	Ruth M. Brend, Tagmemics . . . . .	191
22.	Mark Aronoff, Generative grammar . . . . .	194
<b>IV.</b>	<b>Grundbegriffe</b>	
	<b>Basic concepts</b>	
23.	Pierre Swiggers, Linguistic sign . . . . .	210
24.	Jaap van Marle, Paradigmatic and syntagmatic relations . . . . .	225
25.	José Luis Iturrioz Leza & Stavros Skopeteas, Variation und Invarianz . . . . .	234
26.	Laurie Bauer, Word . . . . .	247
27.	William Croft, Lexical and grammatical meaning . . . . .	257
28.	Andrew Carstairs-McCarthy, Category and feature . . . . .	264
29.	Linda R. Waugh & Barbara A. Lafford, Markedness . . . . .	272
30.	John Haiman, Iconicity . . . . .	281
31.	Wolfgang U. Dressler, Naturalness . . . . .	288
32.	Ferenc Kiefer, Regularity . . . . .	296
33.	Geert Koefoed & Jaap van Marle, Productivity . . . . .	303
<b>V.</b>	<b>Die Rolle der Morphologie in Grammatik und</b>	
	<b>Lexikon</b>	
	<b>The role of morphology in grammar and lexicon</b>	
34.	Andrew Spencer, Morphology and syntax . . . . .	312
35.	Geert Booij, Morphology and phonology . . . . .	335
36.	Mark Aronoff, Morphology between lexicon and grammar . . . .	344
37.	Pius ten Hacken, Derivation and compounding . . . . .	349
38.	Geert Booij, Inflection and derivation . . . . .	360
<b>VI.</b>	<b>Einheiten der morphologischen Struktur</b>	
	<b>Units of morphological structure</b>	
39.	Joan Bybee, Lexical, morphological and syntactic symbolization	370
40.	Hans Basbøll, Word boundaries . . . . .	377
41.	Joel A. Nevis, Clitics . . . . .	388
42.	Rochelle Lieber & Joachim Mugdan, Internal structure of words	404
43.	Elena S. Kubrjakova, Submorphemic Einheiten (Übersetzung und Bearbeitung: Joachim Mugdan) . . . . .	417

44.	Edmund Gussmann & Bogdan Szymanek, Phonotactic properties of morphological units . . . . .	427
45.	Henning Bergenholz & Joachim Mugdan, Nullelemente in der Morphologie . . . . .	435
<b>VII. Allomorphie Allomorphy</b>		
46.	Hans Christian Luschützky, Morphem, Morph und Allomorph . . . . .	451
47.	Martin Neef, Phonologische Konditionierung . . . . .	463
48.	Martin Neef, Morphologische und syntaktische Konditionierung . . . . .	473
49.	Henriette Walter, Fluctuation and free variation (Translation: Marie Landick) . . . . .	484
50.	Tsutomu Akamatsu, Generalized representations . . . . .	489
51.	Edmund Gussmann, Underlying forms . . . . .	499
52.	Igor Mel'čuk, Suppletion . . . . .	510
<b>VIII. Formale Prozesse Formal processes</b>		
53.	Igor Mel'čuk, Morphological Processes . . . . .	523
54.	Christopher J. Hall, Prefixation, suffixation and circumfixation . . . . .	535
55.	Edith A. Moravcsik, Infixation . . . . .	545
56.	Ellen Broselow, Transfixation . . . . .	552
57.	Caroline Wiltshire & Alec Marantz, Reduplication . . . . .	557
58.	Rochelle Lieber, Substitution of segments and features . . . . .	567
59.	Thomas Becker, Metathesis . . . . .	576
60.	Wolfgang U. Dressler, Subtraction . . . . .	581
61.	Larry M. Hyman & William R. Leben, Suprasegmental processes . . . . .	587
<b>IX. Flexion Inflection</b>		
62.	Andrew Carstairs-McCarthy, Lexeme, word-form, paradigm . . . . .	595
63.	Aleksandr V. Bondarko, Meaning vs. use in inflection . . . . .	607
64.	Richard Coates, Exponence . . . . .	616
65.	Andrew Carstairs-McCarthy, Inflection classes . . . . .	630
66.	Silvia Luraghi, Synkretismus (Übersetzung: Simona Fabellini) . . . . .	638
67.	Fred Karlsson, Defectivity . . . . .	647
68.	Martin Haspelmath, Periphrasis . . . . .	654
<b>X. Wortarten Word classes</b>		
69.	Kjell-Åke Forsgren, Wortart, syntaktische Funktion, syntaktische Kategorie . . . . .	665

70.	Clemens Knobloch & Burkhard Schaefer, Kriterien für die Definition von Wortarten . . . . .	674
71.	Barbara Kaltz, Wortartensysteme in der Linguistik . . . . .	693
72.	Nicholas Evans, Word classes in the world's languages . . . . .	708
73.	Christian Lehmann & Edith Moravcsik, Noun . . . . .	732
74.	D. N. S. Bhat & Regina Pustet, Adjective . . . . .	757
75.	Joseph H. Greenberg, Numeral . . . . .	770
76.	Linda Schwartz, Pronoun and article . . . . .	783
77.	Joan Bybee, Verb . . . . .	794
78.	John M. Anderson, Auxiliary . . . . .	808
79.	Casper de Groot, Minor word classes . . . . .	820

**XI. Wortbildung I: Grundprobleme**  
**Word formation I: Fundamental problems**

80.	Laurie Bauer, System vs. norm: coinage and institutionalization	832
81.	Wiecher Zwanenburg, Correspondence between formal and semantic relations . . . . .	840
82.	Jacob Hoeksema, Compositionality of meaning . . . . .	851
83.	Geert Booij, Inheritance . . . . .	857
84.	Claudio Iacobini, Base and direction of derivation . . . . .	865
85.	Franz Rainer, Produktivitätsbeschränkungen . . . . .	877

**XII. Wortbildung II: Prozesse**  
**Word formation II: Processes**

86.	Wolfgang Fleischer, Die Klassifikation von Wortbildungsprozessen . . . . .	886
87.	Susan Olsen, Composition . . . . .	897
88.	Marianne Mithun, Incorporation . . . . .	916
89.	Bernd Naumann & Petra M. Vogel, Derivation . . . . .	929
90.	Jan Don & Mieke Trommelen & Wim Zonneveld, Conversion and category indeterminacy . . . . .	943
91.	Garland Cannon, Blending . . . . .	952
92.	Charles W. Kreidler, Clipping and acronymy . . . . .	956
93.	Philip Baldi & Chantal Dawar, Creative processes . . . . .	963

## XIII. Semantische Kategorien und Operationen in der Morphologie I: Entitätsbegriffe

### Semantic categories and operations in morphology I: Entity concepts

#### 94. Entity concepts

1. Introduction
2. Four orders of entities
3. The hypostatization of qualities
4. Morphological signalling of orders of entities
5. First-order entities
6. Higher-order entities
7. Hypostatized qualities
8. Conclusion
9. References

##### 1. Introduction

The cognitive structure of a situation can be described as a set of – possibly imaginary or abstract – entities and a set of static or dynamic relations between them. In this article we shall be concerned with the orders of entities that are distinguished in natural languages and the main categories of entity concepts which can be marked morphologically. The article will thus serve partially as a preparation for articles 95–103, which deal with various subcategorizations of entity and their morphological consequences.

The situations to which entities belong are invoked linguistically through the co-operative, interpersonal process of reference. Reference involves a fourfold relation between a speaker S, an addressee A, a referent (or set of referents) E, and a referring expression R: {S, A, E, R}. The referents referred to by a speaker are not phenomena of the real world, but “mental phenomena”, phenomena in the mental world. They are introduced into the ongoing discourse, and referred to and talked about irrespective of their status in reality. Illusory, hallucinatory or imaginary phenomena can in principle be talked about just as easily as veridical phenomena, i.e. those corresponding to phenomena in the real world,

and without any consequences for the form of the referring expression itself.

What is more, the linguistic codification of reality is often at odds with phenomena in the real world. One of the most striking aspects of this disparity is the linguistic distinction between countable and uncountable referents, a distinction to which we shall have occasion to return in 6 (see also Art. 101). There is ultimately no basis in external reality for this opposition: to give but one example, the English collective *furniture* denotes an uncountable entity yet corresponds to a plurality of entities in the real world (chairs, tables, bookcases, etc.).

All such referents, be they real or imaginary, countable or uncountable, will in this chapter be termed **entities**. “Entity” will thus here be used to identify the mental unit E that the speaker S, by using a referring expression R, wishes the addressee A to either construe (in the case of first mention) or re-identify (in the case of subsequent mention) – for details, see Dik (1989: 114). This approach entails a broader understanding of the word entity than is normal in ordinary discourse since, as Lyons (1977: 297 f.) points out for pre-theoretical use, “‘entity’ ... covers only countable nouns ... [and there] is no lexeme which is superordinate to all abstract nouns, or all concrete nouns, or all mass nouns, or all the members of any of the major subclasses of nouns ...”. In this article, however, entity will be employed, in a manner that has become standard practice in linguistics, in the general sense outlined above.

This having been said, it should be pointed out that speakers can also refer to phenomena that are not entities. In a sentence such as (1):

- (1) *The South of France was where I spent my summer.*

*the South of France* and *my summer* are referring expressions that identify a place and a period of time, respectively. As argued in Mackenzie (1992), places are in the ontology of the mental world distinct from entities (see also Lyons 1977: 693). One argument for making this distinction is that the standard definition of physical objects, which represent the most basic order of entity (see 2) involves their being located in space, i.e. at some place (Lyons 1977: 443; Dik 1989: 181). If places were entities, the definition would then be circular. Strawson (1959: 37), too, insists that entities and places are not only distinct but also mutually defining:

“[P]laces are defined only by the relations of things; and ... one of the requirements for the identity of a material thing is that its existence, as well as being continuous in time, should be continuous in space.” (Strawson 1959: 37)

As is indeed also apparent from this quotation, expressions like *my summer* in (1), which identify a time, similarly do not refer to an entity. And analogous remarks apply to those expressions, such as *very badly* in (2), which identify the manner in which an action is carried out:

(2) *The way the children behaved was very badly.*

Although such referents as those discussed in the preceding paragraph are non-entities (a fact which has a number of grammatical consequences – see Mackenzie 1992), there is nevertheless a tendency for places, times, and manners to be treated linguistically as though they were entities. This process, known as **hypostatization**, allows qualities, i.e. properties of entities, themselves to be reinterpreted (and consequently linguistically handled) as entities (see 3).

The focus of the present chapter will be on a typology of entity which can be marked morphologically, generally by processes of word formation. Since entity concepts are predominantly designated by nouns (but see 2), this chapter will be fundamentally concerned with processes of **nominalization** across the languages of the world. It behooves us first to present a typology of entities; we shall then progress to an examination of the morphological indicators of the various orders of entity.

## 2. Four orders of entities

Lyons (1977: 442–447) presents a three-way typology of entities, which refines the traditional distinction between concrete and abstract nouns (cf. also Art. 73):

- Entities of the first order are physical objects, i.e. persons, animals, and things. “[U]nder normal conditions, they are relatively constant as to their perceptual properties; ... they are located, at any point in time, in what is, psychologically at least, a three-dimensional space; and ... they are publicly observable” (Lyons 1977: 443). **First-order entities** are evaluated in terms of their existence.
- Entities of the second order are “events, processes, states-of-affairs, etc., which are located in time and which, in English, are said to occur or take place, rather than to exist” (Lyons 1977: 443). **Second-order entities** are evaluated in terms of their reality.
- Entities of the third order are “such abstract entities as propositions, which are outside space and time” (Lyons 1977: 443). **Third-order entities** are evaluated in terms of their truth.

To this tripartition we may add, following Hengeveld (1992: 7), a fourth order of entities, which comprises speech acts. **Fourth-order entities** are located in space and time, and are evaluated in terms of their felicity. The following table gives an overview of the four orders of entity, with examples of corresponding nouns in English:

order	evaluation	examples
first	existence	woman, tortoise, nose, lorry
second	reality	arrival, error, excursion, contest
third	truth	belief, idea, fact, hope
fourth	felicity	statement, question, command

Tab. 94.1: Orders of entities

Correspondingly, Lyons (1977: 446) refers to a noun such as *woman* as a first-order noun, *arrival* as a second-order noun, etc. First-order nouns are universally regarded as the prototypical examples of the category of nouns. This is reflected, for example, in the fact that most higher-order nouns are syn-

chronically or at least diachronically complex in their morphology. Moreover, whereas first-order entities are in general identified exclusively by nouns, second- and third-order entities are also referred to by embedded clauses, and fourth-order entities by quoting direct speech.

Consider now (3), in which both first-order entities referred to are designated by a noun, *boy* and *goat* respectively:

- (3) *The boy looked at the goat.*

In Iroquoian languages, by contrast, there are few true nouns (cf. Sasse 1988). Here, it frequently occurs that first-order entities, too, are referred to by embedded clauses, as in (4) from Tuscarora, the translational equivalent of (3):

- (4) *Ra-kwá.tihs wa-hr-Ø-atkáh-to?*  
 M.SBJ-young PAST-M.SBJ-OBJ-look.at-PNCT  
*ka-téksr-ahs.*  
 N.SBJ-stink-IMPF  
 ‘The boy looked at the goat. (lit. He is young, he looked at it, it stinks.)’ (Mithun Williams 1976: 32)

Tuscarora, it should be said, represents in this respect an extreme type across the languages of the world (cf. Hengeveld 1992: 69).

The possibility of choosing, in reference to higher-order entities, between nominal and clausal expression is familiar from many languages. Consider such pairs as (5a) and (5b), or (6a) and (6b):

- (5) (a) *I witnessed John refuse the offer.*  
 (b) *I witnessed John’s refusal of the offer.*
- (6) (a) *I heard that John had refused the offer.*  
 (b) *I heard about John’s refusal of the offer.*

In (5a), the second-order entity, the event of John’s refusing the offer, is presented as a (non-finite) clause, but in (5b) as a noun phrase; similarly, in (6a) the third-order entity, the fact of John’s refusing the offer, is presented as a (finite) clause, but in (6b) as a noun phrase (identical to the noun phrase in (5b)).

Languages differ from one another in the extent to which nominalizations, such as *refusal* in (5b) and (6b), can be involved in reference to higher-order entities. Mackenzie (1987) distinguishes four “degrees of nominalization” in the 30-language sample he examines:

Degree 1: partial deverbalization, i.e. the loss of otherwise expected verbal proper-

ties or their replacement by indicators of embedded or dependent status;

Degree 2: external trappings of noun phrase status, i.e. some indication (e.g. adpositions, affixes) of the semantic, syntactic or pragmatic function of the referring expression;

Degree 3: optional presence of a possessor or some other dependent term within the referring expression;

Degree 3 (optional): optional or required presence of term operators, i.e. articles, demonstratives, etc. within the referring expression;

Degree 4: high degree of nouniness, i.e. optional presence of adjectives, gender, countability.

The four degrees of nominalization identified by Mackenzie (1987) can be presented as an implicational hierarchy, in the sense that any language displaying the characteristics of degree n also displays the characteristics of degree n-1:

- (7) nouniness (> operators) > possessor/dependent > function marking > deverbalization

Even across languages that display the same degree of nominalization, there may be considerable differences in the uses to which nominalizations with reference to higher-order entities are put in discourse. Whereas nominalizations in English are typically used in reference to discoursally backgrounded events, facts, etc., Nichols (1988) has shown that Russian (also a degree 4 language) readily nominalizes asserted, foregrounded information, such that a literal translation into English, preserving Russian nominalizations as English ones, yields prose that fails to get the point across (because the English reader assumes that the nominalized material is backgrounded).

What is more, within one and the same language, different registers may employ nominal reference to higher-order entities to different extents: in German, for example, *offizialese* makes frequent use of the “Nominalstil”, as in (8):

- (8) *Zur Wiederholung der Aufführung dieses Stücks ist von unserer Seite keine Veranlassung gegeben.*

The same ideas can be expressed without any use of higher-order nouns such as *Wiederholung*, *Aufführung*, and *Veranlassung*:

- (9) *Wir sehen uns nicht veranlaßt, dieses Stück noch einmal aufzuführen.* (example adapted from Duden 1965: 456)

In one and the same language, there may be a range of strategies for reference to higher-order entities that recapitulates the implicational hierarchy shown in (7) above. In English, for example, it has been recognized at least since Ross (1973) that there is a **nouniness squish** in this area which leads from fully verbal to fully nominal expression, through three kinds of gerund (V-ing) with a partially verbal, partially nominal character:

- (10) (a) *That John refused the offer (came as a great surprise).*  
 (b) *John refusing the offer ...*  
 (c) *John's refusing the offer ...*  
 (d) *John's refusing of the offer ...*  
 (e) *John's refusal of the offer ...*

In (10a) and (10b), both arguments of *refuse*, namely *John* (Agent) and *the offer* (Patient), are obligatory, in (10c) only the Patient is grammatically required, and in (10d) and (10e) neither argument is required (cf. Art. 83 on the optionality of inherited arguments). In (10a), the verb can have all finiteness, tense, and aspect properties, whereas in (10b) and (10c) only the tense opposition between Past (*having refused*) and Unspecified (*refusing*) is possible; (10d) and (10e) cannot take any marking for verbal categories. As for markers of nouniness, these are completely lacking in (10a); (10b) can be associated with a preposition (*surprised about John refusing the offer*; cf. Degree 2 above); (10c) accepts a demonstrative (*that refusing the offer of John's*; cf. Degree 3 (optional) above); (10d) allows a definite article (*the/\*a refusing of the offer*), adjectives, and relative clauses; and (10e), finally, shows all the characteristics of a noun phrase (cf. Degree 4). Thus, in terms of argument inheritance, deverbalization, and nominalization proper, there is clear evidence of a gradual transition.

Whereas the relationship between a verb such as *refuse* and a noun such as *refusal* is clearly a matter of derivational morphology, the question arises whether the English gerund should be regarded as displaying inflectional or derivational morphology. This has been a matter of some dispute (cf. the discussion in Koptjevskaja-Tamm 1993: 263–266):

Anderson (1985: 163) regards *-ing* nominalization, involving as it does a shift in word class, as a “derivational formation”, whereas Bybee (1985: 85), pointing to the lexical generality of *-ing* attachment, considers the morpheme to be inflectional. There is a basis for solving the dispute in the work of Booij (1993), who makes a precise distinction between contextual inflection, which is induced by the syntactic context of the word, and inherent inflection, i.e. inflection not so induced (see also Art. 38). Cases of inherent inflection combine the property of lexical generality with the possibility of meaning change (as in plural, comparative/superlative, and tense morphology) and/or change in syntactic category (as in the formation of infinitives, which are at once verbal and nominal, or participles, which are simultaneously verbal and adjectival). In this light, we may regard the English gerund, too, as an instance of inherent inflection. As Booij (1993: 31) points out, whereas there is a sharp distinction between contextual and inherent inflection, the difference between inherent inflection and derivation is gradual: this fact is again demonstrated in the gradual, “squishy” transition from *refuse* through *refusing* to *refusal*.

To summarize, we recognize a basic four-order typology of entities. The referring expressions that are used to identify first-order entities are in almost all languages built around nouns; those used to identify higher-order entities in many languages take the form of embedded clauses, although both within and across languages there are differences in the extent to which there are nominalization processes available whereby reference to those entities may also be constructed around nouns. The nominalizing morphemes involved may be examples of inherent inflection or derivation.

### 3. The hypostatization of qualities

Corresponding to each of the four orders of entities, we may also recognize four orders of qualities (for some discussion, see Weigand 1990: 100 f.). As mentioned above, qualities are not entities and indeed their basic expression is not as nouns; rather, they are signalled in the first instance by adjectives in those languages that possess that category or, where adjectives are lacking, by verbs. Nevertheless, languages have developed nominalization strategies which enable qualities to be talked

about as though they were entities. This hypostatization of qualities falls under what Halliday (1985) recognizes as **grammatical metaphor**: alongside (11a), which Halliday would term “congruent”, in that the quality is expressed adjectivally, we also find (11b), which is its “metaphorical” equivalent and which, in the nominalization *height* (from *high*), expresses the quality in nominal, “reified” form:

- (11) (a) *The Empire State Building is higher than the Chrysler Building.*
- (b) *The height of the Empire State Building exceeds that of the Chrysler Building.*

First-order entities (persons, things, and animals) have qualities that we may refer to as attributes; hypostatized examples include the height of the Empire State Building, the temperature of a liquid, the number of a crowd, the colour of a squirrel. Second-order entities (situations or “states of affairs”; cf. Art. 104) are qualified in terms of their circumstances: these include such qualities as their location in place and time, the manner in which they are conducted, as well as, where relevant, the frequency with which they occur, e.g. the venue of a football match, the time of a train’s arrival, the carelessness of a music performance, the rate of a heartbeat. Third-order entities are qualified in terms of their (epistemic) status, e.g. the plausibility of a belief, the truth of a proposition. And fourth-order entities, finally, have qualities of style, e.g. the appropriacy of a question, the impiousness of an order. The following table gives an overview of the four orders of hypostatized quality, with examples of corresponding nouns in English:

order	quality	examples
first	attribute	temperature, number, colour
second	circumstance	venue, time, carelessness
third	status	plausibility, truth
fourth	style	appropriacy, impiousness

Tab. 94.2: Orders of hypostatized qualities

On the basis of the typologies of entities and hypostatized qualities presented here, we shall now turn to an examination of the mor-

phological marking of nouns around which expressions referring to entities and qualities can be constructed.

#### 4. Morphological signalling of orders of entities

##### 4.1. Basic functions and extensions

Across the languages of the world, we find morphemes that have been specialized as markers of derived nouns for entities of the first, second, and third orders and for hypostatized second-order qualities. Among the types of first-order entities morphologically signalled in this way are Agents, Patients, and Instruments; the relevant second-order entities include actions, events, and states; and the third-order entities are generally facts and propositions. The hypostatized second-order qualities are notably places, manners, and times. What of fourth-order entities and hypostatized first-, third-, and fourth-order qualities? Although these are frequently denoted by nominalizations (cf. *statement* from the verb *state*, or *appropriacy* from the adjective *appropriate*), cross-linguistically there seem not to be regular, productive morphological procedures for their derivation.

Before we progress to a more detailed examination of the various orders and types, however, it is well to note that many morphemes specialized for one order of entity can also be used in contexts where they are to be interpreted as referring to another kind of entity. Generally, such reinterpretations, which may diachronically lead to actual shifts of meaning, work from higher to lower order. Thus, the English morpheme *-ation* is a marker of third-order status:

- (12) *I was most surprised by his resignation.*

Yet it can be reinterpreted as designating a second-order entity (an event) or a first-order entity (a thing):

- (13) (a) *I was present at his resignation.*
- (b) *His resignation was lying on the table for everyone to read.*

Similarly, in Dutch, the infinitive ending *-en* signals a second-order entity, but can in particular contexts also be used for first-order entities:

- (14) (a) *Na het et-en ging-en we after the eat-INF go:PAST-PL we wandel-en.*
- walk-INF
- ‘After the meal we went for a walk.’

- (b) *Snel, het et-en word-t*  
quick the eat-INF become-PRES.3.SG  
*koud.*  
cold  
‘Quick, the food is getting cold.’

In Yagua (Payne & Payne 1990: 352 f.), too, the second-order ‘action nominalization’, which in the Cahocuma dialect has the marker *-janu*, is open to reinterpretation as a first-order or quality noun; some examples:

- (15) *vicha* ‘live’  
> *vichajanu* ‘living’  
*ray-vicha-janu-mu* (*ravyíchqanumu*)  
‘1.SG-live-NR-LOC (to my village)’  
(= reinterpretation as Place)
- (16) *nicyee* ‘speak’  
> *nicyee-janu* ‘speaking’  
*jij-nicyee-janu* (*jiñnicyeejanu*)  
‘2.SG-speak-NR (your language)’  
(= reinterpretation as Patient)

What we in general do not find is that markers specialized to indicate lower-order entities have meaning extensions to higher-order entities.

We thus may not assume a one-to-one relationship between morphological form and order of entity. But even within one and the same order, the reading of a nominalization is not predictable from its form. Consider the English second-order nouns *washing* and *ironing*; these can also be found as first-order nouns, but then *washing* has two readings, as Intended Patient (‘clothes that are to be washed’) or as Affected Patient (‘clothes that have been washed’), whereas *ironing* as a first-order noun only has the Intended Patient sense (‘clothes that are to be ironed’).

Nevertheless, it remains possible to state for English, for example, that *-ing* is fundamentally a marker of higher-order entities and *-er* a marker of first-order entities, and for Yagua that it is basically to second-order entities that *-janu*, which also functions as an infinitive, is suffixed. It is in this spirit that the following sections will deal with the morphological marking across languages of types of first-order entities nameable by means of nominalizations (see 5); types of second- and third-order entities (see 6); and types of hypostatized qualities (see 7).

#### 4.2. Generalized nominalizing markers

Let us first, however, consider instances of **generalized nominalizing markers** that appear to be (largely) neutral as the order or type of

entity they indicate. In Acehnese, as described by Durie (1985: 145–149), there is an infix *-eun-* which is a very general nominalizer, giving a large range of meanings (but apparently excluding Agent):

- (17) (a) *høy* ‘call out to’ > *neuhøy* ‘the act of calling out’ (2nd order; Instance)  
(b) *noh* ‘put in stocks’ > *neunoh* ‘stocks’ (hypostatized quality; Place)  
(c) *karom* ‘brood’ > *keunarom* ‘brooded over eggs’ (1st order; Affected Patient)  
(d) *krut* ‘scrape’ > *keuneurut* ‘dung-hoe’ (1st order; Instrument)  
(e) *pheuet* ‘carve’ > *peuneuheuet* ‘a carving’ (1st order; Affected Patient)

In Mangarayi (Merlan 1982: 172 f.) the suffix *-yinl-/jinl-/njin* yields either higher-order nouns or first-order Agents, so that *ŋuj+bu-njin* may be understood as ‘deception’ or ‘deceiver’ and *ŋuñja-jin* as ‘mockery’ or ‘ mocker’. In Albanian (Camaj 1984: 79), too, there is a suffix *-esë-* which gives meanings ranging across various types of entity:

- (18) (a) *shtoj* ‘expand’ > *shtesëa* ‘expansion’ (2nd order)  
(b) *fshij* ‘sweep’ > *fshesëa* ‘broom’ (1st order; Instrument)  
(c) *kthej* ‘turn’ > *kthesëa* ‘turn, curve, bend’ (hypostatized quality; Place)

In all of these cases, the nominalizer does not contribute any meaning, merely indicating the occurrence of a change of category. The appropriate interpretation of the nominalization is established either by convention or by pragmatic inference.

### 5. First-order entities

The major types of first-order entity marked morphologically as such are Agent, Instrument, and Patient. It is possible for a language to have one marker for all types of first-order entity. Such a language is Babungo (Schaub 1985: 244 f.), in which the prefix *N-* can indicate an Agent, Instrument or Patient:

- (19) (a) *káy* ‘fry’ > *ŋkáy* ‘frier, person frying’ (Agent)  
(b) *tàø* ‘dig’ > *ntàø* ‘digging stick’ (Instrument)  
(c) *sháŋ* ‘chew’ > *nsháŋ* ‘food’ (Patient)

More frequently, however, languages have separate marking for different types of first-order entity.

### 5.1. Agent

Many languages have nominalizing markers that yield names for **Agents**. Typically, these morphemes are also found on names for other comparable roles, such as the sole participant in a Process (Processed) or in a State (Theme); hence the proposal to rename *nominum agentis* as “subject names” (Bauer 1983; Booij 1986) or “first-argument names” (MacKenzie 1990). One such morpheme is *-ni* in Pipil (Campbell 1985: 49):

- (20) (a) *ehku-* ‘arrive’ > *ekhu-ni* ‘arriver; ladino, foreigner’ (Agent)
- (b) *miki-* ‘die’ > *miki-ni* ‘dier; dead person’ (Processed)
- (c) *kuchi-* ‘sleep’ > *kuchi-ni* ‘sleeper; sleepyhead’ (Theme)

Some languages have different morphemes according as the Agent is male or female: in French for example, the suffix *-eur* is used for a male Agent and *-euse* for a female one; in Dutch, the female equivalent of the male Agent suffix *-er* is *-ster*:

- (21) *schrijv-en* ‘write’ > *schrijv-er* ‘male writer’; *schrijf-ster* ‘female writer’

Occasionally, languages will show a subcategorization of the Agent type for tense/aspect, frequency, polarity, etc. In Hixkaryana (Derbyshire 1979: 167), the suffix *-xenyeno* indicates the intransitive-subject (or transitive-object) of a recently performed action: thus *enuxenyeno*, from *onu-* ‘be born’ is a ‘newborn baby’. In West Greenlandic (Fortescue 1984: 319) we find special suffixes for a Frequent Agent (*-gajiuq*), a Skilled Agent (*-llaamak*), and a Début Agent (*-qqaaq*). And Hixkaryana (Derbyshire 1979: 167 f.), finally, has a Negative-Agent suffix *-hini*:

- (22) *i-manho-hini*  
GENERAL.PREFIX-dance-NEG.AG.NR  
‘one who doesn’t dance, a non-dancer’

Many languages make a distinction between Essential and Contingent Agent names. Essential Agent names typically designate professions, whereas Contingent names are used to refer to individuals performing the action on one or more specific occasions. Yagua (Payne & Payne 1990: 354 f.), for example,

marks an Essential Agent with *-ra* and a Contingent Agent with *-nu*:

- (23) *dapuuy* ‘hunt’ > *dapuuy-nù* (*dapúúñu*)  
‘hunter on a specific occasion; one who is hunting’
- dapuuy-ra* (*dapuurya*) ‘(professional) hunter’

In English, the distinction is not marked in the Agent suffix (*-er*), but occasionally in accent placement (see MacKenzie 1990: 137):

- (24) *underTAKE* > *underTAKer* of risky projects (Contingent)  
*Undertaker* (‘funeral director’: Essential)

The difference is also manifest in the greater ease with which the Essential Agent enters into synthetic compounds:

- (25) (a) *He is by profession an opera-singer/ \*singer of operas* (Essential)
- (b) *He is an occasional singer of arias/ ?aria-singer* (Contingent)

### 5.2. Instrument

Another role frequently named by means of derivational morphology is **Instrument**. In English, the suffix *-er*, used for Contingent and Essential Agents, is also the marker of Instruments (*starter, nail clipper, etc.*); for discussion of the overlap of Agent/Instrument morphology, see Dressler (1986). In French, the Instrument suffix *-euse* is homophonous with the female Agent such that *moissonneuse*, from the verb *moissonner* ‘harvest’ is ambiguous between ‘female harvester’ and ‘harvesting machine’.

In many languages, the Instrument form is clearly distinguished. In Breton (Press 1986: 218), for example, the Agent is marked by *-aer*, *-er* or *-our*, but the Instrument by *-ell*:

- (26) *skub-añ* ‘sweep’ > *skub-ell(enn)*  
‘broom’

In Jacalteco (Day 1973: 46), there are two Instrument suffixes according as the Patient of the input verb is named or not:

- (27) *mak'a* ‘hit’ >  
     *mak'b'al*      *chiyo*  
hit:INSTR.NR chicken  
‘chicken-hitting instrument’  
     > *mak'b'anil*  
‘hitting instrument’

### 5.3. Patient

A third type of first-order entity signalled by derivational morphology is the **Patient**. In English, the major use of the suffix *-ee* is to indicate the undergoer of an action: *employee*, *trainee*, *draftee*, etc. In Nkore-Kiga (Taylor 1985: 194), a Patient nominalization is marked by the substitution of *-o* for verb-final *-a*:

- (28) *biba* ‘sow’ >  
*em-bib-o*  
CL-SOW-PATNR  
‘seed’

As with Agent nouns, so also certain languages subcategorize Patients for tense/aspect and/or polarity. Thus Abkhaz (Hewitt 1979: 246) has a suffix *-t°(ə)* for Intended Patient, ‘that which is to be’ (cf. the discussion of English *washing* and *ironing* in 4.1):

- (29) *a-fa-t°*  
NOUN.PREFIX-eat-INTENDED.PATNR  
‘that which is to be eaten; food’

Hixkaryana (Derbyshire 1979: 165) has suffixes for recent past (‘one who recently was ...’), negative past (‘that which was not ...’), negative present (‘that which is not ...’).

The distinction between two types of Patient, those that are affected by the action (*objecum affectum*) and those that are brought into existence by the action (*objecum effectum*), may also be reflected in morphological distinctions between first-order entity denoting nouns. In Abkhaz (Hewitt 1979: 245) for example, the suffix *-mta* indicates the Effected Patient, the object created by the action:

- (30) *a-y°-t̪-mta*  
NOUN.PREFIX-write-AFFECTED.PATNR  
‘written work, writings’

In Finnish, the suffix *-os/-ös* is employed above all for such Effected Patient (“result”) nominals: *piirtää* ‘draw’ > *piirros* ‘drawing’; *kääntää* ‘turn, translate’ > *käännös* ‘translation’ (Karlsson 1983: 199).

## 6. Higher-order entities

As discussed in 2, higher-order entities are either events (second-order) or propositions (third-order). Second-order nouns may be distinguished into those that indicate an instance of the action (*nomina acti*) and those that denote the action in general. In English,

the distinction correlates fairly well with the opposition between conversion and suffixation with *-ing*:

- (31) *dance* (v.)  
> *dance* (n.) ‘instance of dancing’  
> *dancing* ‘the activity’

*Nomina acti* are countable, whereas action nominals are typically mass nouns (*dances*; *\*dancings*). A similar distinction may be observed in Basque (Saltarelli 1988: 258 f.), where the suffix *-t(z)e* forms action nominals, while *nomina acti* result from the suffixation of *-era*, *-penl-men*, *-(k)eta* or *-kuntza*: *etor-* ‘come’ may thus give either *etor-tze-a* ‘coming’ or *etorrera* (an) arrival’. In Gulf Arabic (Qafisheh 1977: 81), too, we observe an analogous opposition:

- (32) *ragaş* ‘dance’ > *ragş* ‘dancing’  
*ragşa* ‘a dance’

Higher-order nominalizations derived from transitive verbs may “inherit” their arguments as (typically) optional adnominal expressions (see Art. 83). These arguments may (a) retain their original formal marking; (b) convert either Agent or Patient to Possessive marking; (c) convert both arguments to Possessive marking. One language which takes the last-mentioned option is Finnish:

- (33) *Aimo-n kana-n syö-minen*  
Aimo-GEN chicken-GEN eat-NR  
‘Aimo’s eating of chicken’

Koptjevskaja-Tamm (1993: 169), whence this example is taken, points out that (33) can also be interpreted as ‘the eating by Aimo’s chicken’ and ‘the eating of Aimo’s chicken’, suggesting that “[a]s a consequence of this, Finnish tries to avoid transitive [action nominals]”. Be this as it may, there is in actual use a tendency for nominalized verbs to display fewer arguments than in corresponding verbal use (Koptjevskaja-Tamm 1993: 261); in certain languages, this may be grammaticalized as a ban on mentioning both Agent and Patient in connection with an action nominal (e.g. Chukchee; Koptjevskaja-Tamm 1993: 194–196) or reflected in valency-reducing morphology on the noun. In Aguacatec (Mackenzie 1985: 39 f.), for example, when a transitive verb undergoes higher-order nominalization it receives the suffix *-l*, which is otherwise used in the language for the valency reduction of verbs:

- (34) *Ma jalchan na chi-b'en cob'*  
 then early ASPECT 3.PL-go two  
*chakum tan b'iy-l-e'n*  
 messenger to kill-VALENCY.REDUCER-NR  
*cne'r Xchimal.*  
 sheep Xchimal  
 'Two messengers go early in the morning to kill a sheep in Xchimal.' (McArthur & McArthur 1966: 158)

The distinction between second-order and third-order entities is reflected in the morphology of Turkish, in which the "infinitive" suffixes *-mek* and *-me* are in general used for events and the "participial" suffixes *-dik* and *-ecek* for propositions:

- (35) *Halil'in gel-eceğ-in-i*  
 Halil:GEN come-FUT.PART-3.SG.POSS-ACC  
*bili-yor-um.*  
 know-PROG-1.SG  
 'I know Halil's future coming, i.e. that Halil will come.' (third order) (Underhill 1976: 322)
- (36) *En sev-me-dığ-im sey,*  
 most like-NEG-PART-1.SG.POSS thing  
*ders çalis-mak-tir.*  
 lesson study-INF-PRED  
 'My most not-liking thing is studying lessons, i.e. The thing which I dislike the most is studying.' (Underhill 1976: 310f.)

Similarly in English, the distinction between (10b–c) and (10d–e) (see 2), is that the former are most regularly employed in reference to third-order entities, whereas the latter fundamentally indicate a second-order entity. Thus (10d–e) are appropriate in contexts where a sequence of events is being described, as in (37a), while (10b–c) prefer epistemic contexts, as in (37b):

- (37) (a) *Peter's resignation came after John's refusing/refusal of the offer/\*after John('s) refusing the offer.*  
 (b) *Are you aware of John('s) refusing the offer/of John's \*refusing?refusal of the offer?*

## 7. Hypostatized qualities

The qualities to which reference can be made with derived nouns are, in the main, places, times, and manners.

Morphological marking of places is fairly general across the languages of the world. In Abkhaz (Hewitt 1979: 244), *-r+ta* derives, for

example, *à-jaxə-r+ta* 'sewing-room' from *jaxə* 'sew'; in Nkore-Kiga (Taylor 1985: 194), *-erol-iro* attaches to, for instance, *-shoma* 'read' to give *ei-shom-ero* 'CL-read-PLACE.NR (school)'. In West Greenlandic (Fortescue 1984: 319) the same suffix, *-vvik*, is used for both places and times.

Specific morphological indicators of times seem to be rarer. In Abkhaz (Hewitt 1979: 245), there is the suffix *-mta*, derived from *aamta* 'time' and homophonous with the Effected Patient suffix discussed in 5: *l-šə-mta* '3.SG.F-kill-TIME (the moment of her killing)'.

Manner nominals, finally, indicating the way in which an action is carried out, may be marked as such by specific morphemes. Koptjevskaja-Tamm (1993: 19) reports that in Amharic, manner nominals are distinct from action nominals:

- (38) (a) *mohed* 'to go, going'  
*akkiyahed* 'the way of going'  
 (b) *mayot* 'to see, seeing'  
*astoyayot* 'the way of seeing'

In Abkhaz (Hewitt 1979: 245), too, there is a special form for manner nouns, with the suffix *-ša*:

- (39) *a-ħ°à-ša*  
 NOUN.PREFIX-say-MANNR  
 'manner of speaking'

In English, by contrast, the manner interpretation of *-ing* is strongly associated with second-order nominalizations, with the lexical context as the guiding factor. *His driving* in (40) is thus ambiguous between a reading as a third-order nominal (= 'the fact that he was driving') or as a second-order nominal, which will then be interpreted as an indication of manner (= 'the way in which he was driving'):

- (40) *What did you think of his driving?*

## 8. Conclusion

The purpose of this article has been to show how morphological processes of word-formation, principally affixation, contribute to extending the range of nouns available for reference to entities and hypostatized qualities. We have observed that such nominalizing morphemes occur in reference to first-order, second-order, and third-order entities and to hypostatized second-order qualities. Whereas some languages have only one generalized

nominalizer, there is a tendency for languages to have a range of such morphemes, often with more than one function each. Where there is an extension of meaning, this generally proceeds from higher to lower orders of entity.

## 9. References

- Anderson, Stephen R. (1985), "Inflectional Morphology". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III*. Cambridge etc.: Cambridge University Press, 150–201
- Bauer, Laurie (1983), *English Word-Formation*. Cambridge etc.: Cambridge University Press
- Booij, Geert (1986), "Form and Meaning in Morphology: The Case of Dutch 'Agent Nouns'". *Linguistics* 24, 503–518
- Booij, Geert (1993), "Against Split Morphology". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology 1993*. Dordrecht etc.: Kluwer, 27–49
- Bybee, Joan (1985), *Morphology*. Amsterdam: Benjamins
- Camaj, Martin (1984), *Albanian Grammar*. Wiesbaden: Harrassowitz
- Campbell, Lyle (1985), *The Pipil Language of El Salvador*. Berlin etc.: Mouton
- Day, Christopher (1973), *The Jacalteco Language*. The Hague: Mouton
- Derbyshire, Desmond C. (1979), *Hixkaryana*. Amsterdam: North-Holland (Lingua Descriptive Studies 1)
- Dik, Simon C. (1989), *The Theory of Functional Grammar, Vol. I*. Dordrecht: Foris
- Dressler, Wolfgang U. (1986), "Explanation in Natural Morphology, Illustrated with Comparative and Agent-Noun Formation". *Linguistics* 24, 519–548
- Duden (1965) = DUDEN *Haupschwierigkeiten der deutschen Sprache*. Mannheim: Bibliographisches Institut
- Durie, Mark (1985), *A Grammar of Acehnese on the Basis of a Dialect of North Aceh*. Dordrecht: Foris
- Fortescue, Michael (1984), *West Greenlandic*. London etc.: Croom Helm
- Halliday, Michael A. K. (1985), *An Introduction to Functional Grammar*. London: Arnold
- Hengeveld, Kees (1992), *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter
- Hewitt, B. George (1979), *Abkhaz*. Amsterdam: North-Holland (Lingua Descriptive Studies 2)
- Karlsson, Fred (1983), *Finnish Grammar*. Porvoo etc.: Söderström
- Koptjevska-Tamm, Maria (1993), *Nominalizations*. London etc.: Routledge
- Lyons, John (1977), *Semantics. Vol. 1–2*. Cambridge: Cambridge University Press
- Mackenzie, J. Lachlan (1985), "Nominalization and Valency Reduction". In: Bolkestein, A. Machelt & Groot, Casper de & Mackenzie, J. Lachlan (eds.), *Predicates and Terms in Functional Grammar*. Dordrecht: Foris, 29–47
- Mackenzie, J. Lachlan (1987), "Nominalization and Basic Constituent Ordering". In: Auwera, Johan van der & Goossens, Louis (eds.), *Ins and Outs of the Predication*. Dordrecht: Foris, 93–105
- Mackenzie, J. Lachlan (1990), "First-Argument Nominalization in a Functional Grammar". *Linguistica Antverpiensia* 24, 119–147
- Mackenzie, J. Lachlan (1992), "Places and Things". In: Fortescue, Michael & Harder, Peter & Kristoffersen, Lars (eds.), *Layered Structure and Reference in a Functional Perspective*. Amsterdam, Philadelphia: Benjamins
- McArthur, Harry S. & McArthur, Lucille E. (1966), "Aguacatec". In: Myers, Marvin K. (ed.), *Languages of Guatemala*. The Hague: Mouton (Janua Linguarum, series practica, 23), 140–165
- Merlan, Francesca (1982), *Mangarayi*. Amsterdam: North-Holland (Lingua Descriptive Studies 4)
- Mithun Williams, Marianne (1976), *A Grammar of Tuscarora*. New York: Garland
- Nichols, Joanna (1988), "Nominalization and Assertion in Scientific Russian prose". In: Haiman, John & Thompson, Sandra A. (eds.), *Clause Combining in Grammar and Discourse*. Amsterdam, Philadelphia: Benjamins, 399–428
- Payne, Doris L. & Payne, Thomas E. (1990), "Yagua". In: Derbyshire, Desmond C. & Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages, Vol. II*. Berlin, New York: Mouton de Gruyter, 249–474
- Press, Ian (1986), *A Grammar of Modern Breton*. Berlin etc.: Mouton de Gruyter
- Qafisheh, Hamdi A. (1977), *A Short Reference Grammar of Gulf Arabic*. Tucson/AZ: University of Arizona Press
- Ross, John R. (1973). "Nouniness". In: Fujimura, Osamu (ed.), *Three Dimensions of Linguistic Theory*. Tokyo: TEC, 137–257

- Saltarelli, Mario (1988), *Basque*. London etc.: Croom Helm
- Sasse, Hans-Jürgen (1988), "Der irokesische Sprachtyp". *Zeitschrift für Sprachwissenschaft* 7, 173–213
- Schaub, Willi (1985), *Babungo*. London etc.: Croom Helm
- Strawson, Peter Frederick (1959), *Individuals*. London: Methuen

- Taylor, Charles (1985), *Nkore-Kiga*. London etc.: Croom Helm
- Underhill, Robert (1976), *Turkish Grammar*. Cambridge/MA: MIT Press
- Weigand, Hans (1990), *Linguistically Motivated Principles of Knowledge Based Systems*. Dordrecht: Foris

*J. Lachlan Mackenzie,  
Amsterdam (Netherlands)*

## 95. Deixis and reference

1. Introduction
2. The situation of speech
3. Discourse roles
4. Deictic dimensions
5. Deictic domains
6. Deictic morphological structures
7. Uncommon abbreviations
8. References

### 1. Introduction

In the field of deixis and reference there are various surveys (Bühler 1934; Fillmore 1975; 1997; Lyons 1977; Klein 1978; Karmiloff-Smith 1979; Tanz 1980; Levinson 1983; Rauh 1983; Geiger 1995; Cornish 1999; Lyons 1999; Brisard 2002), a number of crosslinguistic studies (Jarvelä & Klein 1982; Weissenborn & Klein 1982; Wiesemann 1986; Morel & Danon-Boileau 1992; Diessel 1999), various articles on the typological variety of deictic expressions (Heeschen 1982; Anderson & Keenan 1985; Feuillet 1992; Kryk-Kastovsky 1996 a; Himmelmann 1996; 1997), and some articles on different theoretical and thematical approaches to deixis (Fludernik 1991; Kirsner 1993; 1996; Duchan et al. 1995, eds.; Kryk-Kastovsky 1996 b).

This article will discuss deixis and reference, presenting a wide range of deictic types of expression with two or more formatives or with at least one bound formative (deictic formatives of verbs are discussed in Art. 77). The following sections will present central deictic notions relating to deictic morphological structures.

### 2. The situation of speech

**Deixis** is the linguistic procedure by means of which a speaker relates an entity to the cur-

rent speech situation in such a way that his addressee gains cognitive access to or mental contact with the entity concerned (Hanks 1989: 112; Langacker 1991: 91; for the notion of 'accessibility', see Givón 1984: 401–403; Ariel 1990).

When using the **pronouns** *I* and *you*, a speaker normally refers to him/herself and another person in such a way that he or she relates him/herself to the current situation of speech as the speaker, and the other person as the addressee. By using an expression such as *the x*, a speaker refers to a third party in such a way that he or she relates it uniquely to the current speech situation, while enabling her or his addressee to discern it as the entity referred to (Donnellan 1978; Hawkins 1984; Richter 1988: 133–147). The speech situation to which the entity at issue is related is constituted by what the speaker considers to be the current situation of speech, by what has been raised in the previous discourse, or by the communal background knowledge which the speaker assumes his addressee shares, possibly through inference (see, e.g., Du Bois 1980; Perner & Garnham 1988; Laury 1997: 34–51). In this article the term *deixis* is used in a broad sense. The relationship between deictics, indexicals, and shifters is analyzed by, e.g., Lyons (1977), Richter (1988), Rousseau (1992), Corblin (1992), and Nunberg (1993). The notion of 'grounding' is substituted by Langacker (1987; 1991; 2002) for a subdomain of the notion of 'deixis'.

### 3. Discourse roles

Expressions such as *I*, *you*, and *the x* are analyzed as definite descriptions. What makes

definite article	relational element	suffix	personal pronoun	
<i>t-</i>		<i>-en</i>	<i>ten</i>	1. SG
<i>t-</i>		<i>-óʔon</i>	<i>tóʔon</i>	1. PL
<i>t-</i>		<i>-óʔon (-éʔeš)</i>	<i>tóʔon (-éʔeš)</i>	1. PI
<i>t-</i>		<i>-eč</i>	<i>teč</i>	2. SG
<i>t-</i>		<i>-éʔeš</i>	<i>téʔeš</i>	2. PL
<i>le</i>	<i>ti?</i>		<i>le ti?</i>	3. SG
<i>le</i>	<i>ti?</i>	<i>- óʔob'</i>	<i>le ti?óʔob', le óʔob ti?'</i>	3. PL

Tab. 95.1: Lexical personal pronouns in Yucatec Maya

definite descriptions definite? The use of a definite description requires a cognitive **frame of reference** to which the entity at issue is conceivable as uniquely related (cf. Art. 76; see for the notion of ‘definiteness’, e.g., Hawkins 1984; Lambrecht 1994; Lyons 1999). What is the cognitive frame which enables a speaker to use expressions such as *I* and *you*? In order to answer this question, we have to examine the speech situation, which serves as the central frame of reference to which the entities referred to are related. The structure of the **speech situation** and its role for definite referencing can be elucidated by the personal pronoun system in Maya spoken in Yucatán.

In Yucatec Maya, a speaker and his addressee can be indicated by means of personal pronouns sharing the formative *t-*, whereas third person personal pronouns have the form *le* combined with the form *ti?* (see Tab. 95.1; Hanks 1990: 155f.).

The first and second person pronouns have the base *t-*, which can be analyzed as the formative of the deictic category of ‘core participants’ in the speech situation (see Hanks 1990: 141). Alongside the base *t-*, the grammatical status of the category of core participants can be encoded in various ways. For instance, first and second person pronouns differ from third person pronouns in some of their syntactic and morphological applications, as shown by the two patterns in Tab. 95.2 (Hanks 1990: 164f.).

Tab. 95.2 shows three variations. In the first case, the position of the first and second person pronouns differs from the position of the third person pronoun. In the second, the third person pronoun can be omitted. In the third, the dative prefix of the first and second person pronouns can be omitted. Additionally, only the plurality suffix of the third person pronoun is the same as the optional plu-

ral marker of common nouns (e.g. *le maákóʔob’* ‘the men’).

Godié also features a clear formal division between the group of the first and second person pronouns, which just differ in tone, and a coherent group of third person elements such as pronouns, determiners, and demonstratives (Marchese 1986: 218–220).

In accordance with their grammatical features, the lexical personal pronouns in Yucatec Maya can be analyzed in terms of their discourse-role system as shown in Fig. 95.1 (for related analyses, see Frei 1944; Benveniste 1956; Kuryłowicz 1972; Zwicky 1977; Bátori 1982; Beifuss et al. 1985; Hanks 1989; Nguyễn 1992; Smith 1992).

The deictic system of Fig. 95.1 features two subsystems. Subsystem I shows the relationship between the pronouns of the core participants in the situation of speech, i.e. the speaker and addressee with their possible associate(s). In the phrasing of Hanks (1990: 142; following Goffman 1979): “the speaker projects himself as a ‘figure’, a protagonist in a scene, story”; “the individual acting as a speaker, the one who performs the current act of reference, constitutes himself or herself as an object of description, as well as an actor engaged in executing a speech act” (for special references to core participants, see, e.g., Daalder 1986; Levinson 1988; Irvine 1996; Schegloff 1996). The relevant configuration of construals is illustrated in Fig. 95.2.

Here, region A of the speaker’s mental field of vision is the frame of reference to which the party referred to by means of *I* can be considered to be uniquely related. Region B is the frame of reference to which the party referred to by means of *you* can be considered to be uniquely related. Subsystem II of Fig. 95.1 shows the relationship between the

<i>yàan</i>	( <i>ti?</i> )-X	<i>çíimín</i>	x: a personal pronoun of 1. or 2. person 'I (etc.) have (a) horse(s)'
<i>yàan</i>	<i>çíimín</i>	<i>ti?</i> -(-X)	x: a personal pronoun of 3. person 'he (etc.) has (a) horse(s)'

Tab. 95.2: Grammatical differences between personal pronouns in Yucatec Maya

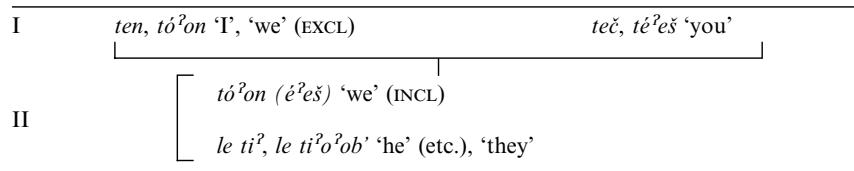
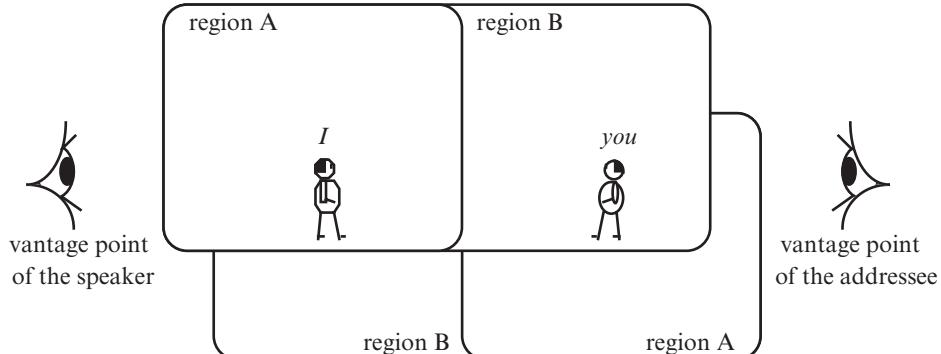


Fig. 95.1: The discourse-role system of the lexical personal pronouns in Yucatec Maya

Fig. 95.2: The cognitive frame of reference of the speech situation in which *I* and *you* are used

pronoun signalling the core participants and the pronouns for the third party. The pronoun *tó?oné?eš*, consisting of the formation *tó?on* 'we' and the suffixal pronoun *-é?eš* 'you', encodes the conjoint character of the discourse role played by speaker and addressee with their possible associate(s) (such a fusion is observed by Jacobsen 1980; for Wargamay, see Dixon 1981: 39; for Yag Dii, Bohnhoff 1986: 104 ff.). It is this role which the speaker assumes is shared by the addressee (Donnellan 1978) and it is in this role that they share a common vantage point (or indexical ground, Hanks 1990: 38) for referencing by means of third-party definite descriptions. Since it is the speaker who assumes the addressee shares his or her vantage point, the common vantage point is often qualified as basically **egocentric** (Hawkins

1984). However, the speaker can be assumed to be cooperative (Grice 1975) and to operate from a recipient design (or even from a reflector character's point of view, see, e.g., Fludernik 1995; Epstein 1996; Grenoble 1998: 29–32, 62–64). Then, the common vantage point can be termed **speaker-hearer centric** (Ebert 1987: 477; see also Roberts 1986: 446) or **sociocentric** (Hanks 1989; 1990: 7, 41–43; Jones 1995: 47).

The interrelation between the internal and external relationships of subsystems I and II can be explained by Fig. 95.3 (Janssen 1995 a; 1995 b); it shows two discourse-role perspectives corresponding to subsystems I and II from Fig. 95.1.

In the 1/2-perspective the speaker sees himself with his possible associate(s) in a region or zone of his mental field of vision dif-

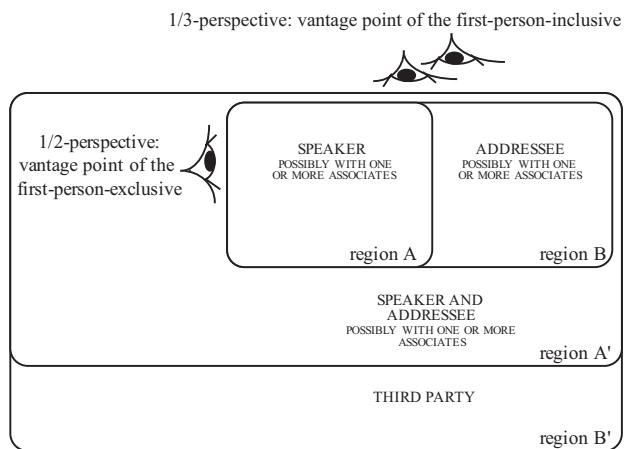


Fig. 95.3: The situation of speech with two discourse-role perspectives

ferent to the one in which he sees his addressee with his possible associate(s). By using **definite descriptions**, the speaker indicates that he divides his central frame of reference, i.e. the situation of speech, into subframes. When using the personal pronoun of the **first person exclusive**, the speaker relates him/herself with his or her possible associate(s) as unique to one of the regions of the deictic dimension (see section 4) which constitute the domain (see section 5) of the core participants of the discourse. When using the **second person pronoun**, the speaker relates the addressee with his or her possible associate(s) as unique to the other region in the dimension of the core-participant domain. Thus, each of the regions serves as a frame of reference to which an entity is uniquely related. It is the entity's unique relationship to one particular region of the mental field of vision that forms the basis for the definiteness of personal pronouns and expressions of the type *the x* is based (Janssen 1995 a; 1995 b).

In the 1/3-perspective the speaker sees him/herself and his or her addressee with their

possible associate(s) in the same region of his or her mental field of vision, and he or she assumes that they share vantage point and mental field of vision. The speaker sees this conjoint party in a region different from the one in which he or she sees the third party. The pronoun for the **first person inclusive** serves to indicate that the conjoint party is uniquely related to one region of the mental field of vision in the 1/3-perspective, and the **third person pronoun** serves to indicate that the third party is uniquely related to the other region.

In a number of languages articles, demonstratives, third person personal pronouns, and anaphors have an overt grammatical relationship (cf. Lyons 1977: 646 f.; Givón 1984: 226 f.; Anderson & Keenan 1985: 260 f.; Croft 1990: 219 f.; proper names or 'proper nominals', e.g. *The Netherlands*, *The Appalachians* can also be classified as third person deictics, cf. Jonasson 1992; Wiederspiel 1992). Consider, for instance, the third person forms in Tab. 95.3 (Hanks 1990: 18 f.).

initial deictic	terminal deictic	definite description	
<i>le</i>		<i>le ti?</i>	'he, she, it, the one'
<i>le</i>		<i>le ti?ó?ob'</i> , <i>le ó?ob' ti?</i>	'they, the ones'
<i>le</i>		<i>le maák</i>	'the man'
<i>le</i>		<i>le maákó?ob'</i>	'the men'
<i>le</i>	<i>a?</i>	<i>le maák a?</i>	'this man'
<i>le</i>	<i>o?</i>	<i>le maák o?</i>	'that man'
<i>le</i>	<i>o?</i>	<i>le maákó?ob'o?</i>	'those men'

Tab. 95.3: Third person forms in Yucatec Maya

In Tab. 95.3 the prenominal article *le* marks all third person descriptions as definite; this implies that they form one natural class in Yucatec Maya. The deictics *a'* and *o'* serve to discern referentially the entities involved. See also the morphological interrelatedness of third person forms (pronouns, anaphoric and cataphoric demonstratives) in Asheninca (Reed & Payne 1986: 325, 330). In Godié third person personal pronouns, determiners, and class-concord suffixes are the same, while demonstratives are morphologically based on these forms (Marchese 1986: 218–220). In Lyélé suffixal definite markers have the same shape as third person object pronouns, but carry a high tone (Sho-walter 1986: 209). In Diyari third person personal pronouns/determiners (see *nhani* in Tab. 95.4), and local adverbs (see *nhingki* in Tab. 95.4) form a morphologically related group, as apparent from the data in Tab. 95.4 (Austin 1982: 274f., 281f.).

The third person system of Wolof is even more extensive, as shown in Tab. 95.5 (Wills 1990; Sauvageot 1992).

base	suffix		
	proximate	proximate	distal
	visible	invisible	
	- <i>rda</i>	- <i>ya</i>	- <i>wa</i>
pronoun			
<i>nhani</i>	<i>nhanirda</i>	<i>nhaniya</i>	<i>nhaniwa</i>
locative			
<i>nhingki</i>	<i>nhingkirda</i>	<i>nhingkiya</i>	<i>nhingkiwa</i>

Tab. 95.4: Morphologically related third person forms in Diyari

Tab. 95.5 shows two types of definite article concurring with two types of gestural and two types of ‘plain’ demonstrative, while the latter two types share a pronominal use. The system of local adverbs also shares the same structure. All the forms have the proximate suffix *-i* and the distal suffix *-a* (allomorph *-e*), with corresponding extensions. The two-fold applicability of the suffix system of Tab. 95.5 suggests not only that categories 1–3 and 4–6 are natural classes, but also that these categories together form a natural class in Wolof. This suggests that entities referred to by means of local adverbs can also be considered third-party entities. This point can be extended to include other types of adverb, such as temporal adverbs (for Hopi, see Malotki 1979: 24f.; for Paumari, Chapman & Derbyshire 1991: 261), and even tensed clauses (Janssen 1991; 1996; 2002; Gildea 1993).

The functional relationship between personal pronouns and demonstratives is morphologically encoded in, e.g., Berik. Here, the form *ai(serem)* serves as the first person singular subject pronoun and the demonstrative ‘this’, while the form *je(serem)* serves as the third person singular subject pronoun and the demonstrative ‘that’ (Westrum & Wiesemann 1986: 43).

#### 4. Deictic dimensions

In many languages the category of deictic forms applicable to third-party entities can be divided into a great variety of subcategories. A regular morphological system of deictic distinctions (with the exception of the form *asoko*) can be found in Japanese, as shown in

base	suffix			
	proximate		distal	
C-	- <i>i</i>	<i>bi</i>	- <i>a</i>	<i>ba</i>
C-	- <i>ii</i>	<i>bii</i>	- <i>ee</i>	<i>bee</i>
C-	- <i>ile</i>	<i>bile</i>	- <i>ale</i>	<i>bale</i>
C- <sup>†</sup>	- <i>ile</i>	<i>kile</i>	- <i>ale</i>	<i>kale</i>
f-	- <i>i</i>	<i>fi</i>	- <i>a</i>	<i>fa</i>
f-	- <i>ii</i>	<i>fii</i>	- <i>ee</i>	<i>fee</i>
f-	- <i>ile</i>	<i>file</i>	- <i>ale</i>	<i>fale</i>
C-:	{ <i>b-</i> , <i>m-</i> , <i>j-</i> , <i>s-</i> , <i>w-</i> , <i>g-</i> , <i>l-</i> , <i>k-</i> }SG, { <i>ñ-</i> , <i>y-</i> }PL			
C- <sup>†</sup> :	{ <i>k-</i> }SG.HUM			

Tab. 95.5: Third person forms in Wolof

base			suffix	
<i>ko-</i>	<i>so-</i>	<i>a-</i>		
1. <i>kono</i> ‘this’	<i>sono</i> ‘that’	<i>ano</i> ‘that’	<i>-no</i>	determiner
2. <i>kore</i> ‘this one’	<i>sore</i> ‘that one’	<i>are</i> ‘that one’	<i>-re(ra)</i>	pronominal INAN
3. <i>koitsu</i> ‘this one, it’	<i>soitsu</i> ‘that one, it’	<i>aitsu</i> ‘that one, it’	<i>-tsu(ra)</i>	pronominal (IN)AN

Tab. 95.6: Deictic dimensions in Japanese

Tab. 95.6 (Coulmas 1982; Kuroda 1985; Tamba 1992).

Each of the rows (1–3) in Tab. 95.6 presents a deictic dimension with three regions or zones (Hanks 1990: 67f., 77f.) signalled by the bases *ko-* (the region closely related to the speaker), *so-* (the region closely related to the addressee), and *a-* (the region not closely related to either speaker or addressee); each region is instrumental to the addressee in gaining access to the entity referred to (see 2; Hanks 1990: 77).

The threefold structure of the deictic dimensions presented by the rows in Tab. 95.6 can be analyzed as “un double système binaire, à quatre valeurs définies par le jeu de deux couplages: celui des formes en *KO* et *A* d’une part; et des formes en *KO* et *so* d’autre part” (Tamba 1992: 191; also Fludernik 1991: 216; for a different view, see Aoyama 1995; see Kuroda 1985 for the anaphoric use of elements based on *so-*; in various languages demonstratives related to the second person also can be used anaphorically; for Mundani, see Parker 1986: 146; for Finnish, Laury 1997). The structure of the “double système binaire” can be conceived of as determined by two discourse-role perspectives, as shown in Fig. 95.4 (Janssen 1995 a; 1995 b).

Apart from Japanese, a number of languages (e.g. Classical Greek, Latin, Italian, Spanish, and Vietnamese, see Nguyễn 1992; and Anderson & Keenan 1985: 282–285) share an overtly morphological system of three demonstratives which can be considered as determined by the double perspective shown in Fig. 95.4. Both perspectives can be applied to a single utterance, such as in sentence (1) from Vietnamese, where *đây* ‘place proximal to the addressee’ is a form of the 1/2-perspective and *kia* is a form of the 1/3-perspective (Nguyễn 1992: 180).

(1) (a) *đây là bà A.*;  
there be Mrs. A.  
(near you)

(b) *kia là bà B.*  
there be Mrs. B.  
(not near me & you)

Fig. 95.4 has to be seen as integral to Fig. 95.3, since the demonstratives are part of the third person region in the domain of the **discourse roles**. Thus, the combination of the 1/2 and 1/3-perspective is a recursive cognitive schema. One could even go further and ask if each of the two perspectival subsystems is an instantiation of a single cognitive schema (Janssen 1995 a). Let us consider this issue regarding demonstrative systems with only two morphological forms to indicate the regions of the deictic dimension involved. In English, for instance, the demonstratives *this* and *that* can have two interpretations; witness texts (2)–(4).

- (2) *I can see everything from this window. And you? Do you have a good view from that window?*
- (3) *Do you also find it so awfully hot in this room?*
- (4) *Did you also find it so awfully hot in that room?*

In text (2) it is most likely that by *this window*, the speaker refers to the window he is looking through, and by *that window*, he refers to the window his addressee is looking through. In this interpretation, the use of *this* and *that* is determined by the 1/2-perspective. In texts (3) and (4) the speaker and his addressee may be in the same place, referred to by means of *this room*. Speaker’s use of *this*

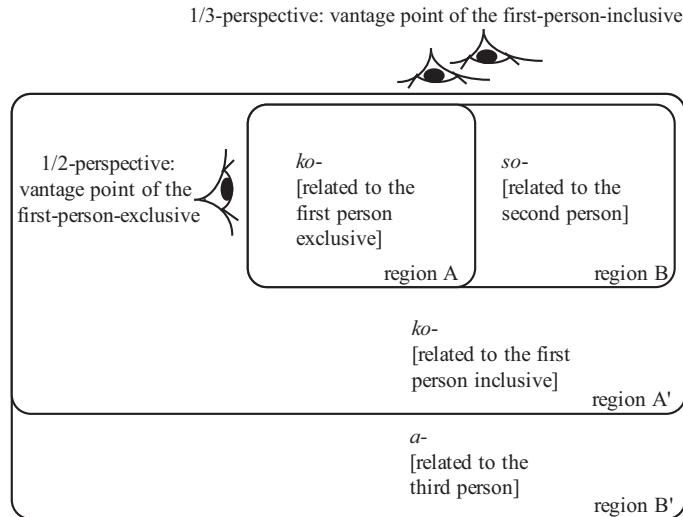


Fig. 95.4: The threefold deictic dimension of Japanese demonstratives

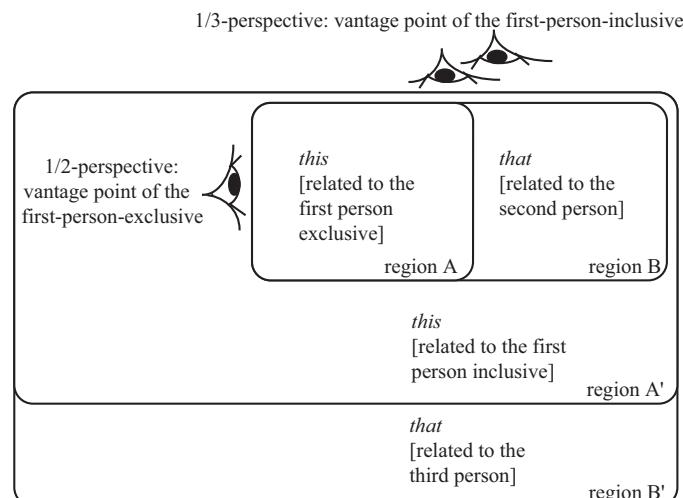


Fig. 95.5: Two possible interpretations of English demonstratives

in text (3) is positively related to the position shared by him and his addressee, while his use of *that* in text (4) is negatively related to their current position. Thus, in texts (3) and (4) both *this* and *that* are based on the 1/3-perspective. Therefore, both *this* and *that* can have two interpretations, one determined by the 1/2 and the other by the 1/3-perspective. The relevant possibilities for interpretation are shown in Fig. 95.5.

The systems of both Fig. 95.4 and 95.5 are elaborative extensions of the third person cat-

egory as displayed in Fig. 95.3. The extension hypothesis explains why a number of languages have an overtly morphological interrelation between definite articles and demonstratives (adnominal, pronominal, and local) and also between categories, such as third person personal pronouns and non-demonstrative local adverbs, as shown in Tab. 95.3–95.5. It is entirely in accordance with this view that the Maya system of discourse roles shown in Fig. 95.1 can be extended to the recursive system in Fig. 95.6.

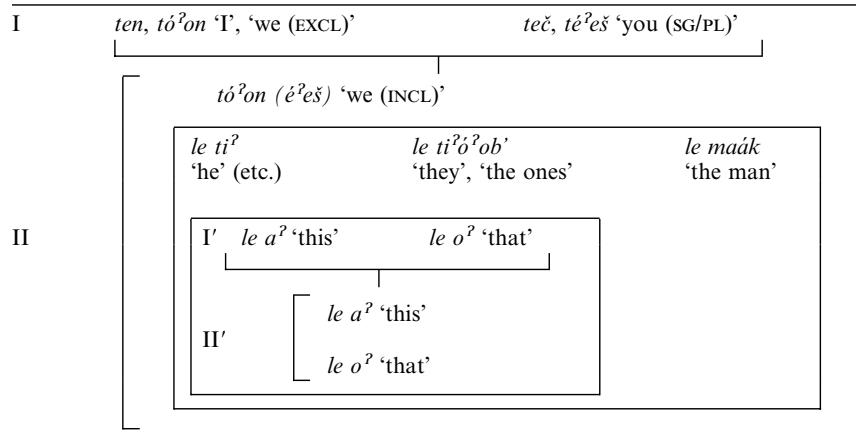


Fig. 95.6: The Yucatec Maya discourse-role system with a recursive extension

The extension is in the inner box; its content has the same structure as the overall schema for subsystems I and II; subsystems I' and II' show the demonstratives in the 1/2 and 1/3-perspectives, respectively. The outer and inner boxes of Fig. 95.6 contain a variety of forms, all with the (p)article *le*. A binary system of demonstratives which share the article-formative can be found in Tahitian, as shown by Fig. 95.7 (Lazard & Peltzer 1992: 210–211).

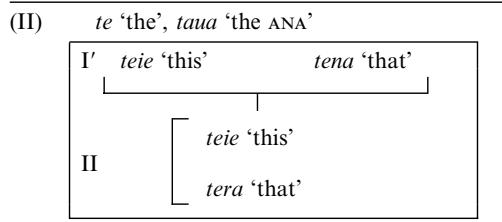


Fig. 95.7: Tahitian third person forms

Tahitian has a separate anaphoric form *taua*, generally combined with the deictic particle *ra* (e.g. *taua ta'ata ra* 'the person at issue', cf. Lazard & Peltzer 1992: 211; anaphora can be considered a type of third-party deixis, cf. Zribi-Hertz 1992; Kleiber 1992).

The strongest evidence for the idea that subsystems I' and II' are recursive applications of a single basic cognitive schema can be found in Palauan, which has a binary third person system with distinct morphological suffixes for its functions, i.e. for both the demonstrative related to the addressee and for the demonstrative related to the

speaker and addressee with their possible associate(s). The data are shown in Tab. 95.7 (Josephs 1978: 344, 360).

When the Palauan demonstratives are divided into subsystems according to two discourse-role perspectives, an interesting analogy emerges, as shown in Fig. 95.8.

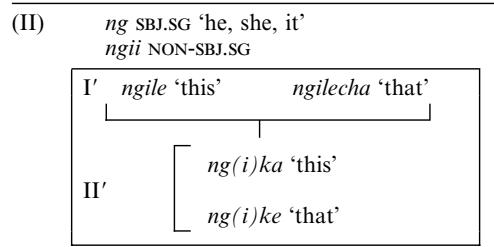


Fig. 95.8: Deictic structure of Palauan third person forms

The suffixes of subsystem I' share the string *-ile-*, and the suffixes of subsystem II' share the string *-(i)k-*. If these strings can be considered to have morphological status, the morphology of each subsystem is marked in a distinctive way. In fact, the suffixes for the counterparts of non-human entities correspond to a large extent to the forms of Tab. 95.7 (Josephs 1978: 360).

The deictic system of Hausa adverbials features an interesting variation on the recursivity of the basic cognitive schema. The relevant data are shown in Tab. 95.8 (Jaggar & Buba 1995: 413).

The deictic adverbials of Hausa have "a tone-meaning interaction whereby F[alling]

base	suffix							
	subject		non-subject		demonstrative			
			1 EXCL	2	-ile	-ilecha	1 INCL	3
<i>ng</i>	<i>ng</i>	<i>ngii</i>		<i>ngile</i>	<i>ngilecha</i>	<i>ng(i)ka</i>	<i>ng(i)ke</i>	SG.HUM
<i>t(ir)</i>	<i>te</i>	<i>tir</i>		<i>tirile</i>	<i>tirilecha</i>	<i>tirka</i>	<i>tirke</i>	PL.HUM

Tab. 95.7: Palauan third person forms

base	superfix				
	fall tone		high tone		
<i>nan</i>	<i>nân</i>	speaker-proximal	<i>nán</i>	hearer-proximal	
<i>can</i>	<i>cân</i>	speaker/hearer-distal	<i>cán</i>	hearer/proximal speaker/hearer-remote	

Tab. 95.8: Deictic adverbials ('here', 'there') in Hausa

tone *nân/cân* signals (relative) proximity to the speaker, and H[igh] *nán/cán* tone maps distance from the speaker, in both spatial and non-spatial (anaphoric, temporal, cognitive) terms" (Jaggar & Buba 1995: 391). The four-fold system can be analyzed as in Fig. 95.9.

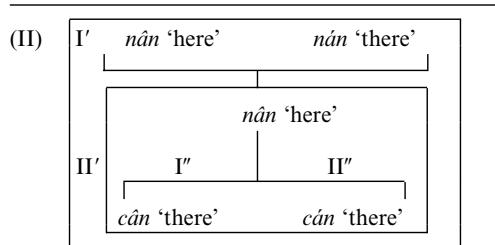


Fig. 95.9: The fourfold system of Hausa deictic adverbials

*Can*-forms refer to entities in the composite third-party region, which is opposed to the speaker/hearer-proximal region. The *can*-subsystem consists of an opposition between a speaker/hearer-distal region and a region which is more remote-distal to speaker and hearer.

Fourfold demonstrative systems are often described as constituting one linear dimension with a range of regions progressively distant from the speaker (exclusive and inclusive, if applicable) and the addressee. Further research is required to determine whether distance can have a cognitive effect as a result

of more basic relations such as those represented by the systems based on the 1/2 and 1/3-perspectives (cf. Anderson & Keenan 1985: 282–285). The feature of distance relative to the speaker is required for deictic dimensions with more than three or four distinct forms (Anderson & Keenan 1985: 286–288) such as, for instance, in Malagasy (Anderson & Keenan 1985: 292 f.); the relevant forms will be shown in the following section.

## 5. Deictic domains

In many languages, the category of deictic forms which can be applied to third-party entities has a great variety of subcategories, as shown in Tab. 95.3, which presents a number of third person forms in Yucatec Maya, and Tab. 95.5, which presents a number of third person forms in Wolof. The Wolof system of third person forms is even more extensive; a few other domains are shown in Tab. 95.9.

Alongside the deictic domains shown in Tab. 95.4 and 95.8, Wolof also has domains for time, manner, and degree, which have the same morphology as these tables, to a large extent (Wills 1990: 198 f.).

Two other extensive and highly regular morphological systems will be discussed here. Consider, first, the demonstrative system of Malagasy presented in Tab. 95.10 (Anderson & Keenan 1985: 292–294; the forms between parentheses have the specialized meaning 'inside of').

base classifier	suffix		anaphora anaphoric local adverb
	proximate	distal	
	-oo...-u	-oo...-a	
C-	<i>boobu</i>	<i>booba</i>	
f-	<i>foofu</i>	<i>foofa</i>	
C-: {b-, m-, j-, s-, w-, g-, l-, k-}SG, {ñ-, y-}PL			

Tab. 95.9: Anaphora in Wolof

prefix	base							
	← closer to speaker				further from speaker →			
	-ty	-to	-o	-tsy	-ny	-roa	-ry	
i-	<i>ity</i>		<i>io</i>	<i>itsy</i>	<i>iny</i>	<i>iroa</i>	<i>iry</i>	DEM (pronominal/ adnominal)
e-	<i>ety</i>	<i>eto</i>	(eo)	<i>etsy</i>	<i>eny</i>	<i>eroa</i>	<i>ery</i>	DEM.LOC.VIS (adverb)
a-	<i>aty</i>	<i>ato</i>	(ao)	<i>atsy</i>	<i>any</i>	<i>aroa</i>	<i>ary</i>	DEM.LOC.INVIS (adverb)

Tab. 95.10: Malagasy demonstrative forms

base	suffix			
	<i>ko-</i>	<i>so-</i>	<i>a-</i>	
<i>koko</i> 'this place'	<i>soko</i> 'that place'	<i>asoko</i> 'that place'	<i>-ko(ra)</i>	noun (local)
<i>kochi</i> 'this direction'	<i>sochi</i> 'that direction'	<i>achi</i> 'that direction'	<i>-ci(ra)</i>	noun (directional)
<i>kô</i> 'this manner'	<i>sô</i> 'that manner'	<i>â</i> 'that manner'	<i>-ôlâ(iu)</i>	modifier (manner)
<i>konna</i> 'such, so many'	<i>sonna</i> 'such' (etc.)	<i>anna</i> 'such' (etc.)	<i>-nna(ni)</i>	modifier (quality/quantity)

Tab. 95.11: Deictic domains in Japanese

The Malagasy system is even more extensive than shown in Tab. 95.10, since the local adverb series, both the one with the feature 'VISIBLE' ('VIS') and the one with the feature 'INVISIBLE' ('INVIS'), can be prefixed by tense forms; the form *t-* signals that the entity modified by the local adverb was at the distance which was relevant in the past; the form *ho-* signals a similar modification in the future (Anderson & Keenan 1985: 293). Furthermore, the *i*-forms of the Malagasy system can have the infix *-za-*, for instance, in *izao* and *izany*; this form signals that the entity referred to is not visible in the situation of speech (Anderson & Keenan 1985: 294).

Japanese is another language with a great variety of deictic domains signalled by a highly regular morphological system. This has partly been shown by Tab. 95.6; see also Tab. 95.11 (Coulmas 1982; Tamba 1992).

Each row of Tab. 95.11 shows a deictic domain made up of a deictic dimension with three regions signalled by the bases *ko-*, *so-*, and *a-*. The forms *kochira* and *sochira* can serve to indicate the speaker and hearer (Coulmas 1982: 213 f.; Tamba 1992: 191; for Gumbaynggir, Eades 1979: 290; for Chinese, Paris 1992: 173; for Vietnamese, Nguyễn 1992: 181). One can assume a metonymic relationship between a place and the central

person at that place; in many languages, however, demonstratives can be analyzed more insightfully as non-local (see, e.g., Lakoff 1974; Kirsner 1979; Janssen 1993; Strauss 1993; Laury 1996; Agha 1996). The form *sochira* can be suffixed by the **honorific formative** *-sama*, which signals respect to the addressee.

A few more types of deictic domain should be mentioned here. There is a wide range of spatially differentiated deictic domains (cf. Denny 1978; Malotki 1979; Heeschen 1982; Mosel 1982; Anderson & Keenan 1985: 295; Feuillet 1992: 242 f.). With regard to the domain of **honorifics**, East Asian languages, in particular, have encoded very richly differentiated systems (cf. Anderson & Keenan 1985: 270–277; Harada 1976). Finally, in some languages (e.g. Hausa, see Jaggar & Buba 1995), demonstratives can be used both spatially and non-spatially, including temporally. The data of Wik-Munkan (Anderson & Keenan 1985: 298) are shown in Tab. 95.12.

base	suffix	
	<i>-man</i>	<i>-pal</i>
<i>in-</i>	<i>inman</i> ‘right here’	<i>inpal</i> ‘from here’
	‘right now’	‘from now’
<i>an-</i>	<i>anman</i> ‘around there’	<i>anpal</i> ‘from there (distant)’
	‘around now’	‘from then (on)’
<i>nan-</i>	<i>nanman</i> ‘there (close)’	<i>nanpal</i> ‘from there (near)’
	‘now (general)’	‘from then (recent)’

Tab. 95.12: Wik-Munkan demonstratives

In languages such as Vietnamese (Nguyễn 1992: 186), Guugu Yimidhirr (Haviland 1979: 72 f.), Hausa (Jaggar & Buba 1995), Mohawk (Bonvillain 1981), and Paumarí (Chapman & Derbyshire 1991: 261), a series of deictic forms can have both a local and a temporal interpretation under certain conditions.

## 6. Deictic morphological structures

In Tab. 95.1 we saw that not only a base and a number of affixes can be deictic but also in a string of two affixes, both can be deictic (e.g. *-ō'ōn* + *é'eš*). In most deictic morphological structures, however, only one of the formatives is deictic. The affix formation

types will be divided into prefix, infix, suffix, superfix, reduplication types (Bauer 1988: 19–29), and vowel lengthening (cf. Ch. VIII).

**Deictic bases.** In Yucatec Maya, the category of lexical first and second person pronouns is formed by the deictic base *t-* and deictic suffixes (Tab. 95.1). Japanese has a threefold system of deictic bases (Tab. 95.6 and 95.11).

**Deictic prefixes.** In Kannada, both singular and plural third person personal pronouns have the prefixes *i-* and *a-*, to indicate proximity and distance (Sridhar 1990: 203, 206). Alongside lexical personal pronouns, Yucatec Maya has a set of affixal ones which can have verbs and nouns as bases; if a noun is the base, the affix is interpreted as referring to the possessor. The affixal system is shown in Tab. 95.13 (Hanks 1990: 156); since the system basically consists of prefixes and suffixes, the combination of a prefix and a suffix should not be analyzed as circumfixes.

prefix	prefix + suffix		person	
	sing	plur	sing	plur
<i>in-</i>	<i>k-</i>	<i>k...é'eš</i>	<i>-en</i>	<i>-ō'ōn</i> 1
? <i>a-</i>	? <i>a...é'eš</i>		? <i>eč</i>	? <i>eš</i> 2
? <i>u-</i>	? <i>u... ó'ob'</i>			? <i>ob'</i> 3

Tab. 95.13: The affixal personal pronoun system in Yucatec Maya

**Deictic infixes.** Local adverbs in Tolai can have the infix *-ba-*, marking them as anaphoric; the infix *-ka-*, signalling ‘further on’; or the infixes can be combined (*-ba* + *ka*) (Mosel 1982: 127, 129).

**Deictic suffixes.** Diyari (see Tab. 95.4) has a threefold deictic suffix system which can mark both pronouns and locatives. In Wolof (see Tab. 95.5) the twofold deictic suffix system consisting of *-i* and *-a* produces definite articles and non-demonstrative local adverbs; both suffixes can be augmented by two other suffixes to form demonstratives, personal pronouns, and demonstrative local adverbs.

**Deictic superfixes.** Hausa (see Tab. 95.8) has pairs of contrastive tone suprasegments marking “relative proximity to the speaker” versus “distance from the speaker” (Jaggar & Buba 1995). In Vietnamese a tone distinction is made between the forms signalling proximity to the speaker (*dày*) and proximity to the addressee (*đáy*) (see (1)); the tone contrast is

indicated by the absence or presence of the diacritical mark “’” (Nguyêñ 1992: 178). In Lyélé tone distinguishes third person object pronouns and definite markers (Showalter 1986: 209). Tone distinguishes first and second person pronouns in Godié. In the phrasing of Marchese (1986: 218–220): “In the singular, first person has a higher tone than the second person, while in the plural, the reverse is true”.

**Deictic reduplication.** Wolof (see Tab. 95.9) has a class of anaphors which is morphologically marked by reduplication (Wills 1990; Sauvageot 1992). The anaphoric forms in Comoro are also reduplicated (Picabia 1992: 160–165). In Godié the independent demonstratives, both the singular and plural forms, are reduplicated (Marchese 1986: 219–220, 233). In Hopi the plural personal pronouns can have an initial reduplication which serves to signal distributivity (Malotki 1982: 228 f.; 245–247).

**Vowel lengthening.** The gestural demonstratives in Wolof have lengthened vowels at the end of words as shown in Tab. 95.5, rows 2 and 5 (Sauvageot 1992: 153).

## 7. Uncommon abbreviations

EXCL	exclusive
INCL	inclusive

## 8. References

- Agha, Asif (1996), “Schema and Superposition in Spatial Deixis”. *Anthropological Linguistics* 38, 643–682
- Anderson, Stephen R. & Keenan, Edward L. (1985), “Deixis”. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*. Cambridge etc.: Cambridge Univ. Press, 259–308
- Aoyama, Takashi (1995), “Deixis and Value: A Semantic Analysis of the Japanese Demonstratives”. In: Contini-Morava & Goldberg (eds.), 289–320
- Ariel, Mira (1990), *Accessing Noun-Phrase Antecedents*. London: Routledge
- Austin, Peter (1982), “The Deictic System of Diyari”. In: Weissenborn & Klein (eds.), 273–284
- Bátori, István (1982), “On Verb Deixis in Hungarian”. In: Weissenborn & Klein (eds.), 155–165
- Bauer, Laurie (1988), *Introducing Linguistic Morphology*. Edinburgh: Edinburgh Univ. Press
- Beifuss, Karin & Czepluch, Hartmut & Tuschinsky, Joachim & Schmid, Wolfgang P. (1985), “Zur Klassifikation pragmatischer Elemente in der grammatischen Beschreibung”. *Sprachwissenschaft* 10, 81–106
- Benveniste, Emile (1956), “La nature des pronoms”. In: Halle, Morris et al. (eds.), *For Roman Jakobson*. The Hague: Mouton, 34–37 [reprinted in: Benveniste, Emile (1966), *Problèmes de linguistique générale, Vol. I*. Paris: Gallimard, 251–57]
- Bohnhoff, Lee E. (1986), “Yag Dii (Duru) Pronouns”. In: Wiesemann (ed.), 103–129
- Bonvillain, Nancy (1981), “Locative Semantics in Mohawk: Time and Space”. *International Journal of American Linguistics* 47, 58–65
- Brisard, Frank (2002, ed.), *Grounding: The Epistemic Footing of Deixis and Reference*. Berlin etc.: Mouton de Gruyter
- Bühler, Karl (1934), *Sprachtheorie: Die Darstellungsfunction der Sprache*. Jena: Fischer [reprinted (1982), Stuttgart etc.: Fischer; translated by Donald F. Goodwin (1990), *Theory of Language. The Representational Function of Language*. Amsterdam: Benjamins]
- Chapman, Shirley & Derbyshire, Desmond C. (1991), “Paumari”. In: Derbyshire, Desmond C. & Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages, Vol. III*. Berlin etc.: Mouton de Gruyter, 161–352
- Contini-Morava, Ellen & Goldberg, Barbara S. with the assistance of Robert S. Kirsner (1994, eds.) *Meaning as Explanation: Advances in Linguistic Sign Theory*. Berlin etc.: Mouton de Gruyter
- Corblin, Francis (1992), “Démonstratif et nomination”. In: Morel & Danon-Boileau (eds.), 439–456
- Cornish, Francis (1999), *Anaphora, Discourse, and Understanding: Evidence from English and French*. Oxford: Clarendon Press
- Coulmas, Florian (1982), “Some Remarks on Japanese Deictics”. In: Weissenborn & Klein (eds.), 209–221
- Croft, William (1990), *Typology and Universals*, Cambridge: Cambridge Univ. Press
- Daalder, Saskia (1986), *First-person deixis*, Duisburg: LAUD
- Denny, J. Peter (1978), “Locating the Universals in Lexical Systems for Spatial Deixis”. In: Farkas, Donka & Jacobsen, Wesley M. & Todrys, Karol W. (eds.), *Papers from the Parasession on the Lexicon. Chicago Linguistic Society*. Chicago/IL: Chicago Linguistic Society, 71–84

- Diessel, Holger (1999), *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam etc.: Benjamins
- Dixon, Robert M. W. & Blake, Barry J. (1979–1991, eds.), *Handbook of Australian Languages, Vol. 1–4*. Amsterdam: Benjamins
- Dixon, Robert M. W. (1981), "Wargamay". In: Dixon & Blake (1981, eds.), xxiv, 1–144
- Donnellan, Keith S. (1978), "Speaker References, Descriptions and Anaphora". In: Cole, Peter (ed.). *Pragmatics*. New York etc.: Academic Press, 47–68
- Du Bois, John W. (1980), "Beyond Definiteness: The Trace of Identity in Discourse". In: Chafe, Wallace L. (ed.), *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood/NJ: Ablex, 203–274
- Duchan, Judith F. & Bruder, Gail A. & Hewitt, Lynne E. (1995, eds.), *Deixis in Narrative: A Cognitive Science Perspective*. Hillsdale/NJ: Lawrence Erlbaum Associates
- Eades, Diana (1979), "Gumbaynggir". In: Dixon & Blake (eds.), 244–361
- Ebert, Karen H. (1987), "Grammatical Marking of Speech Act Participants in Tibeto-Burman". *Journal of Pragmatics* 11, 473–482
- Epstein, Richard (1996), "Viewpoint and the Definite Article". In: Goldberg (ed.), 99–112
- Feuillet, Jack (1992), "La structuration de la deixis spatiale". In: Morel & Danon-Boileau (eds.), 233–243
- Fillmore, Charles J. (1975), *Santa Cruz Lectures on Deixis*. Bloomington/IN: Indiana Univ. Linguistics Club
- Fillmore, Charles J. (1997), *Lectures on Deixis*. Cambridge: Cambridge Univ. Press
- Fludernik, Monika (1991), "Shifters and deixis: Some Reflections on Jakobson, Jespersen, and reference". *Semiotica* 86, 193–230
- Fludernik, Monika (1995), "Pronouns of Address and 'Odd' Third Person Forms: The Mechanics of Involvement in Fiction". In: Green, Keith (ed.), 99–129
- Fox, Barbara (1996, ed.), *Studies in Anaphora*. Amsterdam etc.: Benjamins
- Frei, Henri (1944), "Systèmes de déictiques". *Acta Linguistica* 4, 111–129
- Geiger, Richard A. (1995, ed.), *Reference in Multidisciplinary Perspective: Philosophical Object, Cognitive Subject, Intersubjective Process*. Hildesheim etc.: Georg Olms Verlag
- Geiger, Richard A. & Rudzka-Ostyn, Brygida (1993, eds.), *Conceptualisations and Mental Processing in Language*. Berlin etc.: Mouton de Gruyter
- Gildea, Spike (1993), "The Development of Tense Markers from Demonstrative Pronouns in Panare (Cariban)". *Studies in Language* 17, 53–73
- Givón, Talmy (1984), *Syntax: A Functional-Typological Introduction, Vol. 1–2*. Amsterdam etc.: Benjamins
- Goffman, Erving (1979), "Footing". *Semiotica* 25, 1–29 [reprinted in: *Forms of Talk*, Philadelphia: Univ. of Pennsylvania Press 1981, 124–159]
- Goldberg, Adele E. (1996, ed.), *Conceptual Structure, Discourse and Language*. Stanford/CA: CSLI Publications
- Green, Keith (1995, ed.), *New Essays in Deixis*. Amsterdam, Atlanta: Rodopi
- Grenoble, Lenore A. (1998), *Deixis and Information Packaging in Russian Discourse*. Amsterdam etc.: Benjamins
- Grice, H. Paul (1975), "Logic and Conversation". In: Cole, Peter & Morgan, Jerry L. (eds.), *Speech Acts*. New York etc.: Academic Press, 41–58
- Hanks, William F. (1989), "The Indexical Ground of Deictic Reference". In: Music, Bradley & Graczyk, Randolph & Wiltshire, Caroline (eds.), *Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society. Part Two: Parasession on Language in Context*, 104–122 [reprinted in: Duranti, Alessandro & Goodwin, Charles (1992, eds.), *Rethinking Context. Language as an Interactive Phenomenon*. Cambridge: Cambridge Univ. Press, 43–76]
- Hanks, William F. (1990), *Referential Practice: Language and Lived Space among the Maya*. Chicago etc.: The Univ. of Chicago Press
- Harada, Sin I. (1976), "Honorifics". In: Shibatani, Masayoshi (ed.), *Japanese Generative Grammar*, New York etc.: Academic Press, 499–561
- Haviland, John (1979), "Guugu Yimidhirr". In: Dixon & Blake (eds.), 26–180
- Hawkins, John A. (1984), "A Note on Referent Identifiability and Co-presence". *Journal of Pragmatics* 8, 649–659.
- Heesch, Volker (1982), "Some Systems of Spatial Deixis in Papuan Languages". In: Weissenborn & Klein (eds.), 81–109
- Himmelmann, Nikolaus P. (1996), "Demonstratives in Narrative Discourse: A Taxonomy of Universal Uses". In: Fox (ed.), 205–254
- Himmelmann, Nikolaus P. (1997), *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer

- Irvine, Judith T. (1996), "Shadow Conversations: The Indeterminacy of Participant Roles". In: Silverstein, Michael & Urban, Greg (eds.), *Natural History of Discourse*. Chicago, London: Univ. of Chicago Press, 131–159
- Jacobsen, W[illiam] H. (1980), "Inclusive/Exclusive: A Diffused Pronominal Category in Native Western North America". In: Kreiman, Jody & Ojeda, Almerindo E. (eds.), *Papers from the Parasession on Pronouns and Anaphora*. Chicago Linguistic Society. Chicago/IL: Chicago Linguistic Society, 204–227
- Jaggar, Philip J. & Buba, Malami (1994), "The Space and Time Adverbials Nan/Can in Hausa: Cracking the Deictic Code". *Language Sciences* 16, 387–421
- Janssen, Theo A. J. M. (1991), "Preterit as Definite Description". In: Gvozdanović, Jadranka & Janssen, Theo A. J. M. (eds.), *The Function of Tense in Texts*. Amsterdam: North-Holland, 157–81
- Janssen, Theo A. J. M. (1993), "Tenses and Demonstratives: Conspecific Categories". In: Geiger & Rudzka-Ostyn (eds.), 741–783
- Janssen, Theo A. J. M. (1995a), "Deixis from a Cognitive Point of View". In: Contini-Morava & Goldberg (eds.), 245–270
- Janssen, Theo A. J. M. (1995b), "Cognitive Semantics of Region-based Deixis". In: Geiger (ed.), 141–165
- Janssen, Theo A. J. M. (1996), "Deictic and Anaphoric Referencing of Tenses". In: De Mulder, Walter & Tasmowski-De Ryck, Liliane & Vettens, Carl (eds.), *Anaphores temporelles et (in-)cohérence*. Amsterdam etc.: Rodopi, 79–107
- Janssen, Theo A. J. M. (2002), "Deictic Principles of Pronominals, Demonstratives and Tenses". In: Brisard (ed.), 151–193
- Jarvella, Robert J. & Klein, Wolfgang (1982, eds.), *Speech, Place and Action: Studies in Deixis and Related Topics*. Chichester etc.: Wiley
- Jonasson, Kerstin (1992), "La référence des noms propres relève-t-elle de la deixis?". In: Morel & Danon-Boileau (eds.), 457–463
- Jones, Peter (1995), "Philosophical and Theoretical Issues in the Study of Deixis: A Critique of the Standard Account". In: Green (ed.), 27–48
- Josephs, Lewis S. (1978), "Anaphora in Palauan". In: Hinds, John (ed.), *Anaphora in Discourse*. Edmonton etc.: Linguistic Research, 339–375
- Karmiloff-Smith, Annette (1979), *A Functional Approach to Child Language: A Study of Determiners and Reference*. Cambridge etc.: Cambridge Univ. Press
- Kirsner, Robert S. (1979), "Deixis in Discourse: An Exploratory Quantitative Study of the Modern Dutch Demonstrative Adjectives". In: Givón, Talmy (ed.), *Discourse and Syntax*. New York: Academic Press, 355–375
- Kirsner, Robert S. (1993), "From Meaning to Message in Two Theories: Cognitive and Saussurean Views of the Modern Dutch Demonstratives". In: Geiger & Rudzka-Ostyn (eds.), 81–114
- Kirsner, Robert S. (1996), "The Human Factor and the Insufficiency of Invariant Meanings". In: Andrews, Edna & Tobin, Yishai (eds.), *Towards a Calculus of Meanings: Studies in Markedness, Distinctive Features and Deixis*. Amsterdam etc.: Benjamins, 83–106
- Kleiber, Georges (1992), "Anaphore-deixis: deux approches concurrentes". In: Morel & Danon-Boileau 613–623
- Klein, Wolfgang (1978), "Wo ist hier? Präliminarien zu einer Untersuchung der lokalen Deixis". *Linguistische Berichte* 58, 18–40
- Kryk-Kastovsky, Barbara (1996a), "The Linguistic, Cognitive, and Cultural Variables of the Conceptualization of Space". In: Pütz, Martin & Dirven, René (eds.), *The Construal of Space in Language and Thought*. Berlin etc.: Mouton de Gruyter, 329–344
- Kryk-Kastovsky, Barbara (1996b), "Modern Approaches to Deixis". In: Jankowsky, Kurt R. (ed.), *Multiple Perspectives on the Historical Dimensions of Language*. Münster: Nodus Publikationen, 33–49
- Kuroda, Sige-Yuki (1985), "La sémantique des démonstratifs japonais". *Linguisticae Investigationes* 9, 69–81
- Kuryłowicz, Jerzy (1972), "The Role of Deictic Elements in Linguistic Evolution". *Semiotica* 5, 174–83
- Lakoff, Robin (1974), "Remarks on *This* and *That*". In: La Galy, Michael W. & Fox, Robert A. & Bruck, Anthony (eds.), *Papers from the Tenth Regional Meeting*. Chicago Linguistic Society. Chicago/IL: Univ. of Chicago, 345–356
- Lambrecht, Knud (1994), *Information Structure and Sentence Form*. Cambridge: Cambridge Univ. Press
- Langacker, Ronald W. (1987), *Foundations of Cognitive Grammar; Vol. I: Theoretical Prerequisites*. Stanford: Stanford Univ. Press
- Langacker, Ronald W. (1991), *Foundations of Cognitive Grammar; Vol. II: Descriptive Application*. Stanford: Stanford Univ. Press

- Langacker, Ronald W. (2002), "Remarks on the English Grounding System". In: Brisard (ed.), 29–38
- Laury, Ritva (1996), "Conversational Use and Basic Meaning of Finnish Demonstratives". In: Goldberg (ed.), 303–319
- Laury, Ritva (1997), *Demonstratives in Interaction: The Emergence of a Definite Article in Finnish*. Amsterdam etc.: Benjamins
- Lazard, Gilbert & Peltzer, Louise (1992), "La deixis en tahitien". In: Morel & Danon-Boileau (eds.), 209–219
- Levinson, Stephen C. (1983), *Pragmatics*. Cambridge etc.: Cambridge Univ. Press
- Levinson, Stephen C. (1988), "Putting Linguistics on a Proper Footing: Explorations in Goffman's Concepts of Participation". In: Drew, Paul & Wootton, Anthony (eds.), *Erving Goffman: Exploring the Interaction Order*. Oxford: Polity Press, 161–227
- Lyons, Christopher (1999), *Definiteness*. Cambridge: Cambridge Univ. Press
- Lyons, John (1977), *Semantics*, Vol. 1–2. Cambridge etc.: Cambridge Univ. Press
- Malotki, Ekkehart (1979), *Hopi-Raum: Eine sprachwissenschaftliche Analyse der Raumvorstellungen in der Hopi-Sprache*. Tübingen: Narr Verlag
- Malotki, Ekkehart (1982), "Hopi Person Deixis". In: Weissenborn & Klein (eds.), 222–252
- Marchese, Lynell (1986), "The Pronominal System of Godié". In: Wiesemann (ed.), 218–255
- Morel, Mary-Annick & Danon-Boileau, Laurent (1992, eds.), *La Deixis, Colloque en Sorbonne (8–9 juin 1990)*. Paris: Presses Universitaires de France
- Mosel, Ulrike (1982), "Local Deixis in Tolai". In: Weissenborn & Klein (eds.), 111–132
- Nguyễn, Phu Phong (1992), "La deixis en vietnamien". In: Morel & Danon-Boileau (eds.), 177–186
- Nunberg, Geoffrey (1993), "Indexicality and Deixis". *Linguistics and Philosophy* 16, 1–43
- Paris, Marie-Claude (1992), "Démonstratifs et personnes en chinois standard". In: Morel & Danon-Boileau (eds.), 167–175
- Parker, Elizabeth (1986), "Mundani Pronouns". In: Wiesemann (ed.), 131–165
- Perner, Josef & Garnham, Alan (1988), "Conditions for Mutuality". *Journal of Semantics* 6, 369–385
- Picabia, Lélia (1992), "La Morphologie du démonstratif en comorien [ʃikomor]". In: Morel & Danon-Boileau (eds.), 157–166
- Rauh, Gisa (1983, ed.), *Essays on Deixis*. Tübingen: Narr Verlag
- Reed, Judy & Payne, David L. (1986), "Asheninca (Campa) Pronominals". In: Wiesemann (ed.), 323–331
- Richter, Heide (1988), *Indexikalität: Ihre Behandlung in Philosophie und Sprachwissenschaft*. Tübingen: Niemeyer
- Roberts, Lawrence D. (1986), "The Figure-Ground Model for the Explanation of the Determination of Indexical Reference". *Synthese* 68, 441–486
- Rousseau, André (1992), "La deixis: un problème de logique et de philosophie du langage". In: Morel & Danon-Boileau (eds.), 365–374
- Sauvageot, Serge (1992), "De l'expression de la deixis dans quelques langues du groupe ouest-atlantique (Afrique Noire)". In: Morel & Danon-Boileau (eds.), 151–155
- Schegloff, Emanuel A. (1996), "Some Practices for Referring to Persons in Talk-in-Interaction: A Partial Sketch of a Systematics". In: Fox (ed.), 437–485
- Showalter, Catherine (1986), "Pronouns in Lyélé". In: Wiesemann (ed.), 205–216
- Smith, John Ch. (1992), "Traits, marques et sous-spécification: application à la deixis". In: Morel & Danon-Boileau (eds.), 257–264
- Sridhar, Shikaripur N. (1990), *Kannada*. London etc.: Routledge
- Strauss, Susan (1993), "Why 'This' and 'That' are not Complete without 'It'". In: Beals, Katherine & Cooke, Gina & Kathman, David et al. (eds.), *Papers from the 29th Regional Meeting. Chicago Linguistic Society*. Chicago/IL: Univ. of Chicago, 403–417
- Tamba, Irène (1992), "Démonstratifs et personnels en japonais: Deixis et double structuration de l'espace discursif". In: Morel & Danon-Boileau (eds.), 187–195
- Tanz, Christine (1980), *Studies in the Acquisition of Deictic Terms*. Cambridge etc.: Cambridge Univ. Press
- Weissenborn, Jürgen & Klein, Wolfgang (1982, eds.), *Here and There: Crosslinguistic Studies on Deixis and Demonstration*. Amsterdam etc.: Benjamins
- Westrum, Peter N. & Wiesemann, Ursula (1986), "Berik Pronouns". In: Wiesemann (ed.), 37–46
- Wiederspiel, Brigitte (1992), "Le nom propre: un 3e type de référence?". In: Morel & Danon-Boileau 471–478

- Wiesemann, Ursula (1986, ed.), *Pronominal Systems*. Tübingen: Gunter Narr
- Wills, Dorothy D. (1990), "Indexifiers in Wolof". *Semiotica* 78, 193–218
- Zribi-Hertz, Anne (1992), "De la deixis à l'anaphore: quelques jalons". In: Morel & Danon-Boileau, 603–612

Zwicky, Arnold M. (1977), "Hierarchies of Person". In: Beach, Woodford A. & Fox, Samuel E. & Philosoph, Shulamith (eds.), *Papers from the Thirteenth Regional Meeting. Chicago Linguistic Society*. Chicago/IL: Univ. of Chicago, 714–733

*Theo A. J. M. Janssen, Amsterdam  
(The Netherlands)*

## 96. Person

1. Fundamentals
2. Representation in language structure
3. Person categories in detail
4. Interactions with other categories
5. Geographical distributions
6. Uncommon abbreviations
7. References

### 1. Fundamentals

#### 1.1. Three models of person categories

**Person** is one of the most slippery and elusive of all grammatical categories. Like demonstratives (*this, that, here*), tense (past/present/future), and some lexical adverbs (*tomorrow*), it is an **indexical** category, whose referential values shift with objective changes in the speech situation – but also with conceptual or symbolic modulations thereof (cf. Art. 95).

There are three ways of analysing the main person categories, all of which have Greco-Roman pedigrees, and each of which has some grammatical support. Benveniste (1946) is a classic statement of the **speech-roles** model, identifying person categories with **participant roles**: speaker = first [1st] person; hearer/addressee = second [2nd] person; or as **nonparticipant referents** = third [3rd] person:

(1) 1st, 2nd            (3rd)

This brings out the residual status of 3rd person, Benveniste's "non-person", formally unmarked in many languages. Historical linguists have found that a word form with originally nonzero 3rd person marker is often reinterpreted as a bare stem, on which new 1st/2nd forms are later built (Watkins 1962: 90). In Mara (Australia), in addition to pronominal prefixes for subject and object (omitted below), verbs like *-wu-* 'sit' and *-na-*

'see' distinguish two present stems, one for 3rd and 3rd → 3rd forms and another for any combination involving a 1st or 2nd person (*-lu* and *-mi-* ... *-ja* are lexically specific 'PRES' allomorphs, *-a-* is a '3.PRES' morpheme used with all inflected verbs, and *-yi* and *-ni* are lexically specific '1/2.PRES' allomorphs added to the PRESENT stem).

(2) -a-wu-lu	'(3) sit(s)'
-wu-li-yi	'(1/2) sit'
-a-mi-na-ja	'(3) see(s) (3)'
-mi-na-ji-ni	'(1/2) see (2/1)'
"	'(1/2) see (3)'
"	'(3) see(s) (1/2)'

Furthermore, many hierarchical phenomena treat 1st/2nd as a unit vis-à-vis 3rd person (cf. 4.4, 4.5).

Wilhelm von Humboldt and other early scholars with strong psychological/philosophical interests advocated a **consciousness** model with the major break between 1st person and 2nd/3rd:

(3) 1st            2nd, 3rd

This emphasizes the speaker's unique access to his/her own feelings, memories, and intentions. It explains some grammatical and lexical evidence associated with affect (intimacy) – some kin-term systems have special 1st person forms, presumably because of the speaker's affective relationship toward his/her own kin. It also accounts for evidential modal patterns; in Japanese, reported inner states distinguish, e.g., 'I am hot' from 'You are/He is (apparently) hot', the latter showing derivational modification, cf. Eng. *You look hot* (Aoki 1986). There are also some languages with extensive 2nd/3rd person neutralizations, though one suspects that politeness factors are involved too. The conscious-

ness model also explains why ‘1.SG’ and ‘1.PL’ are so consistently distinguished even in languages that neutralize number in other persons.

A kind of compromise between (1) and (2), not involving binary breaks, thinks of the three persons as an ordered set, and underlies our ordinal terminology for persons. This clinal model easily incorporates further elements, such as a distinction between true 3rd person and a residual “non-person”:

(4) 1st > 2nd > 3rd (> Ø)

It has some support from (infrequent) hierarchical phenomena where 2nd person is in fact intermediate between 1st and 3rd; we may mention a 1st > 2nd > 3rd word order rule for pronominals in some Australian languages (Wurm 1969: 59). There are counter-examples, generally involving morpheme ordering, where 2nd person seems to outrank 1st, but in many hierarchical phenomena 1st and 2nd persons are equivalent. More speculatively, phonological symbolism (e.g., high/mid/low vowels) and lexical evidence are used by Becker & Oka (1974) to argue for a grammatically potent 1st > 2nd > 3rd cline in Old Javanese (and universally). The best evidence for 2nd person as mediator emerges in languages with three spatial deictics (proximal, medial, distal), where 2nd person is often syntactically associated with medial (4.6). As for the distinction between true 3rd person and “nonperson”, many languages such as Manam (Austronesian; Lichtenberk 1983) have a nonzero 3rd person pronoun or affix used only for human reference (perhaps along with some higher animates), versus zero for lower-ranking entities; cf. also obviatives (3.3, below).

Although no-one has advocated it, one can imagine an addressee-centric **pragmatic** model. This would be grounded in the view that speaking is aimed at influencing the addressee’s behavior, and would be at home in contexts involving attempted behavior modification, such as advertising. It would receive grammatical support from imperatives and vocatives.

The use of ordinal terms (1st, 2nd, 3rd; some languages also have a so-called “4th”) for person categories is based rather narrowly on the clinal model (4), as pointed out by Plank (1985). However, an alternative terminology and notation based only on (1) or only on (2) would be equally misleading, so

we will use the well-entrenched 1st/2nd/3rd notation.

Nouns (or noun phrases) usually have 3rd person agreement, but taking 3rd person as intrinsic to nouns is unjustified. In some languages, nouns may be marked affixally for any person category (‘1.PL-man-PL’ for *us men*), and in any event they can be used in address, presupposing 2nd person status (Vairel 1980), whether or not they have special vocative forms (Lat. *Et tu, Brute?*). On the other hand, “pronouns” in some languages (especially in East Asia) have a noun-like character (cf. Eng. *Yours truly* = *I, me*).

In a simple conversation, two individuals (A and B) take turns speaking. A uses ‘1.SG’ pronouns to refer to himself and ‘2.SG’ pronouns to address or refer to B; during B’s turns, she uses ‘1.SG’ for self-reference, ‘2.SG’ to address or refer to A. (Both use nouns or 3rd person pronominals to refer to nonparticipants.). As addressee, A must appreciate that B’s ‘1.SG’ and ‘2.SG’ reflect B’s own vantage point; that is, A must (passively) assimilate B’s **perspective**. The ability to shift between one’s own and the interlocutor’s perspective, while sometimes relevant to other indexicals such as demonstratives and tense, is always critical for personal pronouns because the **deictic center** is reset at every change in speaking turns. Without personal pronouns, discourse would require heavy repetition of names and nouns, with a corresponding reduction in the role of grammatical systems. Perspective shifting, like negation, seems routine to us, but in fact is one of the major differences between human and animal communication. It goes far beyond simple ego/alter differentiation, which can be ascribed even to plants, whose roots can chemically distinguish other same-individual roots from roots of neighboring conspecific plants. In first language acquisition, learning to shift person perspectives (usually by age three) appears to correlate with the ability to appreciate others’ distinct visual perspectives, as revealed in concealment games (Loveland 1984), and is delayed in children born blind (Fraiberg 1977). Correct pronominal usage is also delayed, as we would expect, when the two-year-old child is exposed chiefly to a single caregiver, whose 1st and 2nd person pronouns could be (mis-)interpreted as having absolute reference; the child’s passive observation of other conversational dyads (e.g., father and mother) is helpful in learning to shift perspectives (Oshima-Takane 1988).

The importance of an indexical person category to human speech is supported by surveys (Greenberg 1963; Forchheimer 1953) concluding that all languages distinguish 1st, 2nd, and (perhaps unmarked) 3rd person. It is true that in some languages (e.g., Japanese), there are significant restrictions on the usage of noun-like “pronouns,” but there is still a core system of speech-event participant indexing, most easily visible in unself-monitored registers. One caveat is that person categories cannot always be clearly segregated from spatial deictics, and one can argue whether the deictics are parasitic on person categories, or vice versa (4.6).

In 1.2 to 1.6 we briefly consider factors other than the simple referential function of person indexing – factors that always affect usage (often challenging our conventional “definitions” of persons), and that may crystallize formally (we will look more closely at grammatical categories in 3).

### 1.2. Generalized reference

The conversationalists A and B may speak about **generalized** or unknown (totally indefinite) referents, as in proverbial, gnomic (eternally true), and some future or irrealis statements. Examples are: ‘[ ] can’t fight the government’, and ‘If [ ] build(s) a better mousetrap, the world will beat a path to [ ]’s door’. Since generics do not unambiguously exclude A or B, person categories may break down or undergo neutralization here, and such fuzziness can be put to pragmatic use. Different languages generalize reference using 1st plural, inclusive (cf. 3.2), 2nd singular or plural, 3rd singular or plural, or specialized impersonal markers (cf. 3.5).

### 1.3. Referential tracking

Person categories may be adjusted or expanded to maximize their value in the (horizontal) **tracking** of referents through narrative sequences and establishing (vertical) **coreference** (= coindexation) from matrix to embedded clauses. Consider this narrative chunk, which begins with nouns (names):

- (5) John came over the hill and [ ] noticed George lying under a tree. At [ ]’s approach, [ ] got up. Then [ ] said something to [ ]. [ ] went over to a fallen tree and broke off a branch. [ ] ran at [ ], brandishing [ ]. But [ ] was too fast for [ ] and got away into the scrub.

Insertion of English ‘3.SG’ masculine pronouns (*he, him, his*), along with the one instance of ‘3.SG’ neuter *it* after *brandishing*, would not fully eliminate referential ambiguity, since at certain points we have no way of knowing whether John or George is the referent. The English distinction between masculine and neuter is only marginally helpful in tracking, since it is often redundant due to selectional restrictions on verbs (one brandishes weapons, not people). If we replace *John* with *Jane*, gender marking (*he* vs. *she*, etc.) would be very helpful, but the ambiguity would reappear if the protagonists were *Jane* and *Georgina*. Even with a richer set of noun-class distinctions within the 3rd person (cf. Art. 98), as long as there is just one 3rd masculine pronoun, the passage above will be problematic. In languages lacking gender and number marking in nonparticipant pronominals (e.g., Choctaw) the problem becomes chronic, and a more flexible and explicit co-indexing mechanism is needed. One syntactic solution is deletion under coindexation to a controller; another is switch reference (cf. 4.2), whereby adjacent clauses are specified (e.g., by verbal affixes) as having the same or different subjects.

Similar problems arise in disambiguating embedded structures:

- (6) Betty found Mary and me in the back room. Betty didn’t want [ ] to stay there. Betty asked me to give [ ] [ ]’s medicine.

In English, the first slot is not problematic, since we get an overt pronoun *her* only to refer to Mary, but in languages lacking coreferential deletion we could have ambiguity. Furthermore, Eng. *her* (dative) and *her* (possessive) in the second and third slots may be ambiguous between Betty and Mary.

One way to avoid ambiguity in (5) and (6) is to use special anaphoric pronouns (Lat. *se*, Turk. *kendi*, Basq. *bera*) to indicate coindexation to an antecedent in a preceding or higher clause, in place of the usual 3rd person forms. These special anaphorics may (but need not) be identical to the primary clause-internal reflexive pronouns. An interesting variation on this is the use of logophorics (3.4).

Another strategy is to have one 3rd person referent establish itself as basic (proximate) within a textual span, and to mark other 3rd person forms with a distinct, obviative form (3.3).

#### 1.4. Intersubjectivity (perspective merger)

There are some subtler factors that have a major impact on the usage of person categories, and may be grammaticalized. We noted above that without an ability to shift (passively) from one's own to an interlocutor's perspective, A would not understand B's personal pronouns and so the category of person could not arise. There is nothing to stop passive intersubjectivity from becoming active, and a single speaking turn by B may switch from B's to A's perspective (*If I were you, I'd ...*) and back again. Person-shifting within turns begins with caregivers' one-way speech to prelingual children (Rabain-Jamin & Sa-beau-Jouannet 1989).

The pragmatics of shifting can be illustrated in the exchange (7), where A tries to lighten the burden on B by impersonalizing reference, whereupon B shifts to a generalized but nonetheless personal 2nd person, essentially inviting A to share B's perspective:

- (7) B: I feel like a dog.  
 A: How does a dog feel?  
 B: You feel like everyone's looking for a chance to kick you.

A recurring context for adopting a nonparticipant's perspective is reported speech (or thought), which in English requires a choice between **direct discourse** (retaining the personal pronouns and other deictics of the original utterance), **indirect discourse** (retaining the lexical core of the utterance but substituting here-and-now deictics), and **paraphrase** (reformulation of the lexicogrammatical core). Other languages have unusual **semi-indirect discourse** patterns (Gragg 1972) involving a combination of direct (displaced) and indirect (here-and-now) perspectives. Many African languages have special "person" categories (e.g., logophorics, cf. 3.4) limited to quotative and similar intersubjective contexts, having some characteristics of both 3rd and (embedded) 1st persons.

#### 1.5. Pragmatic masking and skewing

While many preceding examples show the speaker merging his/her perspective into that of the addressee or an external referent, such intrusions can be pragmatically dangerous, and there is a countervailing tendency to de-personalize even the "objective" speaker and addressee (Head 1978). Some shifting patterns (e.g., from 1st person pronoun to impersonal self-reference using own name) oc-

cur in children's speech as early as age two with specific pragmatic functions (Gerhardt 1988: 359 f.).

Devices for masking or blurring reference to the individual addressee are familiar in European languages (*tu/vous*, etc.) and are found around the world (Brown & Gilman 1960). Parallel masking devices for speakers are also common, though perhaps less conspicuously grammaticalized. The most widespread blurring device is the use of plurals for (objectively) singular speaker and/or addressee, which in the case of the speaker means a shift to 1st plural, a category that in many languages does not exclude the addressee. 1st and 2nd person pronominals may be superceded by nominal expressions, with or without a possessive (*Your Worship, My Dear Sir*), in which case pronominal agreement is typically 3rd person and a substantial 2nd/3rd merger may eventually result. For extreme examples of person masking in transitive pronominal clusters, see 4.4.

#### 1.6. Aberrant cases

In complex societies where speakers and addressees are separated and depersonalized by writing and broadcast media, literary and communication theorists agonize over the status of (a) narrative "voice[s]" – the author's?; the narrator's?; the implied omniscient observer's?; the celebrity spokesperson's? – and (b) "reception[s]" or even participation by various professional and lay audiences, often self-hierarchized in some way or differentially recognized/excluded. No aspect of the "speech event" can be taken for granted in such contexts, and mocking or subversion of role structures is often part of the game. Similar remarks apply to comparable genres in tribal societies, such as political oratory (e.g., Rumsey 1989).

Likewise, psychiatric disorders and therapeutic communication reveal many kinds of noncanonical "speaker" and "addressee" roles: split personalities (which occasionally "converse"), identification and projection, extreme narcissism with impoverished perspective-shifting ability, etc. Even such an ordinary phenomenon as talking to oneself (out loud or silently) creates analytical problems.

However, such noncanonical communications have, allowing for some anomalies, an ordinary grammatical expression. We can treat "1st" and "2nd" persons – i.e., the linguistically recognized "speaker" and "addressee" – as abstract interactive roles,

respectively active and receptive, but not always encoding “objective” participant structures.

## 2. Representation in language structure

### 2.1. Morphology

By definition, person is marked (along with other categories) in any system of personal pronouns or pronominal morphemes. Similar sets of forms with no specific marking for speaker and/or addressee are usually labeled “demonstratives,” “noun class markers,” etc. rather than “pronouns.”

**Independent pronouns** are full-fledged words, have their own word prosody, and resemble noun phrases in their positional possibilities. They are usually, but not invariably, morphologically segmentable, e.g., into person, gender, number, case, and perhaps discourse-status markers.

In **pro-drop** languages where pronominal categories are separately expressed by clitics or verbal affixes (see below), independent pronouns are used sparingly. They may be used for minor grammatical roles not otherwise expressed, such as adpositional objects and subjects of nonverbal sentences. Independent pronouns in these languages are also used to mark discourse statuses, linking the information value of their proposition to a larger textual context. Using crude translations, such discourse categories might include contrastive topic (*'he left, but she stayed'*), focus (Q: ‘Who saw it?’; A: ‘*She* did’), unassisted agent (*'you-yourself* built it’), additive (‘He jumped, and *I-too* did’), and sequential (‘He jumped, and *I-too-in-turn* did’). Such pronouns may lack case marking entirely.

In **non-pro-drop** languages like English, where syntactic positions like subject and object are normally nonnull, pronouns have high text frequency, tend to be short and unstressed, and often distinguish at least subject and object cases. From this it is only a short step to reduce the pronominals to **clitics** (cf. Art. 41) that are prosodically parasitic on host words or constituents – **proclitics** if they precede the host, **enclitics** if they follow (the more common pattern). Pronominal clitics tend to cluster together into complexes (e.g., subject and object markers), on which see 4.4. The clitics may attach themselves to other free-standing or floating grammatical formatives, such as tense-aspect-mood

markers, forming complex **auxiliaries**. Clitics often gravitate to second position (following the first free-standing constituent). Alternatively, the pronominal clitics may attach themselves to the predicative nucleus (usually a verb). In this case, they may eventually form an affixal paradigm in a later stage of the language.

Whether in clitic clusters or as verbal affixes, as the pronominals are progressively reduced in form, and stripped of any emphatic, contrastive, or other discourse statuses, they commonly permit co-occurrence with coreferential noun phrases, becoming so-called **agreement** markers, the idea being that the pronominals copy certain features of the corresponding noun phrases. This analysis currently arouses misgivings among linguists, since in most such languages the pronominal marking does not require an overt coreferential noun, and since the actual categories marked by pronominals and nouns may differ. Pronominals “agreeing with” conjoined nouns (‘The man and the woman 3.PL-went’) present further problems for copying analyses.

In northern Basque (Lafon 1959) there is pronominal marking of **allocutive** addressee if there is not already a 2nd person subject or object marker. Thus, ‘I gave it to him’ must be rendered as a verb complex glossable as ‘I-gave-to-you-it-to-him’. Since Basque has several 2nd person forms (familiar masculine and feminine, formal, and plural), these allocutives are essentially sociolinguistic in nature. They may be compared to **ethical datives** of the sort seen (for 1st as well as 2nd person) in Shakespearean English.

Pronominals may also function as parts of a noun phrase. They may be **possessive** (*my father*) or **appositional** (*you guys*). In many languages such examples are expressed by a noun with pronominal affix (sometimes identical to the intransitive or transitive subject affix for verbs). In both functions, there are syntactic matters (agreement, possessor raising) that may treat the pronominal and not the noun as the head of the phrase, and thus increase the scope of person marking in grammar and discourse.

### 2.2. Lexicon

Lexical differences can be relevant, to the extent that they depend on (overt or covert) person categories. An example is lexical suppletion associated with pronominal affixation in verbal paradigms, as in Eng. (I) *am*, (you/

we/they) *are*, (he/she/it) *is*. Note that the 2nd person *you*, even in singular function, still behaves as a (covert) plural, as also in past (*you/we/they*) *were* vs. (*I/he/she/it*) *was*.

In Australian languages, the choice among stems for kin categories may depend on the person category of the **propositus** (this term is used in formal kinship semantics for the “possessor” or ego of reference). Typically there are three stems – for 1st, 2nd, and 3rd person propositus, regardless of gender or number, and merging inclusive and exclusive within 1st, as in Nunggubuyu:

- (8) (a) ‘father’  

<i>baba</i>	‘(1st)’
<i>niñara</i>	‘(2nd)’
<i>niñara-yuŋ</i>	‘(3rd)’
- (b) ‘paternal grandfather’  

<i>muri</i>	‘(1st)’
<i>ra-muru</i>	‘(2nd)’
<i>ra-muru-yuŋ</i>	‘(3rd)’

The major split is between 1st and non-1st, the latter being suppletive (8 a) or derived from 1st by a prefix /raN-/ (8 b). The 3rd person form is the 2nd person stem plus suffix *-yuŋ*. Neither /raN-/ nor *-yuŋ* occurs elsewhere in the language as a person marker and neither is etymologically pronominal. The unusual markedness patterns are compatible with the consciousness model of person (1.1) and may represent lexicalization and grammaticalization of affect categories, cf. Eng. *father*, *Daddy*.

There are also cases where 1st/2nd person kin terms are opposed to 3rd person forms, as in Usan (Papuan), e.g., 1st/2nd *tain* ‘[my/our/your] father’ versus *ur* ‘[his/her/their] father’ (Reesink 1987: 49). This fits with the speech-roles model of person (1.1). There are even Australian cases where the 2nd person kin term is the unmarked form, as in Warndarang (*baba* ‘your father’, *ya-baba* ‘my/our father’, *ña-baba-ña* ‘his/her/their father’), or where the major split is between 2nd and 1st/3rd, as in Dhuwal (*baapa-?miriju* ‘my/his/her father’ vs. *baapa-yali* ‘your father’).

Certain Australian languages such as Warlpiri and Gurindji have yet more elaborate systems including **triangular kin terms** that relate the referent to two separate proposituses (often speaker and addressee). In other words, a single term means ‘your father/my brother’ – or does it mean ‘father of you [who are my nephew]’ or ‘father of me [who am your uncle]’? As in plane geometry,

specifying any two legs of a triangle determines the third leg. Like the allocutive 2nd person, triangular kin terms have a sociolinguistic rather than disambiguating function, forcing speakers to recognize both their own and the addressee’s perspectives.

In languages like Japanese, many verbal and nominal stems can specify **honorific** or **humble** status, whether by suppletion – cf. Eng. *estate/residence/domicile/house/place/pad* – or by affixes. In formal speech situations (which subsume more contexts than in Euro-American society), conversational interlocutors of objectively equal status tend to associate honorific expressions with the addressee ('honorable house', or 'estate', decoded as 'your house') and neutral or humble terms with the speaker ('humble house', or 'place', decoded as 'my house'). Speech levels therefore have a covert person-indexing function.

Another feature of the Japanese lexicon is the role of perspective. The speaker’s own viewpoint can be signalled by using lexical items such as deictically sensitive *kuru* ‘come’ instead of *iku* ‘go’, or particular choices for transactional verbs like ‘give’ and ‘receive’. The speaker may project his/her perspective onto certain other (3rd person) referents, generally sanctioned by specific kinship or other personal associations; hence ‘(He) came to my father’ but ‘(He) went to your father’.

### 3. Person categories in detail

#### 3.1. Minimal systems

It is generally recognized that the minimal set of person categories in human languages is 1st singular, 1st plural (either suppletive, or ‘1.SG’ plus a plural marker), and 2nd persons (Greenberg 1963: 96; Forchheimer 1953; Ingram 1978: 227). Rarely, a language with an (inherently nonsingular) 1st inclusive person (3.2, below) can get by with optional marking of plurality for (exclusive) 1st person, as in Aymara (Hardman 1976).

We can refer to non-1st/2nd person pronouns as “3rd” person. However, in many languages the only 3rd person independent pronouns are demonstratives that do not belong to the same morphological set as the 1st/2nd person pronouns. In other cases, the 3rd person pronoun or pronominal affix/clitic is phonologically null and can be considered the absence of a person marker.

If we accept that the trio of 1st singular, 1st plural, and 2nd persons is universal in

some form, there is still the tricky issue whether 1st singular and 1st plural really belong to a single person category. We discuss the semantics of pronominal number more fully in 4.2, below, but we may here mention that the 1st plural is very often unrelated in form to the 1st singular (Eng. *I, we*), even when 2nd and 3rd persons each have consistent person markers across numbers. This matter is discussed in considerable typological detail by Forchheimer (1953).

### 3.2. Exclusive/Inclusive

In many languages there are two categories corresponding to English 1st plural, one for a group minimally including the speaker and addressee (cf. Eng. *Let's go!*), the other for a group that specifically excludes the addressee. Ordinarily, the exclusive ('1.EX') is the 1st plural form that patterns with the 1st singular (which is, redundantly, exclusive). The inclusive ('1.IN') may distinguish a dual (for speaker and one addressee only) from a plural even if dual is not marked in other persons. In such cases, the '1.IN.DU' may lack nonsingular number marking and function much like the singular of other persons (cf. 4.2). Associating '1.IN.DU' with other singular pronominals provides some morphological evidence for treating inclusive as an entirely distinct category, but even when there is no dual the inclusive form may be distinctive: Chinese *wǒ* '1.SG', *wǒmen* '1.EX.PL', but *zán-men* '1.IN'.

In some languages, a periphrastic '1.IN' is produced by combining a simple (1st/2nd/3rd) person marker with an "inclusive" marker. In some cases the latter is really a cross-cutting person (or person/number) marker, as in many Algonquian languages, where '1.IN' is morphologically '2nd' (prefix) + '1.PL' (suffix), while '1.EX.PL' is '1st' (prefix) plus the same '1.PL' suffix (Proto-Algonquian, however, had a more complex system, Goddard 1967: 68). If '1.IN' is based on a single true person marker, it is usually 1st, but may be 2nd as in Yokuts (California), and may even be 3rd – in Southern Peruvian Quechua, *-chis* (perhaps glossable '2.PL') is added to '2.SG' to produce '2.PL', and is added to '3.SG' to produce a combination expressing 1st inclusive or polite 2nd person. Mannheim (1982: 457) concludes that Quechua has "inclusive number" supplementing 1st/2nd/3rd "persons," while languages where 1st inclusive is morphologically autonomous have true "inclusive person."

On the other hand, in some languages '1.EX' is expressed (perhaps optionally) by adding a morpheme to a simpler form used as '1.IN' or as undifferentiated 1st plural. In Ngiyambaa (Australia, Donaldson 1980: 122f.), '1.IN.DU' *yali:* is the basis for '1.EX.DU' *yali:-na*, where *-na* is identifiable as a '3.SG.INVIS' enclitic.

Instead of "1.EX, 1.IN, 2nd, 3rd" some authors use labels expressing the hybrid structure of '1.IN', hence "1, 12 (= 1.IN), 2, 3" or "I-II, I+II, II, III". However, the previous paragraph warns us that '1.IN' has variable morphological relationships to other persons.

There is a natural association between 1st inclusive and certain other grammatical categories, particularly hortative mood (*Let's ...!*); see 4.3. It may be that languages tend to have either a 1st inclusive pronominal category or a hortative mood marker, but not both. It is interesting to consider how one might express this inclusive hortative in (polite) Japanese, which assiduously maintains the distinction between speaker and addressee perspectives through formality and asymmetric honorifics. A simple *Let's go!* cannot easily be expressed in such a system. This matter deserves far more attention than it has received in studies of Japanese (or other East Asian) **speech levels**, which tend to discuss overt morphological distinctions and thereby miss certain sociolinguistic nuances.

In one dialect of Dyirbal (Australia), there is both an ordinary inclusive ('1.IN.DU', '1.IN.PL') and a special **speaker-and-spouse** inclusive pronoun (Dixon 1972: 50).

An inclusive 2nd plural ('2.IN.PL') 'you-SG + them' is reported for Abkhaz (Northwest Caucasian), distinct from exclusive 2nd plural ('2.EX.PL'), which presumably indicates multiple addressees. The '2.EX.PL' (like the '1.EX.PL') is optional and somewhat infrequent in this language (Hewitt 1979: 156f.). Bhattacharya (1975: 129) speaks of distinct '2.EX' and '2.IN' forms in "many" Munda languages, though the three grammars of individual Munda languages available to this writer failed to turn up examples. In Sierra Popoluca (Mexico), a "limited" '1.IN.PL' denoting speaker plus one or more addressees, excluding nonparticipants, is reportedly distinguished – for pronominal affixes on transitive verbs – both from '1.EX.PL' (speaker plus one or more nonparticipants) and from "general inclusive" plural (speaker, at least one addressee, and at least one nonparticipant) (Foster & Foster 1948: 19).

### 3.3. Proximate and obviative

Proximate/obviative systems, as in Algonquian languages (Bloomfield 1946), allow the speaker to use separate markers for two nonparticipant referents in a textual segment (“span”). In Algonquian, the system applies within the set of relevant 3rd person animates (and sporadically or indirectly to inanimates). The opposition is marked on nouns and on pronominal affixes (subject, object, and possessive). The basic, unmarked form is **proximate**. Additional nonparticipant animate referents in the same span have a distinct **obviative** form (with a partial morphological affinity to the “inanimate” category). “Proximate” is an unfortunate term in Algonquian, where the relevant form has no spatial-deictic value, but in Comanche, (Uto-Aztecán) “proximate” is more appropriate since the form in question is indeed a special usage, limited to narrative, of a proximal demonstrative ‘this’ (Armagost 1985).

In Algonquian, choosing which of the two referents is proximate is roughly determined by the respective syntactic positions of the two at the beginning of the **span** – generally a subject takes precedence over an object, and a possessor takes precedence over the possessed noun, though there are exceptions and the real controlling factor may be a kind of (relative) discourse topicality. Once established, the two referents continue to be marked as respectively proximate and obviative throughout the span, which may include two or more propositions. It is therefore possible for case reversals to occur as the span progresses:

(9) PRX hit OBV, and OBV hit PRX (back).

In (9), the second clause would have (in Algonquian), obviative subject acting on proximate object. However, in practice it seems that spans do not usually extend very far, and there are frequent span-restarting points even in narratives involving the same referents over long stretches.

In some Algonquian languages, there is also a “second” or “farther” obviative (or “surobviative”) for a third referent within the span. It seems to be unproductive and defective, and is formed by reduplicating the obviative morpheme.

Whether proximate and obviative should be considered “persons” is a matter of terminological taste. They are indexical, but their indexing is discourse-internal (horizontal),

referring back to span-initial occurrences. Special obviative forms permit efficient tracking of multiple nonparticipant referents through spans of a few clauses. In this sense, proximate/obviative distinctions are a functional supplement to, rather than a subdivision of, person categories. Nevertheless, proximate and obviative may interact grammatically with person categories, for example in complex pronominal hierarchies (see below, 4.5).

In grammars of some languages with more than one 3rd person category, one of them is called the **fourth person**. In Eskimo, this refers to a kind of reflexive 3rd person, used for example in subordinated clauses to express coindexation to the main-clause subject (Fortescue 1984: 146 f.); in Navaho the term designates a kind of obviative 3rd person (Akmajian & Anderson 1970).

### 3.4. Logophorics

Many African languages also have a difficult-to-describe pronominal category called **logophoric** (after Hagège 1974). (An excellent discussion of logophorics in Ewe is Clements 1975.) It is found most consistently in reported discourse (speech or thought), and serves as an index of the referent whose discourse (or cognitive perspective) is reported. We could think of it as a special form of embedded 1st person (assuming that reported discourse is essentially “direct” except for this substitution). Or, we could think of it as a special variant of the 3rd person (assuming “indirect discourse”), where the current speaker imposes his/her own deictic center but uses the logophoric to mark a 3rd person coindexed to the controlling referent. The two interpretations can be illustrated in (10), where the first line is an interlinear gloss using “LOG” for logophoric and “3rd” for nonlogophoric 3rd person, and the second and third lines are direct and indirect translations, respectively.

(10) 3rd said, LOG hit 3rd  
 ‘She said, “I hit her”.  
 ‘She<sub>i</sub> said she<sub>i</sub> had hit her.’

Assessing how logophoric relates to simple 1st and 3rd persons requires analysis of typologically variable factors. Consider *She said, “I will hit you”*. If this comes out as *She said, LOG will hit you*, the preservation of the original 2nd person object hints that logophoric is (likewise) a preserved person marker (hence 1st), but if it comes out as *She said,*

*LOG would hit him*, the “indirect” replacement of *you* by *him* suggests that logophoric too is converted (i.e., is a special 3rd person). In Babungo, Schaub (1985: 2f.) observes both of these patterns, but presents interesting syntactic evidence (involving plural logophorics) that even in the type *She said, LOG will hit you* the logophoric behaves like a 3rd (not 1st) person pronoun. Specifically, in the construction *X said to Y, “what can we (= X + Y) do?”*, ‘we’ is expressed by the combination of a singular logophoric and a preceding ‘2.PL’ (not ‘1.PL’) pronoun. ‘2.PL’ makes sense, in Babungo as in other languages, for addressee plus nonparticipant, but not for addressee plus speaker (which becomes ‘1.PL’).

Further evidence for a 3rd person connection for logophoric is the fact that logophoric is identical in form to the clause-internal 3rd person reflexive in some African languages. Thus, in Songhay (Mali), the surface form *She<sub>x</sub> said he<sub>y</sub> hit her<sub>x</sub> with LOG’s stick* is ambiguous, since the last logophoric could be coindexed either (logophorically) to ‘she<sub>x</sub>’ or (as a clause-internal reflexive) to ‘he<sub>y</sub>’. Coming from the other direction, some non-African languages make use of reflexive pronouns in certain logophoric-like uses (Latin *se*, Japanese *zibun*).

Logophorics are not generally used when the controlling referent is identical to (coindexed with) the speaker: *I said (that) I came*. In some languages, like Ewe, a logophoric may be used with a 2nd person controller, so that *You said (that) you would come* is expressed as *You said, LOG will come*, but other African languages avoid logophorics with 2nd person controller (Hagège 1974).

In Yaç Dii (Adamawa group, Niger-Congo), a special **second-level logophoric** (LOG<sup>2</sup>) is used to specify coindexation to an already logophoric antecedent in a higher matrix clause (which is, by definition, itself coindexed with an antecedent in a yet higher clause), as in ‘He<sub>x</sub>(3rd) asked why she asked him<sub>x</sub>(LOG) that he<sub>x</sub>(LOG<sup>2</sup>) repay the loan’ (Bohnhoff 1986: 119 f.).

Logophorics have at least two obvious functions. First, they assist with referential tracking, somewhat as the proximate/obviative system does. In a few African languages such as Tuburi, the logophoric form is stretched to the point of occurring in contexts not involving speech or thought, as a tracking device to differentiate two 3rd person referents in a complex sentence or a dis-

course span (Hagège 1974). Secondly, the fact that there are any logophorics in a span (and logophoric spans may comprise entire narratives) provides frequent reminders that the passage represents another person’s speech or thinking, so the current speaker is not (fully) committed to its accuracy. Logophorics are therefore an important bridge between pronominal categories and evidential responsibility; works that lay stress on the micropolitical and perspectival aspects of personal pronouns (e.g., Mühlhäusler & Harré 1990) would benefit from detailed coverage of these African systems.

### 3.5. Impersonals

Whether “**impersonal**” (generalized) pronominals constitute a grammatical category distinct from 3rd person is problematic. To begin with, special impersonal pronominals – whether expressed by an independent, clitic, or affixed form – are often defective (limited to subject, to object, or to some other function): French allows *On m'a vu* ‘Someone saw me’ (= ‘I was seen’) but not *\*J'ai vu on* ‘I saw someone’. Furthermore, an impersonal independent pronoun may have ‘3.SG’ agreement (Eng. *One doesn't eat with his/her fingers*).

In some languages, though, the impersonal is rather more autonomous from the regular 3rd person. Even in English, impersonal *one* is (optionally) carried through to object possessives and tag-question subjects: *One doesn't eat with one's hands, does one?* In some Amerindian languages, the impersonal marker is an affix (typically marking subject or possessor), and therefore need have no connection with regular 3rd person forms.

Of course, many languages simply extend the use of ‘3.PL’ or 2nd for impersonal function (Eng. *They say ...* and *You can't beat that!*). In some African languages, the impersonal use of 2nd person forms is highly developed, to the point where a noun meaning ‘person, man’ regularly takes 2nd person agreement in its impersonal use. Consider these examples from Koyra Chiini (Songhay of Timbuktu, Mali), where the noun stem *doo* functions as a postposition meaning ‘at [\_\_’s] house’ (cf. Fr. *chez*):

- (11) (a) *boro o kaa ni doo*  
man IPFV come 2.SG chez  
'One<sub>x</sub> comes to one<sub>x</sub>'s house.'
- (b) *boro o kaa ngu doo*  
man IPFV come REFL chez  
'A man<sub>x</sub> comes to his<sub>x</sub> (own) house.'

- (c) *boro o kaa a doo*  
 man IPFV come 3.SG chez  
 ‘A man<sub>x</sub> comes to her<sub>y</sub> house.’
- (d) *ni o kaa ni doo*  
 2.SG IPFV come 2.SG chez  
 ‘You come to your house.’

(11a) could also mean ‘Someone comes to your house’, which is close to its literal translation, but commonly the ‘2.SG’ element is interpreted as coindexed to ‘man’, the combination indicating impersonality (generalized reference). This construction differs from that with indefinite (but specific, nongeneralized) ‘man’ as subject, where the postpositional object is coreferential (11b) or noncoreferential (11c) to the subject. It also differs from the fully 2nd singular type (11d), which is not (specifically) impersonal. In a span of several sentences with impersonal subject, the generalized reference is established by one opening sentence of type (11a), with the noun ‘man, person’ as subject and with ‘2.SG’ agreement; following sentences may be of type (11d).

Conversely, true impersonal categories may be used in place of 1st or 2nd persons. While impersonals are most naturally associated with the 3rd person (the residual “non-person”), the very generalizability or indeterminacy of their reference means that inclusion of speaker and/or addressee cannot be ruled out. As a result, impersonals can be used as pragmatic substitutes for 1st or 2nd person. The French type *On arrivera* is commonly used instead of ‘1.PL’ *Nous arriverons* ‘we will arrive’. In Chinook (northwestern United States), an impersonal-subject prefix replaces the usual ‘1.PL’ marker in the presence of a 2nd person object marker.

A superficially similar case is that of **dummy** pronominals with no apparent referent, as with English weather verbs (*It rained*) and some idiomatic “transitives” (*Let’s cool it!*). Such pronominals, typically in (inanimate) ‘3.SG’ form, fill obligatory slots and are best referred to the syntax. They are further evidence of the residual, unmarked status of the 3rd person. However, one should be wary of crosslinguistic generalization; in some languages the ‘3.SG’ pronominal subjects of weather verbs are referential.

#### 4. Interactions with other categories

##### 4.1. Gender and noun-class

**Gender** and **noun class** are more widely marked in the 3rd person than in the speech-participant persons. While speaker and ad-

dressee may be unambiguous in a given speech situation, additional information is required to guarantee referential clarity when multiple nonparticipant referents may occur in a discourse span. One strategy is to add gender or noun-class features to 3rd person – especially ‘3.SG’ – pronominals, increasing the probability that each pronominal in the span will be unique (1.3). Hence English distinguishes gender (*he, she, it*) only in the 3rd singular. Some Australian and African languages have noun classes distinguishing various types of animals and vegetation, but singular 1st and 2nd persons are largely unaffected. (For nonsingular 1st and 2nd persons, see 4.2, below.)

Some languages (such as Arabic) mark gender for 2nd and 3rd but not 1st persons. Basque marks gender (and affective diminution) only in ‘2.SG’ pronominal suffixes, which are often merely allocutive in function. Gender marking of the 2nd person is limited to singular vocatives in Gimira and other Omotic (and Cushitic) languages (Breeze 1986: 49). Marking 2nd person gender is usually best understood not as referential disambiguation, but as overt sociolinguistic marking.

#### 4.2. Number

Although number is nearly always present in pronominal paradigms, the categories may differ from nouns to pronouns and/or between 1st, 2nd, and 3rd persons.

Since the 3rd singular category is residual and open-ended, its plural is semantically straightforward: two or more nonparticipants = ‘3.PL’. Of course, a ‘3.PL’ pronoun may be used with generalized reference (‘They say ...’), as an alternative to an explicitly indefinite pronoun (Eng. *one*, Germ. *man*, Fr. *on*). There are other nuances of 3rd plural usage in some languages (for example, it may be limited to humans or certain animates, while the “3rd singular” is extended to nonsingulars of lower animacy), but these features are usually shared with the plural of nouns (see Art. 100).

A specifically pronominal number crossover occurs in English (where gender-neutral *they* corresponds to singular *he, she*, or *it*). One solution to the *he-or-she* problem of gender agreement for impersonal expressions (*If someone comes, tell him/her to wait*) is to use a ‘3.PL’ pronominal (... *tell them to wait*). Hybrid forms like *themself* instead of *themselves* (as in *Somebody shot themself*) are also

heard, suggesting that the suppletive ‘3.PL’ pronoun is more acceptable than overt plural suffixation.

The 2nd plural is semantically problematic, since it may denote two or more addressees, or one addressee plus at least one nonparticipant. The two readings are almost never obligatorily differentiated in surface linguistic form (but see 3.2, on Abkhaz). This means that in a three-party conversation, a speaker need not alternate between an explicit multiple-addressee pattern and an explicit addressee-plus-nonparticipant pattern (treating one interlocutor as an eavesdropper), a distinction that could be pragmatically dangerous and that would have no reference-disambiguating value. English goes farther by merging singular with plural only in the 2nd person (*you*, but cf. reflexive *yourself* vs. *yourselves*), eliminating social-delicacy problems in the use of 2nd person forms in multiparty conversations. This neutralization rarely causes referential ambiguity, since dual *you two* and the increasingly gender-neutral plural *you guys* are available as explicit nonsingulars.

In languages without a separate inclusive, there is always a category that we may label “1.PL” that represents any combination of the speaker with one or more other referents (speaker plus addressee, speaker plus nonparticipant, speaker plus addressee plus nonparticipant, etc.). For purposes of pluralization, then, we can speak of a hierarchy speaker > addressee > nonparticipant (or ‘1.SG’ > ‘2.SG’ > ...), such that the highest-ranking singular component determines the person of the plural (Zwicky 1977). However, the tendency to have suppletive person markers in the plural is especially strong in the 1st person (Forchheimer 1953: 65). This can be interpreted in various ways; one approach worth considering is that a suppletive ‘1.PL’ mitigates the overt and (hence) pragmatically dangerous egocentricity of a 1st plural containing the ‘1.SG’ form, particularly in contexts involving reference to both speaker and addressee.

In a language that distinguishes inclusive from exclusive, pragmatic difficulties involving joint speaker/addressee reference are resolved. ‘1.EX’, denoting speaker plus at least one nonparticipant, may be morphologically related to ‘1.SG’, while ‘1.IN’ tends to be formally autonomous from the 1st exclusive and 2nd persons. Since the inclusive, denoting at least speaker plus addressee, minimally in-

cludes two referents, the 1st inclusive dual corresponds to other duals in objective reference, but to other singulars in terms of relative quantity within its own (‘1.IN’) subparadigm. Typologically, ‘1.IN.DU’ may be treated as a morphological “singular”, as in Tab. 96.1 (a), in which case its morphological “dual” denotes three individuals. Alternatively, ‘1.IN.DU’ may be treated as a nonsingular: dual if that category exists in the language Tab. 96.1 (b), otherwise plural; see Greenberg (1988). In Tab. 96.1, X, Y, and Z represent person-marking morphemes, and the dash – indicates absence of a uniquely TRIAL pronominal.

Person	Actual Number				
	1	2	3	≥ 3	
(a)	1.EX	X	X-DU	–	X-PL
	1.IN	(n.a.)	Y	Y-DU	Y-PL
	2nd	Z	Z-DU	–	Z-PL
(b)	1.EX	X	X-DU	–	X-PL
	1.IN	(n.a.)	Y-DU	–	Y-PL
	2nd	Z	Z-DU	–	Z-PL

Tab. 96.1: Morphological treatment of ‘1.IN’

McKay (1978; 1979) suggests replacing “singular, dual, plural” with “minimal, unit-augmented, augmented” in type (a), where ‘1.IN.DU’ patterns as a singular (and there is a ‘1.IN.TRL’ patterning as a dual). However, linguistic terminology is sufficiently forbidding without this inelegant addition, and even within a language the ‘1.IN.DU’ may combine singular and nonsingular morphological features. Thus, Nunggubuyu (Australia) has a basically (a) system, but a dual suffix is tacked onto the independent stem for ‘1.IN.DU’. Adjectives agreeing with ‘1.IN.DU’ are always dual (or plural) in form, regardless of the morphology of the ‘1.IN.DU’ pronominal itself.

Plurality also interacts with person in certain syntactic situations where a singular referent X is compared to a group XY that strictly includes it. This is a significant implementational issue in the functioning of **switch-reference** systems, where asymmetrical assignment of “same” vs. “different” depending on the person categories involved has been used to establish a person hierarchy (4.5, below). It is also an issue in the functioning of logophorics (3.4); a sense like *He<sub>x</sub> said, “we<sub>xy</sub> will come”* is expressed variously

in logophoric languages as *He said, we [will] come* or as *He said, LOG-PL will come* (Hagège 1974).

In some Australian languages, nonsingular (especially dual) pronouns are cross-cut by very broad descent categories, e.g., “of same patrilineal moiety” vs. “of opposite patrilineal moieties” (hence “**harmonic/disharmonic**,” Hale 1966). Often the distinctions are the same for all pronominal persons, but in Panyjima only 1st/2nd person forms are affected, and “harmony” interacts in complex ways with ‘1.EX’, ‘1.IN’, and 2nd person categories (Dench 1980–1982).

#### 4.3. Tense, aspect, mood

There seems to be no categorial interaction between person and verbal aspect, but there are interactions between person and either tense or mood.

Imperatives and requests are obviously most natural with the addressee as subject. The hortative mood is predominantly associated with the inclusive person (‘Let’s go!’). In Eskimo, inclusive is distinguished from exclusive only in “imperative/optative” moods (Fortescue 1984: 253). Languages with a specifically hortative mood (as opposed to a broad irrealis) can manage without an inclusive/exclusive distinction, since ‘1.IN’ reference can usually be inferred from hortative. On the other hand, a realis assertion (past or present tense) is most natural with 1st (or 3rd) person subject. Since a reasonably sentient addressee normally knows his/her own recent behavior better than others, assertions like ‘You went there yesterday’ are usually uninformative or accusatory, unlike ‘I went there yesterday’. Conversely, realis interrogatives are somewhat more natural with 2nd person subjects (‘Did you go there yesterday?’).

In languages with high-frequency evidential markers, the confident evidential is appropriate for reports of the speaker’s own inner states, while hedged evidentials are required for reports about other referents (including addressees). On the other hand, if the confident evidential specified ‘eyewitness’ reporting, it may be superfluous in reports of the speaker’s own past actions, and in this event it may be reduced to occasional emphatically assertive uses, as noted for Mari-copa by Gordon (1986: 78). There are also evidential markers, like Aymara *-pi* (Hard-

man 1986: 121) for knowledge that the addressee should share (cf. Eng. *As you know*, ...).

These asymmetries are frequently grammaticalized. Countless languages have a distinctive imperative verb form limited to 2nd person; this may lack the usual 2nd person markers, and may therefore be syncretic with a 3rd person realis verb form (Span. *come* ‘he/she eats’ or [you] eat!). Languages with distinct ‘1.IN’ forms can simply use an irrealis (or future) form in hortative function, and are thus unlikely to have a separate, explicit hortative marker. In polite Japanese speech, the avoidance of overt pronouns is partly compensated for by careful marking of yes/no interrogativity; *went yesterday yes/no?* with (polar) interrogative particle is typically interpreted as ‘Did you go yesterday?’, not ‘Did I...?’. Japanese and other languages also have strongly grammaticalized correlations between person and specific evidential categories.

#### 4.4. Masking (deformation) in pronominal combinations

Perhaps the most interesting interactions involving person are in **bound combinations** of two or more pronominals. Such combinations occur in many polysynthetic languages in which both subjects and objects are represented in agreement subsystems in the verb, or in clitic clusters. (In 4.5, below, we consider linear and other ordering relationships among pronominals, not involving alterations to their forms.) When the subject and object markers are contiguous or otherwise interact closely, important categorial shifts and neutralizations may take place; we can profitably think of them as analogous to chemical reactions. Because the details are invariably language-specific, the generality of the phenomenon has not been recognized.

For our purposes, we may concentrate on **subject-object** person combinations (disregarding whether the “object” is direct [accusative] or dative), represented with arrows leading from subject to object:  $1 \rightarrow 2$ ,  $1 \rightarrow 3$ ,  $2 \rightarrow 3$ , and reciprocally  $3 \rightarrow 2$ ,  $3 \rightarrow 1$ , and  $2 \rightarrow 1$ . The most widespread discrepancy between expected and actual morphological patterns is that expected  $1 \rightarrow 2$  and  $2 \rightarrow 1$  (i.e.,  $1 \leftrightarrow 2$  with double-headed arrow) combinations do not take their expected form (which would contain ‘1.SG’ or ‘1.EX.PL’ and a 2nd person pronominal). Instead, one of the following deformations is applied:

- (a) phonological or allomorphic **disguise** of one or both person markers;
- (b) **zeroing** or **omission** of one or both person markers;
- (c) use of a **portmanteau** not decomposable into recognizable person markers;
- (d) **depersonalization** – replacement of a 1st or 2nd marker by a 3rd person or impersonal marker;
- (e) **inclusive conflation** – replacement of a 1st (exclusive) and/or 2nd person marker by an inclusive ('1.IN') marker;
- (f) **depronominationalization** – substituting an otherwise nonpronominal affix or particle for a particular person marker.

Despite the myriad of formal mechanisms involved, all of these devices involve **masking** (veiling or disguising) the “objective” 1 → 2 or 2 → 1 subject/object relationship. The point of doing so is that such combinations are “dangerous” pragmatically, since they are asymmetrical predication with one participant acting on the other. The important thing is not which veil or disguise is used, rather the mere fact that a transparent subject/object combination is avoided.

This phenomenon occurs widely in Amerindian, Australian, Tibeto-Burman, and other languages with tightly-knit pronominal complexes. Australian examples (from Heath 1991):

- (a) Ngandi ‘1.SG → 2.PL’ and ‘1.PL → 2’ *gura-* (/gur-na-/), with irregular ‘1.EX.PL’ allomorph /gur-/;
- (b) Kunwinyku ‘1.SG → 2.SG’/‘2.PL’ Ø- (neither ‘1.SG’ nor any 2nd person pronominal elsewhere has zero form);
- (c) Alawa ‘1.EX.PL → 2.SG’ *añi-* (undecomposable portmanteau);
- (d) Kunwinyku ‘2 → 1.SG’ *gan-* (identical to ‘3.PL → 1.SG’ and containing an element *ga-* otherwise associated with 3rd person);
- (e) Jawony ‘2.SG → 1.EX.PL’ and ‘2.PL → 1’ *ñanu-* (containing ‘1.IN.PL’ *ña-* and ‘2.PL’ *-nu-*);
- (f) Jawony ‘1.SG → 2.SG’ *wal-* (special use of an irrealis mood marker).

A further deformation is using ‘1.EX.PL’ and/or ‘2.PL’ for expected ‘1.SG’ and ‘2.SG’, respectively, in 1 ↔ 2 combinations (systematic in Ngandi). This is not, strictly, a change in person category, but we have previously noted that plural person markers, particularly for the 1st person, are often morphologically suppletive.

Pronominal combinations therefore show some of the same “pragmatic” distortions elsewhere associated with independent pronouns and pronoun substitutes. This is despite the fact that such clusters are highly fused and grammaticalized.

#### 4.5. Ordering and precedence relations

In addition to the deformations described in the preceding section, combinations of pronominals may have special ordering or (other) **hierarchical asymmetries**.

To the extent that **linear order** is determined by grammatical relations (e.g., subject and object), no further analysis is required. However, in some languages the order of pronominals is fully or partially determined by their semantic features, especially person. The major split is between 1st/2nd and 3rd. In situations where 1st and 2nd must compete for initial ordering, or for a single morphological slot, sometimes 1st is on top of the hierarchy as in some Australian orderings (Wurm 1969), but sometimes 2nd is (Algonquian prefixal slot). Since 1 ↔ 2 combinations are sociolinguistically delicate and are subject to masking (cf. 4.4), simplistic conclusions about the cognitive hierarchization of 1st and 2nd persons should be avoided.

Person-based linear ordering occurs in **direct-inverse** systems, found in more or less full form in Algonquian, Tibeto-Burman (DeLancey 1981), Chukchee (Siberia), and north central Australia (e.g., Ngandi). Here the linear order of subject and object markers depends on pronominal (not case) categories, so that (schematically) both ‘He saw me’ and ‘I saw him’ take the form 1.SG-3.SG-saw. Ambiguities are avoided by various patch-up mechanisms, usually involving a special inverse morpheme inserted among the pronominals where a low-ranking subject acts on a high-ranking object (1.SG-INV-3.SG-saw = ‘He saw me’). Direct-inverse systems locate the more topic-worthy pronominals in salient (usually leftmost) positions, but discourse topicality as such is not centrally involved. In most cases 1st and 2nd persons form one hierarchical set, preceding 3rd person pronominals; further hierarchical divisions involve cross-cutting features (number, animacy, obviation). There is no consistent typological ordering of 1st vis-à-vis 2nd person in these systems. Even Algonquian languages, which have a 2nd > 1st > 3rd hierarchy for the single initial prefix position, do not use the inverse morpheme /-ekw-/ in 1 ↔ 2 forms.

Person, along with animacy features and (occasionally) number, may influence morphological **case** systems or may be a factor in clause-level relational syntax (passive, antipassive). In Australian languages that have independent (not affixed) pronominals, a common pattern is that high-animacy categories have a (nominative-)accusative case system, while low-animacy categories have an ergative(-absolutive) system. Although the split may occur inside 3rd person (i.e., among types of nouns), in some languages the split is defined in terms of person, usually 1st/2nd vs. 3rd, the latter including nouns (Silverstein 1976). Since “nominative” and “absolutive” are unmarked, the effect is that the unmarked transitive sentence is of the type ‘I saw him’ (1.SG-Ø saw 3.SG-Ø), while ‘He saw me’ is highly marked morphologically (3.SG-ERG saw 1.SG-ACC). In some languages there is an intermediate subsystem between the (pure) accusative and (pure) ergative subsystems. In the Giramay dialect of Dyirbal (Dixon 1972: 50), the nonsingular 1st/2nd person pronouns are the accusative subsystem, ‘1.SG’ and ‘2.SG’ are intermediate (mixed), while nouns and demonstratives (which can act as 3rd person pronouns) are predominantly of ergative type.

Careful studies of switch-reference systems, which indicate “same” or “different” subjects of adjoining clauses, have brought out delicate person asymmetries in the treatment of strict-inclusion overlaps (e.g., ‘X came, and X+Y sat’). Reesink (1987: 201 f.) presents data from Usan and Kewa (Papuan) showing that ‘2.SG’ (but not ‘3.SG’) is treated as “same” as a following ‘2.PL’, and that ‘1.SG’ (but not ‘2.SG’) is treated as “same” as a following ‘1.PL’. This suggests a hierarchy 1st > 2nd > 3rd, with a low-ranking singular person not considered “same” as a higher-ranking plural person.

Topicality asymmetries of the 1st/2nd vs. 3rd type also affect passives (and antipassives) in many languages. Basically, these relational rules do for clause-level syntax what direct-inverse systems do for bound pronominal clusters in verbs, viz., they emphasize topicality or subjecthood for the 1st/2nd person marker and decrease them for the 3rd person marker. (In some cases, as with the Chukchee “antipassive,” it is not easy to tell whether we are dealing with a clause-level syntactic rule or a verb-internal morphological rule.) In English, the passive is common with verbs like *surprise* that would otherwise

often take (abstract or nonhuman) 3rd person subject and a human (hence often 1st/2nd person) object: *The fire surprised me* → *I was surprised by the fire*. In other languages, such as Navaho, person asymmetries trigger passive or antipassive rules more strictly.

#### 4.6. Spatial deictics

**Demonstrative** pronouns (*this, that*) and demonstrative adverbs (*here, there*) probably always have at least some association with person categories, since the speaker’s here-and-now is the most common deictic center. The speaker’s position is basic not only for ‘this/that’ distinctions but also for **vertical deixis** (‘above, level, below’), as in Huallaga Quechua (Weber 1989: 38). In English, the connection is weakly developed, but we may mention the replacement of a pronoun by a spatial in the colloquial *Give it here!* (meaning ‘Give it to me!’), and the occasional bipartite vocative type *you there!*. Since person systems start with three values (including the residual 3rd person), a close fit with spatial deictics is most likely when the latter system likewise has three basic values (excluding purely anaphoric demonstratives): proximal, medial, and distant.

In Basque, for example, the *you there!* type is more common than in English. The proximal, medial, and distant demonstrative pronouns are commonly added to or substituted for 1st, 2nd, and 3rd person pronouns, respectively.

#### 4.7. Formal, honorific and diminutive pronouns

In addition to many cases where noun- or noun-phrase-like expressions replace ordinary pronouns (*Yours truly* for *I*; *Her Majesty* for *you*), there are cases where formal/honorific and/or diminutive pronouns are morphologically related to other pronouns. Formality is a complex function of situation and hierarchy (Brown & Gilman 1960 and much subsequent literature). Overt marking of affect or formality is concentrated in 2nd (e.g., vocative or allocutive) and/or 3rd persons, but it can also be marked in the ‘1.SG’ (Indonesian). One possibility for formal (polite) 2nd singular is to use a ‘2.PL’ stem even for singular reference. In this event, true 2nd plural reference is at least optionally differentiated by adding another plural affix, by reduplication, or some other device, as in Gimira (Omotic; Breeze 1986: 49), Basque, and French (*vous* versus *vous autres*). In ex-

treme cases (associated with royal courts), pronouns are frequently omitted, but speaker versus address reference is nonetheless inferable from the use of humble or honorific vocabulary (speech levels, see 2.2).

## 5. Geographical distributions

In the following cursory survey, we indicate the distribution of the major typological variables. We assume a broadly universal skeletal system including ‘1.SG’, ‘1.PL’, 2nd, and (residual) 3rd persons. We focus on the presence of typological variables: inclusive, impersonal, obviative, logophoric, and categorial deformations in subject/object combinations (4.4).

### 5.1. Native America

The languages tend to be polysynthetic, with extensive pronominal affixation on verbs (or pronominal clitics). Inclusive occurs in many families (Jacobsen 1980), as do impersonal and obviative. In languages with bound subject and object pronominals, categorial deformations are moderately common.

The Algonquian languages (Bloomfield 1946) of the U.S. and Canada have the classic obviative system (3.3), which articulates with a direct-inverse system for pronominal subject and object markers. There is a 2nd > 1st > PROXIMATE > OBVIATIVE hierarchy for access to the single prefixal slot (subject or object for verbs, possessor for nouns). Inclusive is expressed by combining the 2nd person prefix with a special nonsingular 1st person suffix.

Athapaskan can be typified by Tututni in Oregon (Golla 1976). The basic system is 1st/2nd/3rd intersected by SG/PL, but there is also a distinct impersonal form in most pronominal series, and in 3rd → 3rd combinations there are special 3rd obviative (SG and PL) object forms. Furthermore, ‘1.PL’ and ‘2.PL’ are formally merged in possessive and object (but not subject) markers.

For the Mayan family, we mention Mam (England 1983), where person is expressed most reliably in verbal agreement, which has ergative and absolutive series. The pronominal prefixes themselves merge 2nd with 3rd person (e.g., ‘2.PL’/‘3.PL’ *ky-* or *chi-*), and do not have a basic inclusive marker. However, an enclitic *-(y)a* is added opportunistically, distinguishing 2nd person from (unmarked)

3rd, and distinguishing exclusive from (unmarked) inclusive; the enclitic is also used redundantly with ‘1.SG’.

### 5.2. Australia

Australian Aboriginal languages typically have elaborate personal prounoun systems including inclusive forms and a SG/DU/PL number system, but lack impersonal, obviative, and logophoric pronominals. In most of the continent, the main prounoun forms are independent or loosely cliticized rather than affixed, though some have affixed possessive pronominals on nouns. Subject and object pronominals are usually case-marked (nominative vs. accusative); in some languages they can fuse together in clitic bundles. In the north and the northwest, however, subject and object pronominals are typically prefixed to verbs, and here we find most of the deformations discussed in 4.4.

### 5.3. Pacific

The Austronesian languages examined have inclusive and full sets of person/number combinations, but no impersonal, obviative, or logophoric. In Kilivila (Senft 1986), the inclusive dual (‘you and I’) is also used for generalized reference (‘one says ...’).

Papuan languages have a basic 1st/2nd/3rd person system, in a few cases augmented with an inclusive. Many of the languages have idiosyncratic person syncretisms, especially 2nd = 3rd, in particular morphological contexts. In Hua (Haiman 1980: 47f., 199), ‘2.PL’ and ‘3.PL’ are broadly merged. In Usan (Reesink 1987: 87), 2nd and 3rd persons are merged in the set of “medial verbs” which indicate that the following sentence begins with a different subject (this switch-reference helps avoid ambiguities). More examples of syncretism are mentioned in Foley (1986: 66–74).

### 5.4. East and Southeast Asia

Such languages as Japanese, Korean, Javanese, and Thai have lexical and grammatical registers known as speech levels, which reflect the elaborate etiquette of present or former court environments (subsequently adopted by government and business elites). Usage of these registers depends on relative social position and various contextual features. These languages offer a variety of sociolinguistically terraced choices for ‘1.SG’ and ‘2.SG’ pronominals. For ‘1.SG’, Japanese men have *watakusi*, *watasi*, *boku*, and *ore*, in roughly descending degree of formality.

Women normally use *atasi* instead of *ore*, and stretch *watasi* to fill the position of *boku* (see Shibatani 1990: 371). Languages like Thai, Burmese, and Vietnamese have large repertoires of noun-like 1st and 2nd singular “pronouns”, some of them used only at court for shades of status differential (long lists in Cooke 1968). However, in actual discourse the frequency of ‘1.sg’ and ‘2.sg’ pronouns can be low, since avoidance of overt reference to speaker or addressee (as such) is favored. Instead, several grammatical features that have other primary functions, but “implicate” speaker or addressee reference, are systematically marked: interrogative vs. assertive, honorific (addressee) vs. humble (speaker), direct vs. indirect evidential, etc.

In these languages, it may be difficult to express inclusive (‘you and I’), since the speech-level system works best when speaker/addressee differentiation is sharp. Some languages have special 3rd person or reflexive pronominals that resemble African logophoric forms in usage.

Most Tibeto-Burman languages (Laos to India) have simple person systems, only a minority having inclusive. In two areas (Nepal and northwestern India) there are compressed subject/object combinations showing some categorial deformations (cf. 4.4).

### 5.5. South Asia

The languages of the area have basic 1st/2nd/3rd person systems. In general, inclusive is found in Dravidian and Munda languages but not in Indo-Aryan, though there are a few exceptions.

### 5.6. Near East and North Africa

Afroasiatic (Hamito-Semitic) languages typically have 1st/2nd/3rd person systems cross-cut by number and (in 3rd and often 2nd persons) gender, without morphological inclusive, impersonal, obviative, or logophoric categories. Subject and object pronominals are morphologically distinct, with few cases of categorial fusion. Some Afroasiatic languages in northeastern Africa (Chadic, Cusitic, Omotic, Ethiopian Semitic) have inclusive and logophoric.

### 5.7. Europe and Central Asia

Indo-European languages typically have simple 1st/2nd/3rd person systems, without a full-fledged inclusive, and without obviative or logophoric forms. Verbs usually have conjugations involving pronominal subject

markers, but object markers are usually a separate subsystem (e.g., clitics). Several languages have an additional impersonal pronominal category (Eng. *one*, Fr. *on*, Germ. *man*), taking ‘3.sg’ agreement (*One has to be careful*).

Other languages or groups showing 1st/2nd/3rd systems include Basque (a relic non-Indo-European language) and Turkic.

The non-Indo-European groups of the Caucasus, of diverse genetic stocks, have distinct inclusive forms in some but by no means all languages. Some have complex subject/object pronominal combinations involving categorial deformation or omission.

### 5.8. Sub-Saharan Africa

The most distinctive feature of many African person systems is logophorics (3.4), which in some languages do double duty as 3rd person reflexive and in others are distinct forms. Inclusive is found in some groups (e.g., some Grasslands Bantu languages), but is far from universal. (Sketches of pronouns in some seven languages or groups, excluding Afroasiatic, are in Wiesemann 1986, ed.).

## 6. Uncommon abbreviations

EX	exclusive
LOG	logophoric
IN	inclusive

## 7. References

- Akmajian, Adrian & Anderson, Stephen (1970), “On the Use of Fourth Person in Navajo, or Navajo Made Harder”. *International Journal of American Linguistics* 36, 1–8
- Aoki, Haruo (1986), “Evidentials in Japanese”. In: Chafe & Nichols (eds.), 223–238
- Armagost, James (1985), “Comanche *ma-*: Undistinguished Deictic, Narrative Obviative”. *International Journal of American Linguistics* 51, 302–310
- Becker, Alton & Oka, I. Gusti Ngurah (1974), “Person in Kawi: Explorations of an Elementary Semantic Dimension”. *Oceanic Linguistics* 13, 229–255
- Benveniste, Émile (1946), “Structure des relations de personne dans le verbe”. *Bulletin de la Société de Linguistique de Paris* 43, 1–12 [reprinted in: Benveniste, Émile (1966), *Problèmes de linguistique générale*. Paris: Gallimard, 225–236]
- Bhattacharya, Sudhibhusan (1975), *Studies in Comparative Munda Linguistics*. Simla: Indian Institute of Advanced Study

- Bloomfield, Leonard (1946), "Algonquian". In: Osgood, Cornelius (ed.), *Linguistic Structures of Native America*. New York: Viking Fund (Viking Fund Publications in Anthropology 6), 85–129
- Bohnhoff, Lee (1986), "Yag Dii (Duru) Pronouns". In: Wiesemann (ed.), 103–129
- Breeze, Mary (1986), "Personal Pronouns in Gimira (Benchnon)". In: Wiesemann (ed.), 47–69
- Brown, Roger & Gilman, Albert (1960), "The Pronouns of Power and Solidarity". In: Sebeok, Thomas (ed.), *Style in Language*. Cambridge/MA: MIT Press, 253–276
- Chafe, Wallace & Nichols, Johanna (1986, eds.), *Evidentiality: The Linguistic Coding of Epistemology*. Norwood/NJ: Ablex
- Clements, George (1975), "The Logophoric Pronoun in Ewe". *Journal of West African Languages* 10, 141–177
- Cooke, Joseph (1968), *Pronominal Reference in Thai, Burmese, and Vietnamese*. Berkeley: Univ. of California Press (Univ. of California Publications in Linguistics 52)
- DeLancey, Scott (1981), "The Category of Direction in Tibeto-Burman". *Linguistics of the Tibeto-Burman Area* (Berkeley) 6, 83–103
- Dench, Alan (1980–1982), "Kin Terms and Pronouns of the Panyjima Language of Northwest Western Australia". *Anthropological Forum* 5.1, 105–120
- Dixon, R[obert] M. W. (1972), *The Dyirbal Language of North Queensland*. Cambridge: Cambridge Univ. Press
- Donaldson, Tamsin (1980), *Ngiyambaa: The Language of the Wangaybuwan*. Cambridge: Cambridge Univ. Press
- England, Nora (1983), *A Grammar of Mam, a Mayan Language*. Austin: Univ. of Texas Press
- Foley, William (1986), *The Papuan Languages of New Guinea*. Cambridge: Cambridge Univ. Press (Cambridge Language Surveys)
- Forchheimer, Paul (1953), *The Category of Person in Language*. Berlin: de Gruyter
- Fortescue, Michael (1984), *West Greenlandic*. London: Croom Helm (Croom Helm Descriptive Grammars)
- Foster, Mary & Foster, George (1948), *Sierra Popoluca Speech*. Washington: U.S. Government (Smithsonian Institution, Institute of Social Anthropology, Publications 8)
- Fraiberg, Selma (1977), *Insights from the Blind: Comparative Studies of Blind and Sighted Infants*. New York: Basic Books
- Gerhardt, Julie (1988), "From Discourse to Semantics: the Development of Verb Morphology and Forms of Self-Reference in the Speech of a Two-Year Old". *Journal of Child Language* 15, 337–393
- Goddard, [R. H.] Ives (1967), "The Algonquian Independent Indicative". *National Museum of Canada, Bulletin* 214, 66–106
- Golla, Victor (1976), "Tututni (Oregon Athapaskan)". *International Journal of American Linguistics* 42, 217–227
- Gordon, Lynn (1986), "The Development of Evidentials in Maricopa". In: Chafe & Nichols (eds.), 75–88
- Gragg, Gene (1972), "Semi-Indirect Discourse and Related Nightmares". *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 8, 75–82
- Greenberg, Joseph (1963), "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". In: Greenberg, Joseph (ed.), *Universals of Grammar*. Cambridge/MA: MIT Press, 73–113
- Greenberg, Joseph (1978, ed.), *Universals of Human Language, Vol. III: Word Structure*. Stanford: Stanford Univ. Press
- Greenberg, Joseph (1988), "The First Person Inclusive Dual as an Ambiguous Category". *Studies in Language* 12, 1–18
- Hagège, Claude (1974), "Les pronoms logophoriques". *Bulletin de la Société de Linguistique de Paris* 69, 287–310
- Haiman, John (1980), *Hua, a Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam: Benjamins
- Hale, Ken (1966), "Kinship Reflections in Syntax". *Word* 22, 318–324
- Hardman, Martha (1976), "Proto-Jaqi: Reconstrucción del sistema de personas gramaticales". *Revista del Museo Nacional* (Lima) 41, 433–456
- Hardman, Martha (1986), "Data-Source Marking in the Jaqi Languages". In: Chafe & Nichols (eds.), 113–136
- Head, Brian (1978), "Respect Degrees in Pronominal Reference". In: Greenberg (ed.), 151–211
- Heath, Jeffrey (1991), "Pragmatic Disguise in Pronominal-Affix Paradigms". In: Plank, Frans (ed.), *Paradigms: The Economy of Inflection*. Berlin: de Gruyter, 75–89
- Hewitt, B[rian] G. (1979), *Abkhaz*. Amsterdam: North-Holland (Lingua Descriptive Studies)

- Ingram, David (1978), "Typology and Universals of Personal Pronouns". In: Greenberg (ed.), 213–247
- Jacobsen, William (1980), "Inclusive/Exclusive: A Diffused Pronominal Category in Native Western North America". *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 16.2 [Parasession on Pronouns and Anaphora], 204–227
- Lafon, René (1959), "Place de la 2e personne du singulier dans la conjugaison basque". *Bulletin de la Société de Linguistique de Paris* 54, 103–129
- Lichtenberk, Frantisek (1983), *A Grammar of Manam*. Honolulu: Univ. of Hawaii Press (Oceanic Linguistics Special Publication)
- Loveland, Katherine (1984), "Learning about Points of View: Spatial Perspective and the Acquisition of 'I/you'". *Journal of Child Language* 11, 535–556
- Mannheim, Bruce (1982), "A Note on 'Inclusive/Exclusive' in Sixteenth-Century Peru". *International Journal of American Linguistics* 48, 450–489
- McKay, Graham (1978), "Pronominal Person and Number Categories in Rembarrnga and Djeebbana". *Oceanic Linguistics* 17, 27–37
- McKay, Graham (1979), "Gender and the Category Unit Augmented". *Oceanic Linguistics* 18, 203–210
- Mühlhäusler, Peter & Harré, Rom (1990), *Pronouns and People: The Linguistic Construction of Social and Personal Identity*. Oxford, Cambridge/MA: Blackwell
- Oshima-Takane, Yuriko (1988), "Children Learn from Speech not Addressed to Them: The Case of Personal Pronouns". *Journal of Child Language* 15, 95–108
- Plank, Frans (1985), "Die Ordnung der Personen". *Folia Linguistica* 19, 111–176
- Rabain-Jamin, Jacqueline & Sabeau-Jouannet, Emilie (1989), "Playing with Pronouns in French Maternal Speech to Prelingual Infants". *Journal of Child Language* 16, 217–238
- Reesink, Ger (1987), *Structures and their Functions in Usan, a Papuan Language of Papua New Guinea*. Amsterdam: Benjamins
- Rumsey, Alan (1989), "Grammatical Person and Social Agency in the New Guinea Highlands". *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 25.2 [Parasession on Language in Context], 242–253
- Schaub, Willi (1985), *Babungo*. London: Croom Helm
- Senft, Gunter (1986), *Kilivila, the Language of the Trobriand Islanders*. Berlin, New York: Mouton de Gruyter
- Shibatani, Masayoshi (1990), *The Languages of Japan*. Cambridge: Univ. of Cambridge Press
- Silverstein, Michael (1976), "Hierarchy of Features and Ergativity". In: Dixon, R[obert] M. W. (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112–171
- Vairel, Hélène (1980), "La personne du nom: 3e personne et absence de personne". *Bulletin de la Société de Linguistique de Paris* 75, 267–283
- Watkins, Calvert (1962), *Indo-European Origins of the Celtic Verb, Vol. I: The Sigmatic Aorist*. Dublin: Dublin Institute for Advanced Studies
- Weber, David (1989), *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley: Univ. of California Press (Univ. of California Publications in Linguistics 112)
- Wiesemann, Ursula (1986, ed.), *Pronominal Systems*. Tübingen: Gunter Narr
- Wurm, S[tefan] (1969), "Person Marker Sequences in Australian Languages". In: Harris, Joy & Wurm, S[tefan] & Laycock, D[onald] (eds.), *Papers in Australian Linguistics, No. 4*. Canberra: Australian National Univ. (Pacific Linguistics A–17), 51–70
- Zwicky, Arnold (1977), "Hierarchies of Person". *Papers from the Annual Regional Meeting of the Chicago Linguistic Society* 13, 714–733

*Jeffrey Heath, Ann Arbor  
(U. S. A.)*

## 97. Classifiers

1. Introduction
2. Semantic categorization
3. A morphosyntactic typology
4. Dynamic dimension of classifier systems
5. Uncommon abbreviations
6. References

### 1. Introduction

Classifier systems are lexico-syntactic systems which provide an overt linguistic categorization of nominals. They come in different types that can be distinguished by their semantics, the size of their inventory, their morpho-syntactic status, and their pragmatic use. Classifier systems *per se* are part of a continuum of nominal categorization systems where they stand in an intermediate position between the two types of nominal classification systems found in European languages: the very grammaticalized gender systems (French *le fils* ‘the son’, *le couteau* ‘the knife’, *la fille* ‘the girl’, *la fourchette* ‘the fork’; cf. Art. 98) and the lexical expressions of measure terms and unit counters (*a piece of paper*, *a cup of milk*, *a handful of candies*, *a pile of clothes*). Classifier systems are found in languages of Asia, Oceania, Australia, Africa, and the Americas.

### 2. Semantic categorization

Classifiers offer a unique window into studying how human beings construct representations of the world and how they encode them into the words of their languages. Beyond a morphosyntactic variation to be discussed later, systems of classifiers share semantic features and categorization principles which yield remarkably similar prototypical members of categories across systems.

Classifier systems select basic universal semantic features of the major word classes from which they originate (mostly nouns and verbs) and extend them metaphorically to overtly express the various types of social, physical and functional interactions that human beings have with their world. The conventionalized perception of the world encoded in classifier systems spans from the most cognitive and universal to the most specific and esoterically cultural points of view.

### 2.1. Universal semantic properties of classifiers

The earlier attempts at identifying universal semantic properties of classifiers are found in Adams & Conklin (1973), Denny (1976), and Allan (1977). Their complementary approaches offer a comprehensive overview of all the semantic features encountered across classifier systems of the world. What follows is a synthesis of their work which takes into account a greater variety yet of classifier systems.

**Animacy** is one of the prevalent semantic features of classifiers. Some languages grant special classification to humans and animals as opposed to objects of the inanimate world, while others treat humans separately and classify animals by their shape together with inanimates. The feature of humanness itself combines with one of two sets of semantic features. One set of features corresponds to the inherent properties of a person, such as sex, age, or kinship. The other corresponds to socio-cultural properties, such as social rank based on wealth, occupation, nobility or sacredness, and is sometimes labelled *features of social function*. Some constraints seem to hold between these sets of features. For instance, the prominent feature of sex does not determine a class by itself in classifier systems proper (as opposed to gender systems) and must combine with other features. Another constraint seems to be on relying either on inherent properties or on social factors when classifying humans, but rarely on both sets at once.

The inanimate domain is classified by a proliferation of semantic features divided between physical and functional sets. The two major features of the physical domain are **material** and **shape**, both cognitively oriented features. Basic material features classify objects by their essence, irrespective of the shape, consistency or use of the object. This type of semantic categorization is very semantically transparent but appears to be rare, having been documented primarily in noun classifier systems of the Americas. In such a classification, for instance, all plants and plant products are classified as plants, all animals and animal products as animals.

The domain labelled *shape* includes inherent and temporary physical characteristics of objects, such as shapes, consistencies, and

configurations. The predisposition of classifier systems to use specifically shape (rather than color, size, weight or smell) reflects the selectiveness found in how humans categorize objects of the world according to basic cognitive categories. The primary semantic features of inherent shape that prevail around the world are the one-dimensional long shape, the two-dimensional flat shape, and the three-dimensional round shape.

The semantic features of **consistencies** (rigid, flexible, soft, hard) are secondary developments; they do not categorize objects by themselves but combine with one of the primary shapes. The feature of rigidity, for instance, can combine with a primary shape feature, yielding the possible combinations: one-dimensional/long + rigid and two-dimensional/flat + rigid. Included under the broad label of shape are **configurations** (in a pile, in a circle, in a straight line, evenly and unevenly scattered, etc.) which are temporary arrangements creating shapes. They, too, have a tendency to combine with other semantic features of shape and consistency, as well as “scores” of other physical characteristics. The result is more and more specific classes which proliferate in the large numeral classifier systems.

The semantic feature of **function** refers to the use to which the objects classified are put rather than to their shape or other physical characteristic. A basic set of functions marked by classifiers would include housing, transportation, edibles, clothing, and tools. Within the domain of function fall most of the categorizations of abstract nouns, such as concepts of time, writing and speech, and much that is very cultural specific and opaque in classifier systems.

## 2.2. Types of classifiers and complexity of classes

The semantic structure of the class defined by a classifier varies from very simple to very complex and heterogeneous, according to the type of classifier that heads it.

**Specific classifiers** are the most common type. They head classes built around prototypical exemplar nouns from which the class generally extends into a more or less heterogeneous whole (see 2.3). The term *specific* is used here primarily to distinguish a type of classifier that covers a large spectrum between two extreme opposite types: unique and general classifiers. Specific classifiers correspond to various levels of specificity of no-

minal taxonomies, including generic and specific nouns. Commonly, the more specific the classifier, the more culturally relevant the class is assumed to be or have been. In Jacaltecan (Mayan) languages, the plant domain includes three classifiers: *te'* is for all the plant world, while *ang* is more specifically for medicinal plants, and *ixim* for corn and corn products.

**Unique classifiers** define the simplest classes, in that they classify only one object, supposedly one that had or still has some particular cultural relevance. In Jacaltecan (Mayan), there is a unique classifier *metx'* for the one noun *tx'i'* ‘dog’, next to the specific classifier *no'* for all the other animals. In Thai, *chyak* is the unique classifier of *čhaan* ‘elephant’ in formal, honorific context, next to *tu*, the specific classifier for elephant and other animals in less formal speech. In Yagua (Macro-Carib) *na* is the unique classifier of standing banana tree trunk and *mu* of standing chambira palm trunk, next to the specific classifier *nu* for all other tree trunks.

**General classifiers** head the most inclusive classes. They are semantically very bleached, generally distinguishing classes on the basis of animacy. They are used as substitutes for more specific classifiers under certain developmental and pragmatic conditions, such as earlier stages of language acquisition and casual adult speech. Increase in the use of general classifiers is also a trait of language loss under language contact with non-classifier languages. General classifiers are most common in large numeral classifier systems, such as Tzotzil, Chinese, Japanese. Of the sixty classifiers of contemporary Indonesian (Austronesian) three are considered to be “general” classifiers: *seorang* ‘human’; *seekor* ‘animal’; *sebuah* ‘the rest’.

Unique, specific, and general classifiers correspond to different stages of evolution of individual classifiers (see 4.1.2). Their existence underscores the extremely variable degree of semantic complexity of the classes of nouns defined by classifiers.

## 2.3. Heterogeneity of a class

The analysis of the semantic structure of heterogeneous classes of nouns headed by specific classifiers often reaches challenging proportions, as illustrated by the example of the *hon* classifier of Japanese, which applies, among other things, to: pencils, sticks, threads, ropes, needles, bananas, carrots, pants, guitars, teeth, cassette tapes, type-

writer ribbons, camera films, telephone calls, letters, movies, TV programs, medical injections, and homeruns in baseball! (Lakoff 1986: 14; Matsumoto 1990: 10f.)

Such heterogeneity is the combined result of various processes of extension which operate in the semantics of classifiers: prototype extension, chaining, and checklist. A checklist model predicts clear cut boundaries between classes and assumes extension of a category to any noun with a set of necessary and sufficient features, such as [+animate] and [-human]. A chaining model, the most prevalent in classifier systems, is based on sets of local analogies that create disjunctive classes with no identifiable common feature. In a prototype model, the members of the class are more or less closely resembling an abstract ideal member with which they share minimally one feature.

The class defined by the Thai classifier *tua* ‘animate quadrupeds’ has the kind of heterogeneity that requires appeal to all three types of extensions (Carpenter 1987: 17). A checklist analysis accounts for the inclusion in the class of all animals. A prototype analysis assigns ‘dog’, ‘cat’, and ‘buffalo’ a more central position in the class than non-limbed animals like ‘snake’ and ‘fish’ on one hand, and the limbed inanimates ‘table’ and ‘trousers’ on the other. But only a chaining analysis can account for the further inclusion in the class of ‘shirt’ (other limbed clothing), ‘dress’ and ‘bathing suit’ (unlimbed clothing). What all models of extension of the classes can do at best is to explain *a posteriori* the inclusion of certain nouns in a class, but they have very limited power of prediction (Allan 1977; Lakoff 1986).

#### 2.4. Complexity of classifier systems

Beyond the varying complexity of the individual classes of a classifier system is the varying semantic complexity of the systems at large. Classifier systems are heterogeneous, non-hierarchical, non-taxonomic organizations which vary idiosyncratically from language to language and culture to culture. Most classifier systems are collections of unique, specific and general classifiers which combine varied semantic categorization. The degree of complexity of classifier systems is also tied to their ability to classify all nouns, including abstract ones such as time expressions and activities.

It is very common in the large numeral classifier systems to have classifiers for an ar-

ray of physical semantic features of basic shape and secondary shapes and consistencies, scores of configurations and multiple functions, with a combination of a few general classifiers, more or less familiar specific classifiers, and a large number of very language specific unique classifiers. This is the state of most large numeral classifiers from Asia.

### 3. A morphosyntactic typology

The literature on the typological diversity of the phenomenon of overt nominal classification has attracted attention to various subsets of classification systems, but a full picture of the phenomenon has yet to be produced (Allan 1977; Craig 1986; 1987; Dixon 1982; 1986; Carlson & Payne 1989; Serzisko 1981). There is so far no agreed upon set of criteria to determine unambiguously which systems qualify as classifier systems per se and which do not. The typology presented here is a synthesis of the early seminal work on classifier systems and more recent case studies of classifier systems and takes an inclusive rather than exclusive approach to the rather fluid phenomenon of nominal classification.

#### 3.1. Rationale for the typology

Various factors contribute to the urgency of the task of establishing a typology. On one hand it could clear up a certain amount of terminological confusion. Numeral classifiers have been variously called *numerative*, *numerical* or *number classifiers* without much danger of confusion does not cause much problem (although the existence of a *number* classifier system distinct from numeral classifiers – admittedly rare but documented in some languages of Meso-America – should warrant standardizing the terminology. More importantly, there is a problem with the use of the expression *noun classifier* which has a variety of referents in the literature. Some use it to refer to all the classifier systems, including, but not restricting it to numeral classifier systems (Denny 1976), others use the expression to refer to numeral classifier systems themselves, while the existence of a type of noun classifiers distinct from numeral classifiers has been argued for (Craig 1986; 1987).

Another rationale for establishing a new **typology of classifiers** is precisely that a considerable amount of new classifier data and

classifier analysis has been produced since the eighties, both new case studies of specific classifier systems, particularly of Middle and South America, and typologizing and theorizing efforts using already published sources, such as the typological project in Cologne (Barron 1982; Serzisko 1981; H. Seiler 1986). A working typology incorporating the recent advances in classifier studies would have to take into account the rich data and amended proposals of such publications as Aikhenvald (2000), Grinevald (2000), and Senft (2000, ed.).

The typology being proposed here is morpho-syntactically based. Although it will consider issues of semantics, pragmatic use and grammaticalization, it identifies the different types of classifiers primarily on the basis of their morpho-syntactic locus. The terminology chosen for this typology responds to the following rationale: it relies as much as possible on currently used terminology in order to avoid the proliferation of new terms, while selecting among various terms in use the one that is most iconic with the morpho-syntactic locus of the classifier.

### 3.2. Nominal classification types

The following typology incorporated classification systems which are not universally considered as belonging to the core of the major classifier systems. The inclusiveness is dictated by a general approach to the study of classifier systems that includes both synchronic and diachronic dimensions and allows for tracing the paths of evolution of such systems, often relating minor or marginal systems to the core ones.

#### 3.2.1. Gender and noun class

Gender is either not included in classifier typology or considered an extreme case of noun class system. **Gender systems**, which are common in European languages, have either two or three classes ('masculine'/feminine'/'neuter') to which most nouns are assigned arbitrarily, beyond the recognition of sex differences for animate nouns (see Art. 98).

**Noun class systems** are more typical of languages of Africa and commonly have between five and twenty classes. More semantic content can be ascribed to noun classes than to genders, although it is more obvious for reconstructed Bantu classes than for most contemporary noun class systems:

(1)	class 1/2	humans
	class 3/4	plants
	class 5/6	fruits
	class 6	liquids
	class 7/8	inanimates
	class 9/10	animals
	class 10/11	elongated objects
	class 12/13	small objects
	class 14	masses
	class 15	infinitive nominalizations
	class 15/6	paired body parts

Gender and noun class systems share the characteristic of forming part of elaborate agreement systems, as illustrated with examples from Europe and Africa:

- (2) Gender in Spanish:

*esta flor roja es bonita*  
this.F flower(F) red.F is pretty.F  
'this red flower is pretty'

- (3) Noun classes in Sesotho (Central Bantu; Demuth et al. 1986: 456):

(a)	<i>mo-tho é-mo-holo ó-rata Ø-ntjá</i>	
	1      1      1      1      9	
	person big      he/she-like dog	
	<i>é-ntle      eá-hae</i>	
	9      9      1	
	beautiful of-his/her	
	'The old man/woman likes his/her	
	beautiful dog.'	
(b)	<i>ba-tho bá-ba-holo ba-rata li-ntjá</i>	
	2      2      2      2      10	
	people big      they-like dogs	
	<i>tsé-ntle tsá-bona</i>	
	10      10      2	
	beautiful of-them	
	'The old people like their beautiful	
	dogs.'	

A characteristic of these systems is that the gender/class may or may not be marked on the noun itself, and may have some derivational uses, such as locative, augmentative and diminutive, derogative or agentive, as well as infinitive (Heine 1982).

#### 3.2.2. Numeral classifiers

**Numeral classifiers** are the most common, the largest and the best known systems of classifiers and are called *numeral* because they appear contiguous to numerals. Morphologically speaking, they may be more or less bound and fused to the numeral:

- (4) Japanese (Matsumoto 1990: 1, 7)

*enpitsu ni-hon ; hon ni-satsu*  
pencil two-CL book two-CL  
'two pencils'      'two books'

- (5) Chinese (Li & Thompson 1981: 105)  
 $sān\text{-}gē$  rén ;  $nèi\text{-}tiáo$  niú;  
 three-CL person that-CL cow  
 ‘three people’ ‘that cow’  
 $nèi\text{-}liù\text{-}běn$  shū  
 that-six-CL book  
 ‘those six books’
- (6) Cabecar (Chibchan; Richards 1983: 6)
- | CLASS:  | ‘one’   | ‘two’  |
|---------|---------|--------|
| humans  | Péklá   | ból    |
| flat    | Pétká   | bótkö  |
| round   | Péklawö | bólwö  |
| long    | Pétabá  | bótabö |
| bundles | Péyeká  | bóykö  |
| trees   | Pélká   | bólká  |

In rare cases, classifiers are morphologically infixes as in Yagua (Macro-Carib), or realized as reduplication, as in:

- (7) Squamish (Salishan; Kuipers 1967):
- |                              |                   |
|------------------------------|-------------------|
| <i>nch'u?</i> + NUM.CL.INAN  | ‘one (inanimate)’ |
| <i>ninch'u?</i>              | ‘one (animal)’    |
| <i>nch'nch'u?</i>            | ‘one (human)’     |
| <i>t'akach</i> + NUM.CL.INAN | ‘six (inanimate)’ |
| <i>t'at'k'ach</i>            | ‘six (animal)’    |
| <i>t'ek'tak'ach</i>          | ‘six (human)’     |

Although labelled *numeral classifiers*, this type of classifiers may also appear on other elements than numerals, such as quantifying words and demonstratives.

A major distinction to be made among numeral classifiers is between sortal and mensural classifiers. **Mensural classifiers** are used for measuring units of both mass and count nouns (cf. Art. 101) and are a fairly large open-ended lexical class. Hundreds have been documented in classifier languages like Tzeltal, Thai, Chinese. They are familiar to everyone because they correspond to the measure terms of non-classifier languages, such as ‘a pound of tobacco’ (weight), ‘a slice of bread’ (shape), ‘a handful of tomatoes’ (contained measure), ‘a sheet/ream of paper’ (quanta), ‘a pile of wood’/‘a line of trees’ (arrangement). Mensural classifiers classify both count and mass nouns. **Sortal classifiers** do not have a direct equivalent in non-classifier languages. They are morphemes that specify units (not quantity) in terms of which the referent of the head noun can be counted, although they may be used in contexts other than quantification (see also Grinevald 2000: 58 f.). They often appear to be semantically redundant, expressing one of the inherent semantic characteristics of the head noun. They may refer to the essence of the object, as in ‘a man car-

penter’, ‘a woman teacher’, ‘an animal dog’, ‘a plant banana’, ‘a liquid river’, or to its shape: ‘a long tree/pencil/bone’, a ‘flat leaf/paper/sheet’, ‘a spherical orange/fist/baby’. They can also refer to its function: ‘a transportation boat’, ‘a drinkable fruit juice’. In some languages they can refer to the social status or kinship relation of humans: ‘honorable Mary’, ‘young male kin Peter’.

It has been argued on the basis of the existence of such sortal classifiers that nouns of classifier languages are neutral with respect to mass and count, being instead ensemble or concept nouns which are individuated by the presence of the classifier in discourse, as patterned below:

	number	individuated sortal classifier	ensemble head nominal
one	animal-unit	(of) dog (kind)	
two	flat-unit	(of) blanket (kind)	
three	round-unit	(of) orange (kind)	

The difference between measure terms and mensural classifiers is not always easy to draw. One defining criteria of mensural classifiers is that they co-exist in the language with sortal ones, in forming a complex and heterogeneous classifier system. By such a criterion English measure terms could not be taken as mensural classifiers, contrary to what some want to argue. In some languages, mensural and sortal classifiers behave differently. In Tzotzil (Mayan), for instance, of the several hundred numeral classifiers identified, only eight are sortal classifiers and both types can be distinguished by their anaphoric behavior: sortal but not mensural classifiers are used anaphorically. In Q'anjob'al (Mayan) the difference appears in agreement. Numbers carry a classifier which agrees with mensural classifiers, if one is present, but with the head noun if a sortal classifier is present (examples from Mayan languages are given in traditional orthography):

- (9) Q'anjob'al (Mayan; Zavala 1989: 282):
- (a) *ox-ep' tinan ep' naj winaj*  
 3-NB.CL MENS.CL PL.CL N.CL noun  
 INAN group HUM man man  
 ‘three groups of men’

- (b) *ox-wan k'itan ep' naj winaj*  
 3-NB.CL SORT.CL PL.CL N.CL noun  
 HUM separated HUM man man  
 'three separated men'

Numerical classifiers come generally in large inventories (from dozens to hundreds) and are found predominantly in Asia, but also in Oceania and the Americas. Some of the major systems of Asia are those of Thai (246 classifiers), Burmese (189 classifiers), Vietnamese (140 classifiers), although such large inventories include a majority of mensural classifiers (Goral 1978). Micronesian languages have from one to several dozen classifiers, while Chibchan languages of Central America like Cabecar (see (6)) have a small system, with a total of six classes.

Numerical classifiers have been the basic material for discussions of classifier systems and are central to any typology. The ones of particular interest are the sortal classifiers, which, unlike the mensural ones, do not have an equivalent in non-classifier languages and raise more questions because of their seemingly redundant semantic function.

### 3.2.3. Noun classifiers

**Noun classifiers** are a much rarer type. In fact, their existence as a distinct type has only been argued for in work on Meso-American languages (see Craig 1986; 1987; Zavala 1989 for descriptions of Q'anjob'alan Mayan systems; de León 1988 for one of Mixtecan systems). Their name comes from the fact that they are most intimately related to the noun in that their presence does not depend on the presence of another element of the noun phrase, such as a numeral or a demonstrative (see numerical classifiers in 3.2.2), or a possessor (see genitive classifiers in 3.2.4). Noun classifiers commonly stand alone with their referent noun, independent of quantification or possession. One of the major functions of noun classifiers is to serve as anaphoric pronouns for the referent nouns, as illustrated in the Jakaltek example below.

- (10) Jakaltek (Craig 1986: 264)

- (a) *xil naj xuwan no'*  
 see.PAST CL.man John CL.animal  
*lab'a*  
 snake  
 '(man) John saw the (animal) snake.'
- (b) *xil naj no'*  
 see.PAST CL.man CL.animal  
 'he saw it (animal)'

- (c) *caj te? tahnaj ixpix*  
 red CL.plant ripe tomato  
 'the ripe (plant) tomato is red'
- (d) *caj te'*  
 red CL.plant  
 'it (plant) is red'

- (11) Coatzoquitengo Mixtec (Otomanguean; de León 1988: 153):

- (a) *xáhni ñá María*  
 kill.PAST CL.woman Mary  
*ti xuhu*  
 CL.animal goat  
 'Mary killed the goat'
- (b) *xahni-ñá-rí*  
 kill.PAST-CL.she-CL.it (animal)  
 'she killed it'

The noun classifiers of Meso-America share several characteristics. They all include specific classifiers for humans, some being limited to them, and all have a prominent pronominal function. They also all classify inanimates primarily by material, a natural kind taxonomy which is not common in numerical classifier systems. Although they have not been identified as such in the literature, some noun classifier systems are found in Australia. Such is the case for the classifiers Dixon calls *generic markers* in Yidiny:

- (12) Yidiny (Australian; Dixon 1982: 185)

- mayi jimirr*  
 CL.vegetable(ABS) yam(ABS)  
*bama-al yaburu-Ngu julaal*  
 CL.person-ERG girl-ERG dig.PAST  
 'the (person) girl dug up the (vegetable) yam'

The identification of a distinct type of classifiers called here *noun classifiers* is practically the main point of the present typology. As already mentioned, not only is the existence of noun classifiers not always recognized, but, in addition, the term *noun classifier* is often used to refer to another type of classifiers, the numerical classifiers of 3.2.2.

### 3.2.4. Genitive classifiers

**Genitive classifiers** are a fairly well established type of classifiers. They have also been called *relational*, *possessive* or *attributive* classifiers. They resemble numerical classifiers in that they piggy-back an element of the noun phrase, in this case the possessor entity of a possessive construction (see Art. 103).

- (13) Ponapean (Micronesian; Rehg 1981: 184)
- (a) *kene-i mwenge*  
CL.edible-GEN/1 food  
'my food'
  - (b) *were-i pwoht*  
CL.transport-GEN/1 boat  
'my boat'

Genitive classifiers constitute one of the major typological characteristics of Micronesian languages where the inventories vary from the two classifiers of Manam (general and edible) to the twenty-one of Ponapean. They are also found in the Americas: in Yuman languages, they are minimal systems with as few as two classes (general and animal/pet). In Yucatec and Mam (Mayan) they are verbal in origin and examples given always deal with edibles. In Tuyuca (Tucanoan) the possessive construction is generally headless, with the possessor classifier attached to the possessor fulfilling a very widespread anaphoric role in discourse:

- (14) Tuyuca (Tucanoan; Barnes 1990: 286)
- (a) *bāriya-ya-da*  
Maria-GEN-CL.long.flexible  
'Mary's (string)'
  - (b) *kīt paki-ya-wi*  
3 father-GEN-CL.hollow  
'his father's (canoe/car/blowgun)'

In all languages, genitive classifier constructions are restricted to a subset of possessed nouns commonly labelled *alienable* (see Art. 103), while the possessive constructions of *inalienable* nouns do not use classifiers. What determines (in)alienability is not always easy to identify, so that the term *alienable* must be taken to refer to a particular grammatical category rather than a semantic one. The categorization of nouns into either the alienable or the inalienable class is a matter of ethnolinguistic research which would aim at drawing the list of which parts of the body, which kinship and social relations, as well as which objects of the world are considered alienable in a particular language. For instance, the list of alienable nouns in Ponapean includes edibles, drinkables, children-pets-domestic animals, vehicles, buildings, hunting bounties, pillows, as well as nouns of humans who are commonly taken to be inalienable in other languages, such as siblings, relatives, maternal uncles, nephews and nieces, clan members. Of the twenty-one genitive classifiers of Panare (Cariban), one is general, three are

the familiar edible, drinkable and vehicle, and the others include the idiosyncratic musical instruments, body paint and artificial light (Carlson & Payne 1989).

### 3.2.5. Verbal classifiers

**Verbal classifiers** are called thus because they are morphologically part of verb words. As systems of nominal classification they rely on the same array of semantic features as the previously described classifier systems. There are various sub-types of verbal classifiers, depending principally on the age of the system and the lexical source of the classifiers. The above mentioned terminological problem surfaces here too. One extreme case of it is found in the literature on Athapaskan languages in which verbal elements with no noun classificatory function have been traditionally called *classifiers*, while others that should be called *classifiers* have been described instead with terms such as *class mark* or *extensor* which hide their classificatory function (Krauss 1968). Verbal classifiers have been documented for many North American languages, as well as Australian and Papuan languages (Mithun 1986; W. Seiler 1986) and American Sign Language (Supalla 1986).

One sub-type of verbal classifiers are **incorporated classifiers** which are still recognizable as incorporated words, generally nouns. Whether specific and generic nouns originally, all take on a generic meaning as classifiers. The free nouns corresponding to incorporated classifiers may still co-occur in the language, or they may have been replaced:

- (15) Cayuga (Iroquian, Ontario; Mithun 1986: 386–388)
- (a) *ohon'atakte: ak-hon'at-a:k*  
it.potato.rotten PAST/I-CL.potato-eat  
'I ate a rotten potato'
  - (b) *so:wa:s akh-nahskw-ae'*  
dog I-CL.domestic.animal-have  
'I have a (pet) dog'
  - (c) *skitu ake'-treh-tae'*  
skidoo I-CL.vehicle-have  
'I have a car'
- (16) Munduruku (Tupian, Brazil; Mithun 1986: 381):
- ti dojot puye, o'-ti-mog*  
water bring when they-CL.water-place
  - ip baseya'a be.*  
they basin in
  - 'when they brought water, they placed it in the basin'

- (17) Ngandi (Australian; Heath 1978: 215; see also Mithun 1986: 389):

*gu-jaŋk-yuŋ*  
gu-water-ABS  
*ba-ga-bun-ŋu-ni*  
3.PL-SUB-CL.water-eat-PCON  
'and they drank water'

The semantics of incorporated verbal classifier systems vary from kinds of entities to qualities (shapes, functions) of objects. In the former they are closest to the semantics of noun classifiers, and in the latter to the semantics of numeral classifier systems. Intermediate systems classify by both kinds of entities and qualities.

The other sub-type of verbal classifiers are classifying verbal affixes which are phonologically very eroded but which have semantics similar to numeral classifier systems, commonly identifying classes of long, round, granular, flexible, liquid objects, for instance:

- (18) Diegueño (Langdon 1970: 80, 87; see also Carlson & Payne 1989):

- (a) *tu-kaṭ*  
CL.round-cut  
'to cut with scissors or adze, to cut in chunks'
- (b) *tu-mar*  
CL.round-cover  
'to cover over a small object'
- (c) *a-kaṭ*  
CL.long-cut  
'to cut with a knife'
- (d) *a-mar*  
CL.long-cover  
'to cover over a long object, to bury someone'

- (19) Imonda (Papuan; W. Seiler 1986: 192 f.)

- (a) *tōbtō kam u-aihu*  
fish me CL.small animal-give  
'give me the fish!'
- (b) *po kam i-aihu*  
water me CL.liquid-give  
'give me some water!'
- (c) *maluō kam lēg-aihu*  
clothes me CL.flat-give  
'give me a piece of clothing!'

One characteristic of these verbal classifiers is that they classify either the subjects of intransitives or the object of transitives, on an absolute basis:

- (20) Eyak (Krauss 1968: 195):

(a) *?u·d də-sətaħt*  
it CL.board-lie  
'it (board) lies there'  
*sič' də-Gəta?*  
to.me CL.board-give  
'give it (board) to me!'

(b) *?u·d xxədə-sətaħt*  
it CL.log-lie  
'it (log) lies there'  
*sič' xədə-Gəta*  
to.me CL.log-give  
'give it (log) to me'

Another characteristic of the more grammaticalized verbal classifiers, besides their more opaque semantics, is the fact that they are associated more or less stringently with certain verbs. The core set of these verbs deals with the concept of manipulation of objects, including the state and position they are in before or after manipulation, hence verbs like: 'lay', 'be in a position or in a specific place', 'handle', 'hold', 'grab', 'pick up', 'push', 'give', 'carry'.

The extreme case of phonological erosion and fusion of verbal classifiers is found in the phenomenon of classificatory verbs in which the shape or position of the subject or object argument is lexicalized into verbal stem paradigms. Cherokee for instance distinguishes five nominal classes through stem variation of a set of basic verbs of position and manipulation verbs:

- (21) Cherokee (Southern Iroquian; Mithun 1986: 392)

*gakaneha*  
'he's giving him a living thing'

*ganehneha*  
'he's giving him some liquid'

*adeha*  
'he's giving him a long, rigid object'

*ganvneha*  
'he's giving him a flexible object'

*ahneha*  
'he's giving it to him (something not contained in one of the above categories)'

The phenomenon of verbal classifiers therefore covers an array of sub-types, in a continuum in which the classifying elements vary from being still close morphologically and semantically to their lexical origins, to systems in which those elements have eroded both semantically and phonologically and have be-

come fused to the verb. The subtype of classificatory verb systems which stands at the margins of a typology of classifiers per se is included in the typology for the same reason gender systems were: they are important for a complete view of the evolutionary process of classifier systems.

### 3.2.6. Marginal types

Even more marginal yet, and not considered a part of the typology, is a type of nominal classification that operates through portmanteau morphemes, expressing deixis and elements of nominal classification. One such example is a classification system found in Toba (Macro-Panoan) which has been labelled *noun classifiers* in the literature. It consists of nominal prefixes which are more similar morphologically to a noun class system than to a noun classifier. They indicate notions of visibility and distance combined with some elements of position somewhat reminiscent of some classifier semantics (Klein 1979). Another marginal type semantically similar to the Toba system are so-called *article* systems of North American languages. One example is found in Squamish (Salishan) where articles combine gender with an elaborate deictic system (Kuipers 1967). Another example are the three articles of Mandan (Siouan) which mark three positions reminiscent of the three basic shapes of numeral classifiers: vertical or one-dimensional/long, horizontal or two-dimensional/flat, and sitting or three-dimensional/round (Barron & Serzisko 1982).

### 3.2.7. Prototypes and mixed types

Gender/noun classes and the different types of classifiers (numeral, noun, genitive, and verbal classifiers) have just been presented as if they were discrete types, although they represent more focal points on various continua than discrete types. In fact many instances of classifier systems do not fit squarely into any of the types considered and how some systems blend one into another through time will be discussed in 4. One of the difficulties in assigning some system to a particular type arises when the classifying elements have other functions than the strict classifier functions just presented, such as multiple inflectional and derivational functions. This is the case for instance in numerous languages of lowland South America (Aikhenvald 2000: 204–241; Derbyshire & Payne 1990; Payne 1987) that might be best reanalyzed as having

noun class systems at various stages of grammaticalization. In Munduruku, for instance, the same set of classifiers is found on verbs, nouns, numerals, demonstratives, and adjectives:

- (22) Munduruku (Tupian; Derbyshire & Payne 1990: 261):

(a) on verb

<i>bekitkit</i>	<i>ako-ba</i>
child	banana-CL.long
<i>o'su-ba-dobuxik</i>	
3-CL.long-find	

‘the child found the banana’

(b) on noun and number

<i>xepxep-</i> 'a	<i>wexik-</i> 'a
two-CL.round	potato-CL.round

‘two potatoes’

(c) on noun and demonstrative

<i>ija-ba</i>	<i>ako-ba</i>
this-CL.long	banana-CL.long

‘this banana’

The Munduruku system is therefore at once like a verbal classifier system – in that it classifies the absolutive argument of the verb –, like a numeral classifier system – in that it classifies by shape and is affixed to number and demonstrative –, and like a noun class system – in that it inflects directly on the noun and functions like an agreement system.

### 3.3. Arguments in support of the typology

With the above caveat that the typology does not mean discrete types of classifier systems, but rather recurrent prototypes, several arguments will be presented in support of the typology, in particular in support of a distinction between the major types found in the noun phrase: numeral, noun, and genitive classifiers.

#### 3.3.1. Argument 1: Co-occurrence of types

The strongest argument to be brought forth is the simple fact that several types of classifiers may co-occur in a single language. Micronesian languages like Ponapean, for instance, have both genitive classifiers and numeral classifiers. The most striking example of multiple classifier systems found within the same language is the case of the K'ajjobalan Mayan languages of the Northwest of Guatemala (Craig 1987; 1990; Zavala 1989; 2000). In these languages a noun may be accompanied by up to four distinct classifying morphemes: [number+number CL + numeral CL + plural CL + noun CL + noun]<sub>NP</sub>

## (23) Q'anjob'al 'classifiers':

- | 1                     | 2             | 3          | 4          |
|-----------------------|---------------|------------|------------|
| (a) <i>ox-ep'</i>     | <i>tinan</i>  | <i>ep'</i> | <i>naj</i> |
| 3-NB.CL               | MENS.CL       | PL.CL      | N.CL noun  |
| INAN                  | group         | HUM        | man man    |
| 'three groups of men' |               |            |            |
| (b) <i>ox-wan</i>     | <i>k'itan</i> | <i>ep'</i> | <i>naj</i> |
| 3-NB.CL               | SORT.CL       | PL.CL      | N.CL noun  |
| HUM                   | separated     | HUM        | man man    |
| 'three separated men' |               |            |            |

The four types of classifiers shown are:

- (a) Fused numeral classifiers, of which there are only three, for human/animal/inanimate nouns, obligatorily suffixed to the number, and maybe derived from the independent numeral classifiers of (b).
- (b) Independent numeral classifiers, of which there are only about a dozen sortal ones (including vertical/circular/round/three-dimensional flat and large classes). The mensural classifiers have the same surface form, but differ by being derived from positional roots and much more numerous, and by controlling different agreement rules.
- (c) Plural classes which inflect on the plural morphemes, for one of two or three classes (human/(animal)/inanimate), depending on the language. This obligatory classification is highly grammaticalized and reminiscent of gender.
- (d) Noun classifiers, which are omnipresent in those languages, functioning as determiners and anaphoric pronouns, totally independently of quantification. There are 12 to 19 noun classifiers per language/dialect, the most common ones being human masculine-feminine/old respected/kin/animal/wood/rock/dirt/plant/corn/water/salt.

### 3.3.2. Argument 2: Matching morpho-syntactic types and semantics

A second argument is semantic and consists in linking each major morpho-syntactic type with a dominant semantic domain. When classifiers are categorized semantically as belonging to one of the three basic semantic domains of classifiers – material, shape and function – a clear alignment emerges. Beyond the great variation in the semantics of numeral classifiers, the semantic domain with the greatest consistency and the highest rate of frequency is that of shape. Meanwhile, the semantics of noun classifiers fall predomi-

nantly into the two semantic domains of material (inherent essence of the objects) and relational status of humans (based on kinship or social status). In contrast, the genitive classifiers are overwhelmingly of the functional kind. This matching of morphosyntax and semantics is illustrated below (Olsen 1990):

CL. type	semantic domain	sample semantic features
numeral	shape	1D/long (tree) 2D/flat (leaf) 3D/round (fruit)
noun	material	man/woman/tree plant/rock/liquid
	status	kin male/non-kin female highly respected/deity
genitive	function	edible/drinkable/vehicle

Tab. 97.1: Morphosyntax and semantics

The correlation is striking in its consistency, considering how the majority of classifier systems are heterogeneous systems that bear the marks of their complex development patterns (see 3.4). What such correlations might mean has been outlined in the literature (see de León 1988 for Tzotzil numeral classifiers, Craig 1986 for Q'anjob'al noun classifiers, and Carlson & Payne 1989 for Oceanic genitive classifiers); the answers are to be found in the real world conditions for the use of each construction type. The use of numeral classifiers primarily for quantification, that prototypically occurs in a marketing context where handling of the objects is paramount, may account for the dominance of a categorization by shape in such systems. On the other hand, the clue for the semantics of genitive classifiers may be found in the fact that they are used specifically in one type of possessive constructions, those involving so-called *alienable* nouns. While the notion of **alienability** may be culturally bound, it is prototypically assigned to objects whose possession is considered acquirable and valued. Therefore, when the possession of such objects is claimed, it is linked to a particular purpose, among the ones that ensure human survival, hence classifiers for edibles and drinkables, housing and clothing, tools and transportation. Hence the fact that more genitive classifiers than other types of classifiers are derived from verbs and the semantics of genitive clas-

sifiers is overwhelmingly of a functional nature.

The case of noun classifiers is different in that their use is not linked to either real world conditions of quantification or possession. They instead have a much closer semantic link to the nouns themselves, forming with them a tighter unit, which is often reflected in their redundant semantics. Noun classifiers are often the nominal superordinates of the nouns they classify, or identify some inherent feature of the noun, such as its essence or material. They are morphologically more often of nominal origin than the other types of classifiers and their role in the language is more intimately identified with that of nominals, as referent tracking devices.

Therefore, while all classifiers may share the function of individuating the nouns to which they refer, the different morphosyntactic types of classifiers are associated with different semantic bases for individuation, and an explanation for the association of a particular semantic domain with a particular classifier type may be found in an analysis of the pragmatic function of the constructions in which each type of classifier occurs.

#### 4. Dynamic dimension of classifier systems

At the juncture of lexical and grammatical systems, classifier systems are interesting for what they can reveal of the processes by which grammatical systems emerge. Behind the absolute uniqueness of every classifier system lie some recognizable patterns of how such systems emerge, evolve, and decay.

##### 4.1. Emergence of classifier systems

Classifiers themselves and classifier systems have multiple origins and paths of development, most established systems being complex patchworks developed and renewed over time.

###### 4.1.1. Source of classifiers

Classifiers have their origin in the major lexical classes, nouns being their most common source. This is best demonstrated by the widespread phenomenon of **repeaters**, a kind of classifiers which are identical to the noun they classify:

- (24) Jacalteco noun classifiers (Q'anjob'alan Mayan):  
*ix ix* ‘CL.woman woman’  
*ch'en ch'en* ‘CL.rock rock’

<i>te' te'</i>	‘CL.plant tree’
<i>ixim ixim</i>	‘CL.corn corn’

The classifiers of Australian languages are said to have come from superordinate nouns (Dixon 1982), those of incorporated verbal classifiers from both generic and specific nouns (Mithun 1986). One of the most striking facts about the origin of classifiers is the very widespread recurrence of a very small and specific set of lexical nouns that has given rise to classifiers all over the world. They are three nouns of the plant domain: tree, leaf, and fruit, the first one being by far the most universal source of classifiers.

Verbs are another source of classifiers, although relatively infrequently when compared to nouns. Scattered examples from numeral and genitive classifiers are:

- (25) Tzotzil numeral classifiers (de León 1988: 55)

verb	classifier
<i>p'as</i> ‘to cut’	<i>p'os</i> ‘short length’
<i>k'as</i> ‘to break’	<i>k'os</i> ‘piece broken off’ (corn on the cob, banana, sugarcane, wood, dry excrement)

- (26) Mam (Mayan; in Zavala 1989): verb of eating

- (a) *chi'* ‘to eat cooked food’

*n-chi'-ye'* *kyix*

1-GEN.CL-POSS fish

‘my (cooked) fish’

- (b) *lo'* ‘to eat fruit’

*n-lo'-ye'* *chul*

1-GEN.CL-POSS zapote

‘my zapote fruit’

- (c) *wa* ‘to eat tubercles’

*n-wa-ye'* *ts'in*

1-GEN.CL-POSS yuca (cassava)

‘my cassava’

A rare instance of verbal classifiers with a verbal source has also been documented:

- (27) Imonda (Papuan; W. Seiler 1986):

verb	class; objects classified
<i>fët</i> ‘to remove from fire’	<i>fët-</i> ‘object going in the fire, wood, food cooked in fire’
<i>puis</i> ‘to cut’	<i>pui-</i> ‘a piece of’
<i>pot</i> ‘to pick fruit’	<i>pot-</i> ‘fruits picked from trees’

As it is, the origin of the vast majority of classifiers cannot be determined from synchronic data. Examples cited above are more the exception than the rule in the languages from which they were extracted. However, enough systematic connections can be gleaned from the various systems to establish with relative certainty the lexical origin of classifiers. The semantic transparency of classifiers and the ability to connect them to lexical sources is in fact one of the criteria used in evaluating the age of a classifier system: the more numerous the connections, the more recent the system is assumed to be.

#### 4.1.2. Semantic evolution of classifiers

From their original lexical semantics, classifiers undergo a process of semantic extension because the kinds of categories that classifiers refer to are very different from the kinds of categories that nouns refer to. While nouns refer to specific entities, the classifiers from which they are derived refer, in an initial classificatory role, to kinds of entities. Examples are ‘berry’ for all fruits in Mohawk (Iroquian), ‘tree’ for all plants in Jakaltek (Mayan), ‘canoe’ for containers in Squamish (Salishan) or ‘sailing vessel’ for all vehicles in American Sign language. This first level semantic extension of classifiers is most typical of noun classifiers and incorporated verbal classifiers, which of all types of classifier are the ones which have the most nominal function.

The majority of numeral classifiers, on the other hand, exhibit a further semantic extension as they metaphorically classify objects by perceived qualities, principally those of shape, secondarily those of consistency. The most widespread examples of such metaphorical extension occur with the cluster of plant part names, which very consistently gives rise to classifiers for the basic one-, two-, and three-dimensional shapes, as in ‘tree’ for long (and rigid), ‘leaf’ for flat (and flexible), ‘fruit’ for (small and) round. A parallel metaphorical extension occurs with the verbal classifiers derived from nouns of body parts. The basic shape classifiers come from one of two sets of body parts, either of human in upright position: ‘arm:long’ and ‘eye:round’, or, as in Tarascan, of four-legged animal: ‘neck/throat:long’, ‘back:flat’, ‘buttock:roundish’. In Mixtec, the noun classifier for animals has evolved to classify round shapes. In Imonda, the verbal classifier for flat and flexible objects originated in the verb ‘to give birth’. Another very common semantic exten-

sion is from material to consistencies, such as water for all liquids; in Imonda, the liquid classifier comes from the verb ‘to scoop water’. In Jakaltek (Mayan) the classifier for rock has come to classify all hard objects, from glass and metal objects to ice. The third type of metaphorical extension is that which takes a classifier of nominal origin to classify objects used for a particular function, increasingly independently of the shape of the object classified, as when an early mode of transportation (canoe, sailing vessel) is the source of a classifier for all modern modes of transportation.

Another perspective on the extension of the semantics of a classifier is how some classifiers progress from unique via specific to general classifier status, from defining a simple class of one member to defining an increasingly complex and heterogeneous one. Chinese provides attestation of a complete scenario of such an evolution which implies the mixing of prototype extension and chaining processes discussed in 2.3:

- (28) Chinese (Erbaugh 1986: 429):
  - ge*: unique classifier for bamboo >
  - specific classifier for bamboo and lengths of bamboo >
  - class extended to include: arrows, candles, dogs, chickens, horses >
  - later extended to include fruit, birds, people >
  - until it is presently a general classifier for people and unclassified objects.

#### 4.1.3. Morphosyntactic paths of evolution of classifiers

One aspect of the origin of classifier systems is the lexical origin of the individual classifiers, another is the grammatical origin of the classifier constructions, which encompasses the story of the emergence of classifier systems from other preexisting morpho-syntactic constructions and the blending through time of some types of classification into others.

Proposed scenarios of the emergence of classifier systems are still speculative, but suggest multiple origins for the various morpho-syntactic types of classifiers. It has been argued that the numeral classifiers of the Thai family have their source in a very productive noun compounding process, and the widespread use of class terms (DeLancey 1986), and that Q'anjob'alan and Mixtec noun classifiers fit in the mold of preexisting

sex markers, honorifics and epithets (Craig 1987; de León 1988). In addition, it has been speculated that some verbal classifiers originate in established processes of noun incorporation (Mithun 1986) – one of which is a common body part incorporation, hence the use of body parts for shape classifiers – while others originate in serial verb constructions (W. Seiler 1986). Another established path of evolution of classifier systems is the outright borrowing through contact, as in the spread of numeral classifier systems in Asia from Thai to Chinese, from Chinese to Japanese. The noun classifiers of Q'anjob'alan languages is a documented case of shared innovation with dialect variation that sprung up in the context of an older and family wide numeral classifier system and has spread through contact to neighboring Mamean languages.

The justification for including some of the most grammaticalized classification systems, such as gender and classificatory verbs, at the margins of a typology of classifier systems was that they have been hypothesized to be the results of evolutionary processes of more prototypical classifier systems. One of the earliest proposals of an evolutionary scenario was that some gender systems have evolved from sets of classifying demonstrative articles which themselves would have come from numeral classifiers (Greenberg 1978). Nominal classification in verbs comes in many shapes in American Indian languages, but can be placed on a continuum of progressive fixation and erosion of what may have originated as a syntactic noun incorporation process. The grammaticalized end of the process is illustrated by phonologically eroded classifying affixes, and their ultimate fusion to classificatory stems (Krauss 1968; Mithun 1986).

#### 4.2. Age and vitality of classifier systems

Beyond the evolutionary processes just discussed, age and vitality are two independent variables of classifier systems which must also be taken into account in any effort at setting up a typology of such systems.

##### 4.2.1. Old and new systems

Many mentions are made in the literature of the relative age of classifier systems. The criteria used for determining the age are by and large criteria discussed when establishing the degree of grammaticalization of the systems, such as the semantic transparency of the classifiers, the phonological erosion of the form

and their morphologically rigid use (see Art. 146). Systems are thought to be recent developments when a substantial number of the classifiers can still be linked to their lexical origin and the classifying system is still largely semantically motivated, so the systems present some coherence. An example of a recent innovative system is the noun classifier system unique to the Q'anjob'alan branch of the Mayan family of languages. The postulated evolution of verbal classifiers from incorporated nominal-like elements to phonologically and semantically eroded elements to fused elements of the classificatory verbal stem systems implies that the latter are the older systems, the former the more recent ones.

Some families of languages exhibit coexisting systems of classification, with a variation across the family as to which system flourished in which language. Proto Athapaskan-Eyak is said to have had two systems of verbal classification, one by stem and one by prefixed classifiers, which themselves had to be old established systems. The stem classification system is the one that flourished in Athapaskan and became vestigial in Eyak, while a well developed classifier system ("class mark prefixes") appeared in Eyak (Krauss 1968). In the Mayan family, the Tzeltalan branch has developed an elaborate numeral classifier system, while the neighboring Q'anjob'alan branch has only a limited numeral system, which coexists with a vestigial number suffixation reminiscent of gender systems, but has developed a full blown new noun classifier system.

##### 4.2.2. Productive, frozen and dying classifier systems

Classifier systems vary also in their level of productivity, largely measured by their ability to deal with new nouns and to allow for the semantics of the categories to be reanalyzed over time. The fact that this variable of vitality is independent of the variable of age can be demonstrated with two cases, Thai and Jakaltek. Thai has an old numeral classifier system which has maintained great vitality and is in constant state of renewal. An example of how the heterogeneity of a class can be reduced over time by the cyclic inclusion and exclusion of items is the case of the *khan* class of Thai (Carpenter 1987: 17).

In contrast, the Jakaltek (Mayan) noun classifier system is a recent system that was frozen until recently. It must have been pro-

ductive at the time of the conquest and the colonization, since it absorbed in the rock category the names of the new artifacts made of metal and glass that were brought in by the Spaniards. But it did not classify newer products made of nylon or plastic or artificial material (Craig 1986). To say that it was frozen does not mean that it was or is decaying; in fact it has become a well used system of pronominal reference through which it occupies a productive place in the grammar (Craig 1987).

Generally systems will decay over time rather than freeze. Decay is what accounts for the existence of classificatory verb systems mentioned earlier. A documented case of decay of a classification system which has reduced its inventory of elements and the range of its semantics is found in Dyirbal. In two decades, the noun class system with four classes (male, female, fruits and vegetables, others) has evolved into a tripartite gender system (male, female, inanimate) (Schmidt 1985: 155).

#### 4.3. Psycholinguistic studies of classifiers

##### 4.3.1. Language acquisition

Acquisition studies of gender, noun class, and numeral classifier data converge on the fact that, no matter how semantically motivated the classification system is, its formal characteristics are mastered before its semantics (Carpenter 1987). The development stages articulated in a detailed study of the acquisition of the numeral classifier system of Thai, a complex, productive and semantically motivated one, shows that the syntactic pattern (classifier suffixed to numeral) is acquired first. It is soon followed by the overuse of a general classifier which marks this relation of adjacency (up to 3/4 years), and later by a main strategy of repeaters indicating that the child now recognizes the relation of the classifier to the non-continuous noun (peak between 4 and 6 years). Semantic overextension later sets in, as the child learns to discriminate the conventional set of classifiers and their semantic motivation. Complex classifier systems such as the Thai one appear to be harder to learn than many other aspects of language and are not mastered by adolescence. Studies of large classifier systems, such as Thai, Burmese, Chinese, Mohawk (verbal incorporation system) all show a great variety of skills in handling the system among adult speakers.

##### 4.3.2. Aphasia

The major study of loss of classifiers in aphasia also dealt with Thai and demonstrated that the semantic errors by adult aphasics were qualitatively and quantitatively similar to those made by children, and were marked by the overuse of repeaters and of a general classifier for inanimates (Gandour et al. 1985).

#### 5. Uncommon abbreviations

MENS.CL	mensural classifier
N.CL	noun classifier
NB.CL	number classifier
NUM.CL	numeral classifier
SORT.CL	sortal classifier

#### 6. References

- Adams, Karen L. & Conklin, Nancy F. (1973), "Towards a Theory of Natural Classification". *Papers from the Ninth Regional Meetings of the Chicago Linguistics Society* 9, 1–10
- Aikhenvald, Alexandra Y. (2000), *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press
- Allan, Keith (1977), "Classifiers". *Language* 53, 285–311
- Barnes, Janet (1990), "Classifiers in Tuyuca". In: Payne (ed.), 273–292
- Barron, Roger (1982), "Das Phänomen klassifikatorischer Verben". In: Seiler & Lehmann (eds.), 133–146
- Barron, Roger & Serzisko, Fritz (1982), "Noun Classifiers in the Siouan Languages". In: Seiler, Hansjakob & Stachowiak, Franz Josef (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen. Teil II: Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen: Gunter Narr, 85–105
- Carlson, Robert & Payne, Doris L. (1989), "Genitive Classifiers". In: Carlson, Robert & DeLancey, Scott & Gildea, Spike & Payne, Doris & Saxena, Anju (eds.), *Proceedings of the Fourth Meeting of the Pacific Linguistics Conference*. Eugene: University of Oregon (Pacific Linguistics Conference 4), 87–119
- Carpenter, Kathie (1987), *How Children Learn to Classify Nouns in Thai*. Ph.D. dissertation, Stanford University
- Craig, Colette G. (1986, ed.), *Noun Classes and Categorization*. Amsterdam, Philadelphia: Benjamins (Typological Studies in Language 7)

- Craig, Colette G. (1986), "Jacalteco Noun Classifiers: A Study in Language and Culture". In: Craig (ed.), 263–293
- Craig, Colette G. (1987), "Jacalteco Noun Classifiers: A Study in Grammaticalization". *Lingua* 71, 241–284
- Craig, Colette G. (1990), "Clasificadores nominales: Una innovación Q'anjob'al". In: England, Nora C. & Elliott, Stephen R. (eds.), *Lecturas sobre la lingüística maya: la antigua*. Guatemala: Centro de investigaciones regionales de Mesoamérica, 253–268
- DeLancey, Scott (1986), "Toward a History of Tai Classifier Systems". In: Craig, Colette G. (ed.), 437–452
- Demuth, Katherine & Faraclas, Nicolas & Marchese, Lynell (1986), "Niger-Congo Noun Class and Agreement Systems in Language Acquisition and Historical Change". In: Craig (ed.), 453–471
- Denny, Peter (1976), "What are Noun Classifiers Good For?". *Papers from the Twelfth Regional Meeting of the Chicago Linguistics Society* 12, 122–132
- Derbyshire, Desmond C. & Payne, Doris L. (1990), "Noun Classification Systems of Amazonian Languages". In: Payne (ed.), 243–271
- Dixon, Robert M. W. (1982), *Where Have All the Adjectives Gone?* The Hague: Mouton
- Dixon, Robert M. W. (1986), "Noun Classes and Noun Classification in Typological Perspective". In: Craig (ed.), 105–112
- Erbaugh, Mary S. (1986), "Taking Stock: The Development of Chinese Noun Classifiers Historically and in Young Children". In: Craig (ed.), 399–436
- Gandour, Jack & Buckingham, Hugh Jr. & Dardarananda, Rochana (1985), "The Dissolution of Numeral Classifiers in Thai". *Linguistics* 23, 547–566
- Goral, Donald R. (1978), "Numeral Classifier Systems: A Southeast Asian Cross Linguistic Analysis". *Linguistics of the Tibeto-Burman Area* 4.1, 1–72
- Greenberg, Joseph H. (1978), "How Does a Language Acquire Gender Markers?". In: Greenberg, Joseph H. & Ferguson, Charles A. & Moravcsik, Edith A. (eds.), *Universals of Human Language, Vol. III*. Stanford: Stanford University Press, 47–81
- Grinevald, Colette (2000), "A Morpho-syntactic Typology of Classifiers". In: Senft (ed.), 50–92
- Heath, Jeffrey (1978), *Ngandi Grammar, Texts, and Dictionary*. Canberra: Australian Institute of Aboriginal Studies
- Heine, Bernd (1982), "African Noun Class Systems". In: Seiler & Lehmann (eds.), 189–216
- Klein, Hamet (1979), "Noun Classifiers in Toba". In: Mathiot, Madeleine (ed.), *Ethnology: Boas, Sapir and Whorf Revisited*. The Hague: Mouton, 85–95
- Krauss, Michael E. (1968), "Noun-Classification Systems in Athapaskan, Eyak, Tlingit, and Haida Verbs". *International Journal of American Linguistics* 34, 194–203
- Kuipers, Aert H. (1967), *The Squamish Language: Grammar, Texts, Dictionary*. The Hague: Mouton
- Lakoff, George (1986), "Classifiers as a Reflection of Mind". In: Craig (ed.), 13–15
- Langdon, Margaret (1970), *A Grammar of Degueño: The Mesa Grande Dialect*. Berkeley etc.: University of California Press
- León, María Lourdes de (1988), *Noun and Numeral Classifiers in Mixtec and Tzotzil: A Referential View*. Ph.D. dissertation, University of Sussex
- Li, Charles N. & Thompson, Sandra A. (1981), *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press
- Matsumoto, Yo (1990), *The Japanese Numeral Classifiers: A Study on Semantic Categories and Lexical Organization*. M.S. thesis, Stanford University
- Mithun, Marianne (1986), "The Convergence of Noun Classification Systems". In: Craig (ed.), 379–397
- Olsen, Gloria (1990), *The Pragmatics of Classifiers*. Ms. University of Oregon
- Payne, Doris L. (1987), "Noun Classification in the Western Amazon". *Language Sciences* 9 [Special issue: Key, M. Ritchie (ed.), *Comparative Linguistics of South American Indian Languages*], 21–44
- Payne, Doris L. (1990, ed.), *Amazonian Linguistics: Studies in Lowland South American Languages*. Austin: University of Texas Press
- Rehg, Kenneth (1981), *Ponapean Reference Grammar Pali Language Texts: Micronesia*. Honolulu: University Press of Hawaii
- Richards, Maña (1983), "Los clasificadores numerales en los dialectos cabecares de Ujarras y Chirripó". *Estudios de Lingüística Chibcha*, Serie A.2, 3–13
- Schmidt, Annette (1985), *Young People's Dyirbal: An Example of Language Death from Australia*. Cambridge: Cambridge University Press
- Seiler, Hansjakob (1986), *Language, Object, and Order: The Universal Dimension of Apprehension*. Tübingen: Gunter Narr

- Seiler, Hansjakob & Lehmann, Christian (1982, eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen. Teil I: Bereich und Ordnung der Phänomene*. Tübingen: Gunter Narr
- Seiler, Walter (1986), "Noun Classificatory Verbal Prefixes as Reanalyzed Serial Verbs". *Lingua* 68, 189–206
- Senft, Gunter (2000, ed.), *Systems of Nominal Classification*. Cambridge: Cambridge University Press
- Serzisko, Fritz (1981), "Gender, Noun Class, and Numerical Classification: A Scale of Classificatory Techniques". In: Dirven, René & Radden, Günter (eds.), *Issues in the Theory of Universal Grammar*. Tübingen: Gunter Narr, 95–123
- Supalla, Ted (1986), "The Classifier System in American Sign Language". In: Craig (ed.), 181–214
- Zavala, Roberto M. (1989), *Los clasificadores nominales del Kanjobal de San Miguel Acatan*. B.A. thesis, Instituto Nacional de Antropología e Historia (Mexico)
- Zavala, Roberto M. (2000), "Multiple Classifier Systems in Akatek (Mayan)". In: Senft (ed.), 114–146

*Colette Grinevald, Lyons (France)*

## 98. Gender and noun class

1. Introduction
2. Formal properties of gender systems
3. Principles of gender assignment
4. Semantics of genders
5. Gender agreement and gender resolution
6. Interrelation with other grammatical categories
7. Diachronic dimensions of gender
8. Uncommon abbreviations
9. References

### 1. Introduction

#### 1.1. Gender and noun class: history and terminology

The term **gender** was first used in the 5th century B.C. by the Greek philosopher Protagoras, when he divided Greek nouns into three classes: 'feminine', 'masculine', and 'inanimate' (nowadays called "neuter"). This was a typical gender system, which is found in many Indo-European languages. Latin had a similar system, but during historical change neuter nouns were redistributed between the other two genders, giving the modern system of masculine and feminine in French and Italian.

When Europeans came to study African languages, they discovered larger gender-like systems with eight or more possibilities in languages like Swahili; these often did not include a masculine – feminine distinction. The term "noun class" came to be used for systems of this type.

**Noun class** and **gender** are often used interchangeably. To avoid confusion, I shall just use the term "gender" in this article.

Since gender systems show some correlation with sex, many non-linguists (and a few linguists) erroneously confuse gender and sex. However, sex represents biological categorization, and gender represents grammatical categorization. Feminine and masculine genders often include inanimate nouns with no connection to female or male sex, e.g. French *maison* 'house' (feminine), *château* 'castle' (masculine).

The terms "gender" and "noun class" have also been used in quite different ways. For instance, in the Athapaskan linguistic tradition the term "gender" is used to refer to verbal classifiers which mark agreement with intransitive subject or transitive object, and characterize the referent noun in terms of shape and form (Aikhenvald 2000: 15).

#### 1.2. Characteristics of gender systems

Gender systems have the following characteristics (cf. Aikhenvald 2000: 18–22).

- (a) There is a limited, countable number of gender classes.
- (b) Each noun in the language belongs to one (or sometimes more than one) class(es).
- (c) There is always some semantic basis to the grouping of nouns into gender classes, but languages vary in how much semantic basis there is. This usually includes animacy, humanness and sex, and sometimes also shape and size.
- (d) Some constituents outside the noun itself must agree in gender with the noun.

Agreement can be with other words in the noun phrase (adjectives, numbers, demonstratives, articles, etc.) and/or with the predicate of the clause or an adverb.

In some languages there is an overt marker of gender on every noun, or on some nouns; in some languages nouns bear no gender marker.

Gender systems are typically found in languages with a fusional or agglutinating (not an isolating) profile. Gender agreement is a major criterion for distinguishing nouns from other word classes. In a language where noun and adjective have similar morphology, an adjective can generally take any gender marking where the noun is normally restricted to one gender class. Languages often have portmanteau morphemes combining information about gender with number, person, case, etc.

Gender is the most grammatical means languages use for the semantic categorization of nouns. Art. 97, on classifiers, discusses other noun categorization mechanisms which are more lexical in nature.

### 1.3. Gender in the languages of the world

The majority of the world's languages have gender or some other noun categorization devices (see Aikhenvald 2000: 77–80).

Many Afroasiatic languages have two genders: masculine and feminine. A system of two or three genders is present in most Indo-European and North-West Caucasian languages. More complicated systems of three to five genders are present in North-East Caucasian and Nakh (Central-Caucasian) languages. The Dravidian languages of South India have two to four genders (Krishnamurti 1975).

In North America, Algonquian languages have animate and inanimate genders. Two genders are also found in Chemakuan, Wakashan, Salishan, and Siouan languages.

In Central America, two genders (animate vs. inanimate) are found in a few Otomanguean languages. More than half of the languages of South America show genders. A system of two genders, masculine and feminine, is characteristic of languages of the Jé, Guahibo and Arawá families, some Arawak languages, and the languages of Gran Choco.

In Africa, East Niloct languages distinguish masculine and feminine gender (Dimmendaal 1983: 211). The majority of Niger-Congo languages have extensive gender sys-

tems (up to 20 agreement classes combined with number).

Genders are widespread in Papuan languages of the Sepik basin and adjoining lowland areas. Lower Sepik languages (Ndu, Ok, Sepik Hill) have two genders, feminine and masculine, which correlate with the shape and size of the referent. The languages of Torricelli and some of those of Lower Sepik have fascinating systems of about a dozen classes (see Foley 1986: 85 ff. for the example of Yimas). Extensive systems with several dozen agreement classes are found in the Papuan languages of Southern Bougainville: Nasioi, Motuna (Foley 1986: 83 ff.; Onishi 1994).

A typical gender system in Australian languages contains four terms which can be broadly labelled as "masculine", "feminine", "vegetable", and "residual" (see Dixon 2002).

There are no genders in the Uralic, Turkic, Tungus-Manchurian, Tibeto-Burman, South-Caucasian (Kartvelian), Eskimo-Aleut, Chukotka-Kamchatkan families, and in most Austronesian languages and the languages of South-East Asia.

## 2. Formal properties of gender systems

### 2.1. Expression of gender

Languages of the world differ in the number of gender classes they have, how much semantic basis there is to gender assignment and the possibility of changing gender assignment to match the semantic characteristics of the referent.

Some, or all, nouns, can have an overt gender marking. In those cases the gender of a noun can be inferred from its form. In Swahili, a noun prefix indicates its gender. In Apurinã (Arawak, Brazil), feminine nouns tend to end in *-ro*, and masculine nouns tend to end in *-ri*. In Portuguese, nouns which end in *-a* are mostly feminine, and those which end in *-o* are mostly masculine (with a couple of exceptions). Some languages seldom or never mark gender on the noun itself, e.g. Ndu languages (East Sepik region of Papua New Guinea) or !Xu (North Khoisan; Heine 1982: 193). This is known as "covert" gender.

Overt and covert marking can be viewed as two extremes of a continuum. In some languages, nouns can be optionally marked for gender. The marked form tends to be more specific than the unmarked, as in Turkana

(East Nilotic: Dimmendaal 1983: 221). Overt gender marking in Alambak (Lower Sepik) is used to focus on the sex of an animate referent, or to indicate a change in size of an inanimate one. Languages with predominantly semantic gender assignment (cf. 3.1) tend not to mark gender on the noun. Only nouns with human and, more rarely, with animate referents have their gender overtly marked, as in Manambu (Ndu family, Papuan).

A hierarchy determining overt gender marking in languages with semantic gender assignment is: human > animal > other animate > other.

The gender realized in agreement can be distinct from the gender marked on the noun itself. In Swahili *ki-boko* ‘hippopotamus’ has the class 7 prefix *ki-*. It belongs to class 1 as far as agreement is concerned, since the agreeing adjective takes the *m-* prefix of class 1: *ki-boko m-bukwa* ‘a big hippopotamus’. These ‘mismatches’ between gender overtly marked on nouns and agreement gender are found in many Australian, Bantu, Afroasiatic, and Indo-European languages. To account for them, Evans (1997) suggested a useful distinction between **head class** (overtly marked on the noun itself) and **agreement class** (realized in agreement).

Nouns which display a conflict between different rules of gender assignment are called **hybrids**. In Russian *mužčina* ‘man’ is feminine by its morphology but masculine by its semantics. In the case of most nouns denoting professions which are morphologically masculine, such as *professor* ‘professor’, the agreement is feminine when focussing on the female sex of the person. In Portuguese, some nouns denoting professional occupations which end in *-a* or *-e* can be assigned either gender, e.g. *dentista* ‘dentist’, *estudante* ‘student’. In Anindilyakwa (Australian), *dirija* ‘dress’ is treated as feminine when conceived as an attribute of a woman, but as a member of inanimate *m*-class when seen as a material (Julie Waddy p.c.). These nouns are said to have **double or multiple gender**, depending on their semantics. In Russian, Portuguese, or Anindilyakwa, only a limited number of nouns can be assigned more than one gender. In other languages, such ‘reclassification’ is much more productive; this typically happens in languages with semantic gender assignment which involves parameters other than just sex or animacy (cf. 3.5).

Some generic terms for animate beings take one agreement form though they denote beings of either sex (termed “epicenes”). It is up to the language as to what gender epicenes are formally assigned to. In Russian, *krolik* ‘rabbit’ takes masculine agreement, and *sobaka* ‘dog’ takes feminine agreement, though both can denote beings of either sex.

Different gender agreement systems for different word classes can coexist in one language (cf. 5.3).

Some gender systems barely satisfy the criteria set out in 1.2. English distinguishes three genders just in 3rd person pronouns, *he/she/it*. They are purely semantic and involve the opposition: male, female, inanimate. There are few exceptions, e.g. ships are referred to with the feminine pronoun *she*. There is no gender agreement within a noun phrase or with a verb in a clause. Gender markers in English simply have an anaphoric function. Similarly, Mam (Mayan) has a dozen pronouns which are only used anaphorically when the head noun is omitted. They involve such meanings as ‘baby’, ‘non-human’, ‘young man’, ‘young woman’, etc. (England 1983: 158–159).

The number of surface realizations of genders (**target genders**) can be different from types of agreement (**controller genders**; see Corbett 1991: 150–160). In Romanian, nouns divide into three gender classes. There are two surface markers of gender in singular and two in plural, but three combinations of these (Corbett 1991: 151). Telugu (Dravidian; Krishnamurti & Gwynn 1985: 56–58) distinguishes two genders in singular: masculine gender covers males, and moon and sun; and feminine/neuter genders covers female referents and all other nouns. In plural, there is a contrast between animate and inanimate gender.

A more complicated system is found in Khinalug (North-East Caucasian; Kibrik et al. 1972: 154–155, 118 ff.). Nouns divide into four classes: I – males; II – females; III – animals and birds (with a few exceptions); IV – abstract nouns (actions and states) (and some others, e.g. *nimts* ‘louse’). Gender agreement is found on demonstrative pronouns, headless adjectives, and verbs. There are two sets of surface agreement markers. Set 1 is used with past and future tenses of resultative and non-resultative aspects. In the singular forms of Set 1, genders II and III have identical marking. Set 2 is used to mark

gender/number agreement on different verbs depending on their morphological class. In the singular forms of Set 2, genders I and IV group together. And in plural, genders I and II, and III and IV have the same marking.

Some languages have genders with a very small number of members. Archi (North-East Caucasian) has an agreement class which can be treated as gender III when singular and I/II when plural; it comprises just two nouns ('people' and 'population'). Tsova-Tush (North-Central Caucasian) has a gender with just four nouns (Corbett 1991: 171).

Further grouping of genders involves **subgenders** and **superclassing**. Subgenders are defined as "minor agreement classes which control minimally different sets of agreement" (Corbett 1991: 163). Innovative subgenders based on the feature of animateness and personality are characteristic of Slavonic languages.

An unmarked agreement gender (see 2.2) may replace the inherent gender of the noun. In Ngalakan (Merlan 1983), demonstratives can agree with the head noun, or can simply show the unmarked prefix *rnu-* 'masculine gender', unless the head noun is feminine. This is described as concordial superclassing (Evans 1997; Sands 1995: 264–265).

## 2.2. Markedness

There are two types of markedness – formal and functional. A term in a system is formally unmarked if it has zero realization. If all terms in a system, save one, are only used in specified circumstances, and the remaining term is used elsewhere, then this is said to be functionally unmarked.

Masculine gender in Russian is formally unmarked, since it usually has a zero-marking, as opposed to *-a* 'feminine' and *-o* 'neuter'. The functionally unmarked gender is the one used (a) for default agreement (Corbett 1991: 206 ff.); (b) as a generic term and for indefinite reference; (c) when the distinction is neutralized or is of no relevance. In traditional English, masculine pronoun *he* is used as a generic term for human reference and also as a subordinate term for male reference (cf. discussion in Alpher 1987).

The functionally unmarked gender is likely to be one of the largest ones. Concordial superclassing provides strong evidence in favour of the unmarkedness of that class which is used as a superclass. In Australian languages, all the languages with superclassing show masculine as superclass. In Warray and

Jawoyn, class prefix (*n*)*a-* 'masculine' can be used on a modifying adjective of any noun under special discourse conditions. In Gaa-gadju, there are two superclasses: masculine with prefix *na-* or *Ø* is the superclass for *animates*, and *ma-* 'plants' is the superclass for *inanimates* (Aikhenvald 2000: 40–55).

In a system of two genders, one may be marked and the other unmarked. The unmarked gender is often masculine, as in Indo-European or Afroasiatic languages; more rarely feminine, as in Guajiro, Lokono (Arawak), Jarawara (Arawá), Wangkumara and Wagaya (Australian).

Alternatively, neither of the two genders can be considered unmarked, as in Ungarinjin, Tiwi, and Alamblik.

Different word classes can display different relations of markedness. In Spanish, masculine is the unmarked form for pronouns, but there is no obvious unmarked class for nouns. In Ngalakan, masculine gender is used as a superclass only for demonstrative modifiers (Merlan 1983).

The issue of markedness is more complicated in systems with three or more genders, where sex-differentiable nouns form a natural subset. In Seneca (Iroquoian), feminine is the unmarked gender for human referents only. Neuter gender is used for animals and inanimates.

Some languages provide contradictory evidence for markedness. The Western Torres Strait language has two semantically assigned "genders", feminine and masculine. The feminine gender is the obvious "candidate" to be considered an unmarked gender, since it is the one used for the majority of instances of default agreement (Bani 1987). All plurals are treated as 'feminine'. However, agreement with an unidentified human being (at a distance) is masculine, irrespective of sex.

Functional and formal markedness do not necessarily coincide. In Dyirbal, residue gender IV is formally unmarked, but none of the four genders can be considered functionally unmarked.

## 3. Principles of gender assignment

The principles by which nouns are allotted to different genders can be governed by semantics (cf. 3.1), or formal morphological (cf. 3.2) or phonological (cf. 3.3) properties of a noun, or a combination of these (cf. 3.4). In a sense all systems of gender assignment are

mixed, since there is always a semantic core which involves universal semantic parameters such as sex, humanness and animacy.

### 3.1. Semantic assignment

In languages with semantic assignment the gender of a noun can be inferred from its meaning. In Tamil all nouns divide into rational and non-rational. Rational nouns include humans, gods, and demons. Malto, Kolami, Ollari, and Parji distinguish male human from the rest. From the North-East Caucasian family, Godoberi, Akhvakh, and Bagvalal distinguish male rational, female rational, and the rest. Diyari (Australian), Kaingang (Jê), and the North-Arawak subgroup divide all nouns into female humans and the rest (Aikhenvald 2000: 22–24).

Semantic assignment can be more complex. Dyirbal (Dixon 1972: 308–312) has four genders. Three of them are associated with one or more concepts: gender 1 – male humans and non-human animates; gender 2 – female humans, water, fire, fighting; gender 3 – non-flesh food. Gender 4 is a residue class, covering everything else. There are also two rules for transferring gender membership. By the first, a referent can be assigned to a gender by its mythological association rather than by its actual form. Birds are classed as feminine by mythological association: women's souls are believed to become birds after death. The second transfer rule is that if a subset of a certain group of objects has a particularly important property, e.g., being dangerous, it can be assigned to a different class from the other nouns in that group. Most trees without edible parts belong to gender 4, but stinging trees are placed in gender 2. In the Western Torres Strait language, all nouns denoting males are masculine, and all the rest are feminine. The moon is masculine, due to mythological association with masculinity. In Ket, all sex-differentiable nouns are masculine or feminine (cf. Art. 128). Among non-sex-differentiable nouns, those which show a higher degree of activity or are particularly important for Ket culture are masculine, e.g. wood, large wooden objects, growing trees. Gender assignment of sun (feminine) and moon (masculine) is determined by their role in myths. Other inanimate nouns are treated as neuter.

It has often been stated that there is no real semantic basis for gender assignment in some well-known Indo-European languages. In a seminal study, Zubin & Köpcke (1986)

provided a semantic rationale for gender assignment of nouns of different semantic groups in German. In agreement with the natural sex principle, masculine and feminine genders mark the terms for male and female adults of each species of domestic and game animals, and neuter is assigned to non-sex-specific generic and juvenile terms. Superordinate terms are very often neuter. Masculine gender is used for types of cloth, types of precipitation and wind, and types of minerals. Types of knowledge and disciplines have feminine gender, and games and types of metal have neuter gender.

### 3.2. Morphological assignment

In a language with a number of nominal declensions, each may typically be assigned a gender. In Russian the dependency on a semantic core is restricted to human and higher animate referents. Otherwise, gender assignment depends on declension: all nouns of declension 1 are masculine, all nouns of declensions 2 and 3 are feminine, and all the rest are neuter. Another morphological basis of gender assignment is found in German where at least some derivational affixes are each associated with one gender, e.g. *-ung* 'action noun' is feminine and *-chen* 'diminutive' is neuter (Aikhenvald 2000: 25).

### 3.3. Phonological assignment

In some languages every noun which ends or begins with a certain vowel or consonant must belong to a particular gender. In Qafar (Saho-Afar, East Cushitic; Aikhenvald 2000: 25), nouns with inanimate referent whose citation form ends in a vowel are feminine, all the rest are masculine. In Hausa, all non-sex differentiable nouns which end in *-aa* are feminine. In Katcha (Kordofanian; Heine 1982: 200), any noun beginning with *m-* belongs to the feminine gender, unless it is "nominally masculine".

### 3.4. Mixed principles of assignment

Many languages display mixed principles of assignment: semantic, morphological, phonological. Irakw (South Cushitic; Heine 1982: 200) has masculine and feminine genders. All nouns denoting singular male and female animates, and male and female noun agents are masculine and feminine respectively. Singular nouns ending in *-mo* or *-ayw* are masculine. Nouns derived from Class I verbs are masculine while those from Class II verbs are feminine.

An interesting interaction of semantic and phonological principles is found in Yimas (Sepik, Papua New Guinea; Foley 1986: 86 ff.), where the first four classes are identified by their semantics: I – human males; II – human females; III – animals; IV – culturally important plants. Classes V–XI are defined phonologically: the agreeing constituent repeats the last consonant of the nominal root.

Morphological and semantic principles interact in German gender assignment for some semantic groups of nouns. Bird names have masculine gender unless they end in a derivational suffix which is feminine. Sailing vessels have feminine gender unless they end in a masculine derivational suffix (Zubin & Köpcke 1986: 175).

Principles of gender assignment can be hierarchically organized. Bowili (Togo Remnant, Eastern Ghana) has 14 classes. In the class assignment of loan words semantic principles are more important than phonological ones. Animate loanwords are always allocated to animate class, e.g. *cicá*, pl. *bacicá* ‘teacher’ (from English). Phonological principles operate with inanimate nouns. These are assigned to genders on the basis of the similarity of their first syllable to a class prefix, e.g. sg. *si-kû*, pl. *fu-kû* ‘school’ (from Ewe *sukû*, which comes from English). All the other loan words are assigned to a special loan-word class (Heine 1982: 199 f.).

There is always some semantic basis to gender assignment. A grammatical system of gender assignment can move towards a more “semantically oriented” one. In Cantabrian Spanish, the semantic features in Table 98.1 have been introduced for the assignment of inanimates to masculine or feminine gender (Holmquist 1991: 69; Aikhenvald 2000: 25–28):

<i>feminine</i>	<i>masculine</i>
female	male
large	small
wide	narrow
horizontal	vertical
squat	tall
supine	phallic
smooth	coarse
light	dark
approbatory, neutral	deprecatory
primary	derived
familiar	occasional

Tab. 98.1: Gender assignment in Cantabrian Spanish

These oppositions are absent from standard Spanish.

For nouns with animate referents, semantic assignment usually overrides morphological assignment. A well-known exception is *Mädchen* ‘girl’ in German which is neuter. Semantics overrides other principles in languages which have predominantly semantic gender assignment. In Alamblak, all nouns which denote females and short, squat or wide objects are feminine and can have a form marked with a feminine suffix. There is one exception, the word for canoe, *doh-t*, which has a feminine marker *-t*, but is always treated as masculine, in agreement with its typical ‘masculine-like’ slender dimensions.

Semantic, morphological, and phonological principles account for the great majority of the assignment of nouns to gender classes in any given language, but there will almost always be a small residue of unexplained exceptions.

### 3.5. Variable gender

Languages with predominantly semantic systems of gender assignment can show variable gender on a number of nouns. The majority of languages allow variable gender on nouns like *baby* and *child*, on some kinterms and some inanimates. In Dyirbal, *bimu* is both ‘father’s elder sister’ (taking feminine gender marker) and ‘father’s elder brother’ (masculine marker). *Jaja* ‘baby’ can be specified as either masculine or feminine. In Ket, a growing tree is masculine, a cut-down tree is inanimate; an upright tree is masculine, a tree with a curved trunk is feminine.

Change of gender in Alamblak implies a change in shape and size of an inanimate object (Bruce 1984). In Kxoe (Khoisan; Heine 1982: 198), an inanimate noun stem can be allocated to masculine or feminine gender depending on its shape: masculine is associated with big, long, rectangular, and feminine with small, round, broad.

Manambu and Iatmul (Papuan) allow variable gender independent of the animacy of the referent noun. Nouns which denote male humans and higher animates, and long and thin inanimate objects are masculine; those which denote female humans and higher animates, and short and round objects are feminine. A smallish woman-like man can be treated as feminine, and a largish woman can be treated as masculine. Reclassification is impossible when the shape cannot be changed (turtles are ‘round’ and always

feminine), or when the “masculinity” is culturally important. Descendance is strictly patrilineal, and so the word *gwalugw* ‘patrilineal clan’ is masculine.

#### 4. Semantics of genders

Semantics of genders in the languages of the world involves the following parameters:

- (a) sex: feminine vs. masculine, as in many Afroasiatic languages, East-Nilotic and Central Khoisan;
- (b) human vs. non-human, as in some Dravidian languages;
- (c) rational (humans, gods, demons) vs. non-rational, as in Tamil and other Dravidian languages;
- (d) animate vs. inanimate, as in Siouan.

The term “neuter” is often used to refer to irrational, inanimate gender, or a residue gender with no clear semantic basis.

Languages can combine these parameters. Zande and Ma (Ubangi, Niger-Congo) distinguish masculine, feminine, non-human animate and inanimate. Godoberi (North-East-Caucasian) has feminine, masculine and non-rational genders.

Semantic parameters of animacy and humanness are universal. In languages with no genders they appear in other systems of noun categorization, i.e. classifiers (Art. 97), number (Art. 100), case (Art. 102) or anaphora (Aikhenvald 2000: 436–441).

Primarily sex-based genders can have additional shape- and size-related meanings. In the languages of the Sepik region, feminine is associated with short, wide, and round, and masculine with long, tall, and narrow objects. Feminine is associated with small size and with diminutives in Afroasiatic and East-Nilotic languages; masculine gender includes long, thick, solid objects. Hollow, round, deep, flat, and thin objects are feminine in Kordofanian and Central Khoisan languages (Heine 1982: 190–191).

Additional semantic oppositions in gender systems include: animal and vegetable food as in Australian (Sands 1995: 317); place, as in Bantu; insects, as in Andian languages. Value distinctions (i.e. endearing or pejorative meanings) are rare. Kuria (Bantu) has a gender with a pejorative meaning. The scope of the semantics of genders in languages with largish systems is comparable to that of classifiers (Art. 97).

Gender is often used metaphorically to describe unusual situations. In Dyirbal, the word *man* can be used with the feminine class marker, instead of masculine, to point out the female characteristics of a hermaphrodite. In Manambu, *head* is usually feminine because of its round shape, but it is treated as masculine when a person has a headache, since then the head feels heavy and unusually big.

In Russian, a dropped knife presages a male guest, and a dropped fork a female one. This is determined by the masculine gender of *nož* ‘knife’ and the feminine of *vilka* ‘fork’.

#### 5. Gender agreement and gender resolution

##### 5.1. Gender agreement

###### 5.1.1. Syntactic and semantic agreement

Agreement in gender can be determined by purely syntactic or semantic parameters. The former is called syntactic, or “mechanical”, agreement. *Mädchen* ‘girl’ in German triggers neuter agreement with modifiers, independently of its semantics. Semantic agreement is found in many Bantu languages of Africa and in Papuan languages. We typically get semantic agreement with animate nouns and syntactic with other nouns, e.g. in Swahili *ki-faru m-bukwa* ‘CL7-rhinoceros CL1-big’ (a big rhinoceros’).

Anaphoric agreement can be more semantic than modifier agreement. In German, *Mädchen* ‘girl’ is sometimes referred to by the feminine anaphoric pronoun *sie*. The word *professeur* in French has masculine agreement, independently of the sex of the person. The conflict between “grammatical” and “natural” gender can be resolved with a sentence like French *Elle est belle, le nouveau professeur* (‘She is beautiful, the new professor’). Almost every language with a grammatical gender may optionally pronominalize in agreement with natural gender of a human referent.

###### 5.1.2. Domain of gender agreement

Languages vary in the domain of gender agreement (Aikhenvald 2000: 30–45). In head-modifier noun phrases, gender agreement can be marked on the following types of modifiers: (i) adjectives, including deverbal adjectives (participles); (ii) demonstratives; (iii) articles; (iv) numbers.

The head determines gender agreement with a noun phrase. Agreement is a major criterion for recognizing the head.

Gender agreement is rarer in possessive constructions. In Swahili the class of the possessed noun is marked on the possessive morpheme. In Shona (Bantu), the possessive morpheme is marked for the class of both possessor and possessed (Welmers 1973: 178), e.g.

- (1) *tu-vana nembwa*  
 PL:CL13-child and+dogs:CL10  
*dz-a-tw-o*  
 CL10-POSS-CL13-their  
 'the little children and their dogs'

*Tuvana* 'children' belongs to class 13 (diminutive plural), and *imbwa* 'dogs' belongs to class 10 (animal plural). Markers of both classes appear on the possessive morpheme.

Possessive constructions show agreement with possessor in Abkhaz (Hewitt 1979: 116). In Chamalal, the possessor agrees in gender with the possessed noun phrase. In Manambu, gender agreement with the possessed noun is marked on the pronominal possessor only.

Agreement in possessive constructions is often linked to the distinction between alienable and inalienable possession. In Jarawara, the possessed noun in an inalienable possessive construction agrees in gender with the possessor (Dixon 1995). In the Australian languages Yanyuwa and Tiwi, body parts agree with the possessor in noun class.

Gender agreement is frequent in predicate-argument constructions. The verb may agree with core constituents: subject and/or direct object.

Bantu languages show gender agreement with subject and direct object – the subject agreement marker comes before tense marker, and the object agreement marker comes between tense and root.

Agreement with a peripheral constituent is less frequent. In Kirundi the verb can agree with locative and with indirect object. Agreement in gender with indirect object is found in Yimas (Foley 1986: 94). In Manambu and Alamblik, agreement in gender with the subject is obligatory, and there can be agreement with one other constituent (direct object, dative, locative, manner, time) if this is in focus.

There can also be gender agreement on adverbs, as in Lak (North-East Caucasian; Khaidakov 1980: 206), and adpositions, as in Abkhaz (North-West Caucasian; Hewitt

1979: 113–114). Complementizers agree in gender with the subject of the complement clause in West Flemish.

Gender is frequently marked on personal, interrogative, and relative pronouns. In many languages interrogative pronouns show an opposition of animate versus inanimate. This may be independent of the existence of a gender system in the language (Aikhenvald 2000: 436–441).

The majority of languages mark gender in more than one place in the clause. Dyirbal marks gender on determiners and interrogatives. Burushaski marks gender on pronouns, adjective modifiers, and verbs. Anindilyakwa (Australian) marks it on head nouns, verbs, adjectives, and all types of pronouns. Bantu, some West-Atlantic languages and the North Kimberley languages of Australia (Worora, Wunambal) mark gender on every type of modifier and on the verb.

#### 5.1.3. Realization of gender agreement

Gender agreement is typically realized via inflectional affixes which can be suffixes or prefixes. Infixed agreement, as in Archi (North-East-Caucasian) or agreement through internal change, as in Jarawara (Arawá) or Marind (Papuan), is rare.

Gender agreement can be realized via **alliterative concord**. In Swahili, the noun itself includes an overt marker identical to the agreement marker. Alliterative concord is found in some languages with phonological principles of gender assignment, as in Yimas.

**Repeaters** found in some South-American languages are a subtype of alliterative concord. The head noun is repeated onto a modifier as an agreement marker (also see Aikhenvald 2000: 61–63). Repeaters can give rise to a closed system of agreement classes.

#### 5.1.4. Restrictions on gender agreement

Gender agreement can be restricted by the topical and referential properties of the referent. In Swedish when the noun phrase is indefinite, a gender agreement marker appears on both article and adjective (2). If it is definite, the head noun and the article show a gender marker, and an adjective is marked for definiteness, but not for gender (3).

- (2) *ett grön-t hus*  
 a:N green-N house  
 'a green house'  
 (3) *det grön-a hus-et*  
 the:N green-DEF house-DEF:N  
 'the green house'

In some Arawak languages of South America, Motuna, and Manambu, gender agreement on the verb is associated with topicality of the head noun.

In North-Arawak languages, and in Abkhaz-Abaza, agreement in gender is neutralized if the subject is preposed to the verb, as a means of focussing the agent.

Members of a single word class can display different agreement restrictions. These lexical restrictions on agreement may be due to a historical accident. In Russian *odin* ‘one’ distinguishes three genders, as do all adjectives, but *dva* ‘two’ and *oba* ‘both’ distinguish only two genders. Historically they are residues of duals, and less gender distinctions were present in dual than in singular. Other numerals do not distinguish genders.

In Nakh languages of the Caucasus (Chechen, Ingush, and Tsova-Tush), only the numeral ‘four’ shows gender agreement. In a number of Dravidian languages (Kolami, Parji, Naiki) which have two genders (male human and the rest), lower numerals have special female human forms. They can be considered lexical exceptions. There is a subclass of adjectives in Latin which do not distinguish gender agreement.

## 5.2. Gender resolution

Languages differ in the ways in which they deal with the coordination of nouns belonging to different genders. This problem is sometimes solved semantically, sometimes syntactically, sometimes by a combination of these methods, and in a few languages it cannot be solved at all.

**Semantic** gender resolution usually involves **superclassing**. In Luganda and Chi-Bemba, conjoined nouns with human referents are treated as class 2 (basically human), even if none of the referents actually belongs to this class. Conjoined nouns with non-human referents are allocated to a ‘default’ class (class 8). Conjoining at least one human with non-humans is treated differently: it is often avoided, or an alternative comitative construction is used, e.g. *a man and a dog* will be translated as *a man with a dog*.

In **syntactic** gender resolution, one of the existing genders will be used, as in Portuguese, French, Spanish, Modern Hebrew, and Hindi. This gender can be considered functionally unmarked (cf. 2.2 and Aikhenvald 2000: 52–54). In Portuguese, when the conjuncts are of different genders, the masculine form of the adjective is used:

- (4) *um menino e uma menina*  
 INDEF:M boy and INDEF:F girl  
*bonit-os*  
 beautiful-PL:M  
 ‘beautiful boy and girl’

Mixed semantic and syntactic gender resolution is found in Latin. Masculine agreement is used when conjuncts of different genders denote persons; otherwise, neutral agreement is used.

In Tamil, gender resolution is restricted: rational nouns (which include humans, gods, and other mythical beings) cannot be conjoined with irrational ones. Many Bantu languages do not have any gender resolution strategy for nouns with non-human referents; these nouns simply cannot be conjoined, just like in Ungarinjin (Australian).

## 5.3. Split agreement systems

Different gender systems can coexist in one language for different types of modifiers. This is called “split agreement”; it is regular and does not depend on lexical choice.

Some languages have two (or sometimes more) gender systems, either for different types of modifiers, or for different types of modifiers and different types of agreement. One system of genders is used with personal, demonstrative and other pronouns, and for verbal cross-referencing; it is called “pro-nominal” gender. The other is used with adjectives (and sometimes other modifiers, such as numerals); it is called “nominal” gender (Heine 1982: 195).

Numerous languages of the Arawak and Tucano families of South America have a small gender system for demonstratives, articles and verbal agreement, and a largish system of genders for adjectival and numeral modifiers. The smaller system involves sex and animacy distinctions. (Baniwa and Tariana (Arawak) have a feminine vs. non-feminine distinction; Palikur and Ignaciano (Arawak) have feminine, masculine and inanimate.) The larger system consists of several dozen classes based on shape and sometimes also animacy and sex. Malak-Malak (Australian) has four genders, with obligatory agreement in some head-modifier and all verb-argument constructions, and a feminine vs. non-feminine distinction in third person singular pronouns.

In Motuna, a system of 51 agreement morphemes is used with a demonstrative or a quantifier as a post-head modifier. A four-

term gender opposition (inanimate, diminutive, masculine and feminine) governs agreement in head-modifier constructions with articles, demonstratives, and some adjectives; it is also used in verb-argument agreement.

Many African languages distinguish nominal and pronominal genders. Mba (Ubangi, Niger-Congo; Heine 1982: 208–209) distinguishes seven gender classes in the singular, and three pronominal genders. Nominal genders have opaque semantics; nouns indicating human beings occur in classes 1, 2, 7; body parts occur in classes 3 and 4. Pronominal genders have a clear semantic basis: (i) masculine, (ii) feminine, (iii) non-human animate and inanimate. Their use is confined to agreement with animate nouns. Agreement in pronominal gender is semantic, and agreement with nominal gender is syntactic.

Two systems of gender marking can co-occur on the same modifier in a noun phrase (Serzisko 1983: 115 ff.). In (5) the numeral ‘one’ has a prefix *bi-* showing pronominal gender agreement followed by a prefix *ú-* showing nominal gender agreement.

- (5) *ju*  
 woman  
*bí-ú-ma*  
 PRON:<sub>1</sub>FEM-<sub>2</sub>NOMINAL.GENDER-one  
 ‘one woman’

Paumari (Arawá; Chapman & Derbyshire 1991: 254 ff. and my field data) has two types of closed agreement classes. One is based on a feminine vs. masculine opposition, feminine being the unmarked term. The other, so called *ka*-noun class, has a partial shape- and structure-based semantics. Containers and things in containers (‘car’, ‘canoe’), extended objects and things which consist of particles are usually *ka*- class. Both occur with non-demonstrative modifiers and on verbs. Demonstratives distinguish only feminine-masculine.

The two types of gender are only partially semantically motivated and constitute distinct agreement systems.

Pronominal genders are more likely to involve animacy/sex/humanness-based semantic oppositions and a small number of agreement classes. Shape-based distinctions are most likely to go with nominal agreement.

## 6. Interrelation with other grammatical categories

Gender interrelates with other grammatical categories in the following ways (Aikhenvald 2000: 243–267):

- (a) Gender distinctions depend on other categories; e.g. there may be different gender distinctions in singular and plural (also see Aikhenvald & Dixon 1998).
- (b) Expression of gender correlates with other categories, e.g. discourse functions of the head noun.
- (c) There can be a portmanteau realization of gender and another category, e.g., gender and number.

### 6.1. Gender and classifiers

Genders and classifiers are two morphosyntactic means for the semantic categorization of nouns (see Art. 97). They can coexist in one language as distinct categories. Malto, a Dravidian language, has two genders in personal pronouns, nouns and verbal agreement markers, and three genders for demonstrative pronouns and several dozen numeral classifiers.

Coexistence of two agreement systems – one primarily sex- and animacy-based, and the other primarily shape-based, was discussed in 5.3.

In ‘Dongo-ko’ (Mba; Pasch 1985: 75 ff.) gender-like sex and animacy distinctions are integrated into the system of possessive classifiers. Numeral classifiers co-occur with gender, in one word, in Achagua (Arawak).

### 6.2. Gender and number

Languages of the world show the following types of interaction between gender and number:

- (a) Expression of number depends on gender: In Bantu languages, gender classes are defined via their correlations with number (Serzisko 1983: 97).
- (b) Gender and animacy distinctions are realized as portmanteau with number marking: Number and gender are realized as portmanteau in many fusional languages, e.g. Indo-European and Afroasiatic, and in predominantly agglutinating or polysynthetic languages, e.g. Arawak or Tucano, from South America.
- (c) There are fewer gender distinctions in non-singular numbers than in singular: All Bantu languages have fewer gender distinctions in singular than in plural. Fula (West-Atlantic) has 23 singular and five plural genders. Seneca (Iroquoian) has three genders in singular and two in dual and plural. In Old Church Slavonic three genders (masculine, feminine, and

- neuter) were distinguished in the singular, two in the dual (masculine and neuter having fallen together), while no gender distinctions were made in plural. There are a few exceptions to the generalization that fewer gender distinctions are made in non-singular than in singular. Tazewalt Shilh (North-Berber, Afroasiatic) distinguishes two genders in first person plural, but not in first person singular (see Aikhenvald 2000: 243–250).
- (d) Some languages have a different semantic basis for genders in singular and plural. Malto (South Dravidian) has male vs. the rest in singular, and human vs. non-human in plural (Mahapatra 1979: 60). Plural genders may involve **superclassing**. Motuna distinguishes four numbers for nouns with animate and human referents (singular, dual, paucal, plural) and two numbers (singular and non-singular) for nouns with inanimate referents.

Greenberg's (1963) Universal 32 states that "whenever the verb agrees with a nominal subject or nominal object in gender, it also agrees in number". There are significant exceptions to this claim, e.g. Caucasian languages, and numerous South-American and Papuan languages.

### 6.3. Gender and possession

There are a few cases of portmanteau realization of gender and possession. In the Western Torres Strait language, gender (feminine or masculine) of possessor is distinguished on the first person singular pronoun (Ford & Ober 1991).

### 6.4. Gender and case

The number of case distinctions may correlate with animacy following the nominal hierarchy formulated by Dixon (1994). These correlations are widespread in Australian languages. Neuter nouns have fewer case distinctions than masculine or feminine in all Indo-European languages.

A few Papuan languages (e.g. Yessan-Mayo; Foley 1986: 101) mark animate transitive objects with dative case. Inanimate transitive objects are unmarked.

In languages with two gender systems, these may relate to different case functions. Paumarí (Arawá) has two types of transitive construction, one of which cross-references the masculine/feminine gender of the subject

and the other that of the object. However, in both constructions, the verb cross-references the shape-based gender of the object.

### 6.5. Gender and person

Gender and person are often realized as portmanteau morphemes. In personal pronouns and person cross-referencing systems, gender distinctions are usually found in third person. If gender oppositions are found in second person, they will also be there in third, and if they are found in first, which is rare, they will also be there in second and third.

A few languages distinguish masculine and feminine genders in all persons in plural pronouns, but only in third person in the singular. In Spanish, first and second person plural pronouns have been derived from a combination of a personal pronoun and the indefinite adjective *otro* 'other', e.g. *nosotros* 'we masculine', *nosotras* 'we feminine'. 'Masculine' plural pronouns are used for default agreement.

### 6.6. Gender and declension

Gender and declension correlate in languages with morphological gender assignment. A few languages which have lost genders, have retained traces of gender-like distinctions in declensions. Examples are Lezgian, a North-East Caucasian language which lost its gender system (Haspelmath 1993: 76–77), and Armenian.

### 6.7. Gender and discourse

Gender marking is often associated with the role a noun has in discourse.

In Alamblik overt gender marking on nouns correlates with focussing a particular, shape-related property of the noun. The use of *kuñ-r* 'house' with a masculine rather than the usual feminine suffix indicates that the house is an unusually long one (Bruce 1984: 97).

The correlation between gender marking and individuation is the basis for tracing gender markers to article-like elements (see 7.1 (d); Greenberg 1978: 61 ff.).

In the Australian languages Nunggubuyu (Heath 1983) and Warray the absence of a noun class prefix indicates that the noun is focussed. Gender markers are used as anaphoric elements for participant tracking in Australian and Papuan languages. This property is shared with classifier systems.

### 6.8. Gender and derivation

Gender markers often combine inflectional (i.e. concordial) and derivational properties (Aikhenvald 2000:30). Gender is productively used in derivation in Brazilian Portuguese, e.g. *ministr-o* ‘minister’, *ministr-a* ‘she-minister’. In Bantu languages, e.g. Swahili, most stems usually occur with a prefix of one class. Prefixes can be substituted to mark a characteristic of a referent. *M-zee* means ‘old person’ and has the human class prefix *m-*. It can be replaced by *ki-* (inanimate class) to yield *ki-zee* ‘scruffy old person’ (Dixon 1982: 166).

Some languages have no gender agreement, but gender can be marked on derivational suffixes, e.g. Estonian (Balto-Finnic) *kuningas* ‘king’, *kuninganna* ‘queen’ (see Aikhenvald 2000: 436–441).

## 7. Diachronic dimensions of gender

### 7.1. The origin and development of gender

The following **language internal sources** can be established for the rise of gender systems.

- (a) Genders, like classifiers (see Art. 97), can originate from lexical nouns: In Mupun (West Chadic), overt gender marking on nouns emerged through lexicalization of nouns ‘man’ and ‘woman’ (Frajzyngier 1993: 49). Diminutive gender marker *pi* in Bantu languages of Cameroon comes from Niger-Congo \**bi* ‘child’ (Heine 1982: 214).
- (b) Genders can originate in classifiers: This scenario is reconstructed for Proto-Australian by Sands (1995). In Ngan’gyem-erri and Marrithiyel the noun class prefixes have developed from generic classifiers via phonological depletion (Dixon 1982: 171; Sands 1995: 253; Reid 1997).
- (c) Genders can originate in a reinterpretation of another nominal category, e.g. number. In Kiowa-Tanoan languages, nouns must have had an inherent number which conditioned the organization of gender classes.
- (d) Gender and gender agreement can originate in gender-marking pronouns or demonstratives: A gender agreement system can arise through the grammaticalization of a deictic element which first functions as a definite article. The stages of development from Proto-Voltaic to modern languages are: definite article > non-ge-

neric article > class prefix (Greenberg 1978: 61 ff.).

New genders can appear through the reanalysis of derivational affixes. Indo-European languages originally had two genders: animate and inanimate (preserved in Hittite). Feminine gender appeared later within the animate through the reanalysis of feminine affix *-a:-, -ya:* (Meillet 1964).

Nominal genders often originate in genders used for pronouns (Zande, Fur; Heine 1982: 215). In Kxoe (Khoisan), gender markers were originally used only with pronouns; they spread to mark gender on nouns.

**Language external sources** for the origin of genders usually involve areal diffusion patterns. Both expansion and reduction of gender and noun class systems very often happen through diffusion. In the Mayali dialect chain (Australian), the western-most dialects are becoming more like their neighbour Jawoyn, with only three, instead of four, agreement classes, and the eastern-most dialects are tending to become more like their neighbours Rembarrnga and Dangbon, both of which lack noun class agreement (Evans 1997). Genders are areal properties of many linguistic areas, e.g. the Sepik region in Papua New Guinea.

Changes in gender systems may involve the creation of new genders. Slavonic languages innovated new subgenders based on animacy opposition (Corbett 1991: 296).

Changes may affect the composition of genders. In Lunda (Bantu), ‘human’ gender 1 became ‘animate’.

Gender systems can shift from more grammatical to more semantic, as in Cantabrian Spanish (cf. 3.4). Loss of formal gender markers can make gender assignment rules opaque. In French, the loss of final *-e*, the feminine marker, has led to the creation of complicated phonological and morphological rules of assignment.

### 7.2. Gender decay and loss

There are the following **language internal** reasons for gender loss:

- (a) Coalescence of paradigms, for phonological reasons: In Indic and Iranian languages, masculine and feminine declension paradigms merged. At first this resulted in a confusion in gender agreement (Kölver 1982), and later, in a complete loss of gender oppositions in a number of Indic and Iranian languages (e.g., As-

- samese, Bengali, Nepali, Oriya; and Persian, Beludzhi and Ossete).
- (b) Loss of formal markers: Neutralization of genders can take place with the loss of pronominal suffixes, as in the North-Arawak language Bahwana.
  - (c) Loss of concord: The loss of gender in Old English is partly due to the loss of agreement with adjectives and demonstratives.
  - (d) Expanding the distribution of an unmarked gender: The non-human/neuter pronoun \**atu* in Proto-Dravidian gradually expanded to cover other classes (Krishnamurti 1975: 345).

Gender and gender agreement are most persistent in personal pronouns (cf. Demuth et al. 1986: 459 for Cross River and Kru languages; Priestly 1983 for Indo-European). Many Niger-Congo languages have lost their nominal gender, but retain pronominal genders (Heine 1982: 196).

The loss of concord systems often involves lexicalization and fossilization of gender markers, as is the case in Australian languages of Arnhem Land (Sands 1995: 255 f.).

When genders are being lost, the agreement rules are more persistent than the formal categorization; thus the gender loss may provide a mirror image of gender genesis (Demuth et al. 1986; Herbert 1991).

When genders are lost, they often leave a trace somewhere in the grammar. In Bengali, the traces of an old masculine/feminine opposition are retained in so called ‘enclitic definitives’ used with numerals in a function similar to numeral classifiers. The feminine form became reinterpreted as an “endearing diminutive”, and the masculine form became unmarked (Onishi p. c.). Kegboid languages, a subgroup of the Benue Congo spoken in Nigeria, lost the Proto-Benue Congo noun class system, but acquired a very peculiar system of numeral classifiers, an unusual phenomenon for African languages.

Loss or reduction of gender classes is often observed in a language death situation (cf. Dorian 1978). Traditional Dyirbal had four semantic genders (cf. 3.1). Young people’s Dyirbal has gradually adjusted its gender system towards that of English and made it more semantically motivated. It lost the gender meaning ‘non-flesh food’. The scope of ‘feminine’ gender 2 became reduced and reserved only for females (it used to include water, fire, and things associated with fight-

ing). Gender assignment by mythical association was lost; exceptions became regularized; and the use of the residue class was expanded (Schmidt 1985).

Gender agreement on adjectives has become optional in Warekena and Bare, endangered North-Arawak languages.

In Paumari, shape-based gender is being lost by younger speakers. The feminine-masculine distinction is more persistent.

Areal influence may be a **language external** source for gender reduction and loss. Absence of genders is an areal feature of a number of linguistic areas, e.g. Meso-America.

## 8. Uncommon abbreviations

PRON pronominal gender

## 9. References

- Aikhena, Alexandra Y. (2000), *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press
- Aikhena, Alexandra Y. & Dixon, R. M. W. (1998), “Dependencies between Grammatical Systems”. *Language* 74, 56–80
- Alpher, Barry (1987), “Feminine as the Unmarked Grammatical Gender: Buffalo Girls are no Fools”. *Australian Journal of Linguistics* 7, 169–187
- Bani, Ephraim (1987), “Masculine and Feminine Grammatical Gender in Kala Lagaw Ya”. *Australian Journal of Linguistics* 7, 189–201
- Bruce, Les (1984), *The Alambak Language of Papua New Guinea (East Sepik)*. Canberra: Pacific Linguistics
- Chapman, Shirley & Derbyshire, Desmond C. (1991), “Paumari”. In: Derbyshire, Desmond C. & Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages, Vol. III*. Berlin: Mouton de Gruyter, 161–352
- Corbett, Greville (1991), *Gender*. Cambridge: Cambridge University Press
- Craig, Colette (1986, ed.), *Noun Classes and Categorization. Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Amsterdam: John Benjamins
- Demuth, Katherine & Faraclas, Nicholas & Marchese, Lynell (1986), “Niger-Congo Noun Class and Agreement Systems in Language Acquisition and Historical Change”. In: Craig (ed.), 453–471
- Dimmendaal, Gerrit J. (1983), *The Turkana Language*. Dordrecht: Foris Publications

- Dixon, R. M. W. (1972), *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press
- Dixon, R. M. W. (1982), *Where Have All the Adjectives Gone? And Other Essays in Semantics and Syntax*. Berlin: Mouton
- Dixon, R. M. W. (1994), *Ergativity*. Cambridge: Cambridge University Press
- Dixon, R. M. W. (1995), "Fusional Development of Gender Marking in Jarawara Possessed Nouns". *International Journal of American Linguistics* 61.3, 263–294
- Dixon, Robert M. W. (2002), *Australian Languages: Their Nature and Development*. Cambridge: Cambridge University Press
- Dorian, Nancy (1978), "The Fate of Morphological Complexity in Language Death". *Language* 54, 590–609
- England, Nora (1983), *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press
- Evans, Nicholas (1997), "Head Classes and Agreement Classes in the Mayali Dialect Chain". In: Harvey & Reid (eds.), 105–146
- Foley, William A. (1986), *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press
- Ford, Kevin & Ober, Dana (1991), "A Sketch of Kalaw Kawaw Ya". In: Romaine, Suzanne (ed.), *Language in Australia*. Cambridge: Cambridge University Press, 118–142
- Frajzyngier, Zygmunt (1993), *A Grammar of Mu-pun*. Berlin: Dietrich Reimer Verlag
- Greenberg, Joseph H. (1963), "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". In: Greenberg, Joseph H. (ed.), *Universals of Language*. Cambridge/MA: M.I.T. Press, 58–90
- Greenberg, Joseph H. (1978), "How Does Language Acquire Gender Markers". Greenberg, Joseph H. & Ferguson, Charles A. & Moravcsik, Edith A. (eds.), *Universals of Human Languages, Vol. III: Word Structure*. Stanford: Stanford University Press, 47–82
- Harvey, Mark & Reid, Nicholas (1997, eds.), *Nominal Classification in Aboriginal Australia*. Amsterdam: John Benjamins
- Haspelmath, Martin (1993), *A Grammar of Lezgian*. Berlin: Mouton de Gruyter
- Heath, Jeffrey (1983), "Referential tracking in Nunggubuyu". In: Haider, John & Munro, Pamela (eds.), *Switch-Reference and Universal Grammar*. Amsterdam: John Benjamins, 129–149
- Heine, Bernd (1982), "African Noun Class Systems". In: Seiler & Lehmann (eds.), 189–216
- Herbert, Robert K. (1991), "Patterns in Language Change, Acquisition and Dissolution: Noun Prefixes and Concords in Bantu". *Anthropological Linguistics* 33, 103–134
- Hewitt, George (1979), *Abkhaz*. Amsterdam: North Holland
- Holmquist, Jonathan Carl (1991), "Semantic Features and Gender Dynamics in Cantabrian Spanish". *Anthropological Linguistics* 33, 57–81
- Khaidakov, Said M. (1980), *Principy imennoj klasifikacii v dagestanskikh jazykakh*. Moscow: Nauka
- Kibrik, Aleksandr E. & Kodzasov, Sandro V. & Olovjannikova, I. P. (1972), *Fragmenty grammatiki hinalugskogo jazyka*. Moscow: Izdateljstvo Moskovskogo Universiteta (Publikacii OSIPL 9)
- Kölver, Ulrike (1982), "Interaktion von nominalen Kategorien am Beispiel der Entwicklung des modernen Bengali". In: Seiler & Lehmann (eds.), 244–251
- Krishnamurti, Bhadriraju (1975), "Gender and Number in Proto-Dravidian". *International Journal of Dravidian Linguistics* 4.2, 328–350
- Krishnamurti, Bhadriraju & Gwynn, John P. L. (1985), *A Grammar of Modern Telugu*. Delhi: Oxford University Press
- Mahapatra, Bijaya P. (1979), *Malto: An Ethnosemantic Study*. Mysore: Central Institute of Indian Languages
- Meillet, Antoine. (1964), "The Feminine Gender in the Indo-European Languages". In: Hymes, D. (ed.), *Language in Culture and Society*. New York: Harper and Row
- Merlan, Francesca (1983), *Ngalkan Grammar, Texts and Vocabulary*. Canberra: Pacific Linguistics
- Onishi, Masayuki (1994), *A Grammar of Motuna (Bougainville, Papua New Guinea)*. Unpublished PhD Thesis. Canberra: ANU
- Pasch, Helma (1985), "Possession and Possessive Classifiers in 'Dongo-ko'". *Afrika und Übersee* 68, 69–85
- Priestly, Tom M. S. (1983), "On 'Drift' in Indo-European Gender Systems". *Journal of Indo-European Studies* 11.3–4, 339–363
- Reid, Nicholas (1997), "Class and Classifier in Ngan'gyemerri". In: Harvey & Reid (eds.), 165–228
- Sands, A. Kristina (1995), "Nominal Classification in Australia". *Anthropological Linguistics* 37, 247–346

- Schmidt, Annette (1985), *Young People's Dyirbal: An Example of Language Death from Australia*. Cambridge: Cambridge University Press
- Seiler, Hansjakob & Lehmann, Christian (1982, eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen, Teil I: Bereich und Ordnung der Phänomene*. Tübingen: Gunter Narr Verlag
- Serzisko, Fritz (1983), "Gender, Noun Class, and Numeral Classification: a Scale of Classificatory Techniques". In: Dirven, René & Radden, Günther (eds.), *Issues in the Theory of Universal Grammar*. Tübingen: Gunter Narr Verlag.
- Tübingen: Gunter Narr Verlag. (Tübinger Beiträge zur Linguistik 196), 92–123
- Welmers, William E. (1973), *African Language Structures*. Berkeley: University of California Press
- Zubin, David & Köpcke, Klaus Michael (1986), "Gender and Folk Taxonomy: the Indexical Relation between Grammatical and Lexical Categorization". In: Craig (ed.), 139–180

*Alexandra Y. Aikhenvald, Melbourne  
(Australia)*

## 99. Diminution and augmentation

1. Terms and definitions
2. Formal aspects
3. Semantic aspects
4. Contextual aspects
5. Etymological aspects
6. References

### 1. Terms and definitions

Broadly speaking, the term **diminutive** refers to any formation in a language expressing the referential meaning ‘small’, and possibly a variety of derived evaluative shades of meaning. The formal devices that may be used in diminution include both analytic (paraphrastic) and synthetic (morphological) diminutives. However, the term “diminutive” is usually restricted to synthetic diminutives only. The **augmentative** is the semantic counterpart of the diminutive: it denotationally expresses the concept ‘big’, and may have derived readings such as evaluative exaggeration and intensification.

### 2. Formal aspects

The main bibliographical keys to the study of diminution and augmentation are Ettinger (1980), who discusses publications from 1900 to 1975, and Dressler & Merlini Barbaresi (1994). Numerous studies have been devoted to the formal, especially phonological properties of diminution, while augmentation has received relatively little attention. Most studies focus on one of the following aspects.

### 2.1. The formation of diminutives and augmentatives

According to the specific kind of the morphological mechanisms used, several subtypes may be identified within the synthetic diminutive (compare Hasselrot 1957: 283–288 for examples). **Prefixation** (cf. Art. 54), for instance, is used in Swahili (Herms 1989) and in Japanese: *kawa* ‘river’ becomes *ko-gawa* ‘small river’. Diminutive **infixation** (cf. Art. 55) is illustrated by the infixes *-ra-* and *-da-* in Hottentot of the Korana and Nama group: *qāb* ‘river’ next to *qarob* ‘brook’, and *qhomī* ‘mountain’ next to *qhomdai* ‘hill’. **Suffixation** (cf. Art. 55) is characteristic for Romance, Slavonic, and Germanic languages (Ettinger 1974; Klimaszewska 1983). **Reduplication** (cf. Art. 57) can be found in Polynesian and Melanesian languages, such as Malay, where *budak* ‘child’ receives the diminutive form *budak-budak* ‘small child’. The mechanisms of diminutive formation may be **submorphemic**, when phonemes are modified in a systematic way to express diminution. An example from Bask is the formation of *txerri* ‘small pig’ from *zerri* ‘pig’ (compare Nichols 1971 for Amerindian examples).

Prefixation, infixation, suffixation, and reduplication can, at least in principle, be accounted for with word formation rules, but this is not possible for submorphemic diminutive formation, which requires a more specific, phonological approach. In the framework of prosodic morphology, a solution is proposed for several languages, such as Arabic (McCarthy & Prince 1990) and Chinese (Yip 1992).

The mechanisms mentioned here do not necessarily exclude each other. Analytic and synthetic diminutives may co-occur, as in German *ein kleines Häuschen* ‘a small small house’. In Italian, suffixation and reduplication are sometimes used simultaneously with an intensifying effect: *piano* ‘slow’ becomes *pianin(o) pianino* ‘very slowly’ by a reduplication of the suffixed diminutive form *pianino*. *Pian(o) pianino*, on the other hand, is formed by applying the diminutive suffix to the reduplicated base form. Also, diminutive suffixes can be used recursively, as in the transition from *corto* ‘short’ to *cortino in oino* ‘very very short’ or in the transition from *tazzo* ‘cup’, to *tazzettina* ‘small cup’, or to *tazzinetta* (with the reverse order of the two diminutive suffixes found in *tazzettina*). While recursiveness and reduplication are marginal processes in Italian morphology (Dressler & Merlini Barbaresi 1994: 99–101), cumulation of diminutive suffixes is crucial to express evaluative meanings in Slavonic languages (Wierzbicka 1991: 52 f.).

The formal mechanisms of **augmentative** formation fall into the same basic categories as the processes of diminutive formation. The augmentative meaning can be expressed analytically with a phrasal paraphrase of the concept ‘big’, and synthetically with morphological devices. In Dutch, augmentatives are derived from nouns and adjectives by adding prefixes such as *reuze-* (*reuzeprobleem* ‘giant problem’), *mega-* (*megadisco* ‘large disco’), and *super-* (*supertanker* ‘huge tanker’). In Italian, augmentatives are formed by suffixation with *-one* (*governone* ‘big government’). Augmentatives are less widespread than diminutives. The two categories are related by an implicational universal: the existence of augmentatives in a language implies the presence of diminutives, but the reverse does not hold. If a language has both categories, diminutives are more frequent and can be formed in more ways than augmentatives (Schneider 1991; Dressler & Merlini Barbaresi 1994: 430 f.).

## 2.2. Restrictions on diminution

The process of diminutive formation may be subject to various types of restrictions. First, we may consider **syntactic restrictions**. Diminutives can be formed on the basis of various syntactic categories. In Dutch, German, and Italian, for instance, they are obtained by adding suffixes to nouns, and much less so, to adjectives and verbs, while derivation

from other categories is rare. The following examples exemplify the wide range of possibilities that exist in Dutch: *tafeltje* ‘small table’ from the noun *tafel*, *nieuwtje* ‘news item, piece of news’ from the adjective *nieuw*, *koopje* ‘bargain’ from the verb *kopen*, *extraatje* ‘something extra’ from the adverb *extra*, *tientje* ‘(a note of) ten guilders’ from the numeral *tien*, *uitje* ‘trip, outing’ from the preposition *uit*, *ietsje* ‘a little something’ from the pronoun *iets*. Additional examples illustrate the phenomenon of **diminutiva tantum**: diminutives that are derived from a virtual or nonexistent base. *Theekrans* is a possible but unattested compound, formed on the basis of *thee* ‘tea’ and *krans* ‘circle’; the diminutive *theekrantsje*, on the other hand, is quite common. *Hachje* in the expression *zijn hachje erbij inschieten* ‘to loose one’s life’ is synchronously untransparent.

Dressler & Merlini Barbaresi (1994: 94, 120–132) try to explain the differences in the productivity of diminution with different word classes on a semantic basis. They point out that only syntactic categories whose referents have gradable dimensions can be diminutivized: nouns referring to objects may have several gradable dimensions, adjectives are mostly gradable by nature, and the intensity of the action denoted by verbs can be diminutivized as well.

A second type of restriction involves the **morphological structure of the base**. In Dutch, words with a certain morphological structure can not be the input for diminutive formation. Derivations on *-de*, *-isme*, *-dom*, *-nis*, and *-schap*, as in *gezegde* ‘saying’, *communisme* ‘communism’, *rijkdom* ‘wealth’, *droefenis* ‘sorrow’, *blijdschap* ‘gladness’ cannot be diminutivized (Klimaszewska 1983: 22 f.). Exceptions such as *ik moet nog even een boodschapje doen* ‘I must do some shopping, must run a small errand’ do exist, but can be explained semantically. Derivations on *-schap*, like those on *-isme*, *-dom*, and *-nis* are abstract terms, which cannot be diminutivized, while *boodschap*(*pen*) ‘errand(s)’ has a concrete meaning and can be diminutivized. The restriction, then, would seem to be of a semantic nature rather than purely formal.

A specific case of morphological restrictions involves the cumulation of diminutive forms. In Italian, evaluative suffixes can be stacked, but the cumulation is restricted. Diminutive and augmentative suffixes cannot be combined freely: the Italian augmentative

suffix *-one* cannot be combined with the diminutive suffixes *-ino* and *-uccio*, while a combination with other diminutive suffixes, such as *-accio*, *-ello*, and *-etto* is possible (*corpaccione* ‘big, ugly body’, *tortellone* ‘big tortellini’, *cassettone* ‘chest of drawers’). Also, a combination of three diminutive suffixes is rare, though not entirely excluded, as in *storiellucciaccia* ‘spicy story, gossip’ and *librettucciaccio* ‘insignificant booklet’ (Ettinger 1974: 162–170).

When languages possess several allo-morphs or separate suffixes to express diminution or augmentation, **phonological restrictions** may be involved in the distribution of the forms. First, the suffixes may be used in a complementary way. For instance, syllable structure plays a role in Rumanian, which has a class of diminutive suffixes (including *-icel* and *uleț*) that occur only after monosyllabic bases, while another class (including *-aș* and *-el*) occurs only after bisyllabic bases (Ettinger 1974: 278, 285). Second, the environments of the suffixes may overlap, as in the distribution of *-chen* and *-lein* in German: *-chen* is used after *l(e)* (*Spielchen* ‘game’), while *-lein* comes after *ch*, *g*, and *ng* (*Bächlein* ‘brook’, *Zweiglein* ‘twig’, *Ringlein* ‘small ring’). The distribution of both suffixes is free in other situations: compare *Brieflein* and *Briefchen* ‘small letter, note’, *Tischlein* and *Tischchen* ‘small table’ (Ettinger 1974: 75f.). Third, the range of one suffix may include that of the other: one of the suffixes has a wider distribution than the other. In Portuguese, the suffix *-inho* can always be used in contexts where the suffix *-ito* can occur, but the first is much more frequent than the last. On the one hand, we find doublets such as *dentitos*, *dentinhos* ‘small teeth’, or *notita*, *notinha* ‘small note’. On the other hand, one of the forms may have a specialized meaning, as in *corpinho* ‘body (garment)’ versus *corpito* ‘small body’ (Ettinger 1974: 203).

We have already encountered examples of **semantic restrictions** on diminution: it was suggested above that some of the syntactic and morphological restrictions may be explained in semantic terms. However, the factors already mentioned do not exhaust the range of possible semantic effects. In German, for instance, it appears that general terms can easily be diminutivized (*Fischlein*, *Fischchen* ‘small fish’), while diminution of taxonomically more specific ones is restricted (\**Bärschlein* ‘small perch’, \**Hechtchen* ‘small

pike’), though not excluded (*Goldfischchen* ‘small goldfish’) (Ettinger 1974: 367). Unfortunately, a typologically adequate overview of the semantic restrictions on diminution is not available.

### 2.3. Diminution between derivation and inflection

Mechanisms of diminution may somewhat obscure the distinction between derivation and inflection. First, in some languages the formal properties of diminution seem to lie in between those of inflections and derivations. Second, there appear to be many cases in which diminution interacts closely with inflectional processes (Bauer 1997).

The intermediate status of diminution involves the question whether diminutive suffixes are heads. In Italian, diminution has characteristics of both derivation and inflection (Scalise 1984: 131–133). On the one hand, diminutives seem to have an overall inflectional status because they do not change the syntactic characteristics of the base. On the other hand, however, this need not always be the case. Sometimes, in fact, the gender is modified: it is possible to derive a masculine diminutive (*festino*) from a feminine noun (*festa* ‘party’), or to derive a feminine diminutive (*carretta*) from a masculine noun (*carro* ‘car’). This can result in duplicates like *casetta* (f.), *casotto* (m.) ‘cottage’ from *casa* (f.) ‘house’ (Pellegrini 1977: 13f.). Scalise (1984) further argues that Italian diminutives have idiosyncratic properties that mark them as a formal category on their own, somewhere in-between inflection and derivation. First, diminution applies after derivational rules and before inflectional ones. In *contrabandierucoli* ‘small-time smugglers’, for instance, the diminutive suffix *-ucolo* is situated after the agentive suffix *-iere*, but before the masculine plural suffix *-i*. Second, diminutives can be formed by the recursive or repeated application of the same word formation rule (as in *cortinoinoino*). However, these formal characteristics are not universal (Stump 1993).

The close connection between diminution and inflection also shows up in the interaction between diminutive formation and inflectional processes (cf. Art. 97, 98, 100). In Fula, for example, the form of diminutives and augmentatives depends on the nominal class and the number of their base (Anderson 1992: 80–82). In Swahili the distribution of the diminutive prefixes *ki-* and *vi-* is governed

by the number of the base, the former prefix being used in the singular, the latter in the plural: *kundi* ‘group’, *kikundi* ‘small group’, *vikundi* ‘small groups’, or *mlima* ‘mountain’, *kilima* ‘small mountain’, *vilima* ‘small mountains’ (Herms 1989).

A special case of the interaction between diminution and inflection is the syncretism between feminization and diminution (cf. Art. 98). Hasselrot (1957: 289–299) explains this as a kind of linguistic anthropomorphism: diminutivized forms have feminine gender because they denote smallness and weakness, features that are associated with women. In Bengali masculine nouns are feminized to express diminution: *dubha* ‘big metallic ball’, *dubhi* ‘small metallic ball’, and *hattaura* ‘big hammer’, *hattauri* ‘small hammer’ exhibit the same opposition as *kora* ‘boy’, *kuri* ‘girl’. In the language of the Massai big things obtain the masculine article *ol-*, whereas small objects get the feminine article *en-*: *ol-gume* ‘big nose’, *en-gume* ‘small nose’; *ol-alem* ‘sword’, *en-alem* ‘knife’. The connection between feminization and diminution does not appear in gender distinctions only. In Dutch, the diminution of first names for boys such as *Dirk* or *Geert* results in names for girls: *Dirkje*, *Geertje*. In Malay, adjectives that express the notions ‘male’ and ‘female’ have an augmentative and a diminutive function respectively: *bah jantan* ‘male river → big river’, *bah betina* ‘female river → small river’.

To be sure, the association between diminutive and feminine forms is not universal. In some languages the opposition between augmentatives and diminutives is not expressed by a male-female contrast, but rather by an adult-child distinction. In Zulu, for example, the diminutive suffix *-ana* means ‘child’, while the augmentative suffix *-kazil kali* refers to ‘mother’. Diminution can also be accompanied by the neutralization of gender, as in Dutch and German, or by a transposition from [+animate] to [−animate], as in Blackfoot, Chinook, and Swahili.

### 3. Semantic aspects

Diminutives and augmentatives are essentially semantic categories, but paradoxically their semantics has not often been investigated in systematic fashion. Even so, the following observations may be derived from the literature.

#### 3.1. Prototype effects

The readings of the diminutive suffix in Dutch can be classified as a prototypical category (Bakema et al. 1993). The core of this structure contains the literal readings of the diminutive, when it functions as a marker of referential smallness. Within this literal, referential core, a distinction is necessary between ‘diminutive’ and ‘explicative’ meanings. Diminutives that name a small exemplar of the category named by the base may be called **diminutive** in the strictest sense, while diminutives that have the same (small) referent as their base, are called **explicative**. *Gebouwtje* ‘small building’ and *tafeltje* ‘small table’ are diminutive in the strict sense, while *peukje* ‘cigarette butt, stub’ and *madeliefje* ‘daisy’ are explicative (*peuk* and *peukje*, then, are synonymous). Both meanings can occur with regard to various (often co-occurring) dimensions, viz. the dimension of space and size (as in *gebouwtje*), time (*reisje* ‘small journey, trip’, next to explicative *ogenblikje* ‘a moment’), intensity (*buitje* ‘a shower which is not heavy’, *kusje* ‘a kiss which is not passionate’ next to explicative *briesje* ‘a slant of wind’), and age (*tijgertje* ‘young tiger’ and *olifantje* ‘young elephant’ next to explicative *veulentje* ‘a foal’).

From these core meanings two groups of meanings are derived by processes of metaphorical and metonymical extension. The first group contains **evaluative** readings. Prominent among these readings are depreciation (*romannetje* ‘insignificant novel’), appreciation and affection (*broertje* ‘dear brother’), approximation (*kilootje* ‘roughly a kilo’) and relativization (*cadeautje* ‘a modest present’). Again, the various readings may co-occur: in *kijk daar, een musje* ‘look there, a (dear little) sparrow’, the use of the diminutive may be simply explicative, but it may at the same time express affection. Wierzbicka (1991: 51) gives the Polish example *weź jeszcze śledzika!* (‘take some more dear-little-herring!’), in which the diminutive is not only appreciative in praising the food, but also minimizes the quantity. Conversely, some of the readings cannot be combined, as they involve incompatible metaphors. The appreciative and the depreciative readings, for instance, involve the general metaphors ‘small is beautiful’ and ‘small is insignificant’ respectively.

The second group of extensions is formed by **partitive** shades of meaning: *chocolaatje* ‘a piece of chocolate’ is a count noun derived from the mass noun *chocolade* ‘chocolate’.

The semantic basis of this kind of use is metonymical: isolating a small quantity from the global mass takes the form of isolating an entity made from that particular substance. This ‘entity-forming’ reading of the diminutive is generalized to cases in which the base is not a mass noun naming a substance: *blondje* ‘little blond one, i.e. blond woman’ is derived from the adjective *blond*, and *weetje* ‘something one knows, i.e. interesting piece of knowledge’ is derived from the verb *weten* ‘to know’.

Finally, a grammaticalized **metasemantic** function may be identified as the most peripheral member of the category. An example is *telefoontje*, which doesn’t mean ‘small telephone’ or ‘dear telephone’ or ‘cursed telephone’, but ‘conversation over the telephone’. In cases such as these, the suffix merely seems to signal that the base must be interpreted figuratively, but it does not seem to impose a particular type of figurative reading on the formation. Another example in Dutch is *hartje* ‘small heart’ in contexts such as *het hartje van de stad* ‘the centre of the city’ and *het hartje van de winter* ‘the dead of winter’ (Dirven 1985).

**Augmentatives** are the semantic counterparts of diminutives: they strengthen rather than diminish the dimensional aspects of their bases. Augmentatives express intensification or exaggeration, as in the Dutch examples *superster* ‘very famous person, superstar’ and *supergevoelig* ‘highly sensitive, touchy in an exaggerating way’. Diminutives and augmentatives alike are Janus-headed categories, since both can have either appreciative or depreciative readings (Hasselrot 1957: 307). Appreciative uses are more characteristic for diminutives than depreciative readings, while augmentatives have the inverse hierarchy (Klimaszewska 1983: 11; Schneider 1991). Diminution of a base that already names something small may intensify the expression of smallness, and may as such take on an ‘augmentative’ function.

### 3.2. Semantic typology

A number of studies may be cited as indicating that the prototypically organized structure described above is not restricted to Dutch only. Rhodes (1990), Tabakowska (1993: 100–110) and Taylor (1989: 144–149), for instance, give short prototypical analyses of diminutives in respectively Ojibwa, Polish, and Italian. Synthetic diminutives in Dutch, German, and Polish share not only the mean-

ing ‘small’ but also evaluative or expressive uses like ‘endearment’, ‘compassion’, ‘understatement’, ‘pejoration’ and ‘irony’ (Klimaszewska 1983). Latin diminutives as well exhibit similar semantic properties as the Dutch ones: *asellus* ‘little donkey’ has a diminutive reading, *aquula* ‘squirt (of water)’ has a partitive meaning, while others express affection or modesty (Fruyt 1989). Finally, on the basis of material taken from 60 languages, Jurafsky (1996) demonstrates that the semantic aspects of the diminutive may be widely distributed among languages of different families.

Obviously, not all languages exhibit exactly the same prototype structure of the diminutive. In some languages only parts of the prototype structure are realized by the diminutive suffix; also, several suffixes may be used to express different parts of the structure. In German, analytic diminutives denote smallness, while the synthetic ones on *-chen* and *-lein* primarily have evaluative uses (Klimaszewska 1983: 53 f., 117–119; Schneider 1991). In Hungarian, diminutives on *-i* only have evaluative functions; as markers of intimacy and jocularity, they determine the social setting of the situation and the relation between speaker and hearer. Diminutives on *-kal-ke* and *-cskal-cske*, however, primarily have core meanings (Dressler & Kiefer 1990). In Fula, appreciative and depreciative shades are expressed with different forms (Anderson 1992: 80).

Somewhat less drastically, differences can exist in the **salience of the parts** in languages with a complete prototypical structure. In the Slavonic languages, the evaluative shades of meaning are foregrounded. In Polish, for instance, expressive affixes constitute a subsystem of their own: a personal name can get several different diminutive derivates, each denoting a different attitude (Wierzbicka 1991: 50–57). Rainer (1993: 540–543, 578–589) lists tens of suffixes to express diminutive meanings in Spanish: most of them share the same meanings, but differ in prominence of these meanings.

### 4. Contextual aspects

Geographic variation of diminutives has been investigated intensively. (Compare, for instance, Seibold 1983 for an overview of the research devoted to diminutives in the German dialects.) Here, we will concentrate on types of variation that have a less language-

specific impact, viz. variation of a pragmatic, stylistic, and sociolinguistic kind. In general, research into the effect of these variational dimensions on the use of diminutives reveals the distributional impact of the semantic aspects of the diminutive described above: if diminutives have evaluative overtones, they tend to occur regularly in specific speech acts, situations, and text types (Dressler & Merlini Barbaresi 1994: 170–394).

#### 4.1. Speech acts

The presence of diminutives correlates with the **illocutionary force** of speech acts. Greek diminutives, for instance, serve as politeness strategies in requests, offers, and compliments; speakers use them to claim common ground, to show solidarity towards the addressee or to display affectionate concern for imposing on the hearer's freedom of action (Sifianou 1992). Diminutives, as features associated with children, signal emotional involvement and are markers of modesty. Conversely, diminutives are hardly used in formal speech acts (such as declaring, condoling, and praying). In addition, diminutives expressing informality and intimacy are inappropriate in situations involving status differences between participants. In other speech acts, the pragmatic effects of diminutives may vary considerably. In euphemistic speech, for instance, the use of the diminutive may mitigate the illocutionary force of a taboo word. The speaker evokes an innocent children's world by the diminution of a vulgar word. In ironic language, the diminutive contradicts a dimension of the referent (for instance, when something big or important is described as tiny or irrelevant), while in sarcastic uses the feature 'non serious' is employed to ridicule someone by creating a distance between speaker and hearer.

#### 4.2. Situations and speech participants

Diminutives are frequent in **situations** with children, in the language of love and in speech directed to pets, and more generally, in situations exhibiting the features 'non serious', 'empathy', 'sympathy', 'familiarity' or 'intimacy'. With the diminutive we express our affection; it is more used in joyful and tender situations than in contexts characterized by anguish and fear.

When it is suggested that women use diminutives more often than men, social factors are mentioned as an explanation: women are considered to be more contact-

oriented than men, to select friends on the basis of personal experience, and to attach only secondary importance to motives such as status or professional interest (Dressler & Merlini Barbaresi 1994: 413). However, empirical verification for Dutch yields no conclusive evidence that women do indeed use diminutives more often than men (Brouwer 1982). Possibly, diminutives are more often used depreciatively towards women (Schneider & Schneider 1991).

An extreme case of situational differences is involved in the claim that the frequency of diminution is **culture-specific**. Languages do not only differ qualitatively in the nature of diminutive formation, but also exhibit quantitative differences in the amount of diminutive formation processes that they allow and the frequency with which they are used. Klimaszewska (1983: 115–117) points out that Dutch and German respectively use just one and two suffixes to create diminutives, while Polish has a number of polyfunctional suffixes. According to Wierzbicka (1991: 53), then, such differences reflect cultural values, because "rich systems of diminutives seem to play a crucial role in cultures in which emotions in general and affection in particular is expected to be shown overtly".

#### 4.3. Text type

The use of diminutives varies from text to text, and from text type to text type; not surprisingly, there is a correlation with some of the factors already identified. The observation that diminutives are very frequent in fairy tales (Klimaszewska 1983: 56, 110) obviously correlates with the fact that diminutives are frequent in speech situations involving children. More specific effects may arise, however. In the fairy tales of the brothers Grimm, for instance, diminutives on *-lein* are more frequent than diminutives on *-chen* (Scheidweiler 1984), which seems to suggest that *-lein* has a more poetic (or perhaps more archaic) stylistic value than *-chen*. In addition, the distribution of diminutives in the fairy tales seems to depend on the gender of the protagonists: in fairy tales in which young men lead an adventurous life, the amount of diminutives is much lower than in stories with female protagonists (Schneider & Schneider 1991).

In Dutch, the use of the diminutive forms of clothing terms correlates with the kind of magazine they appear in (Geeraerts et al. 1994: 181–188). The use of diminutives in-

creases when fashion magazines are more glossy and are more intended for women or young people. In technically specialized magazines and in magazines for a male audience, on the other hand, the use of diminutives decreases. Crucially, these results are statistically controlled for referential differences between the magazines, in the sense that differences in the occurrence of diminutives could be the result of the fact that some sources discuss small items of clothing more often than others. Rather, a stylistic explanation imposes itself: the distribution of diminutives in different types of magazines reflects the informal, emotional shades of the diminutive in Dutch.

## 5. Etymological aspects

Restricting the discussion to affixation, three common sources of diminutives may be identified (compare Hasselrot 1957: 19–43, 300–304). First, they can be derived from words expressing the notion ‘small’ or related concepts. In Swedish, *små*, originally an adjective with the meaning ‘small’, shows up as a prefix in *småkyrka* ‘small church’ and *småbil* ‘small car’. Similarly, augmentatives can originate in nouns expressing the concept ‘big’. For example, the meaning of the first member in a compound like German *Riesenbetrieb* ‘big company’ can be bleached to an augmentative reading; analogously, starting with the Dutch noun *reus* ‘giant’, there is a line of increasing bleaching from *reuzenboom* ‘giant tree’ (huge in a literal sense) over *reuzedrijf* ‘giant company’ (huge in a figurative sense) to *reuzeleuk* ‘gigantically pleasant’ (huge in a purely augmentative sense). Second, diminutives are often derived from nouns that denote kinship relations, such as ‘son’ or ‘child’. In Japanese, the prefix *ko-* is derived from the noun *kō*, which means ‘child’. Even in Chinese, basically an isolating language, some words may be diminutivized by adding the “suffixes” *-érh* and *-tsi*, which denote ‘son’ and ‘child’ respectively. And third, diminutive suffixes may be derived from other suffixes. In the Romance languages, *-in* as a diminutive marker and *-on* as an augmentative marker may be traced to Latin suffixes. The *-in*-forms, as in Portuguese *-inho* and Italian *-ino*, originate from Latin *-inus*, which was used in names for young children to indicate a family relationship. The *-on*-suffixes originally had an indi-

vidualizing function, indicating someone with a striking quality. From this individualizing function, both diminutive and augmentative meanings are derived in later stages. In French the diminutive reading is realized as in *mignon* ‘darling, honey’, whereas in Italian *-one* has an augmentative meaning as in *librone* ‘big book’ (Ettinger 1980: 28–30). So both the formal and the semantic sources of the diminutive suffix may be quite diverse. In general, however, the relative neglect of the semantics of the diminutive shows up again here: a typologically adequate overview of the semantic sources of diminutive forms is not readily available.

## 6. References

- Anderson, Stephen R. (1992), *A-morphous Morphology*. Cambridge etc.: Cambridge Univ. Press (Cambridge Studies in Linguistics 62)
- Bakema, Peter & Defour, Patricia & Geeraerts, Dirk (1993), “De semantische structuur van het diminutief”. *Forum der Letteren* 34, 121–137
- Bauer, Laurie (1997), “Evaluative Morphology: in Search of Universals”. *Studies in Language* 21, 533–575
- Brouwer, Dédé (1982), “The Influence of the Addressee’s Sex on Politeness in Language Use”. *Linguistics* 20, 697–711
- Dirven, René (1985), “Metaphor as a Basic Means for Extending the Lexicon”. In: Paprotté, Wolf & Dirven, René (eds.), *The Ubiquity of Metaphor: Metaphor in Language and Thought*. Amsterdam, Philadelphia: Benjamins (Current Issues in Linguistic Theory 29), 85–119
- Dressler, Wolfgang U. & Kiefer, Ferenc (1990), “Austro-Hungarian Morphopragmatics”. In: Dressler, Wolfgang U. & Luschützky, Hans C. & Pfeiffer, Oskar E. & Rennison, John R. (eds.), *Contemporary Morphology*. Berlin, New York: de Gruyter (Trends in Linguistics. Studies and Monographs 49), 69–77
- Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia (1994), *Morphopragmatics: Diminutives and Intensifiers in Italian, German and Other Languages*. Berlin, New York: de Gruyter (Trends in Linguistics. Studies and Monographs 76)
- Ettinger, Stefan (1974), *Diminutiv- und Augmentativbildung: Regeln und Restriktionen: Morphologische und semantische Probleme der Distribution und der Restriktion bei der Substantivmodifikation im Italienischen, Portugiesischen, Spanischen und Rumänischen*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 54)

- Ettinger, Stefan (1980), *Form und Funktion in der Wortbildung. Die Diminutiv- und Augmentativmodifikation im Lateinischen, Deutschen und Romanischen (Portugiesisch, Spanisch, Italienisch und Rumänisch). Ein kritischer Forschungsbericht 1900–1975*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 47)
- Fruyt, Michèle (1989), “Étude sémantique des ‘diminutifs’ latins: Les suffixes *-ulus*, *-culus*, *-ellus*, *-illus* … dé-substantivaux et dé-adjectivaux”. *Cahiers de l’Institut de Linguistique de Louvain* 15, 127–138
- Geeraerts, Dirk & Grondelaers, Stefan & Bakema, Peter (1994), *The Structure of Lexical Variation: Meaning, Reference and Context*. Berlin, New York: de Gruyter (Cognitive Linguistics Research 5)
- Hasselrot, Bengt (1957), *Études sur la formation diminutive dans les langues romanes*. Uppsala: Lundquistska Bokhandeln; Wiesbaden: Otto Harrassowitz (Uppsala Universitets Årsskrift 11)
- Herms, Irmtraud (1989), “Zur Bildung von Diminutiv- und Augmentativformen im Swahili”. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42, 738–747
- Jurafsky, Dan (1996), “Universal Tendencies in the Semantics of the Diminutive”. *Language* 72, 533–578
- Klimaszewska, Zofia (1983), *Diminutive und augmentative Ausdrucksmöglichkeiten des Niederländischen, Deutschen und Polnischen: Eine konfrontative Darstellung*. Wrocław etc.: Zakład Narodowy Imienia Ossolińskich Wydawnictwo Polskiej Akademii Nauk
- McCarthy, John & Prince, Alan (1990), “Foot and Word in Prosodic Morphology: The Arabic Broken Plural?”. *Natural Language & Linguistic Theory* 8, 209–283
- Nichols, Johanna (1971), “Diminutive Consonant Symbolism in Western North America”. *Language* 47, 826–848
- Pellegrini, Ines Angela (1977), *Die Diminutive im Deutschen und im Italienischen*. Zürich: Juris
- Rainer, Franz (1993), *Spanische Wortbildungslehre*. Tübingen: Niemeyer
- Rhodes, Richard A. (1990), “Lexical Hierarchies and Ojibwa Noun Derivation”. In: Tsouhatzidis, Sava L. (ed.), *Meanings and Prototypes: Studies in Linguistic Categorization*. London, New York: Routledge, 151–158
- Scalise, Sergio (1984), *Generative Morphology*. Dordrecht, Cinnaminson/NJ: Foris
- Scheidweiler, Gaston (1984), “Zur Konnotation der Diminutivsuffixe *-chen* und *-lein*: prosaisch oder poetisch?”. *Muttersprache* 95, 69–79
- Schneider, Iris & Schneider, Klaus P. (1991), “‘Ach Kindchen, davon verstehen Sie nichts!’: Über den sexistischen Gebrauch deutscher Diminutivformen”. In: Feldbusch, Elisabeth & Pogarell, Reiner & Weiss, Cornelia (eds.), *Neue Fragen der Linguistik. Akten des 25. Linguistischen Kolloquiums, Paderborn 1990*. Tübingen: Niemeyer (Linguistische Arbeiten 271), 169–174
- Schneider, Klaus P. (1991), “Affektive Lexik: Kognitive, semantische und morphologische Aspekte”. In: Klein, Eberhard & Pouradier Duteil, Françoise & Wagner, Karl Heinz (eds.), *Betriebslinguistik und Linguistikbetrieb: Akten des 24. Linguistischen Kolloquiums, Universität Bremen 4.–6. September 1989*. Tübingen: Niemeyer (Linguistische Arbeiten 260), 233–241
- Seibold, Elmar (1983), “Diminutivformen in den deutschen Dialekten”. In: Besch, Werner & Knoop, Ulrich & Putschke, Wolfgang & Wiegand, Herbert Ernst (eds.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Berlin: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft 2), 1250–1255
- Sifianou, Maria (1992), “The Use of Diminutives in Expressing Politeness: Modern Greek versus English”. *Journal of Pragmatics* 17, 155–173
- Stump, Gregory (1993), “How Peculiar is Evaluative Morphology?”. *Journal of Linguistics* 29, 1–36
- Tabakowska, Elżbieta (1993), *Cognitive Linguistics and Poetics of Translation*. Tübingen: Narr
- Taylor, John R. (1989), *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Clarendon Press
- Wierzbicka, Anna (1991), *Cross-cultural Pragmatics: The Semantics of Human Interaction*. Berlin, New York: de Gruyter (Trends in Linguistics. Studies and Monographs 53)
- Yip, Moira (1992), “Prosodic Morphology in Four Chinese Dialects”. *Journal of East Asian Linguistics* 1, 1–35

*Peter Bakema, Lint (Belgium)  
Dirk Geeraerts, Leuven (Belgium)*

## 100. Numerus

1. Einleitung
2. Numeruskategorien
3. Numerusfähige Elemente
4. Art der Kodierung
5. Locus der Kodierung
6. Evolution der Numerusmorphologie
7. Unübliche Abkürzungen
8. Zitierte Literatur

### 1. Einleitung

Numerus ist eine grammatische Kategorie, die der Quantifikation von Referenten und Handlungen dient.

Dieselbe Funktion wie der Numerus, nämlich die Quantifikation von Referenten, erfüllen auch andere Elemente wie Numeralien, quantifizierende Adjektive und Adverbien. Die Kategorie Numerus umfaßt jedoch derartige lexikalische Quantoren nicht. Die Kriterien für die Abgrenzung von lexikalischen Quantoren sind distributioneller Art:

- (a) Quantoren stellen keine Numeruskategorien dar, wenn sie mit numeruskodierenden Elementen kombinieren, vgl. die Kombination von adjektivischem Quantor und Pluralaffix im Englischen *some boys* (s. Dryer 1989: 867);
- (b) Numerusmorpheme bilden eine anhand von distributionellen Kriterien definierbare Morphemklasse, die den gemeinsamen Nenner der Kodierung der Quantifikation aufweist.

Numerus wird v. a. mit der Quantifikation von Referenten verbunden. Obwohl Handlungen nur bei der Nominalisierung Referenten darstellen, ist es für viele Sprachen gerechtfertigt, den Begriff "Numerus" auf die Quantifikation der durch das Verb ausgedrückten Handlung(en) auszudehnen (s. 3.1). Handlungsverben sind die prototypische Prädikatsklasse, der Terminus "Handlung" muß hier allerdings als Oberbegriff für Prädikatsbegriffe wie Handlung, Prozeß, Zustand, Eigenschaft usw. verstanden werden. Die Numeruskategorie umfaßt in diesen Fällen Kategorien wie distributiv und iterativ, die sonst nicht als Numerusmorpheme gelten. Sie werden mit denselben Mitteln wie der nominale Numerus ausgedrückt, z. B. Reduplikation, oder sie interagieren funktional mit ihm (s. 3.1).

Greenberg (1963: 96) postuliert, daß die Kategorie Numerus universal vorhanden sei, mindestens bei den Pronomina:

"42. All languages have pronominal categories involving at least three persons and two numbers." (Greenberg 1963: 96)

Inzwischen ist jedoch für einige Sprachen überzeugend dargelegt worden, daß sie über keine Numeruskategorie verfügen (s. auch Corbett 2000: 50 f.). So kann im Frühaltchinesischen die Vielzahl nur durch die Verwendung unterschiedlicher lexikalischer Mittel ausgedrückt werden, die aus Numeralien durch Desemantisierung entstanden sind (s. Dobson 1962: 26 f.), z. B. *bae* 'eine große Anzahl/alle' (< 'hundert'), *san* 'eine kleine Anzahl/alle' (< 'drei'), *i ell* 'ein Paar' (< 'eins zwei'). Insbesondere gibt es einige Sprachen, die Pronomina ohne Numerusmarkierung aufweisen. In manchen dieser Sprachen sind die Pronomina numerosindifferent. Im Golin (Papua New Guinea: Chumbu) werden die Pronomina *na* '1' und *i* '2' sowohl mit singularer als auch mit pluraler Referenz verwendet.

Andere Sprachen verfügen nur über Pronomina mit singularer Referenz, wie z. B. *Pirahã* (Amazonas: Mura; s. Helmbrecht 2004: 4.3.1). Die Quantifikation der pronominalen Referenten wird in diesen Sprachen durch syntaktische Mittel ausgedrückt (s. Helmbrecht 2004: 4.3.1).

In vielen Sprachen ist Numerus eine Flexionskategorie (s. 4), dies ist jedoch keine notwendige Bedingung. Besonders in Sprachen mit isolierender oder analytischer Morphologie wird Numerus nämlich durch ungebundene Morpheme ausgedrückt; so verfügen z. B. zahlreiche südostasiatische und austronesische Sprachen über Pluralwörter (s. 4.2.3).

Im folgenden wird die Kategorie Numerus unter besonderer Berücksichtigung der morphologischen Aspekte dargestellt. Abschnitt 2 bietet einen Überblick über die semantischen Distinktionen der Kategorie, die in den Sprachen der Welt vorkommen. Abschnitt 3 behandelt die Frage, welche Elemente numerusfähig sind. Die Mittel und der Ort der morphologischen Kodierung werden in den Abschnitten 4 und 5 dargestellt. Abschnitt 6 behandelt schließlich Aspekte der Evolution der Numerusmorphologie.

## 2. Numeruskategorien

Das Verhältnis zwischen morphologischen und semantischen Kategorien kann ziemlich komplex sein; von einer eineindeutigen Entsprechung kann nicht die Rede sein, zumal die semantischen Oppositionen inklusiv sein und neutralisiert werden können, aber man kann sich in einer ersten Annäherung auf die prototypischen Bedeutungen konzentrieren. Die Numeruskategorien lassen sich in semantischer Hinsicht danach unterscheiden, ob sie einer absoluten oder einer relativen Quantifikation dienen. In nichtgenerischen Aussagen bezeichnet der Singular einen einzelnen Referenten, der Plural eine Vielzahl von Referenten, die in Opposition zum Singular als '> 1' zu verstehen ist; der Dual bezeichnet zwei Referenten und der Trial drei. Abgesehen von den Funktionen, die bei den einzelnen Substantivklassen vorkommen können, dienen die prototypischen Bedeutungen dieser Numeruskategorien der absoluten Quantifikation, indem sie den Umfang einer Menge von Referenten bestimmen. Relative Numeruskategorien beschreiben die Anzahl der Referenten im Hinblick auf einen expliziten oder impliziten Standard. Der Paukal z. B. bezeichnet eine je nach Kontext variable Anzahl von Referenten, die aber als gering aufgefaßt wird.

### 2.1. Transnumeral

In vielen Sprachen werden bestimmte Nomina transnumeral verwendet, d. h. numerusindifferent; diese bzgl. Numerus unspezifische Kodierung von Referenten wird als **Transnumeral** (s. Biermann 1982: 230) oder auch als "general number" (s. Corbett 2000: 9) bezeichnet. Die Neutralisierung aller Oppositionen bestimmt den Zustand der Nichtindividuiertheit des Nomens. Im Deutschen verfügen z. B. viele Massennomina wie *Wasser* und Abstrakta wie *Frieden* über keine Numerusopposition (s. Biermann 1982: 230; Iturrioz-Leza 1986: 411). Im Huichol (Mexiko: Uto-

aztekisch) sind alle Nomina für unbelebte Referenten im Prinzip transnumeral. Die Unterscheidung zwischen numerusfähigen (individuierten) und nicht-numerusfähigen Referenten ist sprachspezifisch; übereinzelsprachliche Generalisierungen lassen sich jedoch im Zusammenhang mit der Empathiehierarchie feststellen (s. 3.2).

In vielen Sprachen, die eine morphologische Opposition zwischen Transnumeral und anderen Numeri aufweisen, wird der Transnumeral durch den bloßen Nominalstamm kodiert, wie etwa in Bayso (Äthiopien: Kuschitisch):

- (1) (a) *lúban-titi foofe*  
Löwe-SG sah.1.SG  
'ich sah einen Löwen'
- (b) *lúban-jool foofe*  
Löwe-PL sah.1.SG  
'ich sah Löwen'
- (c) *lúban foofe*  
Löwe sah.1.SG  
'ich sah Löwe' (Corbett 2000: 11)

Im Baskischen ist der Transnumeral in den grammatischen Kasus durch die Abwesenheit eines Numerusmorphems gekennzeichnet, in den lokalen Kasus hingegen durch den expliziten Exponenten *-ta*; die eigentlichen Numeri bringen ein Zeichen der Individuation mit sich (*-a* in den grammatischen und *-Ø* in den lokalen Kasus), und der Plural erhält eine zusätzliche Markierung (*-k* im Absolutiv und Ergativ, *-e* in den lokalen Kasus). Die Exponenten der Transnumerallität und der Individuation zeigen komplementäre Verteilung; der einzige Kasus, in dem sowohl der Transnumeral als auch die Individuierung explizit markiert werden, ist der Lokativ, der sich so als der Übergang zwischen beiden Klassen von Kasus erweist (Iturrioz-Leza 1982; 1985). In den lokalen Kasus ist der Transnumeral merkmalhaft. Die Pluralformen der lokalen Kasus leiten sich vom Transnumeral ab.

	Transnumeral	Singular	Plural
Absolutiv	<i>etxe</i> 'Haus'	<i>etxe-a</i> '(ein) Haus'	<i>etxe-a-k</i> 'Häuser'
Lokativ	<i>etxe-ta-n</i> 'zu Hause'	<i>etxe-a-n</i> 'im Haus'	<i>etxe-e-ta-n</i> 'in den Häusern'
Allativ	<i>etxe-ta-ra</i> 'nach Hause'	<i>etxe-ra</i> 'zu Hause'	<i>etxe-e-ta-ra</i> 'zu den Häusern'

Tab. 100.1: Numerusopposition im Baskischen

Man könnte die folgende Hierarchie aufstellen:

	-individuiert transnumeral	+individuiert durch Numerus
gramm. Kasus:	$-\emptyset$	$\begin{array}{c} -a \\ \diagup \quad \diagdown \\ \text{SG: } -\emptyset \quad \text{PL: } -k \end{array}$
lokale Kasus:	$-ta$	$\begin{array}{c} -\emptyset \\ \diagup \quad \diagdown \\ \text{SG: } -\emptyset \quad \text{PL: } -e \end{array}$

Tab. 100.2: Transnumeralität und Individuation

In den lokalen Kasus wird Individuation nicht markiert, in den grammatischen Kasus dagegen ist Transnumeralität das unmarkierte Glied.

## 2.2. Singular

Der **Singular** ist der Numerus, der prototypischerweise für die Bezeichnung einer Einheit verwendet wird. In einigen Sprachen wird die Opposition Singular:Plural durch unterschiedliche Morpheme kodiert (s. die Numerrussuffixe des Bayso in (1a–b)). In vielen Sprachen jedoch ist die Singularform die unmarkierte Form, die durch die Abwesenheit einer Numerusmarkierung gekennzeichnet ist:

- (2) (a) *book*  
(b) *book-s*

Der strukturellen Markiertheit entspricht weiterhin die Bedeutung: Singularformen sind in vielen Kontexten transnumeral, z. B. in *wir waren nur drei Mann*.

Das umgekehrte Muster, nämlich markierter Singular vs. unmarkierter Plural, wird in folgendem Universale ausgeschlossen:

"35. There is no language in which the plural does not have some nonzero allomorphs, whereas there are languages in which the singular is expressed only by zero. The dual and the trial are almost never expressed by zero." (Greenberg 1963: 94)

Die von einer kollektiven bzw. transnumeralen Basis abgeleiteten Singularform heißt **Singulativum** oder auch *nomen unitatis* (s. auch Art. 101). Im klassischen Arabischen bezeichnet das Basiskollektivum eine Gruppe von Referenten oder eine Spezies wie in (3a); in (3b) wird durch Suffigierung mit *-at* ein Singulativum gebildet. Die Opposition Kollek-

tiv: Singulativ ist nicht mit der Opposition Plural: Singular gleichzusetzen, denn das Singulativum kann weiterhin pluralisiert werden, und zwar durch Vokaldehnung wie in (3c).

- (3) (a) *baqar-u-n*  
          ‘Vieh-NOM-INDEF’  
 (b) *baqar-at-u-n*  
          ‘Vieh-SGV-NOM-INDEF (Kuh)’  
 (c) *baqar-aat-u-n*  
          ‘Vieh-SGV.PL-NOM-INDEF (einige Kühe)’  
          (Ratcliffe 1998: 94)

### 2.3. Plural

Der **Plural** wird für die Bezeichnung der Mehrzahl verwendet. Sprachen unterscheiden sich dadurch, wie sie die Mehrzahl definieren. Die Beispiele in (4) zeigen, daß, sobald mehr als ein Referent vorhanden ist, z. B. ab 'ein und ein halbes', im Deutschen der Plural verwendet wird. Im Neugriechischen verweist dagegen der Plural auf eine Vielzahl ganzer Referenten, d. h. ab 'zwei', vgl. (4b) vs. (5b).

- (4) (a) *ein Tag*  
      (b) *anderthalb Tage*  
      (c) *zwei Tage*

(5) (a) *mía*                      *méra*  
       ein:NOM/AKK.SG.F Tag:NOM/AKK.SG.F  
       ‘ein Tag’

(b) *miá-misi*    *méra*  
       ein.F-halb Tag:NOM/AKK.SG.F  
       ‘anderthalb Tage’

(c) *ðío*    *méres*  
       zwei Tag:NOM/AKK.PL.F  
       ‘zwei Tage’

Neben dem allgemeinen Plural für die Bezeichnung der Vielheit kommen weitere Plu-

raltypen vor. Die Begriffe **kollektiver** und **distributiver Plural** sind in erster Linie für zwei Interpretationen einer pluralischen Nominalphrase eingeführt: Eine Prädikation trifft auf ein pluralisches Subjekt kollektiv zu, wenn sie auf die Gesamtheit der Referenten zutrifft (s. Art. 101); sie trifft darauf distributiv zu, wenn sie auf jeden der Referenten einzeln zutrifft. Z. B. wird in *Fünf Möbelpacker tragen drei Klaviere* und *Die Apostel sind zwölf* kollektiv auf das Subjekt referiert; aber in *Die Apostel sind fromm* und *Fünf Möbelpacker tragen drei Namen* wird distributiv referiert.

In vielen Sprachen wird der kollektive Plural mit denselben Mitteln wie der allgemeine Plural ausgedrückt. So drückt die Reduplikation im Warlpiri (Australien: Pama-Nyunga) den allgemeinen Plural bei empathischen (Menschen und Haustiere bezeichnenden) Substantiven (s. (6)) und den kollektiven Plural bei weniger empathischen Substantiven aus (7).

- (6) (a) *kurdu*  
‘Kind’
- (b) *kurdu-kurdu*  
‘Kind-RDP (Kinder)’
- (7) (a) *kiwinyi*  
‘Moskito’
- (b) *kiwinyi-kiwinyi*  
‘Moskito-RDP (Moskitoschwarm)’  
(Fabricius 1998: 67)

Manche Sprachen verfügen über eine Opposition zwischen allgemeinem und kollektivem Plural. Im Hatam (Westpapua: Indonesien) wird der kollektive Plural durch das Suffix *-bat* ‘KOLL’ ausgedrückt, z. B. *ni-kwohop-bat* ‘3.SG-Schwester-KOLL (die Gruppe seiner Schwestern)’ (s. Reesink 1999: 51; vgl. das Morphem für den allgemeinen Plural in (33)).

Im Huasteca Nahuatl wird der allgemeine Plural beim possessiven Suffix ausgedrückt (vgl. (8 a) und (8 b); s. auch Art. 133). Durch partielle Reduplikation kann weiterhin ein distributiver Plural gebildet werden (8 c).

- (8) (a) *kone-*  
‘Kind-ABSL (Kind)’
- (b) *kone-me*  
‘Kind-PL (Kinder)’
- (c) *inīn-kone-wah*  
‘POSS.3.PL-Kind-RELL.PL  
(ihre Kinder)’
- (d) *inīn-koh-kone-wah*  
‘POSS.3.PL-RDP-Kind-RELL.PL  
(ihre verschiedenen Kinder)’  
(Kimball 1990: 204)

Der **sortale Plural** (auch *pluralis generalis*) bezeichnet mehrere Arten von Referenten bzw. Referenten verschiedener Arten (vgl. den sortalen Plural im Deutschen: *Kraut* vs. *Kräuter*, *Tuch* vs. *Tücher*; s. Iturrioz-Leza 1987 b: 99 f.; vgl. auch die Bildung des sortalen Plurals durch rekursive Pluralisierung im Hui-chol, 4.4). Maltesisch weist genauso wie das Arabische (s. 2.2; 2.6) die Opposition zwischen Basiskollektivum und Singulativum auf. Eine häufige Bedeutung der Pluralformen, die auf der Basis des Kollektivums gebildet werden, ist der sortale Plural, z. B. /*a'għin* ‘Pasta’ vs. /*a'għeyen*/ ‘Pasta-PL (Sorten von Pasta)’ (Mifsud 1993).

Der **assoziative Plural** (auch *cohort plural*) kodiert wie der kollektive Plural eine Menge von Referenten, zeichnet sich aber dadurch aus, daß nur ein salienter Referent aus der Menge im Ausdruck vorkommt. Dies wird durch (9) aus dem Baskischen illustriert.

- (9) *Koldo eta etorri d-i-ra.*  
Koldo und komm:PART ABS.3-PRÄS-PL  
‘Koldo und sein Trupp sind gekommen.’

Im Türkischen wird dieselbe Funktion durch das Pluralmorphem erfüllt (Gerd Jendraschek, p. K.):

- (10) *Koldo-lar gel-di*  
Koldo-PL komm-PRÄT(3.SG)  
‘Koldo und sein Trupp kamen.’

#### 2.4. Arithmetische Numeri

Die arithmetischen Numeri bezeichnen eine genaue Anzahl von Referenten. Solche sind der Dual, der Trial und der Quadral. Der **Dual** ist die Numeruskategorie für die Bezeichnung zweier Referenten. In den älteren indogermanischen Sprachen ist der Dual ein wichtiger Bestandteil des Numerussystems. Er ist u. a. im Altgriechischen, im Altindischen und im Gotischen gut belegt, wo sowohl das Nomen als auch das Verb nach der dreifachen Opposition Singular : Plural : Dual flektiert werden. Einen nominalen Dual besitzen außerdem das Litauische, das Slowenische und das Sorbische. Der Dual ist auch außerhalb der indogermanischen Sprachen sehr verbreitet. Er kommt z. B. in finno-ugrischen und semitischen Sprachen sowie in vielen Sprachen Australiens und Austroneasiens vor (s. Rukeyser 1997).

Der **Paritätsdual** umfaßt zwei Referenten, die ein Paar bilden (auch *Paral(is)*). Zu unterscheiden ist hierbei zwischen solchen Paaren, die zwei gleichartige Referenten umfas-

sen (**Äquivalenzdual**; s. Rukeyser 1997: § 2.1), z. B. Altgriechisch *kheîre* ‘Hand:NOM/AKK/VOK.DU.F’ und solchen, die zwei verscheidenartige Referenten umfassen, die jedoch einander zugeordnet sind (**Oppositionsdual**; s. Rukeyser 1997: § 2.1), z. B. Sanskrit *dyāvā* ‘Himmel.DU (Himmel und Erde)’. Beim Oppositionsdual wird genauso wie beim assoziativen Plural (s. 2.3) nur ein salienter Referent aus der bezeichneten Referentenmenge kodiert. So bezeichnen im Sanskrit beide Ausdrücke in (11) dieselben Referenten.

- (11) (a) *pitár-ā*  
‘Vater-NOM.DU.M (Vater und Mutter)’  
(b) *mātár-ā*  
‘Mutter-NOM.DU.F (Mutter und Vater)’

Eine besondere Instanz des Oppositionsduals in australischen Sprachen ist der Dual für die Bezeichnung dyadischer Verwandtschaftsverhältnisse. Im Kayardild wird ein Morphem für den arbiträren Dual (s. u.) verwendet (12 a) und ein weiteres Morphem für die Bezeichnung von zwei Leuten, die in konverser Verwandtschaftsbeziehung stehen, wie in (12 b) (Evans 1995: 190 f.).

- (12) (a) *kularrin-jiyarrng-ka*  
‘Schwester-DU-NOM (zwei Schwestern)’  
(b) *ngamathu-ngarrba*  
‘Mutter-DYAD (Mutter und Kind)’  
(Evans 1995: 184; 190 f.)

Im Gegensatz zum Paritätsdual steht der **arbiträre Dual** (auch *Ambal(is)*), der für zwei beliebige Referenten, die kein Paar darstellen, verwendet wird (s. Plank 1989: 308 f. über das implikative Verhältnis zwischen beiden Dualarten). Ein extremes Beispiel ist Altgriechisch *aiante* ‘Ajax-NOM/AKK.DU.M (die beiden Männer namens Ajax)’. Viele australische Sprachen verfügen über eine Opposition zwischen Paritäts- und arbiträrem Dual (s. Rukeyser 1997: § 3.3). (13) illustriert diese Opposition im Kayardild (Australien: Pama-Nyungan).

- (13) (a) *tar̪ka-yarŋ*  
‘Mann-DU<sub>1</sub> (zwei Männer)’  
(b) *maya-yarpa*  
‘Freund-DU<sub>2</sub> (ein Freundespaar)’

Der **Trial** bezeichnet eine Menge von drei Referenten. Im Larike (Indonesien: Malayo-Polynesisch) wird bei den Pronomina ein Trial gebildet (s. 14).

- (14) *Kalu iridu-ta-?eu*,  
wenn 2.TRIAL-NEG-geh  
*au-na-wela*  
1.SG-IRR-nach.Hause.geh  
‘wenn ihr drei nicht geht, dann gehe ich nach Hause’ (Corbett 2000: 21)

Ein **Quadrals** für die Bezeichnung von vier Referenten wird für manche Sprachen angenommen; jedoch zeigt Corbett (2000: 26), daß die entsprechenden Numerusmorpheme keine klaren Fälle von Quadrals sind, sondern Instanzen des Paukals (s. 2.5), die etymologisch aus dem Numeral ‘vier’ stammen.

## 2.5. Paukal

Der **Paukal** bezeichnet eine kleine Menge von Referenten. Die genaue Anzahl ist von Sprache zu Sprache verschieden, und auch innerhalb derselben Sprache wird sie relativ zum Kontext bestimmt. (15) illustriert die Verwendung des Paukals im Bayso (vgl. (1)):

- (15) *ħibān-jaa foofे*  
Löwe-PAUKAL sah.1.SG  
‘Ich sah ein paar Löwen.’  
(Corbett 2000: 11)

## 2.6. Großer Plural

Einige Sprachen unterscheiden zwischen einfacher Mehrzahl und größerer Zahl. Bei einigen Kollektiva wird im Arabischen neben der Singulativform (s. 2.2) auch eine Pluralform gebildet, die nun in zwei Funktionen auftritt, nämlich als sortaler Plural (s. 2.3) oder als **Abundanzplural**, der eine große Anzahl von Referenten bezeichnet (s. Ratcliffe 1998: 69 f.; Corbett 2000: 32). Im Syrienarabischen wird aus der Basis *dabbān* ‘Fliegen (Kollektiv)’ ein Singulativum mit einer Singularform *dabbāne* ‘Fliege’ und einer Pluralform *dabbānat* ‘Fliegen’ gebildet. Aus der kollektiven Basis wird auch eine weitere Pluralform *dababīt* gebildet, die entweder ‘viele Fliegen’ (Abundanzplural) oder ‘verschiedene Fliegen’ (sortaler Plural) bezeichnet (Corbett 2000: 32).

## 2.7. Hierarchie der Numeruskategorien

Die Markiertheitshierarchie der Numeruskategorien wird in bezug auf das Vorkommen der einzelnen Kategorien in den Sprachen der Welt wie folgt beschrieben:

“34. No language has a trial number unless it has a dual. No language has a dual unless it has a plural.” (Greenberg 1963: 94)

Außerdem wird die Markiertheitshierarchie durch die Merkmallosigkeit des Singulars

bestimmt (s. 2.2). Weitere Evidenz für ein Markiertheitsverhältnis bieten die Neutralisierungen:

“37. A language never has more gender categories in nonsingular numbers than in the singular.” (Greenberg 1963: 95)

Die Genusopposition im Singular des Personalpronomens des Dänischen wird im Plural neutralisiert (s. Plank & Schellinger 1997: 54), einige Nominalklassen des Kinyarwanda werden im Plural neutralisiert (s. Art. 141), Neutralisierung der Kasusoppositionen (Genitiv vs. Akkusativ) des Singulars und Plurals kommt im Dual des klassischen Arabischen vor. Es gibt zwar Gegenbeispiele zu Universal 37 (s. das Nominalklassensystem des Huichol in Iturrioz-Leza et al. 1986 a; s. auch Plank & Schellinger 1997). Aber die universellen Tendenzen der Neutralisierung folgen doch der in (16) wiedergegebenen Hierarchie der Numeruskategorien:

- (16) Singular > Plural > Dual > Trial

Die Stellung des Paukals innerhalb der Markiertheithierarchie variiert in unterschiedlichen Analysen (s. Corbett 2000: 39). Sein Vorkommen in den Sprachen der Welt setzt jedenfalls die Opposition Singular : Plural voraus, er hat jedoch kein eindeutiges implikatives Verhältnis zu den arithmetischen Numeri. In den Sprachen der Welt kommen sowohl Systeme des Typs Singular : Plural : Dual (Sanskrit) als auch Systeme des Typs Singular : Plural : Paukal (Bayso) vor. In diesem Sinne sind die implikativen Verhältnisse nicht in einer unilinearischen Hierarchie anzutragen:

Dual > Trial

- (17) Singular > Plural >  
Paukal

### 3. Numerusfähige Elemente

#### 3.1. Referenten- vs. Handlungsnummerus

Der Numerus gilt als typische Kategorie des Nomens. Numerusmorpheme bei Adjektiven und Verben dienen in vielen Sprachen ebenfalls der Quantifikation nominaler Referenten, unabhängig davon, ob die Markierung des Numerus am Nomen vorhanden ist (s. 5). Numerusmarkierung kommt jedoch auch am Verb für die Quantifikation der durch das Verb kodierten Handlung(en) vor. Dieser **Handlungsnummerus** wird immer am Verb kodiert (s. Corbett 2000: 251) und umfaßt verbale Aktionsarten wie Iterativ und Habilita-

tiv, die der Quantifikation der Handlung dienen (s. Art. 109). Für den Handlungsnummerus wird auch der Terminus “verbaler Numerus” verwendet, der wiederum gelegentlich auch zur Bezeichnung des am Verb kodierten Referentennummerus verwendet wird.

Die Pluralisierung von Handlungen (*pluractionality*, s. Yu 2003: 289) interagiert in vielen Sprachen mit dem nominalen Numerus. Im Tschetschenischen wird die pluralische Verbform vorgezogen, wenn das Argument im Absolutiv pluralisch ist (s. Yu 2003: 295). (18 a) und (18 b) sind beide wohlgeformt; allerdings wird das pluralische Verb in (18 b) gegenüber dem singularischen in (18 a) vorgezogen, und insofern korreliert der verbale mit dem nominalen Numerus. Der verbale Plural kommt jedoch auch in Sätzen wie (18 c) vor, in denen das absolutive Argument singularisch ist. In diesen Fällen bezieht sich die Pluralisierung auf die Wiederholung derselben Handlung desselben Partizipanten.

- (18) (a) *beer-ash suuna mar’iaqq-ira*  
Kind-PL<sub>1</sub> 1.SG.DAT umarm-PFV  
‘die Kinder umarmten mich’  
(Yu 2003: 296)
- (b) *beer-ash suuna marlilkh-ira*  
Kind-PL<sub>1</sub> 1.SG.DAT umarm.PL<sub>2</sub>-PFV  
‘die Kinder umarmten mich’  
(Yu 2003: 296)
- (c) *beer suuna marlilkh-ira*  
Kind 1.SG.DAT umarm.PL<sub>2</sub>-PFV  
‘das Kind umarmte mich immer wieder’ (Yu 2003: 296)

Diese Beispiele zeigen, daß der Handlungsnummerus kein Kongruenzmarker ist, sondern unabhängig vom Numerus der Referenten verwendet wird (s. Frajzyngier 1985; Durie 1986; Iturrioz-Leza et al. 1986 b; Mithun 1988; Corbett 2000; Yu 2003).

Besonders in Sprachen, in denen der Ausdruck des Numerus beim Nomen nicht vorhanden oder fakultativ ist, dient der Handlungsplural als Basis für die Inferenz der Anzahl der Referenten. Im Zentral-Pomo (Kalifornien: Hoka), wo die Kodierung des Numerus bei Belebten fakultativ ist (s. Mithun 1988: 226), drückt der verbale Numerus den Plural der Handlung aus und läßt auf dieser Basis inferieren, daß es sich um eine Gruppe von Partizipanten handelt, wie in (19 b) illustriert wird.

- (19) (a) *yóohłow čaac’ waáda*  
Süden:von Person SG:geh.KONT  
‘Von Süden kommt eine Person.’  
(Mithun 1988: 226)

- (b) *yóohłow čaač' hlaáda*  
 Süden:von Person PL:geh.KONT  
 'Von Süden kommt eine Gruppe Personen.'  
 (Mithun 1988: 226)

Im Niue (Niue, Neuseeland: Austronesisch) werden unterschiedliche verbale Numeri durch mono- und bimoraische Reduplikation ausgedrückt. So können aus dem Basisverb in (20 a) die Form in (20 b), die die pluralische Handlung bezeichnet, und die Form in (20 c), die die singularische Handlung bezeichnet, abgeleitet werden (s. Haji-Abdolhosseini et al. 2002: 481).

- (20) (a) *fue*  
 Insekt.wegjag  
 (b) *fue-fue*  
 RDP<sub>1</sub>-Insekt.wegjag  
 'Insekten mit unterschiedlichen Bewegungen wegjagen'  
 (c) *fu-fue*  
 RDP<sub>2</sub>-Insekt.wegjag  
 'Insekten mit einer einzigen Bewegung wegjagen'

In einigen Sprachen kommen Handlungs- und Referentennummerus in Kombination vor. Huichol hat ein sehr komplexes Numerussystem, in dem ein nominales Numerussystem, bestehend aus acht Nominalklassen, und ein verbales Numerussystem auf vielfältige Weise interagieren. Das verbale System besteht aus 10 verschiedenen Morphemklassen, von flexivischen nichtkongruenten Pronominalaffixen für Subjekt und Objekt bis hin zu verschiedenen derivativen Morphemen, die primär Aktionsarten ausdrücken. Das verbale System ist dominant; während der nominale Plural weitgehend fakultativ ist, kann bzw. muß Numerus am Verb gleichzeitig, aber ohne Redundanz an verschiedenen Stellen ausgedrückt werden (s. Iturrioz-Leza 1986 a; 1986 b; 2004).

- (21) *Tstiikt-ri yeutá-ri*  
*Hund:PL-PL Wildtier-PL*  
*me-pi-te-wa-ta-ku-kuya.*  
*SBJ.3.PL-ASRT-GENUS-OBJ.PL-KOMPL-RDP-*  
*töt:PL*  
 'Hunde töten Wildtiere.'

Eine kollektive Bedeutung wird erst durch die Verbindung eines nominalen Plurals mit einem verbalen Singular kodiert:

- (22) *Tsiimani-xi pi-ta<sup>2</sup>a.*  
*Gestirn-PL ASRT-leucht*  
 'Die Sterne des Großen Bären leuchten zusammen.'

Sogar bei den Pronominalaffixen des Subjekts kommen alle vier logischen Möglichkeiten vor, so daß von Kongruenz im strikten Sinne des Wortes nicht die Rede sein kann:

Nomen	Verbalform
+	-
+	+
-	-
-	+

Tab. 100.3: Numeruskodierung im Huichol

- (23) (a) *Tsiimani-xi*  
*Gestirn-PL*  
*me-pi-ta-tika.*  
*SBJ.3.PL-ASRT-leucht:PL-SBJ.PL*  
 'Die Sterne des Großen Bären leuchten.'  
 (b) *Kaxu p-a-mie.*  
*Wagen ASRT-CISL-geh*  
 'Da kommt ein/der Wagen.'  
 (c) *?Iri ?ena*  
*Pfeil da*  
*me-m-eu-ti-hu-kai*  
*SBJ.3.PL-ASRT-DURCH-HINAUF-geh:PL-IPFV*  
 'Die Pfeile (Vorfahren, durch einen Pfeil repräsentiert) stiegen da hinauf.'

Nominaler und verbaler Numerus im Huichol unterscheiden sich hinsichtlich verschiedener Parameter. Das verbale Numerussystem ist komplexer, weil es mehr morphologische Mittel entfaltet und mehr semantische Dimensionen einbezieht (Individuitheit, Aktionsarten, Klassen von Prädikaten usw.); der verbale Numerus ist dominant, weil er den nominalen Numerus weitgehend verzichtbar macht und seine Interpretation bestimmt, z. B. als individuativ oder sortal; die verbalen Mittel sind in Texten auch sehr viel häufiger.

### 3.2. Empathiehierarchie

Die Empathiehierarchie bildet die Grundlage für verschiedene Typen von gespaltener Kodierung des Numerus (s. Smith-Stark 1974; Corbett 2000: 54–132; Gil 2003: 475 f.). (24) stellt eine vereinfachte Version der Empathiehierarchie dar (s. eine ausführliche Darstellung in Art. 73):

- (24) Sprechaktteilnehmer > menschlich > belebt > Individuum > Entität

Wenn eine Sprache eine Numeruskategorie für eine Klasse von Nomina aufweist, so wird dieselbe Kategorie auch bei den mehr empathischen Nomina auftreten, also denjenigen, die jeweils links in (24) stehen. Weiterhin werden durch dieselbe implikative Skala verschiedene Eigenschaften der Numeruskodierung konditioniert, wie die Obligatorität, die Kongruenz u. a. (s. Corbett 2000: 54–132).

Sprechakteteilnehmer (s. (24)) werden in der Grammatik durch die Personalpronomina vertreten. Viele Sprachen verfügen über Numeruskategorien im Pronominalsystem, die bei den Nomina nicht vorhanden sind (s. auch Universal 42 in 1 dazu). Im Finnischen gibt es z. B. einen Dual bei den Pronomina, jedoch nicht bei den Nomina (s. weiterhin über die interne Hierarchie der Personalpronomina in Smith-Stark 1974: 665; Corbett 2000: 61–66; Harley & Ritter 2002; Helmbrecht 2004: § 4.3.2; s. auch Art. 96).

Menschliche Wesen genießen einen Sonderstatus innerhalb der Gruppe der Belebten. Im Mayali (Australien: Gunwinjguan) wird der Plural am Verb nur bei menschlichen Partizipanten kodiert (vgl. (25 a) und (25 b)).

- (25) (a) *Abanmani-na-ng*  
SBJ.1.SG&OBJ.3.PL-seh-PRÄT.PFV  
*bininj*  
Mann  
'Ich sah Männer.'
- (b) *duruk ginga*  
Hund Krokodil  
*ba-bayeng*  
SBJ.3&OBJ.3-beiß.PRÄT.PFV  
*ba-ngune-ng*  
SBJ.3&OBJ.3-ess-PRÄT.PFV  
'Das Krokodil/die Krokodile hat/  
haben den Hund/die Hunde gebis-  
sen.'

Im Arabischen kommt die Pluralform von Verben, Adjektiven und Pronomina nur dann vor, wenn der Referent ein menschliches Wesen ist. Bei nicht-menschlichen Wesen im Plural wird Femininum Singular am kongruierenden Element kodiert (s. Fradkin 1991: 621). Auch weitere Spaltungen innerhalb der menschlichen Referenten kommen vor, die z. B. die Verwandschaftstermini (s. australische Sprachen in Rukeyser 1997: § 4.0) oder die rationalen Wesen (s. Tamil in Smith-Stark 1974: 662) als salientere menschliche Referenten aussondern.

Die Spaltung in der Kodierung des Numerus in Belebte und Unbelebte kommt sehr

häufig vor (s. Smith-Stark 1974: 663; Corbett 2000: 59 f.). Im Huichol wird das pluralische Pronominalaffix für das Subjekt am Verb nur bei Belebten verwendet (s. Iturrioz-Leza et al. 1986 b). Im Japanischen ist die Pluralmarkierung nur bei Belebten möglich (vgl. (26)–(27) und (28)). Bei Appellativa ist sie noch fakultativ, wie die möglichen Plurallesarten von (26 a)–(27 b) unter geeigneten Kontextbedingungen zeigen. Bei Personalpronomina und Personeneigennamen ist die Kodierung des Plurals dagegen obligatorisch (s. Corbett 2000: 74).

- (26) (a) *Syoonen-ga warat-ta.*  
Junge-NOM lach-PRÄ  
'Der Junge lachte.'  
(im passenden Kontext auch: "Die  
Jungen lachten.")
- (b) *Syoonen-tati-ga warat-ta.*  
Junge-PL-NOM lach-PRÄT  
'Die Jungen lachten.'
- (27) (a) *Tori-ga nai-ta.*  
Vogel-NOM sing-PRÄT  
'Der Vogel sang.'  
(im passenden Kontext auch: 'Die  
Vögel sangen.')
- (b) *Tori-tati-ga nai-ta.*  
Vogel-PL-NOM sing-PRÄT  
'Die Vögel sangen.'
- (28) *Teebaru-ga koware-ta.*  
Tisch-NOM brech-PRÄT  
'Der Tisch brach./Die Tische brachen.'  
(Shinichiro Ishihara, p. K.)

Der letzte Typ von Spaltung, der aus der Empathiehierarchie in (24) zu erwarten ist, faßt die Individualnomina gegenüber weiteren Entitäten wie Massennomina und Abstrakta zusammen. Im Huasteca Nahuatl kann ein Plural nur aus Individualnomina gebildet werden, wobei diese wiederum ein gespaltenes System mit einem Suffix für die Pluralisierung von Belebten und einem anderen Suffix für die Pluralisierung von Unbelebten aufweisen (s. Kimball 1990: 199). In bezug auf die Zählbarkeit unterscheidet sich die Verwendung des Plurals in vielen indogermanischen Sprachen, vgl. Englisch *water* vs. *waters* (s. Allan 1980).

### 3.3. Referentielle Eigenschaften

Die Kodierung des Numerus hängt in vielen Sprachen mit den referentiellen Eigenschaften des Nomens zusammen, insbesondere Spezifität und Definitheit (s. Biermann 1982:

232 f.; Plank 1987; Gil 2003: 473–475). In verschiedenen Sprachen wird die Numerusopposition bei Nomina mit generischer oder mit indefiniter Referenz neutralisiert (s. 2.1).

Im Persischen wird der Plural nur bei definiter Referenz markiert (s. (29 b–c) vs. (29 a); vgl. auch die Kodierung des Plurals beim Objekt eines transitiven Verbs im Türkischen im Art. 126).

- (29) (a) *sæg did-aem*  
Hund seh.PRÄT-1.SG  
'ich sah einen Hund/Hunde'
- (b) *sæg-o did-aem*  
Hund-AKK seh.PRÄT-1.SG  
'ich sah den Hund'
- (c) *sæg-a-ro did-aem*  
Hund-PL-AKK seh.PRÄT-1.SG  
'ich sah die Hunde'  
(Ghomeshi 2003: 48)

Im Baskischen hängt die am Verb kodierte Numeruskongruenz mit der Definitheit zusammen. Wenn das Objekt indefinit wie in (31) ist, so ist die Kongruenz im Ostbaskischen nicht möglich, im Zentralbaskischen obligatorisch und im Westbaskischen heute obligatorisch, aber in früheren Texten optional (Trask 1996: 241):

- (30) (a) *neska ikusi d-ut*  
Mädchen:DEF gesehen AUX-SBJ.1.SG  
'ich sah das Mädchen'
- (b) *neska-k ikusi*  
Mädchen:DEF-PL gesehen  
*d-it-ut*  
AUX-OBJ.3.PL-SBJ.1.SG  
'ich sah die Mädchen'
- (31) (a) *neska bat ikusi*  
Mädchen INDEF gesehen  
*d-ut*  
AUX-SBJ.1.SG  
'ich sah ein Mädchen'
- (b) *neska bat-zuk ikusi*  
Mädchen INDEF-PL gesehen  
*d(-it)-ut*  
AUX-OBJ.3.PL-SBJ.1.SG  
'ich sah Mädchen' (Trask 1996: 241)

## 4. Art der Kodierung

### 4.1. Allgemeines

In diesem Abschnitt werden die Mittel dargestellt, die für die Kodierung von Numeruskategorien verwendet werden. Numerusmorpheme kommen gelegentlich auch unabhän-

gig von der Numerusopposition vor. So sind **Pluralia Tantum** Pluralformen (z. B. Hebräisch *'atik-ót* 'Altertum-PL (Altertum)'), **Dualia Tantum** Dualformen (z. B. Sanskrit *aśvīnā* 'Asvina.DU' als Bezeichnung für die göttlichen Zwillinge) und **Singularia Tantum** Singularformen (z. B. Deutsch *Frieden*), die kein Gegenstück in den anderen Numeri haben. In einigen dieser Fälle ist die morphologische Besonderheit dieser Substantive semantisch motiviert, z. B. in der Dualform *aśvīnā* 'Asvina.DU' oder in der Abwesenheit einer Pluralform für Abstrakte wie *Frieden*. In vielen Fällen jedoch entspricht die Numerusmorphologie keinerlei semantischen Eigenschaften des Referenten wie z. B. im Hebräischen *pifiy-ót* 'Schwert-PL (Schwert)'.

### 4.2. Stammexterne Kodierung

#### 4.2.1. Reduplikation

Die **Reduplikation** dient in vielen Sprachen der Pluralisierung. Fabricius (1998: 60) betrachtet Pluralisierung als die häufigste Funktion der Reduplikation von Nomina in australischen Sprachen. Kiyomi (1995: 1154 f.) berechnet in einer Stichprobe von 23 malayopolynesischen Sprachen, daß 8 Sprachen u. a. den allgemeinen Plural und genauso viele Sprachen den distributiven Plural durch Reduplikation ausdrücken. (32 a) illustriert den Ausdruck des allgemeinen Plurals im Ilokano (Philippinen: Austronesisch), und (32 b) illustriert den Ausdruck des distributiven Plurals im Hiligaynon (Philippinen: Austronesisch).

- (32) (a) *áso*  
'Hund'  
*as-áso*  
'RDP-Hund (Hunde)'
- (b) *baláy*  
'Haus'  
*baláy-baláy*  
'Haus-RDP (jedes Haus)'  
(Kiyomi 1995: 1154 f.)

Im Niue werden unterschiedliche Subtypen von Reduplikation für den Ausdruck des verbalen Numerus verwendet: monomorische und bimorische Reduplikation, die bei einigen Nomina in Opposition vorkommen (s. (20) in 3.1). Die verschiedenen Typen der Reduplikation werden für die Bezeichnung unterschiedlicher verbaler Numeri verwendet (Haji-Abdolhosseini et al. 2002: 477).

#### 4.2.2. Affigierung

**Suffixe** für die Kodierung des Numerus kommen in sehr vielen Sprachen vor, z. B. Englisch *book-s* oder Huichol *tuutú* 'Blume' vs. *tu-*

*utú-ri* Blume-PL' (Iturrioz-Leza et al. 1986 a: 314). Numerus kodierende **Präfixe** kommen häufig in Sprachen mit präfigierten Nominalklassen vor, wie z. B. Swahili *ki-su* 'KL7-Messer (Messer (sg.))' vs. *vi-su* 'KL8-Messer (Messer (pl.))' (s. auch Kinyarwanda in Art. 141, Twi in Art. 140 und Hunzib in Art. 127).

Beim Nomen werden Numerusaffixe gewöhnlich näher am Stamm als Kasusaffixe kodiert:

"39. Where morphemes of both number and case are present and both follow or both precede the noun base, the expression of number almost always comes between the noun base and the expression of case." (Greenberg 1963:95)

Zwischen Stamm und Numerusaffix können natürlich Morpheme anderer Kategorien vorkommen. Ein extremes Beispiel stellt Hatam (Westpapua: Indonesien) dar, wo das optionale Morphem *-nya* 'PL' immer am Ende einer Nominalphrase suffigiert wird.

- (33) (a) *sop-nya*  
Frau-PL
- (b) *munggwom cin pi-ma-nya*  
Kind Paar ANA-dies-PL  
'dieses (vorerwähnte) Paar Kinder'
- (c) *krau misien ni-de-nya*  
fessel Hund 3.SG-POSS-PL  
'er fesselte seine Hunde'  
(s. Reesink 1999:50)

#### 4.2.3. Numeruswörter

Numerus kann auch durch **freie Morpheme** bzw. Wörter ausgedrückt werden (s. Dryer 1989; Macaulay 1989; Corbett 2000: 134). Im Yap (Yapinsel: Austronesisch) werden Singular, Dual und Plural durch Numeruswörter kodiert. Diese Morpheme werden anhand von distributionellen Kriterien von den Numeralien in dieser Sprache unterschieden (s. Dryer 1989: 868).

- (34) (a) *ea rea kaarroo neey*  
KONN SG Auto dies  
'dieses Auto'
- (b) *ea gal kaarroo neey*  
KONN DU Auto dies  
'diese zwei Autos'
- (c) *ea pi kaarroo neey*  
KONN PL Auto dies  
'diese Autos' (Dryer 1989: 868)

#### 4.3. Stamminterne Kodierung

##### 4.3.1. Stammalternation

Unterschiedliche Formen der **Stammalternation** zum Ausdruck des Numerus kommen in den Sprachen der Welt vor:

- Qualitativer Vokalwechsel, z. B. Ablaut des zweiten Vokals in Nomina mit der phonologischen Struktur (C)VCVC im Kaschmiri (Kaschmir: Indogermanisch): *gagur* 'Maus:SG.NOM' vs. *gagar* 'Maus: PL.NOM' (Wali & Koul 1997: 190). Im Tschechischen wird der verbale Numerus ebenfalls durch Ablaut gebildet, z. B. *tatta* 'schieben' vs. *titta* 'schieb.PL (wiederholt schieben)' (s. Yu 2003: 292).
- Qualitativer Konsonantenwechsel, z. B. Palatalisierung für den Ausdruck des Plurals im Kaschmiri: *kul* 'Baum:NOM.SG' vs. *kul'* 'Baum:NOM.PL' (Wali & Koul 1997: 190).
- Quantitativer Vokalwechsel, z. B. Pluralisierung durch Vokaldehnung im Arabischen: vom Singulativum *samak-at* 'Fisch-SGV' wird der Plural *samak-āt* 'Fisch-SGV.PL' gebildet.

Im Arabischen gibt es zwei Muster für die Pluralbildung: den stamminternen Plural (auch *pluralis fractus* oder "gebrochenen Plural"), der durch Stammveränderung gebildet wird, und den stammexternen Plural (auch *pluralis sanus* oder "gesunden Plural"), der auf Affigierung basiert (s. Ratcliffe 1998: 71). Die Bildung des stamminternen Plurals ist in (35) illustriert: aus einer Basis CVCC wird ein Plural mit der Struktur CVCVVC gebildet.

- (35) (a) *qalb-u-n*  
'Herz-NOM-INDEF'
- (b) *quluub-u-n*  
'Herz.PL-NOM-INDEF' (s. Ratcliffe 1998: 71)

##### 4.3.2. Suppletion

Ein typisches Beispiel für **Suppletion** zur Bildung des Numerus ist das Paar *čelovek* 'Mensch' vs. *ljudi* 'Menschen' aus dem Russischen. Im Huichol wird *nunutsi* 'Kind' durch Suppletion und gleichzeitige Affigierung pluralisiert: *tií-rí* 'Kind.PL-PL'.

##### 4.3.3. Prosodische Kodierung

Ausschließlich durch prosodische Kodierung, und zwar durch den **Ton**, wird die Numerusopposition im Shilluk (Sudan: Nilo-Sahara-nisch) ausgedrückt.

- (36) (a) *kíu*  
'essbar.Pflanze:SG'
- (b) *kíu*  
'essbar.Pflanze:PL'  
(Corbett 2000: 146)

**Akzentverschiebung** neben Affigierung wird im Russischen für die Bildung der Numerusoppositionen verwendet.

- (37) (a) *peró*  
          ‘Feder:NOM.SG.N’  
       (b) *pér'ja*  
          ‘Feder:NOM.PL.N’
- (38) (a) *zvjenó*  
          ‘Kettenglied:NOM.SG.N’  
       (b) *zvjén'ja*  
          ‘Kettenglied:NOM.PL.N’

#### 4.4. Mehrfache Kodierung

In einigen Sprachen wird der Numerus durch die gleichzeitige Verwendung verschiedener morphologischer Prozesse kodiert, z. B. Deutsch *Fuß* vs. *Füße* (Pluralisierung durch das Suffix *-e* und den Umlaut). In diachroner Hinsicht kann die mehrfache Kodierung durch Verstärkung entstehen, z. B. Niederländisch *kinderen* ‘Kind:PL’ auf der Basis des früheren *kinder* ‘Kind:PL’. Dieser Typ von mehrfacher Kodierung ist nicht kompositionell (s. Corbett 2000:154).

Die mehrfache Kodierung ist jedoch dann kompositionell, wenn ein Numerusmorphem ein anderes Numerusmorphem modifiziert (s. Corbett 2000: 36–38). In einigen Sprachen werden Dualformen oder Kollektiva pluralisiert, wie in (39) anhand des Bretonischen illustriert wird (s. auch Reesink 1986: 50 über die Kombination des Kollektivmorphems mit dem Pluralmorphem im Hatam).

- (39) (a) *skouarn*  
          ‘Ohr’  
       (b) *div-skouarn*  
          ‘DU-Ohr (zwei Ohren/Ohrenpaar)’  
       (c) *div-skouarn-ou*  
          ‘DU-Ohr-PL (Ohrenpaare)’  
          (Press 1986: 71)

Im Huichol wird rekursive Pluralisierung mit verschiedenen Bedeutungen angewandt. In *tuutú-ri-te* ‘Blume-PL<sub>1</sub>-PL<sub>2</sub> (Blumen verschiedener Sorten)’ fügt das zweite Pluralaffix eine sortale Bedeutung hinzu. In *wakat-tsixi-te* ‘Kuhfigur-PL<sub>3</sub>-PL<sub>2</sub>’ gibt *-te* zu verstehen, daß es sich um materielle Repräsentationen von Tieren handelt. In *ne-tuxú-ri-ma* ‘POSS.1-Schwein-PL<sub>1</sub>-PL<sub>4</sub> (meine Schweinchen)’ drückt das Suffix *-ma* (Pluralzeichen für Verwandtschaftsbezeichnungen und ähnliche Begriffe, die Nähe zum Ego beinhalten) eine affektive Beziehung aus. Ein weiterer Typ liegt im zusammengesetzten Wort *paapá tii-ri-xi* ‘Maisfladen Kind.PL-PL<sub>1</sub>-PL<sub>5</sub> (kleine rituelle Mais-

fladen)’ vor; die neue lexikalische Einheit *paapá tii-ri* (Plural von *paapá nunutsi*) wird erneut pluralisiert: [*paapá tii-ri*]-xi; *-xi* bringt zum Ausdruck, daß es sich um einen sakralen, symbolischen Gegenstand handelt, ähnlich wie in *?iiri-te-xi* ‘Pfeil-PL<sub>2</sub>-PL<sub>5</sub> (sakrale Pfeile)’.

#### 5. Locus der Kodierung

Die Sprachen der Welt unterscheiden sich darin, wie sie den Numerus eines Referenten in syntaktischen Konstruktionen kodieren. Tab. 100.4 zeigt alle Möglichkeiten der Numeruskodierung bei der Kombination von zwei verschiedenen Satzgliedern; die Konstruktionen der ersten und der zweiten Zeile sind disjunktiv zu verstehen. Im Kodierungstyp 3 kann Kongruenz vorliegen, aber Kongruenz soll nicht mit Kokurrenz gleichgesetzt werden, sie besteht vielmehr in einer asymmetrischen Dependenzrelation.

	Nominalphrase Nomen	Verb Attribut
Typ 1	a	Ø
Typ 2	Ø	a
Typ 3	a	a
Typ 4	Ø	Ø
Typ 5	a	b

Tab. 100.4: Kodierung des Numerus

Alle Möglichkeiten in Tab. 100.4 kommen in den Sprachen der Welt vor. Der erste Kodierungstyp läßt sich anhand des Japanischen illustrieren. Der Numerus kann nur beim Substantiv ausgedrückt werden, wie in (26)–(27) illustriert (dasselbe Kodierungsmuster kommt in der Nominalphrase des Englischen vor: *the red house* vs. *the red houses*).

Beim zweiten Kodierungstyp von Tab. 100.4 wird der Numerus nicht beim Substantiv kodiert. Dieses Muster kommt im gesprochenen Französischen vor. Hier wird die Numerusopposition nicht durch das Substantiv, sondern durch den Artikel ausgedrückt: *l'enfant* /l̪ ðfã/ vs. *les enfants* /lez ðfã/. Im Hualapai (Kalifornien: Hoka) wird der Plural obligatorisch am Verb und nur optional am Nomen kodiert (s. (40); Langdon 1992: 413–417).

- (40) *wal-ch nya mako-l*  
       Feder-SBJ mein Rücken-in  
       *disha-j(i)-v(i)-k-yu*  
       3.häng-PL-ZUSTAND-SS-AUX  
       ‘die Feder hängen an meinem Rücken’  
       (Langdon 1992: 414)

Ein sehr verbreitetes Muster ist die Nicht-Kodierung des Plurals beim Nomen, wenn dieses durch Numeralia oder andere Quantoren modifiziert wird. Dies kommt im Ungarischen (s. (41)), Koreanischen, Türkischen, Baskischen u. a. vor.

- (41) *hat artista*  
sechs Akrobat(sG)  
'sechs Akrobaten' (Hurford 2003: 567)

Der Typ 3 in Tab. 100.4 wird durch die Beispiele (42) und (43) aus dem Deutschen illustriert, die eine mehrfache Kodierung desselben Numerus aufweisen.

- (42) (a) *Der Junge schläft.*  
(b) *Die Jungen schlafen.*  
(43) (a) *der freundliche Junge*  
(b) *die freundlichen Jungen*

Der Typ 4 von Tab. 100.4 bezeichnet völliges Fehlen der Numeruskodierung. Das betrifft sowohl Sprachen, die gar keinen Numerus haben (s. 1), als auch solche, die ihn fakultativ kodieren (s. (27a) für Japanisch).

Der letzte Typ bezieht sich auf die Kodierung unterschiedlicher Numeri in Nominalphrase und Verb oder beim Substantiv und Attribut. Dieses Kodierungsmuster kommt in unterschiedlichen Fällen vor:

- (a) Die Numeruskodierung am Nomen entspricht nicht der Anzahl der bezeichneten Referenten. (44a–b) illustrieren den Fall von Nomina im Singular mit kollektiver Referenz (s. Iturrioz-Leza 1987a), während das Prädikat im Plural steht.
- (44) (a) Span. *La mayoría opinan así.*  
(b) Lat. *exercitus venerunt.*

- (b) Unterschiedliche Numeruskodierung kann vorkommen, wenn ein Satzglied die markierte Numeruskategorie und ein anderes die unmarkierte Kategorie kodiert. Im Homerischen Griechisch kommen die dualen Subjekte häufig mit Verb im Plural vor wie in (45).

- (45) *tòn ouk edúnanto*  
3.SG:AKK.M NEG kann:PRÄT.3.PL  
*dúō*  
zwei.DU.NOM/AKK  
*Aíante* ...  
Aias.DU.NOM/AKK.M  
*deidíksastʰai*  
fürcht:PASS.AOR.INF  
'Ihn konnten die zwei Personen namens Aias nicht ängstigen.'  
(Homer, *Ilias* 18.163 f.)

(c) Unterschiedliche Numeri werden kodiert, wenn sich die Numeruskategorie am Nomen und diejenige am Verb nicht auf denselben Referenten beziehen. Dies kommt z. B. dann vor, wenn das Subjekt von einem durch den Komitativ lierten Partizipanten begleitet wird wie in (46) aus dem Hualapai (s. auch den assoziativen Plural in (9) aus dem Baskischen). Die Pluralmarkierung am Verb bezieht sich auf beide Referenten, die durch den Subjekt und den Komitativ referiert werden.

- (46) *ba:-h-ch hme:-h(a)-m*  
Mann-DEM-NOM Junge-DEM-mit  
*swa:d(j(i)-k-i*  
3.sing-PL-SS-AUX  
'Der Mann und der Junge singen zusammen.' (Langdon 1992: 414)

(d) Aus der Kombination der unterschiedlichen Numeruskategorien ergibt sich ein nicht-kompositionaler Numerus (*constructed number* in Corbett 2000: 169). Im Hopi (U.S.A.: Uto-aztekisch) entsteht ein Dual aus der Kombination einer pronominalen Form im Plural und einer verbalen Form im Singular.

- (47) (a) *pam wari*  
D2.SG run.PFV.SG  
'er/sie lief'  
(b) *puma yìutu*  
D2.PL run.PFV.PL  
'sie liefen'  
(c) *puma wari*  
D2.PL run.PFV.SG  
'beide liefen'  
(Corbett 2000: 169)

## 6. Evolution der Numerusmorphologie

Pluralmorpheme stammen häufig von Morphemen mit kollektiver Bedeutung. Das indogermanische Suffix *-a* der kollektiven Feminina wurde als neutrales Pluralmorphem reanalysiert. Das Pluralsuffix *-k* des Ungarischen geht auch aus einem Kollektivsuffix hervor. Ähnliches geschah im Türkischen, Armenischen, Bulgarischen u. a. (s. Biermann 1982: 237).

Eine mögliche Quelle für Numerusaffixe sind Personalpronomina. Im ersten Stadium werden die Nomina zum Zwecke der emphatischen Kodierung des Numerus von Pronomina begleitet. In der nächsten Phase können aus Pronomina Numerusaffixe entstehen (s.

Lehmann 1995: 57). Daher sind Numerusaffixe häufig den Personalaffixen des Verbs ähnlich, z.B. Yukatekisch *ch'ich'-o'b* 'Vogel-PL' und *bin'-o'b* 'geh-3.PL' (s. Lehmann 1995: 57).

Die Morpheme der arithmetischen Numeri stammen häufig aus Numeralia (s. Corbett 2000: 266). Im Bretonischen ist z.B. das Dualpräfix *daou-* identisch mit der Kardinalzahl für 'zwei' (s. Press 1986: 86). Durch Desemantisierung entwickeln sich die arithmetischen Numeri zu nicht-arithmetischen Numeri. Der Wandel vom Dual zum Plural ist in einigen arabischen Dialekten dokumentiert (s. Blanc 1970). Der Wandel vom Quadral zum Paukal wird in Corbett (2000: 26–30) behandelt.

Morpheme, die ursprünglich der Kodierung des Handlungsnummerus dienten, werden in vielen amerikanischen Sprachen auch für die Kodierung des Referentennummerus verwendet. Nach Mithun (1988: 224–228) behalten diese Morpheme in den amerikanischen Sprachen die distributive Funktion auch in ihrer Verwendung für die Quantifikation von Referenten und entwickeln sich nur in wenigen Sprachen zum allgemeinen Plural.

## 7. Unübliche Abkürzungen

CISL	cislativ
KONT	kontinuativ
EZF	Ezafe
SGV	Singulativ

## 8. Zitierte Literatur

- Allan, Keith (1980), "Nouns and Countability". *Language* 56.3, 541–567
- Biermann, Anna (1982), "Die grammatische Kategorie Numerus". In: Seiler, Hansjakob & Lehmann, Christian (Hrsg.), *Apprehension: Das sprachliche Erfassen von Gegenständen I: Bereich und Ordnung der Phänomene*. Tübingen: Narr, 229–243
- Blanc, Haim (1970), "Dual and Pseudo-Dual in the Arabic Dialects". *Language* 46, 42–57
- Corbett, Greville G. (2000), *Number*. Cambridge: Cambridge University Press
- Dobson, W. A. C. H. (1962), *Early Archaic Chinese: A Descriptive Grammar*. Toronto: University of Toronto Press
- Dryer, Matthew 1989, "Plural Words". *Linguistics* 27, 865–895
- Durie, Mark (1986), "The Grammaticalization of Number as a Verbal Category". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 12, 355–370
- Evans, Nicholas D. (1995), *A Grammar of Kayardild*. Berlin; New York: Mouton De Gruyter
- Fabricius, Anne H. (1998), *A Comparative Study of Reduplication in Australian Languages*. München: Lincom Europa
- Fradkin, Robert A. (1991), "Marking, Markedness, and Person-Gender-Number Patterning in the Arabic Tenses and Moods". *Folia Linguistica* 25.3/4, 609–651
- Frajzyngier, Zygmunt (1985), "Ergativity, Number, and Agreement". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 11, 96–106
- Ghomeshi, Jila (2003), "Plural Marking, Indefiniteness, and the Noun Phrase". *Studia Linguistica* 57.2, 47–74
- Gil, David (2003), "English Goes Asian: Number and (In)definiteness in the Singlish Noun Phrase". In: Plank, Frans (Hrsg.), 467–514
- Greenberg, Joseph H. (1963), "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". In: Greenberg, Joseph H. (Hrsg.), *Universals of Language*. Cambridge/Mass.: The MIT Press
- Haji-Abdolhosseini, Mohammad & Massam, Diane & Oda, Kenji (2002), "Number and Events: Verbal Reduplication in Niuean". *Oceanic Linguistics* 41.2, 475–492
- Harley, Heidi & Ritter, Elizabeth (2002), "Person and Number in Pronouns: A Feature-Geometric Analysis". *Language* 78.3, 482–526
- Helmbrecht, Johannes (2004), *Personal Pronouns: Form, Function, and Use*. Habilitationsschrift, Universität Erfurt
- Hurford, James R. (2003), "The Interaction between Numerals and Nouns". In: Plank, Frans (Hrsg.), 561–620
- Iturrioz-Leza, José Luis (1982), "Apprehension im Baskischen". Seiler, Hansjakob & Stachowiak, Franz J. (Hrsg.), *APPREHENSION. Part II. Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen, Gunter Narr, 1–43
- Iturrioz-Leza, José Luis (1985), "La función de los afijos vascos *-al-ta* a la luz de la dimensión de INDIVIDUACION". *Euskera* 30, 175–213
- Iturrioz-Leza, José Luis (1986), "Grammatical Number, Nominal Subcategorization, and Techniques for the Apprehension of Objects I: Objections to the Traditional Approach". *Función* 1.3, 391–421
- Iturrioz-Leza, José Luis (1987 a), "Grammatical Number, Nominal Subcategorization, and Tech-

- niques for the Apprehension of Objects II: The Operational Approach". *Función* 2.1, 12–53
- Iturrioz-Leza, José Luis (1987b), "Breve historia de la investigación en torno al número gramatical". *Función* 2.1, 54–153
- Iturrioz-Leza, José Luis (2000/2001), "Diversas aproximaciones a la nominalización: De las abstracciones a las macrooperaciones textuales". *Función* 21.4, 21–127
- Iturrioz-Leza, José Luis (2004), "El número grammatical en huichol". *Revista Mexicana de Língüística* 5 (im Druck)
- Iturrioz-Leza, José Luis & Gómez López, Paula & Ramírez de la Cruz, Rritákame & Leal Carretero, Silvia (1986a), "Individuación en Huichol I: Morfología y semántica de las clases nominales". *Función* 1.2, 309–354
- Iturrioz-Leza, José Luis & Gómez López, Paula & Leal Carretero, Silvia & Ramírez de la Cruz, Rritákame (1986b), "Individuación en Huichol II: Aspectos morfológicos y sintácticos de las clases nominales". *Función* 1.3, 422–462
- Kimball, Geoffrey (1990), "Noun Pluralization in Eastern Huasteca Nahuatl". *International Journal of American Linguistics* 56.2, 196–216
- Kiyomi, Setsuko (1995), "A New Approach to Reduplication: A Semantic Study of Noun and Verb Reduplication in the Malayo-Polynesian Languages". *Linguistics* 33, 1145–1167
- Langdon, Margaret (1992), "Yuman Plurals: From Derivation to Inflection to Noun Agreement". *International Journal of American Linguistics* 58.4, 405–424
- Lehmann, Christian (1995), *Thoughts on Grammaticalization*. München: Lincom Europa
- Macaulay, Monica (1989), "The Plural Word in Chalcatongo Mixtec". *Chicago Linguistics Society: Papers from the Annual Regional Meeting* 25, 288–299
- Mifsud, Manuel (1993), "The Collective in Maltese". Manuskript
- Mithun, Marriane (1988), "Lexical Categories and the Evolution of Number Marking". In: Hammond, Michael & Noonan, Michael (Hrsg.), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego u. a.: Academic Press, 211–234
- Plank, Frans (1987), "Number Neutralization in Old English: Failure or Functionalism?". In: Koopman, Willem F. & van der Leek, Frederike & Fischer, Olga & Eaton, Roger (Hrsg.), *Explanation and Linguistic Change*. Amsterdam: Benjamins, 177–238
- Plank, Frans (1989), "On Humboldt on the Dual". In: Corrigan, Roberta & Eckman, Fred & Noonan, Michael (Hrsg.), *Linguistic Categorization*. Amsterdam; Philadelphia: Benjamins, 293–333
- Plank, Frans (2003, Hrsg.), *Noun Phrase Structure in the Languages of Europe*. Berlin; New York: Mouton De Gruyter
- Plank, Frans & Schellinger, Wolfgang (1997), "The Uneven Distribution of Genders over Numbers: Greenberg Nos. 37 and 45". *Linguistic Typology* 1.1, 53–101
- Press, Ian (1986), *A Grammar of Modern Breton*. Berlin; New York: De Gruyter
- Ratcliffe, Robert R. (1998), *The 'Broken' Plural Problem in Arabic and Comparative Semitic*. Amsterdam/Philadelphia: John Benjamins
- Reesink, Ger P. (1999), *A Grammar of Hatam*. Canberra: Pacific Linguistics
- Rukeyser, Alison (1997), *A Typology of the Nominal Dual: Evidence from Indo-European, Finno-Ugric, Semitic and Australian Languages*. Davis: The UC Davis Department of Linguistics (Davis Working Papers in Linguistics, 6)
- Smith-Stark, T. Cedric (1974), "The Plurality Split". *Chicago Linguistic Society: Papers from the Annual Meeting* 10, 657–671
- Trask, R. Larry (1996), *Historical Linguistics*. London: Arnold; New York: Oxford University Press
- Wali, Kashi & Koul, Omkar N. (1997), *Kashmiri: A Cognitive-Descriptive Grammar*. London: Routledge
- Yu, Alan C. L. (2003), "Pluractionality in Chechen". *Natural Language Semantics*, 289–321

*José Luis Iturrioz-Leza, Guadalajara  
(Mexico)  
Stavros Skopeteas, Erfurt/Potsdam  
(Germany)*

## 101. Mass and collection

1. Introduction
2. Semantics of mass nouns
3. Morphological categories on mass nouns
4. Is the count-mass distinction a linguistic universal?
5. Collections and collectives
6. Operations relating collectives and collections, mass and individuals
7. References

### 1. Introduction

Grammatical descriptions of European languages normally classify nouns into **count nouns**, such as *cow*, vs. **mass nouns**, such as *petrol* (Germ. *Individualnomina – Massennomina*, Russ. *isčisljaemye, diskretnye – neiscisljaemye, nediskretnye imena*).

In semantics, the difference is between denoting (or referring to) discrete entities with a well-defined shape and precise limits vs. homogeneous undifferentiated stuff without any certain shape or precise limits.

In syntax, the two differ in the contexts where they can be used and in their combinability with other expressions. Thus, the noun *cow* easily combines with numerals, e.g. *one cow, three cows*, etc., with the indefinite article *a* and with the denumerating measure words *many, several*, etc. Measuring stuff presupposes a procedure of dividing it up in portions, which are further counted – e.g. *three gallons of petrol* etc. (cf. 6). The mass term itself lacks a built-in mode of dividing reference, to use Quine's (1960: 91) formulation, with the effect that it cannot combine with numerals and with the indefinite article *a*, and takes the measure word *much* instead of *many*.

Finally, morphologically mass nouns lack the number opposition otherwise typical of English nouns (cf. *cow* vs. *cows* and *petrol* vs. *\*petrols*) and have thus a **defective paradigm**.

Some of these features are clearly language-specific. Thus, the “many-vs.-much” distinction is absent from many languages, e.g. Russian, French, and Basque. An indefinite article exists only in the tiny minority of the world's languages, and even there, it is not necessarily restricted to count nouns, as, e.g., in Bavarian (Kolmer 1999: 37). Likewise in many languages, number marking is only relevant for a subset of nominals referring to animals and humans or even just humans (Corbett 2000).

In the next example, one and the same word *cake* is used both countably (1 a) and uncountably (1 b).

- (1) (a) *Hetty likes to gorge herself on cake.*  
      (b) *Whenever Hetty gobbles down a cake,  
          her diet ‘starts tomorrow’*  
      (Allan 1980: 546).

Given that many English nouns (and their correspondences in the familiar European languages) can be used in both countable and uncountable contexts, a reasonable conclusion is that count-mass distinction is primarily manifested in noun occurrences (Pelletier 1975; Bunt 1985; Allan 1980; Gil 1996: 58). Philosophers have therefore suggested various hypothetical machines, or **rules of construal**, for dealing with the occurrences of one and the same noun in different contexts. Thus, the **Universal Grinder** (Pelletier 1975) can chop any object into a homogeneous mass. Similarly, any “mass noun” can be used as a count noun due to the **Universal Sorter** (Bunt 1985: 11), which issues qualifications like ‘this is an excellent wine’. The **Universal Packager** (cf. Jackendoff 1991: 24 on the unclear attribution of the term) is responsible for cases like *I'll have a coffee/three coffees*.

The universality of the hypothetical machines, however, turns out to be much restricted. Thus, different nouns show preferences for occurring in different contexts, or to have a typical countable or mass reading in contexts which allow for both, cf. *there is no wood in the yard* (MASS) vs. *there is no child in the yard* (COUNT).

One solution is to retain the count-mass opposition in the lexicon, but to complement it with the various recategorization rules (Jackendoff 1991). In a more radical way Allan (1980) takes into account the preferences that English nouns show for occurring in four different environments for countability and arrives at 8 different classes of nouns. See Pelletier & Schubert (1989) and in particular Behrens (1995) for the various interpretations of the mass-count distinction in the philosophical and linguistic literature and the various criteria used for distinguishing the different classes of nouns.

The mass/count distinction is normally assumed to be a semantically basic binary opposition. Behrens (1995) suggests instead that

the generalized mass/count distinction in English is the result of the collapse of two more or less independent semantic distinctions – the distinction between SHAPE and SUBSTANCE and the one between OBJECTS and QUALITIES. Other languages do not need to collapse the two dimensions.

## 2. Semantics of mass nouns

The traditional understanding of the semantic difference underlying the count-mass distinction is very much in line with Jespersen: “There are many words which do not call up the idea of some definite thing with a certain shape or precise limits. I call these ‘mass words’” (Jespersen 1924: 198).

The key notion for characterizing the main semantic difference between count and mass nouns has been termed “**bounding**” (e.g. Jackendoff 1991; Langacker 1991: 63–70) or “shape” (Rijkhoff 2002: 50–59), i.e. the existence vs. non-existence of precise limits of an entity referred to by a count or a mass noun respectively. Bounding can apply to various domains – a cow, as a physical entity, has spatial limits, a beep is bound both in time and pitch and a chapter is bound within a written work (Langacker 1991: 63–70). The unbounded character of mass nouns shows itself in several properties, primarily in “the **cumulative reference condition**”, e.g. any sum of parts which are water is water (Quine 1960: 91), and in “the **distributive reference condition**”, e.g. any part of water is water (Bunt 1985: 20; expansibility/contractibility vs. replicability in Langacker 1991: 70–72). Another property of mass nouns is their “**homogeneous reference**”: “mass nouns provide a way of speaking about things as if they do not consist of discrete parts” (Bunt 1985: 45; cf. also Langacker 1991; Rijkhoff 2002: 50–59).

Languages such as English, French or Russian show strong correlations between the properties of referents and their proneness to be referred to by a count or a mass noun, more or less in accordance with the **animacy hierarchy** (cf. Art. 73, 94). Thus, humans and animals favour a high degree of individuation and are normally coded by count nouns, while stuffs are more often than not expressed by mass nouns (cf. Sasse 1993). However, discreteness (boundedness) of count-noun referents and homogeneity of mass-noun referents reflect not so much the

objective reality itself, but rather its construals by language speakers. Thus, under appropriate circumstances several cows can alternatively be referred to as (*several*) *cows*, as *a herd (of cows)* by means of a **collective count noun**, or as *cattle*, by means of a **collective mass noun**.

There are amazing cross-linguistic differences in the lexical preferences for count or mass uses even within the European Indo-European languages, e.g.

	count (+) or mass (−)	Russian	German	Swedish	Italian
strawberry	—	+ <i>klubnika</i>	+ <i>Erdbeere</i>	+	+
fruit	+	—	+/-	+	<i>frutto</i>
hair	+	+/-	—	+	<i>ca-</i> <i>pello</i>
furniture	—	+	+	+	<i>mobile</i>
		<i>mebel'</i>	<i>Möbel</i>	<i>möbel</i>	

Tab. 101.1: Examples of mass/count nouns in Russian, German, Swedish, and Italian

It is also amazing how little work has so far been done for identifying such lexical preferences and for pinpointing and explaining the similarities and differences even among the otherwise best described languages. Among the exceptions are works focusing on the treatment of berries, fruits, and vegetables (often as uncountables) in Slavonic (cf. Corbett 2000: 80 for some references), Behrens (1995) on German, English, and Hungarian, and Wierzbicka (1988) on English, Russian, and Polish as count or mass nouns (cf. 3.1).

## 3. Morphological categories on mass nouns

### 3.1. Number

The category of number (cf. Art. 100) per se presupposes an opposition between two or more values, which in turn corresponds in more or less transparent ways to the distinction between one and several entities of a certain sort. For mass nouns number is not obvious – there is a long discussion among philosophers on whether a mass noun such as *petrol* refers to a single, though scattered entity or to a multiplicity of objects, a class of all portions of petrol. In English, and many other languages, both interpretations find lin-

guistic correlates. Thus, for the purposes of agreement *petrol* is undoubtedly singular. On the other hand, bare (articleless) mass nouns and bare plurals, to the exclusion of singulars, occur in sentences referring to distributive location, cf. *there was petroll\*a cow all over the meadow* and *there were cows all over the meadow*.

Plurals share with masses the properties of cumulative and distributive reference, at least to a certain degree: thus, a sum of cows and cows will still be cows, and a part of cows are still cows (with the limiting case of one cow). The referent of plurals is thus unbounded, but also entailing “a medium comprising a multiplicity of distinguishable individuals” (Jackendoff 1991; cf. also Langacker 1991: 121). Mass nouns and plurals can show various similarities in behaviour, e.g. the use of the **partitive articles** *dul/de laldes* in French or the occurrence as **genitively marked objects** to non-negated verbs in Russian.

Although mass nouns show some semantic similarities both with count nouns in the singular and in the plural, at the same time they differ from either, are, strictly speaking, outside of the usual number opposition and should deserve to be marked as such. A distinct **mass number** different from both singular and plural is reported for various Spanish dialects. Thus, in Lena dialect (Hualde 1992 quoted in Corbett 2000: 124), some nouns have different endings for the singular, plural, and mass, e.g. for ‘hair’ – *pílu* (singular) vs. *pélos* (plural) vs. *pélo* (mass). Also adjectives have a separate mass agreement form, which can be used even with those nouns that, for purely morphological reasons, lack a special mass form, e.g. *la maéra tábá séko* ‘the wood (MASS) was dry’ vs. *la maéra tábá séka* ‘the (piece of) wood was dry’. A similar distinction (interpreted in the literature as a gender and not a number one) exists in many Italian dialects. Its manifestations differ considerably across dialects and include different endings for mass nouns, different agreement forms for articles and demonstratives, and a lengthening of the initial consonant of a mass noun when it is preceded by a determiner (cf. Puglian *u ‘pa:nə* ‘the loaf’ vs. *u p’pa:nə* ‘bread’, Maiden 1995: 248).

Mass forms opposed both to the regular singular and plural count forms are marginally found in other languages as well. In Swedish a few words have an uncountable form, e.g. *mygga* ‘mosquito:SG’ vs. *myggor* ‘mosquito: PL’ vs. *mygg* ‘mosquito: MASS.PL’

and *ärta* ‘pea:SG’ vs. *ärter* ‘pea:PL’ vs. *ärter* ‘pea:MASS.PL’ (Delsing 1993: 45); cf. also 5 on Arabic.

Having a regular mass form seems to be rare (Corbett 2000: 124), and mass nouns are normally treated as having one of the numbers typical of the count nouns in the language, in most cases the singular (**singularia tantum**). However, mass nouns as basically plural (**pluralia tantum**) are attested in various languages. For Russian, Polish, and English, Wierzbicka (1988) discerns several classes of entities which tend to be lexicalized as basically count vs. mass nouns and as basically singularia tantum, pluralia tantum or fully countables. Thus, names of homogeneous substances (*wine, butter*) and names of substances with a minimal unit (*rice, sand*) tend to be lexicalized as singularia-tantum mass nouns, and the same goes for names of heterogeneous classes of objects (*furniture*). Names of substances “composed of particles, not too many for anyone to be able to count but too many for anyone to want to count” (*oats*), as well as names of groups of objects and/or ‘stuff’ (*leftovers*) are often lexicalized as pluralia-tantum mass nouns. Membership of nouns in one or another category depends partly on the inherent properties of their referents and partly on cultural usage. Thus, oats consist of much more clearly perceivable separate particles than flour – therefore in English the former is conceptualized as a multiple entity and the latter as homogeneous stuff. Berries can be fairly big (e.g. strawberries), but in the Russian culture they are eaten as stuff rather than separately, therefore they are lexicalized as singular mass nouns. Wierzbicka’s explanations, although both insightful and inspiring, cannot account for all the concrete cases of mass vs. count and singular vs. plural distinctions in a particular language and, most importantly, across languages. For instance, Russians and Swedes have both more or less the same repertoire of berries and fairly similar habits in eating them, but only in Russian their names are singularia-tantum mass nouns.

Slavic and Baltic are notorious for their numerous pluralia tantum including mass nouns, cf. *slivki* ‘cream’ and *černila* ‘ink’ in Russian, or *putos* ‘smoke’, *miltai* ‘flour’ and *solotos* ‘salad’ in Lithuanian. In Baltic, plurality of mass nouns in certain semantic categories is so pervasive that originally singular loanwords may become plural there, e.g. in Latvian *ziepes* ‘soap:PL’ > Middle German

*se(e)pe* ‘soap:SG’, and *kaposti* ‘cabbage:PL’ > Old Russian *kapusta* ‘cabbage:PL’. Conversely, due to the intensive language contacts, the plural character of some mass nouns in Latvian has been transferred to the corresponding words in Latvian Romani and Livonian, cf. *thuvá* ‘smoke:PL’ and *firjá* ‘foam:PL’ in Latvian Romani.

Many mass nouns may show number oppositions, often with non-trivial semantic effects. Thus, for the basically singular mass nouns, the plural form frequently has the meanings of abundance (normally, referring to an extensive geographic area covered by the stuff in question), e.g. Russ. *peski* ‘a large area covered by sand’; occurrence at different places, e.g. Chirag (Dagestanian) *k'um-re* (lit. ‘sour.cream-PL’) ‘sour cream in different vessels’ (Kibrik 1992); several portions, e.g. *three beers*, or different sorts, kinds, e.g. *wines*. Occasionally the plural form of a noun referring to a typical material can refer to things made of it – e.g., Russ. *žemčuga* lit. ‘pearls’ used for ‘jewelry’. Names of stuffs in the plural may sometimes refer to a special subtype of the stuff. For instance, Russ. *vody* lit. ‘waters’ often refers to mineral waters or spas. A possibility itself to use a basically singular mass noun in the plural and the resulting meanings are lexically restricted. Conversely, the plural form of count nouns may get lexicalized as a mass noun referring to a special kind of the stuff designated by the regular plural, e.g. *drops* in the pharmaceutical context vs. *drops* in general.

Luganda and other Bantu languages demonstrate a particularly clear connection between the plural and mass nouns. Mass nouns are included in the *ma*-class, which is otherwise reserved for the plural of the nouns with the prefix *li*- in the singular, as *likongó* ‘a/the spear’ vs. *makongó* ‘(the) spears’. For some *li*-words, such as *lilálá* ‘one orange’ (fruit, the *ma*-form alternates between the count-plural and the mass interpretation, e.g. *malálá* ‘more than one orange or mass of orange’. Other words most commonly occur in the *ma*-form, e.g. *makayábo* ‘mass of salted fish’. Their *li*-form refers to a unit/measure, e.g. *likayábo* ‘one unit/cut of salted fish’ (Mufwene 1980).

### 3.2. Gender/noun class

An interesting phenomenon, which is clearly sensitive to the count-mass distinctions, is **gender assignment** of nouns in various Germanic varieties (cf. 3.1 on mass gender/

number in Romance). In West Somerset English, Newfoundland English, and Tasmanian English, where pronouns regularly distinguish between three genders, the use of the neuter is mainly restricted to uncountable and abstract nouns (Siemund forthc.).

In Standard Danish and Swedish names of some liquids can alternate between the neuter gender for reference to the stuff itself and the common gender for reference to a portion of it, e.g. Swedish *en kaffe* ‘one(COMMON) coffee’ vs. *kaffe-t* ‘coffee-DEF.N’ (‘the coffee’). The West Jutish dialects of Danish distinguish between anaphoric and deictic pronouns for countables (*den*) and those for uncountables and abstracts (*det*). Related phenomena are found in Frisian and Dutch.

### 4. Is the count-mass distinction a linguistic universal?

At least two types of languages are frequently quoted as challenging the universality of count-mass division in nominals, because they combine all nouns with numerals in more or less the same fashion, without distinguishing, say, between “three cows” and “three glasses of water”. First, numeral-classifier languages like Chinese or Vietnamese where all nouns in numeral constructions must be accompanied by a classifier. Second, some languages, like Tolai (Mosel 1982) or Tagalog (Gil 1996: 59), where all nouns, even the most “mass-like” ones, seem to be able to combine with numerals and plural markers directly – the most “mass-like” ones are then understood as referring to portions of the mass. This is sometimes interpreted as the absence of mass-count distinctions in those languages and, in most extreme cases, as the existence of only mass nouns in the former language type (Chierchia 1998; Krifka 1991: 401) and of only count nouns in the latter one (Krifka 1991). Both this conclusion and the evidence underlying it have been questioned at least for some classifier languages. The usual counter-argument boils down to there being two types of **classifiers** – **sortal** and **mensural** ones – which differ in semantics and combinability with various kinds of nominals (cf. Art. 97). However, in spite of the extensive literature on classifiers and the well-established tradition to distinguish the two classifier types, there is relatively little done on their morphosyntactic differences (cf. Art. 97; Aikhenvald 2000: 114–120 and the references there).

## 5. Collections and collectives

In linguistic literature the terms “collection” and “collective” are used in several meanings (see Gil 1996: 66–68 for the list of uses of “collective”). “**Collective nouns**” often refer to words like *herd* and *family*, or *cattle* and *furniture*, which “denote collections or groups, of persons or objects” (Lyons 1977: 309–311). Collective nouns are semantically dual – they refer to a multiplicity of clearly discernible entities “hidden” in a collection, which is either bounded (*herd*, *family*) or unbounded (*cattle*, *furniture*) at a higher level of abstraction. Some works distinguish between “collective nouns”, like *cattle*, and “**collections**” like *herd* or *forest*, but this distinction is far from being generally accepted.

Collectives on the whole are hardly recognizable on formal grounds, but some of them might have certain peculiarities in behaviour or marking. Thus, a well-known property of certain collectives, sometimes called “**corporate nouns**”, in British and New Zealand English is the ability to trigger both singular and plural agreement, e.g. *the committee has/have chosen*. Plural agreement strongly correlates with the animacy hierarchy: it is found primarily with collections of humans, is excluded with inanimates (\**the forest are ...*) and is possible, but dispreferred with non-human animates ((?)*the herd are grazing*) (Corbett 2000: 188–190). Similar alternations occur in Spanish, Old Church Slavonic, Swedish, Paumari (an Aruan language of Brazil), Samoan (Polynesian), and Kabardian (North-West Caucasian). Number ambivalence typical of collectives is also witnessed by the historical developments of erstwhile collectives into regular plurals, well known e.g. from Slavonic.

Collective nouns, formally in the singular, can serve as arguments to verbs which otherwise require plurals, and occur in other “plural” contexts, e.g. *die Polizei umzingelte das Gebäude* ‘the police surrounded the building.’

The Arabic tradition distinguishes two kinds of collectives – “nouns of collections”, like ‘herd’ and ‘company’, and “nouns of collective kinds”. The latter comprises, e.g. many nouns denoting fruit, vegetables, flowers, grains, insects, and birds denoted without specifying number. These forms serve as the basis for the derivation of **singulatives (unit nouns)**, which, in turn, can be pluralized, e.g. Maltese *baqar* (collective) ‘the species “cow”’, ‘cows’ vs. *baqra* (singulative singular) ‘a single cow’ vs. *baqriet* (singulative plu-

ral) ‘several single cows’ (Gil 1996: 72). The collective is not completely parallel to the European collectives of the *cattle*-type because it does not presuppose several entities and can therefore be characterized as a form of the general number (Corbett 2000: 9–19, 129–132).

Also in Celtic, “collective nouns” are outside of the regular singular-plural opposition, whereby the plural is marked with respect to the singular. “Collectives” serve as the basis for deriving singulatives, but, in contrast to singulatives in Arabic, the latter cannot further be pluralized, e.g. in Welsh *plu* (collective) vs. *pluen* (singulative) ‘feather’, *adar* (collective) vs. *aderyn* ‘bird’ (singulative). In practice, “singulatives” and “collectives” build number oppositions comparable to other nouns, differing mainly in the relative markedness of their members.

The term “collective” is often opposed to “distributive”. Thus, many languages, in particular in North America, but also elsewhere, distinguish between “**collective number**”, used for referring to several entities viewed as a group, and “**distributive number**”, used for referring to several entities viewed individually, e.g. in Budukh (Dagestanian) *t'ilibər* (collective) ‘fingers (of one hand)’ vs. *t'ilimber* (distributive) ‘fingers (of several hands)’ (Kibrik 1992: 15; cf. Corbett 2000: 111–120).

“**Collective**”, or “**set interpretation**” vs. “**distributive**”, or “**individual interpretation**” may be illustrated by the sentence “all the children in the classroom are reading a book”. In the former interpretation the children are reading one and the same book, whereas in the latter case each of them is reading a separate book.

“**Collective numerals**” are found, for instance, in Slavic (e.g., *dvoe* (collective) ‘two’ and *troje* (collective) ‘three’, as opposed to *dvaldve* and *tri* in Russian), where they are traditionally characterized as referring to several entities, preferably humans viewed as a group. The semantic difference between collective and cardinal numerals is fairly blurred, but some contexts are restricted to only one of the types. Thus, only collective numerals combine with countable pluralia tantum, e.g. *dvoe očkov* ‘two (collective) glasses’.

## 6. Operations relating collectives and collections, mass and individuals

Collective nouns may often be derived from nouns denoting individuals, e.g. with a prefix in *Ge-birge* from *Berg* ‘mountain’ and a suffix in *Mensch-heit* from *Mensch* ‘human being’

in German, or with a circumfix in *kapuluan* ‘archipelago’ from *pulo* ‘island’ in Tagalog (Gil 1996: 63). Collectives can sometimes have an additional pejorative connotation, e.g. *trjap'e* ‘finery’ from *trjapka* ‘rag’ in Russian. A number of languages use compounding for deriving collective nouns from two individual nouns, each of which denotes a member of the collective, e.g. Chuvash *ulma-syrla* ‘apple-berry(fruit)’.

Singulatives (unit nouns) are derived from collective and mass nouns and denote individuals, which constitute the collective, or discernible particles, which constitute the mass (see 5 for Arabic and Celtic), e.g. by compounding in Swedish *saltkorn* ‘a grain of salt’ from *salt* ‘salt’ + *korn* ‘grain’ or by suffixation in Russian *gorošina* ‘a pea’ from *gorox* ‘peas’ (mass). Collections and units/ portions of mass/collectives can also be referred to by means of **mensurative (pseudo-partitive) constructions**, in which one nominal quantifies over the kind of entity indicated by the other nominal. Nominal quantifiers originate as concrete nouns, are used for creating a unit of measure which may further be counted and show a variety of types, e.g. conventionalized measures (*a litre of milk*), containers (*a cup of tea*), fractions (*a slice of bread*), quanta (*a lump of sugar*), collections (*a group of students*), and forms (*a pile of sand*).

Cross-linguistically, the by far dominating technique in such constructions consists of merely juxtaposing the nominal quantifier and its complement, e.g. *eine Flasche Wein*. However, many languages involve some construction markers for relating the nominal quantifier and its complement to each other, which are either a case-marking on the quantified nominal, as in Russian *bokal vina* – ‘glass wine:GEN (a glass of wine)’ and Finnish *säkki perunoita* ‘sack:NOM potato:PRTV.PL (a sack of potatoes)’ or a preposition introducing it, as in their English correspondence. These construction markers originate from several sources, primarily from possessives, ablatives (e.g., due to the reanalysis of sentences referring to physical separation of a part from an object, such as “give me a part from the cake”) and comitatives (cf. Koptjevskaia-Tamm 2001 for a large-scale synchronic and diachronic cross-linguistic study).

## 7. References

- Aikhenvald, Alexandra Y. (2000), *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press
- Allan, Keith (1980), “Nouns and Countability”. *Language* 56, 541–567
- Behrens, Leila (1995), “Categorizing between Lexicon and Grammar: The MASS/COUNT Distinction in a Cross-Linguistic Perspective”. *Lexicology* 1.1, 1–112
- Bunt, Harry C. (1985), *Mass Terms and Model-Theoretic Semantics*. Cambridge: Cambridge University Press
- Chierchia, Gennaro (1998), “Reference to Kinds across Languages”. *Natural Language Semantics* 6, 339–405
- Corbett, Greville G. (2000), *Number*. Cambridge: Cambridge University Press
- Delsing, Lars-Olof (1993), *The Internal Structure of Noun Phrases in the Scandinavian Languages: A Comparative Study*. Ph.D. dissertation, University of Lund
- Gil, David (1996), “Maltese ‘Collective Nouns’: A Typological Perspective”. *Rivista di linguistica* 8.1, 53–88
- Hualde, José Ignacio (1992), “Metaphony and Count/Mass Morphology in Asturian and Cantabrian Dialects”. In: Laeufer, Christiane & Terrell A. Morgan (eds.), *Theoretical Analyses in Romance Linguistics: Selected Papers from the Nineteenth Linguistic Symposium on Romance Languages, the Ohio State University, 21–23 April 1989*. Amsterdam: John Benjamins (Current Issues in Linguistic Theory 74), 99–114
- Jackendoff, Ray (1991), “Parts and Boundaries”. *Cognition* 41, 9–45
- Jespersen, Otto (1924), *The Philosophy of Grammar*. London: Allen & Unwin
- Kibrik, Aleksandr E. (1992), “Defective Paradigms: Number in Daghestanian”. European Science Foundation, Programme in Language Typology (EuroTyp), Noun Phrase Structure, Working Paper 16
- Kolmer, Agnes (1999), “Zur MASS/COUNT-Distinktion im Bairischen: Artikel und Quantifizierung”. Cologne: Universität zu Köln, Institut für Sprachwissenschaft, Arbeitspapier 34 (Neue Folge)
- Koptjevskaia-Tamm, Maria (2001), “‘A Piece of the Cake’ and ‘a Cup of Tea’: Partitive and Pseudo-Partitive Nominal Constructions in the Circum-Baltic Languages”. In: Dahl, Östen & Koptjevskaia-Tamm, Maria (eds.), *Circum-Baltic Languages, Vol. II*. Amsterdam, Philadelphia: John Benjamins, 523–568
- Krifka, Manfred (1991), “Massennomina”. In: Stechow, Arnim von & Wunderlich, Dieter (eds.), *Semantics: An International Handbook of Contem-*

- porary Research. Berlin: Walter de Gruyter, 399–417
- Langacker, Ronald W. (1991), *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter
- Lyons, John (1977), *Semantics, Vol. I*. Cambridge: Cambridge University Press
- Maiden, Martin (1995), *A Linguistic History of Italian*. London, New York: Longman
- Mosel, Ulrike (1982), “Number, Collection and Mass in Tolai”. In: Seiler, Hansjakob & Stachowiak, Franz Josef (eds.), *APPREHENSION: Das sprachliche Erfassen von Gegenständen. Teil II: Die Techniken und ihr Zusammenhang in Einzelsprachen*. Tübingen: Narr, 123–154
- Mufwene, Salikoko S. (1980), “Number, Countability and Markedness in Lingala LI-/MA- Noun Class”. *Linguistics* 18, 1019–1052
- Pelletier, Francis J. (1975), “Non-Singular Reference: Some Preliminaries”. *Philosophia* 5, 451–465
- Pelletier, Francis J. & Schubert, Lenhart K. (1989), “Mass Expression”. In: Gabbay, Dov & Guenther, Franz (eds.), *Handbook of Philosophical Logic, Vol. IV*. Dordrecht: D. Reidel, 327–407
- Quine, Willard V. O. (1960), *Word and Object*. Cambridge, Mass: MIT Press
- Rijkhoff, Jan (2002), *The Noun Phrase*. Oxford: Oxford University Press
- Sasse, Hans-Jürgen (1993), “Syntactic Categories and Subcategories”. In: Jacobs, Joachim & Stechow, Arnim von & Sternfeld, Wolfgang & Vennemann, Theo (eds.), *Syntax: An International Handbook of Contemporary Research*. Berlin: Mouton de Gruyter, 646–686
- Siemund, Peter (forthc.), *Pronominal Gender in English: A Study of English Varieties from a Cross-Linguistic Perspective*. London: Routledge
- Wierzbicka, Anna (1988), *The Semantics of Grammar*. Amsterdam, Philadelphia: John Benjamins

*Maria Koptjevska-Tamm, Stockholm  
(Sweden)*

## 102. Case

1. A case system
2. Problems of description
3. Realisation
4. Survey of case marking
5. Diachrony of case
6. References

### 1. A case system

**Case** in its most central manifestation is an inflectional system of marking noun phrases for their relationship as dependents to their heads. Typically case marks the relationship of a noun phrase to a verb as in the following example from Turkish (cf. Lewis 1967: 25–52).

- (1) *Mehmet adam-a elma-lar-i*  
Mehmet(NOM) man-DAT apple-PL-ACC  
*ver-di.*  
give-PAST.3.SG  
'Mehmet gave the apples to the man.'

*Mehmet* is in the nominative case which indicates the subject. *Adam* is in the dative case which indicates that it is the recipient of *vermek* ‘give’. *Elmalar* is in the accusative which indicates that it is the direct object. The accu-

sative also indicates that it is definite or at least specific, since in Turkish only specific direct objects are marked by the accusative. Non-specific direct objects are put in the nominative.

Turkish has a system of six cases as in Tab. 102.1. The genitive marks an adnominal, dependent noun phrase as in *adam-in ev-i* ‘man-GEN house-3.SG.POSS (of the man his house)’, i.e. ‘the man’s house’. In Turkish the possessor, if definite, is cross-referenced by a possessive pronominal suffix on the possessed noun, hence the *-i*. The locative marks location as in *Istanbul-da* ‘in Istanbul’, and the ablative indicates ‘from’ or ‘out of’ as in *Ankara-dan* ‘from Ankara’.

---

nominative	<i>adam</i>
accusative	<i>adami</i>
genitive	<i>adamin</i>
dative	<i>adama</i>
locative	<i>adamda</i>
ablative	<i>adamdan</i>

---

Tab. 102.1: Turkish case system

In (1) the cases are determined or governed by the verb. *Vermek* ‘to give’ requires a subject in the nominative, an indirect object in the dative and a direct object in the accusative or nominative. Cases can also be governed by prepositions or postpositions. Turkish has postpositions which govern the ablative like *dolayı* ‘because of’: *toplanti-dan dolayı* ‘meeting-ABL because.of (because of the meeting)’. It also has postpositions that govern the genitive with pronouns but the nominative with nouns. Such a postposition is *gibi* ‘like’: *kim-in gibi* ‘who-GEN like (like whom)’, but *bu adam gibi* ‘this man(NOM) like (like this man)’.

The word forms displayed in Tab. 102.1 make up a table or paradigm of case forms. In Turkish one could say that there is only one paradigm. It is true that noun stems of different shapes take different inflectional suffixes, but all these differences are phonologically conditioned by principles of vowel harmony and the like. The locative, for instance, has the form *-da* following stems with back vowels and *-de* following stems with front vowels. The *d* of this suffix devoices to *t* following a stem-final voiceless consonant: *kitap-ta* ‘on (the) book’. One could refer to *-da*, *-de*, *-ta*, and *-te* as case markers or one could consider that at a more abstract level there was only one locative case marker. We need to make a distinction between **cases**, (of which there are six in a system of oppositions) **case markers** or **case forms** (the affixes or word forms respectively that realise the oppositions) and **case relations** (the syntactic or semantic relations that the cases express). Case relations, also referred to by other labels such as grammatical relations or functions, need not be in a one-for-one correspondence with cases. In Turkish the nominative expresses the subject, but not all noun phrases in the nominative are subject since the nominative also marks a non-specific direct object of a transitive verb (see (13)).

The term “case” is also used for the phenomenon of having a case system and a language with such a system is sometimes referred to as a **case language**.

## 2. Problems of description

Turkish is a convenient language with which to illustrate a case system since there is in a sense only one paradigm and case is repre-

sented by a set of separable markers. However, the traditional study of case in the West has been concentrated on Ancient Greek and more particularly Latin where a number of complications are found. In Latin, for instance, it is not possible to separate number marking from case marking. The two categories have fused representation or **cumulative exponence** (Matthews 1974). This means separate paradigms for the two number categories, singular and plural. Moreover, there are different case/number markers for different stem classes. Traditionally five such classes are recognised and there are also variations within the classes. These inflectional classes are known as declensions from *declīnātiō*, literally a bending down or aside. This refers to the notion of varying from the basic form, the nominative. In fact the word *case* itself is from Latin *cāsus* ‘a falling’ which is in turn calqued on the Greek *ptōsis*. Again the notion is of falling away from the *ptōsis orthē*, Latin *cāsus rectus*, ‘the upright case’. The non-nominative cases were traditionally the *ptōseis plagiai*, or *cāsūs obliquū* ‘the oblique cases’, literally, the leaning or sloping cases (Sittig 1931; Robins 1967: 29).

Three declensions are illustrated in Tab. 102.2: the first declension (-*ā* stems), second declension (-*o* stems), and third declension (consonant stems). The third declension also contains -*i* stems, which are not illustrated, nor are two other minor declensions, the fourth (-*u* stems) and fifth (-*ē* stems). The distinctions are a reality, but the labels are not synchronically transparent and are to some extent based on reconstruction.

In Latin there is also a three-way gender distinction: masculine, feminine, and neuter. With a few exceptions male creatures are masculine and females feminine, but inanimates are scattered over all three genders (though almost all neuter nouns are inanimate). There is a partial association of form and gender in that -*ā* stems are almost all feminine and -*o* stems mostly masculine (except for a sub-class of neuters represented by *bellum* in Tab. 102.2). This means that there can be fusion of gender, number, and case. The point is illustrated in Tab. 102.2 where we have *domina* ‘mistress of a household’ illustrating feminine -*ā* stems and *dominus* ‘master of a household’, which is based on the same root, representing masculine -*o* stems. As can be seen from Tab. 102.2 *domina*

	1 -ā stems	2 -o stems	3 cons. stems	
	feminine	masculine	neuter	
	<i>domina</i> ‘mistress’	<i>dominus</i> ‘master’	<i>bellum</i> ‘war’	<i>cōnsul</i> ‘consul’
<b>singular</b>				
nominative	<i>domina</i>	<i>dominus</i>	<i>bellum</i>	<i>cōnsul</i>
vocative	<i>domina</i>	<i>domine</i>	<i>bellum</i>	<i>cōnsul</i>
accusative	<i>dominam</i>	<i>dominum</i>	<i>bellum</i>	<i>cōnsulem</i>
genitive	<i>dominae</i>	<i>dominī</i>	<i>bellī</i>	<i>cōnsulis</i>
dative	<i>dominae</i>	<i>dominō</i>	<i>bellō</i>	<i>cōnsulī</i>
ablative	<i>dominā</i>	<i>dominō</i>	<i>bellō</i>	<i>cōnsule</i>
<b>plural</b>				
nominative	<i>dominae</i>	<i>dominī</i>	<i>bella</i>	<i>cōsulēs</i>
vocative	<i>dominae</i>	<i>dominī</i>	<i>bella</i>	<i>cōsulēs</i>
accusative	<i>dominās</i>	<i>dominōs</i>	<i>bella</i>	<i>cōsulēs</i>
genitive	<i>dominārum</i>	<i>dominōrum</i>	<i>bellōrum</i>	<i>cōsulum</i>
dative	<i>dominīs</i>	<i>dominīs</i>	<i>bellīs</i>	<i>cōsulibus</i>
ablative	<i>dominīs</i>	<i>dominīs</i>	<i>bellīs</i>	<i>cōsulibus</i>

Tab. 102.2: Latin case paradigms

simultaneously represents nominative case, feminine gender and singular number, *dominum* represents accusative case, masculine gender and singular number, and similarly with other suffixes.

In Latin there is concord between a noun and an attributive or predicative adjective. This concord is sensitive to case and number and those adjectives that belong to the first and second declension are sensitive to gender so we find *domina bona* ‘good mistress’ and *nauta bonus* ‘good sailor’ where *nauta* is a noun of masculine gender in the first declension. With adjectives of the first and second declensions the inflections simultaneously represent case, number, and gender without exception.

As can be seen six cases are recognised: nominative, vocative, accusative, genitive, dative, and ablative; however, no paradigm exhibits six different forms. In the traditional descriptions a case is established wherever there is a distinction for any single class of nominals. The vocative, the case used in forms of address, has a distinctive form only in the singular of the second declension. Elsewhere there is a common form for the nominative and vocative, i.e. there is syncretism between the two cases. However, distinct no-

nominative and vocative cases are recognised for all paradigms.

The expression of location (‘place at’) is interesting with respect to the principles of establishing cases. Location is usually expressed by a combination of a preposition like *in* and the ablative case: *in urbe* ‘in the city’. However, with towns and small islands (which were considered points rather than areas) no preposition was used and the form coincided with the genitive for names belonging to the first and second declension singular and with the ablative otherwise (e.g. *Karthagine* ‘in Carthage’), with a few alternative dative-like forms in the third declension (e.g. *in Karthaginī* ‘in Carthage’). There were also a few fossilised forms like *bellō* ‘at war’. Here a formal distinction correlates with a function, but there is no unique locative form for any class of nominal. Strict adherence to the principle of establishing a separate case whenever a formal distinction correlates with a function would demand setting up a locative case as in Tab. 102.3. In practice traditional Latin grammars sometimes talk about a locative case, but exclude it from their paradigms. However, this is perhaps to be regarded as the result of a space saving strategy rather than of a theory based decision.

		singular	plural
first declination	genitive	<i>Rōmae</i>	<i>Athēnārum</i>
	locative	<i>Rōmae</i>	<i>Athēnīs</i>
	ablative	<i>Rōmā</i>	<i>Athēnīs</i>
second declination	genitive	<i>Milētī</i>	<i>Delphōrum</i>
	locative	<i>Milētī</i>	<i>Delphīs</i>
	ablative	<i>Milētō</i>	<i>Delphīs</i>

Tab. 102.3: Locative case in Latin

The practice of recognising a case on the basis of a distinction in one paradigm is widespread but not universal. Another descriptive practice is to recognise different sets of cases in different paradigms, in particular to use a common label for case markers or case forms that neutralise distinctions made in other paradigms. In Estonian, for instance, there is less differentiation in the case marking of the attributive adjective than in the noun, with a common form corresponding to several distinct forms in the noun. This has been referred to as an **oblique case** (cf. Comrie 1981: 121; see Tab. 102.4 for weak adjectives in German). Applying this practice to Latin we might say that there is a **direct case** for neuter nouns corresponding to the nominative and accusative of non-neuter nouns. However, to be consistent we would have to find common labels for all other syncretisms and all these deviations from the ‘maximal differentiation’ principle complicate the description of concord (cf. Goddard 1982 on Australianist practice).

Describing the meanings and functions of the cases in Latin and other Indo-European languages has long been a matter of dispute. The standard treatment in reference grammars is to list a number of meanings or functions for each case. A typical Latin grammar, for instance, would recognise an ablative of separation in *Athēnīs redēo* ‘I return from Athens’, a locative ablative in *Athēnīs habitō* ‘I live in Athens’, and an instrumental ablative in *gladiō ferīre* ‘to wound with a sword’, with many other semantic subdivisions besides. The Roman writer Quintilian suggested splitting the ablative into an ablative proper and an instrumental (*Institutio oratoria*, 1.4.26). He appears to have been influenced by the fact that the separation and instrumental senses are differentiated in Ancient Greek, the separation or source sense being expressed via the genitive and the instrumental sense via the dative (Greek had no abla-

tive) (Hjelmslev 1935: 16). Since the nineteenth century it has been commonplace to point out that the separation, locative, and instrumental senses reflect three separate cases in Indo-European which syncretised in Latin, although, as noted above, it is still possible to pick out a locative in Latin by considering the fact that the locative meaning is expressed by genitive forms with some place names and ablative elsewhere. There is another approach to describing the meaning of cases and that is to characterise each in terms of a generalised meaning (**Gesamtbedeutung**) which minimally distinguishes one case from another. This notion goes back a long way. It can be found, for instance, in the work of the Byzantine grammarian Maximus Planudes (1260–c. 1310), and is prominent in the writings of the scholastic grammarians of the twelfth and thirteenth centuries (Hjelmslev 1935: 11 f.; Serbat 1981: 24–26). The most prominent exponents of this approach in this century have been Hjelmslev (1935; 1937) and Jakobson (1936), and the idea still has its proponents such as Rubio (1966) and de Carvalho (1980) (see also the discussion in Touratier 1978). To maintain the generalised meaning or function one needs to attribute differences of semantic role to the choice of lexical filler or choice of governor (verb or adposition). Applying this latter principle to the Latin ablatives quoted above we can see that the separation sense of the ablative in *Athēnīs redēo* can be attributed to the meaning of *redire* ‘to return’, the locative sense of the ablative in *Athēnīs habitō* to the meaning of *habitare* ‘to dwell in’, and so on (Rubio 1966: 155–187). This might be called an emic approach since it seeks to establish that the etic varieties of meaning are predictable from the environment or it seeks to dismiss them as uses of cases (cf. Hjelmslev 1935: 90 f.).

A prominent theme in establishing **Gesamtbedeutungen** is the idea of characterising cases in terms of local notions of position and direction. This has become known as the **localistic** approach to case and it goes back at least to Maximus Planudes. Obviously an ablative can be characterised in terms of ‘source’ or ‘origin’, but so can a nominative at a more abstract level, since it can be thought of as expressing the source of the action. Similarly an accusative can be thought of as expressing the destination of the activity or process denoted by the verb (cf. Serbat 1981: 26).

Such an approach has its critics. Wierzbicka, for example, argues that the generalised definitions of Jakobson are too broad to be predictive and that one cannot learn to use the Russian cases from his formulas (Wierzbicka 1980: xv). However, Jakobson was able to analyse his generalised case meanings in terms of distinctive features and these facilitate generalisations over groups of cases. In Russian there is syncretism of nominative and accusative in some nouns but syncretism of accusative and genitive in others. Comrie (1986: 102 f.) demonstrates that one can capture the grouping nominative + accusative by means of a feature [direct] and accusative + genitive by means of a feature [+ objective]. Thus the form *stol* which is the nominative and accusative of *stol* ‘table’ can be described as a [+ direct] form and the form *slona* which is the accusative and genitive of *slon* ‘elephant’ can be described as [+ objective]. Comrie’s features are not Jakobson’s and they are syntactically rather than semantically based.

### 3. Realisation

#### 3.1. Within the clause

In 1, case was described as an inflectional system of marking the relationship of nominal dependents to heads. This describes case in languages like Ancient Greek and Latin where the case marking is integrated into the host word and where it marks the relationship of a nominal to a governing verb or preposition. The same applies to Turkish except that Turkish has postpositions rather than prepositions. In many languages, however, there are adpositions but no case inflection. In Korean and Japanese case relations are marked by postpositions. In Japanese the subject is marked by *ga*, the direct object by *o* and the indirect object by *ni*.

- (2) *Sensei ga Tasaku ni hon o yat-ta.*  
teacher SBJ Tasaku IO book OBJ give-PAST  
'The teacher gave Tasaku a book.'

In some languages the core grammatical relations are distinguished by word order as in English or by the use of bound pronominal forms, either clitics or inflectional affixes. These pronominal forms are usually affixed to the verb as in the Bantu languages or enclitic to the first constituent of the clause as in some Uto-Aztecán languages. With the non-core cases these alternatives do not exist

(but see 4.2.2 on alternatives to the genitive). One finds either case inflection, usually with adpositions as well, or simply adpositions. Another widespread category used to mark the relationship of a noun phrase to a verb is a specialised **relator noun** to express relative position, that is, notions like ‘over’, ‘under’, ‘beside’, etc. English has examples such as *top* in *on top of* (note that *top* cannot be modified by a determiner or adjective when functioning as a relator noun).

If we take the term “case” to embrace all these means of expressing the relationship of noun phrases to their governors, then all languages will be case languages and the notion becomes vacuous as a characterisation of means of expression. Case languages in the traditional sense, that is languages with inflectional case, are characteristic of Eurasia and Australia. Indo-European had case though this has been almost entirely lost in English and the western Romance languages. Case is characteristic of the Uralic, Altaic, Caucasian, Eskimo-Aleut, and Dravidian families. In Australia the Pama-Nyungan languages which cover most of the mainland have case as do a majority of the families found in the north of the continent. In Semitic there are up to three cases (nominative, accusative, and genitive) (Hetzron 1987: 659) and small systems are found in Cushitic and Nilo-Saharan languages (Tucker & Bryan 1966). In the Americas a number of languages have case systems, the Penutian, Yuman, and Uto-Aztecán languages, for instance, exhibit case as do Tarascan and Zoque. Some have the minimal number of cases, namely two, as in some Uto-Aztecán languages such as Yaqui, Chemehuevi, and Comanche (Sherzer 1976; Suárez 1983: 87).

#### 3.2. Within the noun phrase

##### 3.2.1. Internal relations

The head of a noun phrase will normally be able to take further noun phrases as dependents. The relationship of these to the head can typically be expressed by a variety of indirect cases (cases other than those used for subject and object; see 4), but often there is a special adnominal case. In Latin the genitive has this function (though it also marks the complement of a handful of verbs). It marks the possessor as in *cōsulis villa* ‘the consul’s villa’, but it also marks a variety of types of complement as in the following phrase from Caesar (*de Bello Gallico*, 1.30.2).

- (3) ... *prō veter-ibus Helvēti-ōrum*  
 for old-ABL.PL Helvetii-GEN.PL  
*iniūri-īs popul-īt Romān-īt.*  
 injury-ABL.PL people-GEN Roman-GEN  
 ‘... for the old injuries inflicted by the  
 Helvetii on the Roman people.’

The first genitive here is of a type traditionally called **subjective genitive** and the second **objective genitive**. They express complements corresponding to the subject and object of a verbal predicate meaning ‘to injure’. This development of a special adnominal grammatical case has parallels in a number of languages (see also 4.2.2).

### 3.2.2. External relations

There are two common distributions of case marking within the noun phrase. In one type case appears not only on the head nominal in the noun phrase but also on the determiner and usually the adjective. This system is familiar from Indo-European languages and can be illustrated from Ancient Greek. In (4) the determiner and attributive adjective show agreement for case, number, and gender. *Bios* in (4) is masculine.

- (4) *ho anexéast-os*  
 the:M.NOM.SG unexamined-M.NOM.SG  
*bí-os*  
 life-M.NOM.SG  
 ‘the unexamined life’

Concord is also found in Balto-Finnic, Semitic, in some Pama-Nyungan languages and elsewhere, but it is a minority system.

In the other type, case marking appears on the final word in the noun phrase. We can distinguish two sub-types. In the first the final word is the head noun in the noun phrase. This is the system found in the Altaic, Dravidian languages, and Papuan languages. The following example is from the Papuan language Fore (Scott 1978: 110).

- (5) *Pí 'kiná'mi'-ti waye.*  
 that being-ALL go.3.SG  
 ‘He goes to those people.’

In the other sub-type the final word in the noun phrase is not always the head. This is the situation in various Australian and Amazonian languages (Derbyshire & Pullum 1986; 1989; 1991). The phenomenon also occurs in Basque: *etxe zaharr-etan* ‘house old-PL.LOC (in old houses)’ (Saltarelli et al. 1988: 77).

The difference between postpositions and phrase-final inflection is not always clear. In

Korean it is normal to speak of postpositions, but the forms in question have variants for vowel-final stems and consonant-final stems: *can mul ul* ‘cold water OBJ’, but *can ooyoo lul* ‘cold milk OBJ’. At the other extreme is a language like Diyari (cf. Art. 137; Austin 1981a) where the marking consists of a series of morphologically determined variants plus some phonologically determined alternants that cannot be related synchronically to a single basic form. It would be difficult to maintain that sets of alternants like these were postpositions.

Among Australian languages there are some that need mark only one constituent of the noun phrase but not necessarily the last word in the phrase. In Nyigina and Goonyandi (McGregor 1990) discourse principles play some part. In Nyigina, for instance, the marking is often on the first word: *gudyarra-ni wamba* ‘two-ERG man (two men)’, or frequently on the word that is most significant: *ginya marninga Wurrawurra-ni* ‘that woman Wurrawurra-ERG (that Wurrawurra woman)’ (Stokes 1982: 59 f.). The ergative suffix marks the agent of a transitive verb. In another Australian language Uradhi the case marking applies obligatorily to the head, but only optionally to the dependents (Crowley 1983: 371 f.).

- (6) *Utagha-mpu amanyma(-mpu)*  
 dog-ERG big(-ERG)  
*udhumpuyn ighanhanga-n.*  
 back:ACC break-PAST  
 ‘The big dog broke [the other dog’s]  
 back.’

In some languages there is concord between the determiner and what one would take to be the head noun, but not between the adjective and head. This is the situation in Yaqui (Uto-Aztecán), for instance (Lindenfeld 1973: 60). In the Indo-Aryan languages adjectives exhibit fewer case distinctions than nouns and in some, including Bengali (Klaiman 1987: 499), they exhibit no concord at all. In Germanic languages that exhibit a case system it is noticeable that in noun phrases with a determiner, an adjective, and a noun it is the determiner that displays the maximum amount of differentiation. A paradigm of German nouns is given in Tab. 102.4. It shows a masculine singular noun of the strong declension with an adjective and determiner.

nominative	<i>der liebe Mann</i>
accusative	<i>den lieben Mann</i>
genitive	<i>des lieben Mannes</i>
dative	<i>dem lieben Mann(e)</i>

Tab.102.4: German case inflection

In Maba (Nilo-Saharan) a nominative-accusative distinction is made by tone only on the article: *àmárà-gù* ‘lion-ART.NOM’ versus *àmárà-gú* ‘lion-ART.ACC’ (Tucker & Bryan 1966: 199).

In Georgian and Hurrian concord extends from the head of a noun phrase to dependent noun phrases bearing adnominal case marking (Plank 1990). The effect of this is to produce double case marking. (7) is from Old Georgian (Mel'čuk 1986: 69).

- (7) *sarel-itá man-isa-jta*  
name-INSTR father-GEN-INSTR  
'with father's name'

Double case marking is not uncommon in Australia. It is found not only in languages with concord but in languages that mark case phrase-finally. In (8), which is from Alyawarra, *-kinh* marks the relationship of *artwa ampu* to *ayliyla*, and *-ila* marks the relationship of the phrase as a whole (Yallop 1977: 117f.).

- (8) *ayliyla artwa ampu-kinh-ila*  
boomerang man old-GEN-INSTR  
'with the old man's boomerang'

### 3.3. Within the subordinate clause

#### 3.3.1. Internal relations

In general case marking in a subordinate clause is the same as in an independent clause, particularly if the clause is finite. However, where a subordinate verb is nominalised it may take complements in the favourite adnominal case, i.e. the genitive. In Latin, for instance, the present participle may take a genitive complement corresponding to what would have been an accusative (or dative, ablative or even genitive) complement with the verb: *appetentes gloriae* ‘seek:PRES.PART.NOM.PL glory:GEN (seeking glory)’ but *appetunt gloriam* ‘seek:3.PL glory:ACC (they seek glory)’.

In Turkish the genitive is used to mark the subject of a nominalised verb. The object of such a verb if present takes the normal case marking (Kornfilt 1987: 640).

- (9) *Ahmed-i ben-i sev-diğ-in-i*  
Ahmed-GEN 1.SG-ACC love-NR-3.SG-ACC  
*bil-iyor-um.*  
know-PRES-1.SG  
'I know that Ahmed loves me.'

The form *-in* is a third singular genitive form in cross-reference with *Ahmed-i*. In Turkish noun possessors are cross-referenced on possessed nouns: *Biz-im heykel-imiz* ‘we-GEN statue-1.PL.GEN (our statue)’. The accusative on the nominalised verb marks it as the complement of *biliyorum*, and the accusative on *ben* marks it as the complement of *sevmek*.

#### 3.3.2. External relations

Where a clause rather than a noun phrase is a dependent, the same possibilities for case marking arise. In most instances the case marker appears only on the head of the clause, namely the verb, as in the following example from Finnish (Uralic) where the translative case (which expresses the notion of ‘to’ with nouns) is found on the infinitive and indicates purpose (Branch 1987: 615).

- (10) *Osti-n karttakirja-n*  
bought-1.SG atlas-ACC  
*suunnitella-kse-ni automatka-n.*  
plan-TRNSL-1.SG.POSS car.trip-ACC  
'I bought an atlas in order to plan a car journey.'

Another possibility is for the case marking of a dependent verb to spread to its dependents by concord. The following example is from Yukulta (Northern Australian). Note that the dative which is appropriate to the verb *warratj-* spreads to the allative-marked complement to yield a second layer of case marking (Keen 1972: 270).

- (11) *Taamitya-ngandi*  
ask-SBJ.1.SG.&OBJ.3.SG.FUT.AUX  
*natha-rul-ngkurlu warratj-urlu.*  
camp-ALL-DAT go-DAT  
'I'll ask him to come to the camp.'

This spreading may also occur where a clause modifies a noun phrase. In the following example from the Pama-Nyungan language Panyjima all words in the modifying clause show concord with the accusative-marked object of the main clause (Dench & Evans 1988: 28).

- (12) *Ngatha wiya-rna ngunha-yu marlpa-yu*  
I:NOM see-PAST that-ACC man-ACC  
*paka-lalha-ku nharniwalk-ku*  
come-PFV-ACC hither-ACC  
*warrungkamu-la-ku.*  
morning-LOC-ACC

'I saw that man who came this way this morning.'

Concord in dependent clauses is not nearly so common as concord in noun phrases; the former implies the latter.

### 3.4. Within the word

Case is most often realised via suffixes. The almost total absence of prefixes is striking (cf. Cutler et al. 1985). Squamish (Salishan) is reported to have case prefixes, and Coos and Siuslaw (Penutian) have a subject prefix in addition to cases marked by suffix and vowel alternation (Sherzer 1976: 69 f., 261 f.). Another example is to be found in Nungali, a Northern language of Australia (Hoddinott & Kofod 1976: 397). In this language there are noun classes. Prefixes mark case and class simultaneously (Tab. 102.5).

class	absolutive	oblique
1	<i>di-gal</i>	<i>nyi-gal</i>
2	<i>nya-ngarrung</i>	<i>ganyi-ngarrung</i>
3	<i>nu-ngulud</i>	<i>nyu-ngulud</i>
4	<i>ma-yadayn</i>	<i>nyi-yadayn</i>

Tab. 102.5: Nungali class/case prefixes

In some languages the case marking is not linearly separable from the stem. Consider the umlauted datives of Germanic, for instance. In Old English the dative of *fōt* 'foot' is *fēt* which derives from \**fōti* via *fāti* and *fāt*, the original *-i* suffix having induced an assimilatory fronting in the stem vowel before being lost. In Serbo-Croat the genitive/dative and locative in i-stems are distinguished by tone: *stvāri* 'thing(GEN/DAT)' versus *stvári* 'thing(LOC)' (Corbett 1987: 398 f.). In Russian the genitive singular and nominative plural are distinguished in some instances by stress: *ruki* 'hand.GEN.SG' versus *rúki* 'hand.NOM.PL' (Comrie 1987: 333). The ultimate union of stem and case is suppletion as evidenced, for example, in the pronouns of Indo-European languages.

In some languages there is what is called an **oblique stem**. In Tamil, for instance, singular nouns have a stem distinct from the nominative to which all oblique case marking is added. *Maram* 'tree', for example, has a nominative *maram* and an oblique stem *maratt-* to which the case markers are suffixed: *maratt-ai* 'tree:OBL-ACC', *maratt-ukku* 'tree: OBL-DAT', etc. (Steever 1987: 737). Sometimes

the differentiation is the result of phonological processes as in Ancient Greek where the oblique stem is original and the nominative results from cluster simplification, e.g. *elpis* 'hope:NOM' has an oblique stem *elpid-* as in *elpid-os* 'hope:OBL-GEN', *elpid-i* 'hope:OBL-DAT'. In the nominative *d* has been lost before *s*. In other languages the oblique stem consists of the base plus a non-functioning case marker. In Lezgian the ergative of *lam* 'donkey' is *lamra* which form serves as a stem for the genitive *lamran*, dative *lamraz*, etc. (Melčuk 1986: 63).

Another category of inflection found frequently on nouns is number. The marking for number, like case marking, is usually via suffixation, and the number suffix normally precedes the case suffix as in Tamil: *āRu-kał-ai* 'river-PL-ACC (rivers)' (Steever 1987: 737).

Where pronominal possessors are marked on the noun these usually appear before the case marking as in Turkish where they appear between the number marking and the case marking; similarly in Hungarian: *hajó-i-m-on* 'ship-PL-1.SG.POSS-LOC (on my ships)' (Abondolo 1987: 584). In the Balto-Finnic languages, however, the possessor marking usually follows the case marking. In Finnish, for instance, we find: *kirkko-lla-mme* 'church-ADESS-1.PL.POSS (at our church)' (Branch 1987: 610). The **adessive** case expresses the sense of 'near' or 'at'.

Another category sometimes marked on nouns, and often on their dependents, is gender or class. Where there is a separable suffix as in Dravidian or Semitic, this usually appears before the case marking as in Arabic: *mudarris-at-u-n* 'teacher-F-NOM-INDEF (a female teacher)' (cf. Kaye 1987: 672). Dyirbal (Pama-Nyungan) is exceptional in marking noun class by post-case suffixation on determiners: *ba-gu-l* 'that-DAT-M' versus *ba-gu-n* 'that-DAT-F' (Dixon 1972: 44).

## 4. Survey of case marking

In describing types of case marking it is useful to separate those cases that primarily mark the core or nuclear grammatical relations (subject and direct object) from the others. The former are sometimes called **direct** cases and the latter **oblique**, but this use of **oblique** is confusing since, as noted in 2, the term **oblique** is traditionally used to mean non-nominative. It would seem preferable to speak of **direct** and **indirect** cases. Kuryłowicz

(1949:40; 1964:179) made a distinction between grammatical and concrete cases. With the former he held the syntactic function to be primary and with the latter the semantic function. Among the grammatical cases he included the genitive (see 3.2.1 and 4.2.2). It is difficult to maintain a strict distinction between grammatical and concrete cases (Nichols 1983), but there are useful generalisations to be made about the marking of direct cases as opposed to indirect ones, and one can make further generalisations about a subgroup of indirect cases, namely the **local** cases. These are the cases that express location, destination, relative position ('above', 'inside'), etc.

#### 4.1. Direct cases

Languages differ in the way they use case marking to distinguish the agent and patient complements of predicates. **Agent** in this context covers not only the semantic role of agent but also other roles that are treated in the same way syntactically, similarly with **patient**. In the majority system, which is exemplified by Latin, the agent of a two-place transitive verb is encoded in the same way as the sole argument of a one-place predicate, namely in the nominative case. On the other hand the patient of a transitive verb is encoded in the accusative case. This system has come to be called an **accusative system** and languages with this system are called **accusative languages** (cf. Dixon 1972; 1994).

In many languages this pattern of marking does not hold if the patient is indefinite or non-specific. In these circumstances the patient remains unmarked. As noted in 1 this is the situation in Turkish.

- (13) *Bir adam gör-dü-m.*  
a man see-PAST-1.SG  
'I saw a man.'
- (14) *Adam-i gör-dü-m.*  
man-ACC see-PAST-1.SG  
'I saw the man.'

There is another case treatment of agent and patient which is not so common but quite widespread. In this system the patient of a transitive verb is treated in the same way as the sole argument of a one-place predicate. The case covering these functions is usually the zero case and is referred to by some as the nominative and by others as the **absolutive**. The agent of a two-place transitive verb is encoded in what is usually called the **erga-**

**tive case**. This system is referred to as an **ergative system** and languages with this system are referred to as **ergative languages** (Comrie 1978; Dixon 1994; Plank 1979; Givón 1984: 151). The Caucasian language Avar provides an example (Ebeling 1966: 77).

- (15) *W-as w-ekér-ula.*  
M-child(NOM) M-run-PRES  
'The boy runs.'
- (16) *Inssu-cca j-as*  
(M)father-ERG F-child(NOM)  
*j-écc-ula.*  
F-praise-PRES  
'Father praises the girl.'

In Avar, as in ergative languages generally, the nominative (or absolutive) is unmarked. The case marker glossed as "ERG(ATIVE)" is the one that also indicates instruments and could be labelled more accurately "ergative-instrumental". It is common to find that in ergative languages the so called ergative marks a peripheral (adjunct) function such as genitive, locative or instrumental as well as agent of a transitive verb. The *w-* in (15) is a masculine class marking prefix and the *j-* in (16) a feminine class marker. Note that the verb cross-references the absolute.

In some languages case marking is sensitive to the relative position of the agent and patient on a hierarchy. In Fore (Papuan), for instance, the use of the ergative case is determined by the following hierarchy: pronoun, personal name, kin term > human > animate > inanimate (cf. Tab. 102.6). If an entity higher on the hierarchy acts on an entity lower on the hierarchy, no case marking is used, but if a lower entity acts on a higher, the lower appears in the ergative. So if a man kills a pig (higher acting on lower), then no ergative is used, but if a pig kills a man (lower acting on higher) then the ergative marker *-wama* is used (Scott 1978: 100–117).

- (17) *Yagaa wá aegúye.*  
pig man 3.SG:hit:3.SG  
'The man kills the pig.'
- (18) *Yagaa-wama wá aegúye.*  
pig-ERG man 3.SG:hit:3.SG  
'The pig kills the man.'

In languages with ergative case marking on nouns it is true more often than not that the ergative marking is lacking from first and second person pronouns and sometimes from third. This latter situation is found in Yup'ik Eskimo, for instance, where nouns exhibit

	1, 2, 3	kinship personal names	human	animate	inanimate
ergative					
accusative					

Tab. 102.6: Case marking hierarchy

nominative-ergative marking but pronouns remain unmarked for all core functions (Reed et al. 1977). In Kiranti and Gyarong (both Tibeto-Burman) first and second person pronouns lack ergative marking, but all third persons exhibit it (Delancey 1987: 806).

A few languages, perhaps only a score or so, organise their core grammar on an agent versus patient basis irrespective of transitivity. These have been called **active languages** (Klimov 1973) or **split-S languages** (where S stands for intransitive subject) (Dixon 1994). In these languages agent complements of one-place verbs are marked in the same way as agent complements of two-place verbs while the patient complements of one-place verbs are marked like the patients of two-place verbs. Such languages include the American Indian languages Dakota, Arikara, and Seneca, where the marking shows up only in pronominal affixes, and the Austronesian language Acehnese, where agent clitics precede the verb and patient clitics follow. The agent in a transitive clause is marked by a preposition *lē* (Durie 1985: 180–195). In the Kartvelian languages of the Caucasus the active pattern is to be found in the case marking of Georgian and Laz (Harris 1981; 1985). The following examples are from Laz (Harris 1985: 52 f.).

- (19) *Bere-k imgars.*  
child-ERG 3.SG:cry  
'The child cries.'
- (20) *Bere oxori-s doskidu.*  
child house-DAT 3.SG:stay  
'The child stayed in the house.'
- (21) *Baba-k meçcaps skiri-s exeni.*  
father-ERG SBJ.3.SG:give:OBJ.3.SG&IO.3.SG  
child-DAT horse  
'The father gives a horse to his child.'

In Georgian this pattern applies only to certain classes of verbs in the aorist tense group. In the present tense all subjects are in the no-

nominative case and the direct object in the dative (Harris 1981: 46).

Some languages employ both ergative and accusative case. This is the situation in Nez Perce though first and second persons lack ergative marking. The following examples illustrate the nominative-ergative-accusative opposition (Rude 1985: 83, 228).

- (22) *Hi-páay-na háama.*  
3.SG-arrive-PFV man  
'The man arrived.'
- (23) *Híamap-im 'áyato-na pée-'nehne-ne.*  
man-ERG woman-ACC  
3.SG&3.SG-take-PFV  
'The man took the woman away.'

Interpretations of the distribution of core case marking have been popular since the mid 1970s. Silverstein (1976) pointed out that ergative and accusative marking show regularities of distribution with respect to the following hierarchy (1 stands for first person, 2 for second, and 3 for third).

Accusative marking may be restricted to first and second person pronouns, to all pronouns, to pronouns and personal names, and so on. If it appears at any point on the hierarchy, it appears on all points to the left. Conversely ergative marking may appear on inanimates (as in the northern Australian language Mangarayi; cf. Merlan 1982: 57), on all nouns, all nouns and third person pronouns, or all nominals. If ergative marking appears at any point on the scale, it appears on all points to the right.

It should be noted that the three categories at the top (left) of the hierarchy are inherently definite and that in quite a number of languages patients in the three categories at the lower end of the hierarchy are marked when in direct object function only if definite or at least specific (cf. (13), (14)). Another point is that the same marking is often used for direct objects as is used for indirect objects. Spanish provides a well known example

*deseo un empleado* ‘want:1.SG an employee’ means ‘I want an employee’, any one will do. However, *deseo a un empleado* with the patient marked by the preposition *a* ‘to’, which also marks indirect objects, means ‘I want a (specific) employee’. In Spanish specific animate objects are marked by *a*.

Comrie (1978: 384–388) and Dixon (1979: 69; 1994) emphasise the discriminatory function of direct case marking. With a two-place verb only one argument needs to be marked and with one-place verbs no marking is needed. Languages with both ergative and accusative marking usually employ them in a complementary way. Silverstein, Comrie and Dixon see motivation for the economy here in that agent marking (ergative) is restricted to nominals for which the agent role is less expected and patient marking (accusative) to nominals for which the patient role is less expected (but see Silverstein 1981: 241).

Mallinson & Blake (1981: 86–91) draw attention to the fact that the hierarchy (Tab. 102.6) above relates not just to the markedness of a nominal to appear as agent or patient, but to topic propensity. Entities to the left of the hierarchy are unmarked choices for topic whereas entities to the right, particularly inanimates, are marked choices. Mallinson & Blake (1981: 103–106) endorse the view of Wierzbicka (1981: 66–68) that first and second person pronouns tend to avoid ergative marking because of their propensity to be topic. An ergative marked noun phrase is not an ideal means of expression for a topic because of the presence of agent marking. Topics are normally presented without semantic marking. Interestingly ergative marked noun phrases are uncommon in discourse. A transitive subject is very frequently given rather than new and tends to be represented by zero anaphora, clitic pronouns or

affixes, or by first or second person pronouns which more often than not lack ergative marking.

#### 4.2. Indirect cases

##### 4.2.1. Local cases

The term **local** in this context refers to ‘place’. Local cases express notions of location (‘at’), destination (‘to’), source (‘from’), and path ‘through’. Usually at least two different cases are employed to make local distinctions, though not necessarily two exclusively local cases. Indo-European languages originally used two local cases, locative and ablative, with the accusative expressing destination and path. In Turkish there is a locative and an ablative, with destination expressed via the dative. A case expressing destination is found in a number of languages including Balto-Finnic and Altaic languages. A separate case for path is not so common but such a case is found in a few Australian languages (Blake 1987: 40), and is part of the local case system in Avar (Tab. 102.7).

Where languages have large case systems it is always through the elaboration of the local cases. This is a feature of the Finno-Ugric branch of the Uralic family and of the Northeast Caucasian languages. It arises from a combination of markers for relative orientation (‘above’, ‘beside’, etc.) with markers for location, destination, source, and path. One dialect of Tabassaran (Northeast Caucasian) is reported to have 53 cases (Comrie 1981: 209). In Avar (also Northeast Caucasian) there are 27 cases, including 20 local cases deriving from a combination of four ‘cases’ and five orientation markers. The combinations of orientation markers and ‘case’ markers are morphologically agglutinative and semantically transparent. In a language like this one could simply recognize a

	locative (location)	allative (destination)	ablative (source)	perlative (path)
on (top) of	-da	-d-e	-da-ssa	-da-ssa-n
at	-q	-q-e	-q-a	-q-a-n
under	-x'	-x'-e	-x'-a	-x'-a-n
in, among	-x	-x-e	-x-a	-x-a-n
in a hollow object	-Ø	-Ø-e	-Ø-ssa	-Ø-ssa-n

Tab. 102.7: Avar local cases

layer of orientation markers plus a layer of ‘case’ markers. Tab. 102.7 is based on Ebeling (1966). If one analyses a further outermost layer of forms as case markers rather than as clitic particles, the system is even larger (Hjelmslev 1937: 2–25).

The method of labelling local cases builds on the model of the Latin *ablātīvus* ‘ablative’ which is made up of the preposition *ab* ‘from’ plus a stem *lātīvus* the root of which is *lāt*, the suppletive root of the perfect participle of *ferō* ‘I bear’. Other labels are formed by varying the preposition or by combining prepositions with the stem *essīvus* from *esse* ‘to be’ (cf. Mel’čuk 1986: 72–75).

<i>ad</i> ‘to’	allative: <i>to(wards)</i> ( <i>the exterior of</i> )
<i>in</i> ‘into’	illative: <i>into</i>
<i>ab</i> ‘from’	ablative: <i>from (the exterior of)</i>
<i>e(x)</i> ‘out of’	elative: <i>from (the inside of)</i>
<i>super</i> ‘above’	superlative: <i>on, above</i>
<i>trans</i> ‘through’	translative: <i>through</i>
<i>per</i> ‘through’	perative: <i>through, along</i>
<i>in</i> ‘in’	essive: <i>at</i> (cf. locative)
<i>ad</i> ‘to’	inessive: <i>in(side)</i> adessive: <i>at</i>

Tab. 102.8: Local cases

More often than not there is some fusion and/or semantic shift that obscures the transparency of the system. In Finnish, for instance, there are 15 cases of which nine are local. As a first approximation one can say that there are three local cases for location, destination, and source, which can appear on their own or appear with interior (‘inside’) and exterior markers (‘outside’). These are displayed in Tab. 102.9.

As can be seen, the system is only partly agglutinative. Moreover, the meanings can-

not be derived from a combination of the labels for the columns and rows. The *-tta* form, for instance, has developed a number of functions including a partitive one (cf. (26), (27) below).

#### 4.2.2. Other indirect cases

In Latin there are three indirect cases: the ablative, the genitive, and the dative. As noted in 2 the ablative represents a syncretisation of three cases in Indo-European: two local cases (ablative proper and locative (see 4.2.1)) and instrumental. All these cases have analogues in a variety of languages and counterparts naturally tend to be named on the basis of Indo-European models.

In Latin the dative marks the indirect object of most three-place verbs including *dāre* ‘to give’ whence it derives its name: *Cassius Brūtō librum dat* ‘Cassius:NOM Brutus:DAT book:ACC give:3.SG (Cassius gives a book to Brutus)’. It also marks the complement of a score or so of two-place verbs such as *auxiliōr* ‘I help’, *ignōscō* ‘I pardon’, and *placeō* ‘I please’. All these verbs are non-activity, non-impingement verbs, though not all such verbs take a dative complement; *dēlectō* ‘I please’, for instance, takes an accusative complement. A case with this kind of distribution and function is found in numerous languages including other Indo-European languages, in Turkish and other Altaic languages, in Caucasian languages, and in many Australian languages. In some there is an alternation between transitive verbs and derived intransitive verbs with a dative or other indirect complement. The derived intransitive is used where there is reduced semantic transitivity along one or more parameters (cf. Hopper & Thompson 1980). In Kalkatungu (Pama-Nyungan), for instance, the agent of a transitive verb is in the ergative and the patient in the

	location	source	destination
general	<i>-na</i>	<i>-tta</i>	<i>-ksi</i>
	essive	partitive	translative
interior	<i>-ssa</i>	<i>-sta</i>	<i>-(h)Vn</i>
	inessive	elative	illative
	‘in’	‘from(inside)’	‘into’
exterior	<i>-lla</i>	<i>-lta</i>	<i>-lle</i>
	adessive	ablative	allative
	‘at’	‘from(outside)’	‘to(wards)’

Tab. 102.9: Finnish local cases

nominative (24). A handful of two-place verbs such as *yakapi* ‘to understand’ and *uth-arrā* ‘to wait for’ take a nominative actor and a dative patient/theme. However, for any transitive verb a derived intransitive alternative is possible under certain conditions. In the imperfective aspect, for instance, the agent is in the nominative and the patient in the dative (25) (cf. Blake 1979: 44 f., 55).

- (24) *Nyin-ti lhai yurru.*  
you-ERG hit man:NOM  
'You hit the man.'
- (25) *Nyini lhai-minha-n yurr-ku.*  
you:NOM hit-IPFV-2.SG man-DAT  
'You are hitting the man.'

In eastern Europe it is common to find that partly affected patients are put in the **partitive** case as in Hungarian (Moravcsik 1978: 261).

- (26) *Olvasta a könyvet.*  
read:3.SG the book:ACC  
'He read the book.'
- (27) *Olvasott a könyvböl.*  
read:3.SG the book:PRTV  
'He read some of the book.'

In Estonian and Finnish the partitive is used for the patient if it represents part of a whole or an indefinite quantity, or if the action is incomplete, or if the polarity of the clause is negative. In Polish and Russian similar conditions determine the choice of genitive as opposed to accusative (Moravcsik 1978: 265, 269; see also Plank 1984, ed.).

In Latin the dative is used to encode the agent with the gerundive form of the verb as in *pax mihi petenda est* ‘peace:NOM me:DAT seek:GER be:3.SG (peace is to be sought by me)’. The dative is used under certain conditions to mark the agent in a range of other languages including Icelandic and several Indic languages. In Georgian an agent appears in the dative in the evidential mode (Harris 1981: 117–127, 247–249). These dative-marked agents often display some syntactic properties of subjects and are sometimes referred to as **indirect subjects** (cf. Klaiman 1987: 507 on Bengali). Some Northeast Caucasian languages have an **affective** case exclusive to this function (Comrie 1981: 223).

A genitive case is widespread. On the basis of Latin one would normally ascribe the label genitive to the most common or unmarked adnominal case, although one would not expect such a case to be exclusively adnominal. In Latin the genitive marks the complement

of a handful of verbs such as *oblivisci* ‘to forget’ and *miserērī* ‘to pity’. The unmarked adnominal case normally covers the sense of possessor, and the label **possessive** case is a common alternative. A genitive is found not only in Indo-European languages, but also in Uralic, Caucasian, Altaic, Dravidian languages, and Semitic. In Australian languages a single case covers the range of the Latin dative and genitive cases and the label dative is preferred. In Australian languages the dative often expresses purpose as in ‘he went for fish’ or a beneficiary as in ‘he got the medicine for father’. However, in some of these languages there is a separate case labelled purposive or benefactive to cover one or both of these senses. Basque has both a purposive and a benefactive as well as a dative (Saltarelli et al. 1988: 156–166).

A large number of languages have possessive affixes for first, second, and third person. In many of these the possessive affix cross-references the possessor. An example from Turkish was given in 1, namely *adam-in ev-i* ‘man-GEN house-3.SG.POSS (the man’s house)’. Here of course there is a genitive, but in many languages the cross-referencing is the sole means of marking possession.

A variety of languages including Basque, Zoque, Ossete (Indo-Iranian), and Finnish have a **comitative** case expressing accompaniment.

An **instrumental** case encodes the instrument with which an action is carried out as in ‘she wiped the screen with a cloth’. It is found in the Indo-European family (e.g. Sanskrit), Uralic (e.g. Komi-Permyak), Altaic (e.g. Nanai), Dravidian (e.g. Tamil), Basque, Tonkawa (Uto-Aztec), Tarascan, Australian, and elsewhere. It is sometimes used to encode the agent of the passive in an accusative language (e.g. Russian) and the agent of a transitive verb in an ergative language (as in Avar, Tibetan, and the majority of Pama-Nyungan languages).

In the Uralic languages a case called **abessive** (Latin *ab-esse* ‘to be absent’) or **privative** (Latin *prīvare* ‘to deprive’) is found. It means ‘lacking’, ‘not having’. In Finnish, for instance, *rahta-tta* is ‘money-PRIV (moneyless)’. This case is also found in Australia where it is matched by a ‘having’ case called the **con-comitant** or **proprietary**. These two categories tend to show up in many Australian languages in lexicalised formations. In Kalkatungu, for instance, *putu-yan* ‘stomach-PROPR’ means ‘pregnant’. In light of this and

other evidence the ‘having’ and ‘lacking’ suffixes have been taken to be derivational. On the other hand they exhibit concord in languages with case concord which suggests they are inflectional and should be considered cases (Blake 1987: 87f.; Dench & Evans 1988: 7–13).

**Aversive** (lit. turning from), **evitative** (avoiding), and **causal** have been used for a case category common in Australian languages. It indicates what is to be feared or avoided as in Kalkatungu: *yanyi-ngungu rumpi* ‘ghost-AVERS fear (He is afraid of ghosts)’.

Where a case covers a variety of indirect functions it is often labelled **oblique** (see 2 and 4). Yaqui, for example, has only two cases: nominative and oblique. A parallel system occurred in Old French.

In the traditional description of Ancient Greek and Latin a **vocative** case appears (Tab. 102.2). The vocative is used as a form of address. In Latin, for instance, *domine* is the form used to address a *dominus* as in *Quōd vādis, domine?* ‘whither go:2.SG lord:voc (Where are you going, master?)’. Vocatives do not enter into constructions, but rather stand outside them or are inserted parenthetically. They are unlike other cases in that they do not mark the relation of dependents to heads. For these reasons vocatives have not always been considered cases (Hjelmslev 1935: 4). In Ancient Greek and Latin the vocative’s claim to being a case is structural. The vocative is a word-final suffix like the recognised case suffixes. However, modified forms of nouns used as forms of address also occur in languages that do not have case inflection.

## 5. Diachrony of case

### 5.1. Origin of case marking

In many areas case marking is of long ancestry as witnessed by formal complexity (suppletive variants and fusion) and semantic complexity. This is true of case marking in Indo-European languages, for instance, and in such a situation the prospects for determining the origin of particular case markers are not too promising. On the other hand there are some languages including the Indo-Iranian branch of Indo-European where the development of ‘new’ case markers is attested. Moreover, there are many languages in which we can plausibly reconstruct the origin of postpositions and phrase-final postpo-

sition-like case markers, and there is evidence pointing to the fact that the latter derive from the former.

There are two common origins for adpositions and case markers, one nominal and the other verbal (cf. Givón 1984: 228–232). Locative markers can often be shown to derive from a noun. The Sinhalese dative *-ṭ* and locative *-ge* come from Sanskrit *artha* ‘aim, wealth’ and *grha* ‘house, place, town’ respectively, and the locative suffix in Sora derives from *ba* ‘place’ (Starosta 1985: 111, 122). The Persian suffix *-ra*, used for marking specific objects, derives from a noun *rādi* ‘aim, goal’ (Windfuhr 1987: 541). With adpositions and inflectional case markers of nominal origin it appears to be often the case that the original noun is a head with an adnominal dependent. The comitative in Karelian, for instance, is marked by *-ke* as in *velle-ŋ-ke* ‘with the brother’ where *-ke* derives from *\*kerða-lla* ‘at a turn, on time of’ used with genitive attributes as in Finnish *koira-n keralla* ‘dog-GEN with (with the dog)’. The genitive *-n* is reflected as *-ŋ* in *velleŋke* (Anttila 1972: 149). In developments like this the original head (*\*kerða-lla* in this instance) eventually becomes an affix and the original dependent a stem.

Verbs, particularly in their non-finite forms, frequently give rise to adpositions as with the English preposition *regarding* which obviously derives from the present participle of *regard*. In Sanskrit the past participle of *kr* ‘to do, make’ gave rise to a postposition governing the genitive: *mama kṛte* ‘me:GEN for (for me)’ (Andersen 1979: 26).

Within systems of case markers there is frequent development from concrete to grammatical. The Latin preposition *ad* ‘to’ develops into Spanish *a* which not only marks indirect objects, but also specific, animate objects: *vi a Juan* ‘I saw John’ (cf. 4.1 above). Similarly another Latin preposition *per* ‘through’ develops into Rumanian *pe* which means ‘on’ but which also marks direct objects that are animate or pronominal (Mallinson 1987: 315f.). Within inflectional case marking systems one notes that often markers of grammatical cases are shorter than markers of concrete cases, a sign of a longer span of development taking in one extra stage (Lehmann 1982: 85). There are also numerous instances of the same marker having concrete and syntactic functions. For instance, the accusative in Latin covers the local function of destination as well as that of

direct object. Given that in attested instances the direction of development is regularly from concrete to abstract one would expect the local function to have developed into the grammatical. In Nez Perce the ergative marker *-nim* can be traced to a local form meaning 'hither' (Rude 1988: 1).

### 5.2. Loss of case marking

Case systems can wither away, their functions taken over by other grammatical mechanisms. Proto-Indo-European is reconstructed as having eight cases but most of its daughter languages have fewer: Latin six (cf. 2), Ancient Greek five and also Old English (if we count the vestigial instrumental separately from the dative), Modern German four. Rumanian has three cases, nominative/accusative, genitive/dative, and vocative (Mallinson 1987: 310–312), but the Western Romance languages have lost the entire Indo-European case system save for distinct pronominal forms, mostly clitic, for subject, object, and indirect object. In Western Europe subject-verb-object word order has taken over the main function of nominative and accusative, namely to mark subject and object, and prepositions have supplanted the non-core cases.

In some languages, particularly in the Americas and in northern Australia, two or sometimes three arguments of the verb are represented by clitic forms or by inflection on the verb. The forms involved are often identifiable diachronically if not synchronically with case marked pronouns. In many of these languages there is no case marking on core noun phrases and one would assume that case marking has been lost from these noun phrases (cf. Blake 1988: 31–37).

### 5.3. Developments of case markers

Whether a case system withers away or not certain markers are likely to develop new functions. It is common to find case marking on adverbs as in English *homewards* or the parallel German formation *heimwärts* where the *-s* is of genitive origin. Dependent verbs are often marked for case (as illustrated in (9) and (10)) and such case-marked verbs often develop new functions. In Latin the active infinitive derives from a locative-marked form (e.g. *age-re* 'to do' < \**-si* 'LOC') and the passive infinitive derives from a dative-marked form (e.g. *ag-i* 'to be done' < \**-ei* 'DAT') (Coleman 1985: 308). In several Australian languages dative-marked forms serve as sub-

ordinate purposive forms and as independent purposive or desiderative forms. The development seems to be from noun + dative to dependent verb + dative to independent verb with the dative reanalysed. The three stages can be illustrated from Kalkatungu (cf. Blake 1979: 56). In (28) there is a noun in the dative expressing purpose. In (29) a nominalised verb taking the dative expresses purpose. In (30) a form consisting of a verb stem, a nominaliser, and a dative serves as a main verb.

- (28) *Ngai ingka-nha natha-aya.*  
I:NOM go-PAST nurse-DAT  
'I went for the nurse.'
- (29) *Ngai ingka-nha natha-ngku*  
I:NOM go-PAST nurse-ERG  
*nanyi-nytya-aya.*  
see-NR-DAT  
'I went so the nurse could have a look [at me].'
- (30) *Natha-ngku ngai nanyi-nytyaaya.*  
nurse-ERG I:NOM see-PURP  
'The nurse is going to have a look at me.'

The use of an originally subordinate verb form as a main verb probably arises from the ellipsis of semantically weak governing verbs with meanings like 'go' (cf. the modal-like force of *go* in English *I'm going to wash*). Where this ellipsis occurs in a language that exhibits concord from the verb to the other constituents of a dependent clause, the result is marking covering the whole verb phrase. An example of a language with such concord was given above from Yukulta (11), a Tangkic language of northern Australia. The following example is from another Tangkic language called Kayardild (cf. Dench & Evans 1988: 24; Evans 1985: 332–385).

- (31) *Nyingka kurri-nang-ku niwan-ju*  
you:NOM see-NEG-FUT 3.SG-FUT  
*balmbi-wu.*  
tomorrow-FUT  
'You will not see her tomorrow.'

The set of forms *-ku* ~ *-ju* ~ *-wu* has dative-like functions and a proprietive or 'having' function when used as a normal case marker. In examples like (31) it has what Evans calls a **modal** function, and it has an unusual distribution in that it is found on the verb and all dependents within the verb phrase.

In some languages there is formal identity of case markers and switch reference markers. This identity can be seen in the

Muskogean, Yuman, and Uto-Aztecán languages (Jacobsen 1983: 151) as well as in Australian languages (Austin 1981b). The following examples are from the Yuman language Diegueño (Langdon 1970: 150–154). In (32) -(v)c marks the subject; in (33) it marks the fact that the unexpressed subject is the same as the subject of its sister clause and is glossed “S(AME)S(UBJECT)”.

- (32) *Sin<sup>v</sup>-c      ?əcwəyu·w-m*  
       woman-NOM 3.SG:sing-DS  
       ?əyip-s.  
       SBJ.1.SG&OBJ.3.SG:hear-ASRT  
       ‘I heard the woman sing.’
- (33) *?amp      n<sup>v</sup>a-ta?am-c*  
       1.SG:walk when-1.SG:be.around-ss  
       ?əwu·w-s.  
       SBJ.1.SG&OBJ.3.SG:saw-ASRT  
       ‘As I was walking, I saw him.’

The suffix -(v)m marks direction away from the point of reference and toward the object as in ?əwa·vəm ‘away to the house’ as well as instrumental and accompaniment. It is also used to indicate that a verb has a different subject from that of a sister clause. This is illustrated in (32) where it is glossed “D(IF-FERENT)S(UBJECT)”. There is general agreement that same-subject and different-subject markers derive from case marking forms.

For further information on case consult Blake (2002).

## 6. References

- Abondolo, Daniel (1987), “Hungarian”. In: Comrie (1987, ed.), 577–592
- Andersen, Paul Kent (1979), “Word Order Typology and Prepositions in Old Indic”. In: Brogyanyi, Bella (ed.), *Studies in Diachronic, Synchronic and Typological Linguistics, Vol. II*. Amsterdam: Benjamins, 23–34
- Anttila, Raimo (1972), *An Introduction to Historical and Comparative Linguistics*. New York: Macmillan
- Austin, Peter (1981a), *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press
- Austin, Peter (1981b), “Switch-reference in Australia”. *Language* 57, 309–334
- Blake, B[arry] J. (1979), *A Kalkatungu Grammar*. Canberra: Pacific Linguistics
- Blake, Barry J. (1987), *Australian Aboriginal Grammar*. London: Croom Helm
- Blake, Barry J. (1988), “Redefining Pama-Nyungan: Towards the Prehistory of Australian Languages”. In: Evans, Nicholas & Johnson, Steve (eds.), *Aboriginal Linguistics I*. Armidale/NSW: University of New England, Department of Linguistics, 1–90
- Blake, Barry J. (2002), *Case*. Cambridge: Cambridge University Press
- Branch, Michael (1987), “Finnish”. In: Comrie (1987, ed.), 593–617
- Brecht, Richard D. & Levine, James S. (1986, eds.), *Case in Slavic*. Columbus, Ohio: Slavica
- Carvalho, Paulo de (1980), “Réflexions sur les cas: Vers une théorie des cas latins”. *L'information grammaticale* 7, 3–11
- Coleman, R[obert] G. G. (1985), “The Indo-European Origins and Latin Development of the Accusative with Infinitive Construction”. In: Touratier, Christian (ed.), *Syntaxe et Latin, Actes du IIème Congrès International de Linguistique Latine, Aix-en-Provence, 28–31 Mars 1983*. Aix-en-Provence: Université de Provence, 307–344
- Comrie, Bernard (1978), “Ergativity”. In: Lehmann, Winfred P. (ed.), *Syntactic Typology*. Sussex: The Harvester Press, 329–394
- Comrie, Bernard (1981), *The Languages of the Soviet Union*. Cambridge, London: Cambridge University Press
- Comrie, Bernard (1986), “On Delimiting Cases”. In: Brecht & Levine (1986, eds.), 86–105
- Comrie, Bernard (1987, ed.), *The World's Major Languages*. London, Sydney: Croom Helm, 322–328
- Comrie, Bernard (1987), “Russian”. In: Comrie (1987, ed.), 329–347
- Corbett, Greville (1987), “Serbo-Croat”. In: Comrie (1987, ed.), 391–409
- Crowley, Terry (1983), “Uradhi”. In: Dixon, R[obert] M. W. & Blake, B[arry] J. (eds.), *Handbook of Australian Languages, Vol. III*. Canberra: Australian National University Press, 306–428
- Cutler, Anne & Hawkins, John A. & Gilligan, Gary (1985), “The Suffixing Preference: A Processing Explanation”. *Linguistics* 23, 723–758
- Delancey, Scott (1987), “Sino-Tibetan Languages”. In: Comrie (1987, ed.), 797–810
- Dench, Alan & Evans, Nicholas (1988), “Multiple Case-Marking in Australian Languages”. *Australian Journal of Linguistics* 8.1, 1–47
- Derbyshire, Desmond C. & Pullum, Geoffrey K. (1986; 1989; 1991), *Handbook of Amazonian Languages, Vol. 1–3*. Berlin: Mouton de Gruyter

- Dixon, R[obert] M. W. (1972), *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press (Cambridge Studies in Linguistics 9)
- Dixon, R[obert] M. W. (1976, ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies and Atlantic Highlands New Jersey: Humanities Press
- Dixon, R[obert] M. W. (1979), "Ergativity". *Language* 55, 59–138
- Dixon, R[obert] M. W. (1994), *Ergativity*. Cambridge: Cambridge University Press
- Durie, Mark (1985), *A Grammar of Acehnese on the Basis of a Dialect of North Acehnese*. Dordrecht: Foris
- Ebeling, Carl L. (1966), "Review of Chikobava and Cercvadze's 'The Grammar of Literary Avar'". *Studia Caucasic 2*, 58–100
- Evans, Nicholas R. D. (1985), *Kayardild, the Language of the Bentinck Islanders of North West Queensland*. Ph.D. dissertation, The Australian National University (Canberra)
- Givón, Talmy (1984), *Syntax: A Functional-Typological Introduction, Vol. I*. Amsterdam, Philadelphia: Benjamins
- Goddard, Cliff (1982), "Case Systems and Case Marking in Australian Languages". *Australian Journal of Linguistics* 2, 167–196
- Harris, Alice C. (1981), *Georgian Syntax: A Study in Relational Grammar*. Cambridge: Cambridge University Press
- Harris, Alice C. (1985), *Diachronic Syntax*. Orlando: Academic Press (Syntax and Semantics 18)
- Hetzron, Robert (1987), "Semitic Languages". In: Comrie (1987, ed.), 654–663
- Hjelmslev, Louis (1935), *La catégorie des cas: Étude de grammaire générale, Vol. I*. Kobenhavn: Munksgaard (Acta Jutlandica: Aarsskrift for Aarhus Universitet 7.1)
- Hjelmslev, Louis (1937), *La catégorie des cas: Étude de grammaire générale, Vol. II*. Kobenhavn: Munksgaard (Acta Jutlandica: Aarsskrift for Aarhus Universitet 9.3)
- Hoddinott, William G. & Kofod, Frances M. (1976), "Djamindjungan". In: Dixon (ed.), 397–401
- Hopper, Paul & Thompson, Sandra A. (1980), "Transitivity in Grammar and Discourse". *Language* 56, 251–299
- Jacobsen, William Jr (1983), "Typological and Genetic Notes on Switch-reference Systems in North American Indian Languages". In: Haiman, John & Munro, Pamela (eds.), *Switch-Reference and Universal Grammar*, 151–183
- Jakobson, Roman (1936), "Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus". [reprinted in: Jakobson, Roman (1971), *Selected Writings II: Word and Language*. The Hague: Mouton, 23–71]
- Kaye, Alan S. (1987), "Arabic". In: Comrie (1987, ed.), 664–685
- Keen, Sandra L. (1972), *A Description of the Yuktula Language: An Australian Aboriginal Language of Northwest Queensland*. M.A. Thesis, Monash University (Melbourne)
- Klaiman, M[imi] H. (1987), "Bengali". In: Comrie (1987, ed.), 490–513
- Klimov, Georgij A. (1973), *Očerk obščej teorii erativnosti*. Moscow: Nauka
- Kornfilt, Jaklin (1987), "Turkish and the Turkic Languages". In: Comrie (1987, ed.), 619–644
- Kuryłowicz, Jerzy (1949), "Le problème du classement des cas". *Bulletyn PTJ* 9, 20–43 [reprinted in: Kuryłowicz, Jerzy (1960), *Esquisses linguistiques*. Wrocław-Krakow: Ossolineum, 131–150]
- Kuryłowicz, Jerzy (1964), *The Inflectional Categories of Indo-European*. Heidelberg: Winter
- Langdon, Margaret (1970), *A Grammar of Diegueño: The Mesa Grande Dialect*. Berkeley: University of California Press
- Lehmann, Christian (1982), *Thoughts on Grammaticalization: A Programmatic Sketch*. Köln: Institut für Sprachwissenschaft, Universität Köln (Arbeiten des Kölner Universalien-Projekts 48.1)
- Lewis, Geoffrey L. (1967), *Turkish Grammar*. Oxford: Oxford University Press
- Lindenfeld, Jacqueline (1973), *Yaqui Syntax*. Berkeley: University of California Press
- Mallinson, Graham (1987), "Rumanian". In: Comrie (1987, ed.), 303–321
- Mallinson, Graham & Blake, Barry J. (1981), *Language Typology: Cross-Linguistic Studies in Syntax*. Amsterdam: North-Holland (North-Holland Linguistic Series 46)
- Matthews, P[eter] H. (1974), *Morphology: An Introduction to the Theory of Word-Structure*. Cambridge etc.: Cambridge University Press
- McGregor, William (1990), *A Functional Grammar of Gooniyandi*. Amsterdam: Benjamins
- Mel'čuk, Igor A. (1986), "Toward a Definition of Case". In: Brecht et al. (1986, eds.), 35–85
- Merlan, Francesca (1982), *Mangarayi*. Amsterdam: North Holland (Lingua Descriptive Series 4)

- Moravcsik, Edith A. (1978), "On the Case Marking of Objects". In: Greenberg, Joseph H. (ed.), *Universals of Human Language, Vol. IV: Syntax*. Stanford: Stanford University Press, 249–289
- Nichols, Johanna (1983), "On Direct and Oblique Cases". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 9, 170–192
- Plank, F[rans] (1979), "Bibliography on Ergativity". In: Plank, Frans (ed.), *Ergativity: Towards a Theory of Grammatical Relations*. London etc.: Academic Press, 511–554
- Plank, Frans (1984, ed.), *Objects: Towards a Theory of Grammatical Relations*. London: Academic Press
- Plank, Frans (1990), "Suffix Copying as a Mirror-Image Phenomenon". *European Science Foundation, Programme in Language Typology: Working Paper* 1, 1–11
- Reed, Irene & Miyaoka, Osahito & Jacobsen, Steven & Afcan, Paschal & Krauss, Michael (1977), *Yup'ik Eskimo Grammar*. University of Alaska: Alaska Native Language Center and the Yup'ik Language Workshop
- Robins, R[obert] H. (1967), *A Short History of Linguistics*. London: Longman
- Rubio, Lisardo (1966), *Introducción a la Sintaxis Estructural del Latín, Vol. I: Casos y Preposiciones*. Barcelona: Ariel (Convivium 4)
- Rude, Noel (1985), *Studies in Nez Perce Grammar and Discourse*. Ph.D. dissertation, University of Oregon (Oregon)
- Rude, Noel (1988), "On the Origin of the Nez Perce Ergative NP Suffix". [Conference on Grammaticalization, University of Oregon, 12–15 May 1988]
- Saltarelli, Mario et al. (1988), *Basque*. London etc.: Croom Helm (Croom Helm Descriptive Grammars)
- Scott, Graham (1978), *The Fore Language of Papua New Guinea*. Canberra: Pacific Linguistics
- Serbat, Guy (1981), *Cas et Fonctions*. Paris: Presses Universitaires de France
- Sherzer, Joel (1976), *An Areal-Typological Study of American Indian Languages North of Mexico*. Amsterdams, Oxford: North-Holland (North-Holland Linguistic Series 20)
- Silverstein, Michael (1976), "Hierarchy of Features and Ergativity". In: Dixon (ed.), 112–171
- Silverstein, Michael (1981), "Case Marking and the Nature of Language". *Australian Journal of Linguistics* 1.2, 227–244
- Sittig, Ernst (1931), *Das Alter der Anordnung unserer Kasus und der Ursprung ihrer Bezeichnung als "Fälle"*. Stuttgart: Kohlhammer (Tübinger Beiträge zur Altertumswissenschaft 13)
- Starosta, Stanley (1985), "Relator Nouns as a Source of Case Inflections". In: Leed, Richard L. & Acson, Veneeta Z. (eds.), *For Gordon Fairbanks*. Honolulu: University of Hawaii Press, 111–133 (Oceanic Linguistics Special Publications 20)
- Steever, Sanford B. (1987), "Tamil and the Dravidian Languages". In: Comrie (1987, ed.), 725–746
- Stokes, Bronwyn (1982), *A Description of Nyigina, a Language of the West Kimberley, Western Australia*. Ph.D. dissertation, The Australian National University (Canberra)
- Suárez, Jorge A. (1983), *The Mesoamerican Indian Languages*. Cambridge: Cambridge University Press
- Touratier, Christian (1978), "Quelques principes pour l'étude des cas (avec application à l'ablatif latin)". *Languages* 50, 98–116
- Tucker, A[rchibald] N. & Bryan, Margaret A. (1966), *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa*. Oxford: Oxford University Press
- Wierzbicka, Anna (1980), *The Case for Surface Case*. Ann Arbor/MI: Karoma (Linguistica Extranea. Studia 9)
- Wierzbicka, Anna (1981), "Case Marking and Human Nature". *Australian Journal of Linguistics* 1.1, 43–80
- Windfuhr, Gernot L. (1987), "Persian". In: Comrie (1987, ed.), 523–546
- Yallop, Colin (1977), *Alyawarra: An Aboriginal Language of Central Australia*. Canberra: Australian Institute of Aboriginal Studies

*Barry Blake, Bundoora (Australia)*

## 103. Possession

1. Definition und Abgrenzung
2. Prädikative Possession = Etablierung
3. Attributive Konstruktionen
4. Inhärenz der Possessiv-Relation
5. Nominaler vs. pronominaler Possessor
6. Unübliche Abkürzungen
7. Zitierte Literatur

### 1. Definition und Abgrenzung

Unter **Possession** verstehen wir die sprachliche Repräsentation einer Zugehörigkeitsrelation zwischen zwei Entitäten, von denen die eine den **Possessor** repräsentiert, die andere das **Possessum** (vgl. Seiler 1983: 4).

Die Zugehörigkeitsrelationen können jeweils unterschiedlicher Art sein. Bei Teil-Ganzen- und Verwandtschaftsbeziehungen z. B. handelt es sich um inhärente Beziehungen, da ein Teil nicht ohne Bezug auf das Ganze konzeptualisiert werden kann, und ein Vater ist immer ein Vater von jemandem. Besitzrelationen dagegen, die die prinzipielle Veräußerbarkeit einer Habe implizieren (vgl. Taylor 1989: § 11.2), sind nicht inhärent und müssen daher vorgängig etabliert werden. Dabei ist zu beachten, daß auch der Besitzbegriff unterschiedlich interpretiert werden kann, je nach dem, ob der jeweilige Besitzer einen Rechtsanspruch hat (inhärenter Besitz) oder nur einen Nutzungsanspruch (akzidentieller Besitz). Des weiteren spielt es eine Rolle, ob der Besitz verfügbar ist (physikalischer Besitz; vgl. Miller & Johnson-Laird 1976: 558–568).

Possession kann von daher nicht als „eine allgemeine und abstrakte Beziehung schlechthin gesehen werden“, sondern es müssen je nach der „größeren ‘Nähe’ oder ‘Ferne’ zum Ich“ (Cassirer 1973/1977: 229) verschiedene Stufengrade unterschieden werden. Diese Distanz (‘conceptual distance’ Haiman 1985: § 2.2) bezeichnen wir im folgenden als **Inhärenz** der Possessivrelation, die gekennzeichnet ist durch die beiden Eckpunkte **inhärent** und **etabliert**.

In der Linguistik wird dieser Unterschied üblicherweise mit den von Lévy-Bruhl (1914) eingeführten Termini *alienabel* vs. *inalienabel* bezeichnet. Diese kategorische Einteilung wird den sprachlichen Gegebenheiten jedoch nicht gerecht (vgl. Seiler 1983: Kap. 5). Inhärenz ist vielmehr als ein kontinuierliches Konzept aufzufassen, das sich in den Spra-

chen in den unterschiedlich expliziten Possessivkonstruktionen zeigt. Dabei gibt es eine ikonische Beziehung dergestalt, daß die inhärentere Beziehung auch durch die einfachere Konstruktion ausgedrückt wird.

### 2. Prädikative Possession = Etablierung

Je weniger inhärent die Zugehörigkeitsbeziehung zwischen zwei Entitäten ist, um so eher erlaubt sie eine explizite Etablierung. Relationen wie *mein Vater* oder *mein Bein* bedürfen in aller Regel keiner Etablierung, und Äußerungen der Form: *Ich habe einen Vater* bzw. *Ich habe ein Bein* implizieren, daß diese Relation in Frage stand, oder aber daß es sich nicht um eine inhärente Relation handelt.

Formal können die prädiktiven Possessivkonstruktionen danach unterschieden werden, wie die Relation markiert wird. Sie kann durch ein possessives Prädikat (2.1) ausgedrückt werden, am Possessum (2.2) oder am Possessor (2.3) markiert sein.

#### 2.1. Possessive Prädikate

**Possessive Prädikate** sind z. B. dt. *haben/gehören*, span. *tener*, Krongo *áná*. Diese Verben gehen in aller Regel auf ein aktives, transitives Verb zurück (‘erfassen’, ‘halten’), woher sich die transitive Struktur der Possessiv-Prädikation erklärt (vgl. Givón 1984: 103 f.; Isachenko 1974; Boeder 1980; Bendix 1966; Mey 1969). Prädikationsbasis ist in diesen Fällen immer der Possessor, der typischerweise [+belebt] ist, wie in den folgenden Beispielen aus dem Deutschen (1), dem Quechua (2) und dem Krongo (3):

- (1) *Ich habe ein Auto.*
- (2) *Marya-ka atalpa-ta chari-n*  
Maria-TOP Henne-AKK hab-3.SG  
‘Maria hat eine Henne.’ (Cole 1982: 93)
- (3) *N-áná à?àŋ kitáabà*  
1/2-IPFV.hab ich Buch  
*η-âdeelá.*  
KONN.M-IPFV:schön.sein  
‘Ich habe ein schönes Buch.’ (Reh 1985: 324)

Ein explizites ‘haben’-Verb ist jedoch eher selten in den Sprachen der Welt. Statt dessen werden eine Reihe unterschiedlicher Konstruktionen verwendet, die im nachfolgenden

exemplifiziert werden. Heine (1997:47) hat gezeigt, daß der prädikative Ausdruck von Possession auf eine begrenzte Menge von kognitiven Schemata zurückgeführt werden kann: (a) Aktion: X nimmt Y, (b) Lokation: Y ist lokalisiert bei X, (c) Begleitung: X ist mit Y usw.

## 2.2. Possession am Possessum markiert

Wird die Relation am Possessum markiert, dann bildet ebenfalls der Possessor die Prädikationsbasis, und das Possessum wird in der Regel mit einem **Komitativ** angeschlossen (vgl. Creissels 1979), wie z. B. im Bemba:

- (4) *N-di no-omuana.*

1.SG-sein mit-Kind

‘Ich habe ein Kind.’ (wörtl. ‘Ich bin mit Kind’) (Givón 1984: 104)

Hierunter fallen auch die mit dem ‘haben’-Affix gebildeten Formen in australischen Sprachen, die in vielen Fällen einer ‘haben’-Konstruktion entsprechen, wie im Guugu Yimidhirr.

- (5) *Ngayu galga-dhurr ngayu*

1.SG.NOM Speer-KOMIT.ABS 1.SG.NOM

*buurray-irr.*

Wasser-KOMIT.ABS

‘Ich habe einen Speer. Ich habe Wasser.’ (wörtl. ‘Ich bin mit Speer. Ich bin mit Wasser.’) (Haviland 1979: 58)

## 2.3. Possession am Possessor markiert

Dies geschieht zumeist in der Form, daß der Possessor im **Dativ** oder in einem **Lokalkasus** steht, wie in den folgenden Beispielen aus dem Hebräischen (6) und dem Ik (7):

- (6) *Le-yóav hayú harbé.*

zu-Yoav waren viele.Freunde

‘Yoav hatte viele Freunde.’ (Givón 1984: 104)

- (7) *ia yoka ntsi-k<sup>e</sup>*

sein Hund:NOM 3.SG-DAT

‘Er hat einen/den Hund.’ (wörtl. ‘Der Hund ist bei ihm.’)

Im weitesten Sinne gehören hierzu auch Konstruktionen, bei denen eine Possessivphrase mit einem Existenzprädikat versehen wird und so die Besitzrelation etabliert wird. Dies ist z. B. im Guugu Yimidhirr (Australien) der Fall:

- (8) *Yarrga-wi galga wu-naa.*

Junge-GEN.ABS(DAT) Speer.ABS EXIST-NPRT

‘Der Junge hat einen Speer.’ (wörtl. ‘Des Jungen Speer existiert.’, oder ‘Dem Jungen existiert ein Speer.’) (Haviland 1979: 58)

Am explizitesten ausgeprägt finden wir diesen Konstruktionstyp im Yaqui, wo das Possessum-Nomen direkt die prädikative Kennzeichnung in Form von Tempus/Aspekt/Modus-Affixen trägt:

- (9) *Vempo b<sup>w</sup>e'ere-m kaar-k.*

sie groß-PL Auto-PFV

‘Sie hatten große Autos.’

- (10) *'inepo b<sup>w</sup>e'ere-m kaaro-m vica-k.*

ich groß-PL Auto-PL seh-PFV

‘Ich sah (die) große(n) Autos.’ (Jelinek & Escalante 1988: 412)

Die Inhärenz der Possessivrelation findet in der prädiktiven Konstruktion in unterschiedlicher Weise ihren morphologischen Ausdruck. So wird z. B. im Ewe bei alienabler Possession ein lokales Prädikat mit dem Nomen für ‘Hand’ verwendet: ‘haben’ = ‘in der Hand sein’ (ebenso im Bengali; vgl. Masica 1991: 359), bei inalienabler Possession dagegen eine statische Konstruktion mit dem Verb für ‘geben’: ‘haben’ = ‘gegeben sein’ (siehe Heine 1997; Ameka 1996).

- (11) *Ga le ási nye.*

Geld ist Hand mein

‘Ich habe Geld.’

- (12) *Ta le ná-m.*

Kopf ist geben-mich

‘Ich habe einen Kopf.’

Im Bambara, ebenso wie in den anderen Mande-Sprachen, wird bei der Prädikation zwischen inhärentem, akzidentiellem und physikalischem Besitz unterschieden.

- (13) *Wari' be Bàba fe/bolo/kùn.*

Geld ART AUX Baba POSTPOSITION

‘Baba hat das Geld.’ (wörtl. ‘Das Geld ist bei Baba’)

Durch die unterschiedlichen Postpositionen wird angezeigt, daß Baba der rechtmäßige Eigentümer ist (*fe* ‘Kopf’), und *kùn* ‘Hand’ impliziert, daß er das Geld bei sich trägt (vgl. Kastenholz 1988); *bolo* ist dagegen unspezifiziert.

Eine vergleichbare Unterscheidung findet sich im Imnambura Quechua, wo durch unterschiedliche prädiktive Possessivkonstruktionen angezeigt wird, ob es sich um permanenten, neutralen oder temporären Besitz handelt (Cole 1982: 94):

- (14) *Juzi iskay kaballu-ta chari-n.*

José zwei Pferd-AKK hab-3.SG

‘José hat zwei Pferde.’

- (15) *Pay-ka kaballu-yuj ka-rka.*  
 er-TOP Pferd-POSSESSOR sein-PRÄT:3.SG  
 'Er hatte ein Pferd; er war ein Pferdebesitzer.'

Inhärenz spielt auch bei dem Unterschied in der Zuordnungsrichtung, wie er sich im Deutschen zwischen *Ich habe ein Auto* und *Das Auto gehört/ist mir* zeigt, eine Rolle. Dieser hängt mit der Thematisierung und der Definitheit des Possessums zusammen und wird in Sprachen in unterschiedlicher Weise angezeigt (vgl. Clark 1978: 113). Im Somali, wo der Possessor die Prädikationsbasis darstellt, wird der Unterschied z. B. durch Fokus und Wortstellung angezeigt:

- (16) *Nin-kii babuur ay-uu*  
 Mann-ART.M Auto FOK-3.SG.M  
*leh-yahay.*  
 hab-KOP:3.SG.M  
 'Der Mann hat ein Auto.'

- (17) *Babuur-ka nin-kii ayaa leh.*  
 Auto-ART.M Mann-ART.M FOK hab  
 'Das Auto gehört dem Mann.'

Daß es sich hierbei tatsächlich um ein Äquivalent zur 'gehören'-Konstruktion des Deutschen handelt, wird daran deutlich, daß inhärent relationale Nomina nicht in dieser Weise konstruiert werden:

- (18) *Nin-kii ínan ay-uu*  
 Mann-ART.M Mädchen FOK-3.SG.M  
*leh-yahay.*  
 hab-KOP:3.SG.M  
 'Der Mann hat eine Tochter (Mädchen).'

- (19) *\*ínan-ta nin-kii ayaa leh.*  
 Mädchen-ART.F Mann-ART.M FOK hab  
 ? 'Die Tochter gehört dem Mann.'

### 3. Attributive Konstruktionen

**Attributive Possessivkonstruktionen** sind endozentrische Nominalsyntagmen, in denen das Possessum den Nukleus repräsentiert und der Possessor als Determinierer fungiert (s. Koptjevskaia-Tamm 2002; Ultan 1978). Die syntaktische Relation kann wiederum in unterschiedlicher Weise markiert sein, und wir unterscheiden insgesamt vier Typen, je nach dem, ob die Relation ausschließlich durch die Wortstellung (3.1) oder durch einen Linker (3.2) markiert ist. Wird die Relation an einem der beiden Nomen markiert, dann unterscheiden wir danach, ob das Possessum die Mar-

kierung trägt (*head marking*, vgl. Nichols 1988) (3.3) oder der Possessor (*dependent marking*) (3.4).

#### 3.1. Juxtaposition: N N

Die Kennzeichnung der possessiven Relation erfolgt ausschließlich über die Wortstellung, wie z. B. im Somali:

- (20) *faras-ka nin-ka*  
 Pferd-ART.SG.M Mann-ART.SG.M  
 'das Pferd des Mannes'

In vielen Sprachen ist die **Juxtaposition** als alternative Konstruktion zu einer expliziteren Possessiv-Konstruktion möglich, insbesondere zum Ausdruck inhärenter Relationen. In diesen Fällen erweist sich die Abgrenzung zur Komposition jedoch als problematisch (vgl. Chappell & McGregor 1989; Kay & Zimmer 1976). Im Somali, wo beide Nomina mit dem definiten Artikel stehen (21), wird der Unterschied zur Komposition nur dadurch deutlich, daß die invertierte Konstruktion, bei der der Possessor am Possessum pronominal wieder aufgenommen wird (s. 3.4), nicht möglich ist (22).

- (21) *warshad-da hilis-ka*  
 Fabrik-ART.SG.M Fleisch-ART.SG.M  
 'die Fleischfabrik'

- (22) *\*hilis-ka*  
 Fleisch-ART.SG.M  
*warshad-diis-a*  
 Fabrik-POSS.3.SG.F-ART  
 'dem Fleisch seine Fabrik'

#### 3.2. Linker-Konstruktion

Die Relation wird durch ein separates morphologisches Element angezeigt, daß zwischen N und N steht. Wir können zwischen unveränderlichen **Konnektoren** und solchen Konnektoren, die mit dem Possessum kongruieren, unterscheiden.

##### 3.2.1. Unveränderliche Konnektoren

Klassisches Beispiel hierfür ist die *Izafe*-Konstruktion im Persischen, bei der ein unbetontes *e*, nach Vokal *ye*, an den Nukleus eines attributiven Syntagmas tritt.

- (23) *pa-ye-man*  
 Fuß-IZAFE-mein  
 'mein Fuß'  
*ash-e-pedar*  
 Pferd-IZAFE-Vater  
 'Vaters Pferd'

Im Japanischen finden wir ein Element *no*, das in den Grammatiken als Postposition oder als Konnektor bezeichnet wird.

- (24) *anata no musume*  
 du KONN Tochter  
 ‘deine Tochter’  
*koobe no sensei*  
 Kobe KONN Lehrer  
 ‘der Lehrer von Kobe’  
 (Hinds 1986: 136)

Im Hebräischen finden wir eine Postposition *shel*, die sich herleitet aus dem Relativpronom *sher* + Dativ *l*.

- (25) *ha sefer shel ha yeled*  
 ART Buch von ART Junge  
 ‘das Buch des Jungen’

Im Ewe leitet sich der Konnektor von einem Nomen für ‘Platz’ her (s. (26); Claudi & Heine 1989: 14), und in den tschadischen Sprachen geht der Konnektor auf ein ursprüngliches Demonstrativum zurück (vgl. Schuh 1981).

- (26) *kofi phé ta'*  
 Kofi Platz Kopf  
 ‘Kofis Kopf’

Linker-Konstruktionen finden sich außerdem in ozeanischen Sprachen, hier ein Beispiel aus dem Houailou (27), und auch das engl. Genitivklotikum *-s* wird dazu gerechnet (vgl. Croft 1990: 32).

- (27) *nevā i mōrua i wiʔ-a?*  
 Land von Großvater von Mann-dies  
 ‘dieses Mannes Großvaters Land’  
 (Wilson 1982: 95)

Der Konnektor kann auch suprasegmental realisiert werden, wie z. B. im Yoruba (Kwa). Der finale Vokal des voranstehenden Possessums wird gelängt:

- (28) *filà Tündé* > *filáa Tündé*  
 ‘Tundes Mütze’  
*ilé Bísí* > *ilée Bísí*  
 ‘Bisis Haus’

Die Abgrenzung von Adpositionen zu Linkern ist nicht immer eindeutig, vor allem da Linker dazu tendieren, sich enklitisch an ein Nomen anzuschließen. Bei den unveränderlichen Formen ist dies das voranstehende Nomen, wie im Kresh.

- (29) *neté ká lukpigi* > *netá(a) lukpigi*  
 ‘Arbeit von Frau’ (Santandrea 1976: 115)

In Sprachen, in denen der Linker mit dem Possessum kongruiert, schließt er dagegen an das nachfolgende Nomen an.

### 3.2.2. Konnektor kongruiert mit Possessum

Dieser Konstruktionstyp ist charakteristisch für die Klassensprachen in den Niger-Kongo-Sprachen, wie z. B. im Swahili:

- (30) (a) *mwana wa kipofu*  
 KL1:Sohn KL1:KONN Binder  
 ‘des Blinden Sohn’  
 (b) *kisu cha mpishi*  
 KL7:Messer KL7:KONN Koch  
 ‘das Messer des Kochs’

Es besteht die Tendenz, den Konnektor enklitisch an das nachstehende Nomen anzuschließen, so daß eine Konstruktion entsteht, in der der Possessor mit dem Possessum kongruiert. Die folgenden Beispiele stammen aus dem Herero:

- (31) (a) *omo-atje uo-mu-serandu*  
 KL1.SG-Kind KL1.SG-KL1.SG-Nama  
 ‘das Kind des Nama’  
 (b) *ova-natje vo-va-serandu*  
 KL1.PL-Kind KL1.PL-KL1.PL-Nama  
 ‘die Kinder der Nama’  
 (c) *e-ke r-omo-atje*  
 KL4.SG-Hand KL4.SG-KL1.SG-Kind  
 ‘die Hand des Kindes’  
 (d) *oma-ke uo-mu-atje*  
 KL4.PL-Hand KL4.PL-KL1.SG-Kind  
 ‘die Hände des Kindes’ (Meinhof 1937: 13 ff.)

Kongruenz zwischen Possessum und Konnektor findet sich auch im Hindi, einer Genus-Sprache. Der Konnektor, der in der Grammatik als Postposition bezeichnet wird, kongruiert mit dem Possessum in Genus und Numerus.

- (32) (a) *kamre kā darvāzā*  
 Zimmer KONN Tür  
 ‘die Tür des Zimmers’  
 (b) *kamre ki- divār*  
 ‘die Wand des Zimmers’  
 (c) *kamre ke darvāze*  
 ‘die Türen des Zimmers’ (Porízka 1972: 70f.)

### 3.2.3. Konnektor als Klassifikator

Vor allem in austronesischen Sprachen, aber auch in einigen anderen Sprachen (Jacalteco, Mba-Sprachen (vgl. Pasch 1986: Kap. 4)) finden wir Konnektoren, die als **Klassifikatoren**

(vgl. Art. 97) fungieren und das Possessum näher kennzeichnen. Nachfolgend ein Beispiel aus dem Hawaianischen:

- (33) (a) *he pāpale ā-na*  
          ART Hut POSS<sub>1</sub>-sie  
          ‘Ein Hut von ihr, sie machte ihn.’
- (b) *he pāpale o-na*  
          ART Hut POSS<sub>2</sub>-sie  
          ‘Ein Hut von ihr, sie trägt ihn.’ (Wilson 1982: 1)

Die Klassifikatorposition kann aber auch von anderen Elementen eingenommen werden, die nicht das Possessum klassifizieren, sondern die possessive Relation selbst modifizieren.

- (34) (a) *ka i'a a ka wahine*  
          ART Fisch POSS ART Frau  
          ‘der Fisch der Frau’
- (b) *ka i'a na ka wahine*  
          ART Fisch gehör ART Frau  
          ‘der Fisch, der der Frau gehört’  
          (Wilson 1982: 49)

### 3.3. Relation am Possessor markiert

Die Possessivrelation kann durch eine Adposition markiert werden, wie im Irischen, wo wir als Alternative zur Genitivkonstruktion eine Präpositionalkonstruktion finden, die die ältere Genitivkonstruktion ersetzt (vgl. McCaughey 1976).

- (35) (a) *Tá a theach nua réidh*  
          sein POSS.3.SG.M Haus neu fertig  
          anois.  
          jetzt  
          ‘Sein neues Haus ist jetzt fertig.’
- (b) *Ta an theach aige réidh*  
          sein ART Haus bei:3.SG.M fertig  
          anois.  
          jetzt  
          ‘Sein neues Haus ist jetzt fertig.’

In den Zentralsudansprachen wird die Relation durch Postpositionen am Possessor angezeigt, hier zwei Beispiele aus dem Logo:

- (36) *dili logo a'dia*  
          Speer Logo POSTPOSITION  
          ‘ein Logo Speer’
- (37) *fa kokia a'dia*  
          Knochen Hund POSTPOSITION  
          ‘ein Knochen des Hundes’ (Tucker 1940: 164)

In den australischen Sprachen wird die Possessivrelation durch ein Kasusaffix am No-

men angezeigt, das in vielen Fällen identisch ist mit dem Dativ. Das folgende Beispiel stammt aus dem Gumbaynggir:

- (38) *Yaraŋ gamay ni:garundi.*  
          DEM Speer:S Mann:GEN:S  
          ‘Dieser Speer ist des Mannes.’ (Eades 1979: 276)

Inhärenz spielt auch bei Kasus eine Rolle, wie das Beispiel des Krongo zeigt. Hier werden zwei adnominale Kasus unterschieden: (a) Der Genitiv, der bei unbelebten Possessoren und bei inhärenten Relationen verwendet wird, und (b) der Possessiv, der mit belebten Possessoren verwendet wird.

- (39) (a) *shàaká má-nímyà*  
          Milch GEN-Frau  
          ‘die [Mutter-]Milch der Frau’
- (b) *shàaká kà-nímyà*  
          Milch POSS-Frau  
          ‘die [Kuh/Ziegen-]Milch der Frau’  
          (Reh 1985: 317)

Der Possessiv ist von dem Lokalkasus abgeleitet, ist aber distributiv verschieden, so daß die Unterscheidung gerechtfertigt ist.

Possessor	[+ belebt]	[− belebt]
lokal	<i>kà</i>	<i>kí</i>
possessiv	<i>kà</i>	<i>má</i>

Tab. 103.1: Possessiv und Lokalkasus im Krongo

Die folgenden Beispiele demonstrieren die Verwendung sowohl des Lokatifs als auch des Genitivs/Possessivs.

- (40) (a) *m-àssàlà kà-bílyáatà kà-káaw*  
          F-PFV:seh LOK-Kind POSS-Person  
          *m-àanímyà*  
          KONN:F-weiblich  
          ‘Sie hat sich das Kind der Frau angesehen.’
- (b) *m-àssàlà kí-bólóocój má-coori*  
          F-PFV:seh LOK-Tür GEN-Haus  
          ‘Sie hat sich die Tür des Hauses angesehen.’ (Reh 1985: 154)

### 3.4. Relation am Possessum markiert

Hierunter fassen wir zwei unterschiedliche Konstruktionstypen: (a) den **status constructus**, wie wir ihn z. B. in den semitischen Sprachen vorfinden, und (b) Konstruktionen, in denen der Possessor am Possessum **pronominal** angezeigt wird.

### 3.4.1. Status constructus des Semitischen

Im Hebräischen steht das Possessum in einer sog. Konstruktform, nachfolgend am Beispiel *dod* 'Onkel, Tante' demonstriert:

	Konstruktform	Freie Form
M.SG	<i>dod</i>	<i>dod</i>
F.SG	<i>dod-at</i>	<i>dod-a</i>
M.PL	<i>dod-ey</i>	<i>dod-im</i>
F.PL	<i>dod-ot</i>	<i>dod-ot</i>

Tab. 103.2: Konstruktform im Hebräischen (nach Glinert 1989: 30)

Der Unterschied wird in den folgenden Beispielen deutlich:

- (41) *bigd-ey ha-tinok*  
Kleid-KONSTR ART-Baby  
'die Kleider des Babys'  
*ha-bgad-im shel ha-tinok*  
ART-Kleid-FREI VON ART-Baby  
'die Kleider des Babys'  
(Glinert 1989: 24)

Das Nomen in der Konstruktform und das nominale Attribut bilden eine Einheit, die in der Regel nicht aufgebrochen werden kann. So stehen Adjektive z. B. nach der Konstruktphrase (42); gleiches gilt auch für das Arabische (43).

- (42) *shur-at konim aruka*  
Linie-KONSTR Kunde lang  
'eine lange Schlange von Kunden'  
(Glinert 1989: 26)
- (43) *kitāb er-rāgil el-kebīr*  
Buch ART-Mann ART-groß  
'des Mannes großes Buch'  
(O'Leary 1955: 82)

Diese enge Beziehung zwischen Possessum und Possessor ist die Ursache dafür, daß sich in den semitischen Sprachen eine Linker-Konstruktion entwickelt hat, bei der ein explizites Besitznomen die Possessivrelation anzeigt. So stellt Brockelmann (1908: 238) fest, daß in den neuarabischen Dialekten vielfach als Apposition zum Possessum Wörter mit der allgemeinen Bedeutung 'Besitz, Eigentum' o. ä. auftreten, die nun in engere Beziehung zum Genitiv treten und so dem Possessum seine in den altsemitischen Sprachen verlorene syntaktische Beweglichkeit wiedergeben.

Diese Entwicklung zeigen die folgenden Beispiele aus dem Klassischen Arabisch (44)

und dem neuarabischen Dialekt von Damaskus (45–46):

- (44) *bayt-u t-tāğiri*  
Haus-KONSTR ART-Händler
- (45) *bēt ət-tāżer*  
Haus ART-Händler
- (46) *əl-bēt taba' ət-tāżer*  
ART-Haus Besitz ART-Händler  
'das Haus des Händlers' (Diem 1986)

### 3.4.2. Pronominale Kennzeichnung am Possessum

Eine in den Sprachen der Welt weit verbreitete Konstruktion besteht darin, daß der Possessor am Possessum pronominal angezeigt wird. Dieser Konstruktionstyp findet sich vor allem in den Kaukasus-Sprachen (Abkasisch (47)), den Indianersprachen Nord- und Mittelamerikas (Cree (48)) und als alternative Konstruktion im Afroasiatischen (Somali (49), Hebräisch (50)).

- (47) *sarà sə-y'oñā*  
ich mein-Haus  
'mein Haus'  
*á-č'kō'ən ə-y'oñā*  
ART-Junge sein-Haus  
'das Haus des Jungen' (Hewitt 1979: 116)
- (48) *cān o-mōhkomān*  
John 3.SG-Messer  
'Johns Messer' (Wolfart & Carroll 1981: 42)
- (49) *nin-ka faras-kiisa*  
Mann-ART.3.SG.M Pferd-POSS.3.SG.M  
'das Pferd des Mannes'
- (50) *sifr-o shel ha yeled*  
Buch-POSS.3.SG VON ART Junge  
'das Buch des Jungen'

Der Possessor steht in dieser Konstruktion in der Regel in der unmarkierten Form. In den nordwestkaukasischen Sprachen, z. B. Adyge, steht der Possessor jedoch im Genitiv:

- (51) *tə-m ə-nă*  
Mann-GEN 3.SG-Auge  
'das Auge des Mannes'  
(Comrie 1981: 222)

Im Türkischen steht er nur dann im Genitiv, wenn es sich um eine definite Konstruktion handelt, ansonsten bleibt er unmarkiert.

- (52) (a) *üniversite-nin professör-ler-i*  
Universität-GEN Professor-PL-3.POSS  
'die Professoren der Universität'

- (b) *üniversite professör-ler-i*  
 Universität Professor-PL-3.POSS  
 'Professoren der Universität' (Lewis 1985: 42f.)

Im Ungarischen kann der Possessor im Nominativ oder im Dativ stehen (53). Im letzten Falle brauchen die beiden Nominalphrasen nicht juxtaponiert zu sein.

- (53) (a) *a férfi-Ø ház-a*  
 ART Mann-NOM Haus-POSS.3.SG  
 'das Haus des Mannes'  
 (b) *a férfi-nak a ház-a*  
 ART Mann-DAT ART Haus-POSS.3.SG  
 'dem Mann sein Haus' (Biermann 1985: 7)

In einigen uto-aztekischen Sprachen steht der Possessor im Akkusativ:

- (54) *i-koonya-y tiw'aya'-at*  
 mein-Ehemann-AKK Nichte-sein  
 'meines Ehemannes Nichte' (Langacker 1977: 90)

#### 4. Inhärenz der Possessiv-Relation

Sprachen zeigen in aller Regel Variation in der attributiven Possessivkonstruktion: An- bzw. Abwesenheit eines Linkers, Verwendung einer Genitiv- oder Präpositionalkonstruktion: *des Mannes Haus* vs. *das Haus von dem Mann*. So werden z. B. die folgenden Konstruktionen im Tamil als bedeutungsgleich bezeichnet:

- (55) *itu en makan-ø pā*  
*itu en makan-in pā*  
*itu en makan-uṭaiya pā*  
*itu en makan-in-uṭaiya pā*  
 PROX:N 1.SG.OBL Sohn-KAT-GEN Matte  
 'Dies ist die Matte meines Sohnes.' (Kukuczka 1982: 22)

Im Tamil wird Inhärenz formal nicht unterschieden (Kukuczka 1982: 27). In anderen Sprachen entspricht die Variation jedoch normalerweise einer Differenzierung von Inhärenzgraden.

Um eine explizitere Konstruktion handelt es sich z. B. im To'abaita, wo der pronomiale Possessor suffigiert wird, wenn es sich um eine inalienable Besitzrelation handelt, der Possessor aber durch ein selbständiges Pronomen realisiert wird, wenn die Relation alienabel ist.

- (56) *gwau-na* 'sein Kopf'  
*gwau-ku* 'mein Kopf'

- luma nia* 'sein Haus'  
*luma nau* 'mein Haus'  
 (Simons 1986: 25 ff.)

In den australischen Sprachen wird die alienable Relation durch die optionale Verwendung des Genitivsuffixes markiert, die inalienable Relation bleibt unmarkiert, wie z. B. im Gumbaynggir:

- (57) *Gula:du ma:nij nuju:gundi*  
 er:AGT nehm:PRÄT Känguruh:GEN.OBJ  
*dara.*  
 Bein:OBJ  
 'Er nahm des Känguruhs Bein.'  
*Nuju: gaygin dawgar*  
 Känguruh.OBJ schneid:PRÄT PART  
*da:ra.*  
 leg.OBJ  
 'Und so schnitt er dem Känguru ein  
 Bein ab.' (Eades 1979: 317)

Wie das Beispiel zeigt, müssen Possessum und Possessor jedoch nicht juxtaponiert sein. Dies verweist darauf, daß beide in vergleichbarer Weise an der Handlung beteiligt sind. Im Macushi (Amazonas) werden alienable Relationen durch ein Possessivsuffix gekennzeichnet:

- (58) *u-yewi'*  
 1.SG-Haus  
 'mein Haus'  
*u-ye'ma-ri*  
 1.SG-Weg-POSS  
 'mein Weg'  
 (Abbott 1991: 85f.)

Für die ozeanischen Sprachen gilt, daß, wenn eine Sprache zwischen einer direkten (Juxtaposition) und einer indirekten Konstruktion (Linker) unterscheidet, die direkte Konstruktion immer für inalienable Relationen verwendet wird (Lichtenberk 1985); Nichols (1988) zeigt, daß in den nordamerikanischen Indianersprachen die Markierung bei inalienablen Relationen entweder älter oder morphologisch enger ist. In den Sprachen, für die ein Unterschied zwischen alienabler und inalienabler Possession postuliert wird, gilt somit, daß die formale Explizitheit in der Regel mit der Inhärenz der Relation korreliert. Es zeigt sich somit eine Ikonizität von Form und Inhalt dergestalt, daß eine engere/inhärentere Beziehung zwischen Possessor und Possessum normalerweise auch durch eine 'engere' Konstruktion realisiert wird (vgl. Haiman 1985: § 2.2.3; Seiler 1983). Problematisch für die generelle Hypothese sind jedoch Fälle wie

das Krongo (s. o. 3.3), wo zwei Kasus unterschieden werden, und Sprachen wie das Seneca, in dem inalienable Relationen mit den subjektiven Verbalaffixen, alienable aber mit den objektiven Verbalaffixen gebildet werden:

- (59) *ha-hsi?ta?* 'sein Fuß' (inalienabel)  
*ho-?nɔ?* 'sein Pfeil' (alienabel)  
 (Chafe 1963: 12–15)

#### 4.1. Relationale Nomina

Inhärenz ist in vielen Sprachen eine verdeckte Kategorie, die sich nur darin zeigt, daß inhärent **relationale Nomina** normalerweise mit einem Possessor verwendet werden müssen. Sie wird jedoch dann zu einer offenen Kategorie, wenn Sprachen die absolute Verwendung eines semantisch relationalen Nomens obligatorisch morphologisch kennzeichnen. Dies ist vor allem in den uto-aztekischen Sprachen der Fall. So wird im Tzotzil (Delgaty 1961), wie insgesamt in den Maya-Sprachen, bei Nomina unterschieden, (a) ob sie obligatorisch mit einem Possessivaffix vorkommen: *s-bek' mayil* 'es.ist-Samen Kürbis (Kürbissamen)', (b) ob sie optional ein Suffix haben können: *na* 'Haus' (*h-na* 'mein Haus', *s-na* 'sein Haus'), und (c) ob sie nie mit einem Possessivaffix vorkommen, *te'ik* 'Holz' oder *ni-cin* 'Blume'. Außerdem gibt es eine Klasse von Nomina, die dann, wenn sie mit einem Possessivaffix verbunden werden, das Absolutivsuffix *-Vl* nehmen:

- (60) *nc* 'Laus/Läuse'    *knc-il* 'meine Läuse'  
*ho* 'Wasser'            *ka'al* 'mein Wasser'

und solche, die das Absolutivsuffix nehmen, wenn sie ohne Possessivaffix verwendet werden:

- (61) *hvis*            'meine ältere Schwester'  
*vis-il*            'ältere Schwester'  
*hc-ek*            'mein Rock'  
*c-ek-il*            'Rock'

Bei den Beispielen unter (60) handelt es sich um einen Fall von **Relationierung**. Ein absolutes Nomen, das in der unmarkierten Form nicht besessen werden kann, muß zunächst relationiert werden, bevor es mit einem Possessivaffix verbunden werden kann.

Bei den Beispielen unter (61) hat das Suffix jedoch derelationierende Funktion. Die offene Argumentstelle des relationalen Nomens muß zunächst besetzt werden, bevor das Nomen absolut verwendet werden kann.

In den Sprachen, in denen Relationalität bei Nomina lexikalisch verankert ist, finden

sich in der Regel Verfahren, diese Relationalität aufzuheben. Im Navaho z. B. werden relationale Nominalstämme (Körperteile vor allem), wenn sie absolut, d. h. ohne Possessor, verwendet werden, mit dem generellen Possessivpronomen *a-* 'genereller Possessor unbelebt' verbunden, Verwandtschaftsterme mit *ha-* 'genereller Possessor belebt'. Durch die Verwendung von *a-* kann differenziert werden, ob es sich um einen eigenen oder einen fremden Teil handelt.

- (62) *'a-be*            'Milch' (absolut)  
*bi-be*            'ihre Milch, Muttermilch'  
*be-'a-be*            'ihre (gekauft) Milch'  
 (Young & Morgan 1980: 7)

Derelationierung relationaler und Relationierung absoluter Nomina finden sich auch in anderen Sprachen. So müssen im Hixkaryana Körperteilbezeichnungen, wenn sie absolut verwendet werden, mit dem derelationierenden Affix *-nano* versehen werden.

- (63) *rownari*            'meine Nase'  
*owna-nano*            'Nasen'

Diese Formen werden jedoch selten verwendet. Statt dessen wird die 1. Person Plural als generalisierte Form gebraucht: *rownari romo* 'unsere Nasen' im Sinne von 'Nasen von Menschen' (Derbyshire 1979: 69).

Im Ik können relationale Nomina, die als Postpositionen fungieren und in einer Nominalphrase nach ihrem Bezugsnomen juxtapaniert stehen, auch vorangestellt oder alleinstehend verwendet werden. In diesem Falle müssen sie jedoch mit der semantisch weitgehend leeren Proform *-de* verbunden werden, wodurch die offene Argumentstelle besetzt wird; das Bezugsnomen wird mittels Genitiv angeschlossen.

- (64) (a) *K'e-es-ia ho=ak'o-ke.*  
 geh-INT-1.SG Haus=innen-DAT  
 'Ich werde in das Haus gehen.'  
 (b) *K'e-es-ia ak'w-ed-e ho-e.*  
 geh-INT-1.SG innen-Ø-DAT Haus-GEN  
 'Ich werde in den Innenraum des Hauses gehen.' (Serzisko 1992: 192)

Dieses Element kann aber auch mit absoluten Nomina verwendet werden und zeigt dann Relationalität an. So wird aus dem absoluten Nomen *ama* 'Mensch' das relationale Nomen *am-ed* 'Besitzer' abgeleitet.

#### 4.2. Verwandtschaftsbezeichnungen

Eine besondere Klasse relationaler Terme stellen die Verwandtschaftsbezeichnungen dar, die sich u. a. dadurch auszeichnen, daß

die inhärente Possessivrelation in vielen Sprachen lexikalisiert ist. Dies ist u. a. der Fall im Ik (Serzisko 1992: 191):

	Vater	Mutter	Bruder	Schwester
mein	<i>abap</i>	<i>yaj</i>	<i>ede</i>	<i>eya</i>
dein	<i>baba</i>	<i>yo</i>	<i>leo</i>	<i>eayo</i>
sein	<i>babt</i>	<i>γwaat</i>	<i>leaat</i>	<i>eayaat</i>

Tab. 103.3: Verwandtschaftsbezeichnungen im Ik

Vergleichbares findet sich auch im Dakota und im Mbangi. Verwandtschaftsbezeichnungen weisen in vielen Sprachen Eigenschaften auf, die sie von inalienablen Teil-Ganzes-Beziehungen unterscheiden. So werden sie z. B. im Guugui Yimidhurr nicht inalienabel konstruiert, und es erscheint von daher sinnvoll, zwischen diesen beiden Arten von Relationen deutlich zu unterscheiden (vgl. hierzu auch Voeltz 1976).

#### 4.3. Teil-Ganzes-Beziehungen

Mit dem Problem der Inalienabilität von Teil-Ganzes-Beziehungen, insbesondere mit Bezug auf Körperteile, hat sich eine umfangreiche Literatur beschäftigt (vgl. Fillmore 1968; Frei 1939; Fox 1981; Isacenko 1965; Bierwisch 1965). Dabei geht es jedoch zumeist um syntaktische Phänomene wie z. B. die Beziehung zwischen Possession und Dativ (externe Possession; s. Payne & Barshi 1999, eds.). In vielen Sprachen wird der Possessor als direktes oder indirektes Objekt enkodiert, wenn er durch die Handlung affiziert wird: *Ich habe mir das Bein gebrochen*. Im Somali ist eine Äußerung wie (65) nicht möglich.

- (65) \**Lug-tii inan-ka ayaa*  
 Bein-ART.F Junge-ART.M FOK  
*jabtey.*  
 brech:3.SG.F  
 ‘Das Bein des Jungen ist gebrochen.’

Statt dessen muß entweder der Junge (66) oder aber das Bein (67) als Subjekt konstruiert werden und der Possessor resp. das Possessum als Partizipant – eine Tatsache, die Hyman (1977) als “possessor promotion” und Fox (1981) als “possessor ascension” bezeichnet hat.

- (66) *Inan-kii lug ayaa jabtey.*  
 Junge-ART.M Bein.F FOK brech:3.SG.F  
 ‘Ein Bein ist dem Jungen gebrochen.’

- (67) *Inan-ku lug ay-uu ka*  
 Junge-ART.M Bein.F FOK-3.SG.F ABL  
*jabtey.*  
 brech:3.SG.F  
 ‘Der Junge ist am Bein gebrochen.’  
 (= Der Junge hat ein gebrochenes Bein.) (Serzisko 1983: § 3.2; vgl. Hyman 1977; Hinnebusch & Kirsner 1980; Claudi & Serzisko 1985)

Der Possessor wird getilgt (*Possessor-deletion*), wenn er identisch ist mit dem Subjekt, wie in dem folgenden Beispiel aus dem Haya.

- (68) *N-k-ógy’ émikono.*  
 ich-PAST.3-wasch Hände  
 ‘Ich wusch meine Hände.’ (wörtl. ‘Ich wusch Hände’) (Hyman 1977: 100)

#### 5. Nominaler vs. pronominaler Possessor

Der pronominale Possessor kann durch eigenständige **Pronomina** oder durch **Possessivaffixe** gekennzeichnet werden. Was die Beziehung zwischen den adverbialen und den possessiven Formen anbetrifft, so finden wir die folgenden Regularitäten:

- (a) Pronomina werden wie Nomina behandelt und von daher konstruiert wie andere N-N-Konstruktionen auch. Ein Beispiel hierfür ist das Ik, in dem das Pronomen im Genitiv als Possessivpronomen fungiert. Darüber hinaus kann das Pronomen auch in einem Nominalkompositum verwendet werden, was seinen nominalen Status unterstreicht.

- (69) (a) *ima ntsi-e*  
 Kind 3.SG-GEN  
 ‘sein Kind’

- (b) *ntsi-ima*  
 3.SG-Kind  
 ‘sein Kind’

- (b) Es gibt eine eigene Reihe von Possessivpronomina, wie z. B. die Possessivadjektive im Deutschen oder die Possessivsuffixe im Somali. In beiden Fällen kongruieren die Pronomina mit dem Possessum in Genus und Numerus.

- (70) (a) *faras-k-ay-gii*  
 Pferd-M-POSS.1.SG-DEM.M  
 ‘(dies) mein Pferd’

- (b) *naag-t-ay-dii*  
 Frau-F-POSS.1.SG-DEM.F  
 ‘(diese) meine Frau’

Auffällig ist, daß die Possessivformen der ersten und der zweiten Person in vielen Sprachen eigene Formen haben, die dritte Person aber häufig direkt von den Personalpronomina abgeleitet ist; vgl. die Paradigmen des Deutschen und des Norwegischen:

(71) <i>ich/jeg</i>	<i>dul/du</i>	<i>er/han</i>
<i>mir/mig</i>	<i>dir/dig</i>	<i>ihm/han</i>
<i>mich/min</i>	<i>dich/dig</i>	<i>ihn/han</i>
<i>mein/min</i>	<i>dein/din</i>	<i>sein/hans</i>

Es gibt im Norwegischen zwar auch die Form *sin*, sie wird jedoch nur reflexiv verwendet.

- (72) *Han bad hende at hun<sub>i</sub> skulde hjälpe hans<sub>j</sub> soster.*  
 ‘Er bat sie, sie möge seiner Schwester helfen.’

- (73) *Han bad hende at hun<sub>i</sub> skulde hjälpe sin<sub>j</sub> soster.*  
 ‘Er bat sie, sie möge ihrer Schwester helfen.’

(c) In Sprachen, die eine Subjekt- und eine Objektreihe von Personalpronomina bzw.-affixen haben, zeigt sich, daß Possessivpronomina in der Regel Affinitäten zu der Objektreihe aufweisen. So sind z. B. im Nera (Nilosaharanisch) die Possessivpronomina identisch mit den Objektpronomina:

(74)	Subjekt	Objekt	Possessiv
	<i>ag</i>	<i>wo</i>	<i>wo-ga</i>
	<i>ipja</i>	<i>nga</i>	<i>nga-ga</i>
	<i>tib/teb</i>	<i>te</i>	<i>te-ga</i>

Der Possessor kann vor- oder nachgestellt werden. Steht er vor dem Possessum, dann sind Possessiv- und Objektpronomen identisch.

- (75) *wo wal* ‘mein Haus’  
*wal wo-ga* ‘mein Haus’  
*hasan horge* ‘Hasans Esel’  
*horge hasan-ga* ‘Hasans Esel’  
 (Thompson 1976: 486)

Im Berik (Papua) werden die Possessivpronomina von der Objektform gebildet:

(76)	Subjekt	Objekt	Possessiv
	<i>ai</i>	<i>am</i>	<i>amna</i>
	<i>ne</i>	<i>nem</i>	<i>nemena</i>
	<i>aame</i>	<i>im</i>	<i>imna</i>
	<i>je</i>	<i>jemi</i>	<i>jemena</i>

(Westrum & Wiesemann 1986: 39)

In anderen Sprachen, z. B. Navaho (77) und Arabisch (78), entsprechen sie den Objektaffixen am Verb:

- (77) *si-jaad* ‘mein Fuß’  
*n-si-nliteeh* ‘du stellst mich nieder’  
*mi-jaad* ‘dein Fuß’  
*n-ni-lteeh* ‘er stellt dich nieder’

- (78) *bet-i* ‘mein Haus’  
*bet-ek* ‘dein Haus’  
*bet-uh* ‘sein Haus’  
*darab-ni* ‘er schlägt mich’  
*darab-ek* ‘er schlägt dich’  
*darabn-ah* ‘er schlägt ihn’

Im Murle sind die Possessivpronomina identisch mit den Dativpronomina (Arensen 1982: 98):

(79)	‘mein’	‘dein’
NOM	<i>naana</i>	<i>niina</i>
AKK	<i>aneeta</i>	<i>ineeta</i>
DAT	<i>yaat-aan</i>	<i>yaat-un</i>
GEN	<i>c-an/on-an</i>	<i>c-un/un-un</i>

In Aktiv-Sprachen, in denen aktive und inaktive Pronomina unterschieden werden, werden am Nomen die inaktiven Formen verwendet, wie z. B. im Dakota (Pustet 1985: 5, 11):

- (80) *ma-ya'-kte*  
 1.SG.INAKT-2.SG.AKT-töt  
 ‘du tötest mich’  
*ma-i'* ‘mein Mund’  
*ni-i'* ‘dein Mund’

Seltener ist dagegen der Fall, daß die Possessivpronomina mit der Reihe der Pronomina übereinstimmen, die im transitiven Satz den **Aktor** kennzeichnen, so z. B. in den Maya-Sprachen, wo wir eine ergativische und eine absolute Pronominalreihe am Verb vorfinden. Am Nomen wird ausschließlich die ergativische Reihe verwendet, wie in dem folgenden Beispiel aus dem Mam (England 1983: 58, 66):

- (81) (a) *ma n-tzeeq'a-ya*  
 ASPEKT 1.SG.ERG-stoß-1.SG  
 ‘Ich stoße es.’  
 (b) *ma t-tzeeq'a*  
 ASPEKT 3.SG.ERG-stoß  
 ‘Er/Sie stößt es.’  
 (c) *n-jaa-ya* ‘mein Haus’  
*t-jaa* ‘sein Haus’

Dies ist typischerweise in Ergativsprachen wie z. B. Bella Coola, Quechua, Nenets und Takelma der Fall, so daß wir generalisieren können, daß Possessivpronomina in den Sprachen, wo zwei Reihen vorliegen, in aller Regel von der markierten Reihe genommen werden. Unterscheidet die Sprache formal zwischen alienablen und inalienablen Relationen, dann wird für die inalienable Relation die unmarkierte Reihe gewählt (vgl. hierzu auch Seiler 1982 a; 1982 b).

## 6. Unübliche Abkürzungen

FREI	freie Form
KAT	katalytisches Affix
KONSTR	Konstruktform

## 7. Zitierte Literatur

- Abbott, Miriam (1991), *Macushi*. In: Derbyshire, Desmond C. & Pullum, Geoffrey K. (Hrsg.), *Handbook of Amazonian Languages Vol. III*. Berlin, New York: Mouton de Gruyter, 23–160
- Ameka, Felix (1996), “Body Parts in Ewe Grammar”. In: Chappell, Hilary & McGregor, William (Hrsg.), *The Grammar of Inalienability: A Typological Perspective on Body Parts and Part-Whole Relation*. Berlin, New York: Mouton de Gruyter, 783–840
- Arensen, Jonathan E. (1982), *Murle Grammar*. Juba: Summer Institute of Linguistics and University of Juba (Occasional Papers in the Study of Sudanese Languages 2)
- Baron, Irene et al. (2001, Hrsg.), *Dimensions of Possession*. Amsterdam: Benjamins
- Bendix, Edward M. (1966), *Componential Analysis of General Vocabulary*. Mouton: Den Haag
- Biermann, Anna (1985), *Possession und Zuschreibung im Ungarischen*. Tübingen: Gunter Narr Verlag (Continuum 4)
- Bierwisch, Manfred (1965), “Eine Hierarchie syntaktisch-semantischer Merkmale”. *Studia grammatica 5*, 29–86
- Boeder, Winfried (1980), “‘Haben’ in den Kartwelsprachen”. In: Lehmann, Christian & Brettschneider, Gunter (Hrsg.), *Wege der Universalienforschung. Festschrift für Hansjakob Seiler zum 60. Geburtstag*. Tübingen: Gunter Narr Verlag, 207–217
- Brockelmann, Carl (1908), *Grundriß der vergleichenden Grammatik der semitischen Sprachen, Bd. I: Laut- und Formenlehre*. Berlin: Reuther & Reichard
- Cassirer, Ernst (1973/1977), *Philosophie der symbolischen Formen. Erster Teil: Die Sprache*. Darmstadt: Wissenschaftliche Buchgemeinschaft
- Chafe, Wallace L. (1963), *Handbook of the Seneca Language*. Albany: The University of the State New York
- Chappell, Hilary & McGregor, William (1989), “Alienability, Inalienability and Nominal Classification”. *Berkeley Linguistics Society: Proceedings of the Annual Meeting 15*, 24–36
- Clark, Eve V. (1978), “Locutionals: Existential, Locative and Possessive constructions”. In: Greenberg (Hrsg.), *Vol. IV*, 85–126
- Claudi, Ulrike & Heine, Bernd (1989), “On the Nominal Morphology of ‘Alienability’ in Some African Languages”. In: Newman, Paul & Botne, Robert D. (Hrsg.), *Current Approaches to African Linguistics, Vol. V*. Dordrecht: Foris, 3–19
- Claudi, Ulrike & Serzisko, Fritz (1985), “Possession in Dizi: Inalienable or not?”. *Journal of African Languages and Linguistics 7.2*, 131–154
- Cole, Peter (1982), *Imbambura Quechua*. Amsterdam: North-Holland
- Comrie, Bernard (1981), *The Languages of the Soviet Union*. Cambridge: CUP
- Creissels, Denis (1979), “Le comitatif, la coordination et les constructions dites ‘possessives’ dans quelques langues africaines”. *Annales de l’Université d’Abidjan, Serie H. Linguistique, Tome XII, Fascicule 1*
- Croft, William (1990), *Typology and Universals*. Cambridge: University Press
- Delgaty, Alfa (1961), “Notes on Dependent versus Independent Nouns in Tzotzil”. In: *A William Cameron Townsend en el vigésimoquinto aniversario del Instituto Lingüístico de Verano*. Mexiko D. F., 413–419
- Derbyshire, Desmond C. (1979), *Hixkaryana*. Amsterdam: North-Holland (Lingua Descriptive Studies 1)
- Diem, Werner (1986), “Alienable und Inalienable Possession im Semitischen”. *Zeitschrift der Deutschen Morgenländischen Gesellschaft 136.2*, 227–291
- Dixon, Robert M. W. & Blake, Barry (1979, Hrsg.), *Handbook of Australian Languages*. Canberra: The Australian National University
- Eades, Diane (1979), “Gumbaynggir”. In: Dixon & Blake (Hrsg.), 244–361
- England, Nora C. (1983), *A Grammar of Mam: A Mayan Language*. Austin: University of Texas Press

- Fillmore, Charles J. (1968), "The Case for Case". In: Bach, Emmon & Harms, Robert T. (Hrsg.), *Universals in Linguistic Theory*. New York: Holt, Rinehart & Winston, 1–88
- Fox, Barbara (1981), "Body Part Syntax: Towards a Universal Characterization". *Studies in Language* 5.3, 323–342
- Frei, Henri (1939), "Sylvie est jolie des yeux". In: Sechehaye, Albert et al. (Hrsg.), *Mélanges de linguistique offerts à Charles Bally*. Genève: Georg, 185–194
- Givón, Talmy (1984), *Syntax: A Functional-typological Introduction, Bd. I*. Amsterdam usw.: Benjamins
- Glinert, Lewis (1989), *A Grammar of Modern Hebrew*. Cambridge: University Press
- Greenberg, Joseph H. (1978, Hrsg.), *Universals of human language, Vol. 1–4*. Stanford: Stanford University Press
- Haiman, John (1985), *Natural Syntax: Iconicity and Erosion*. Cambridge: University Press
- Haviland, John (1979), "Guugu Yimidhirr". In: Dixon & Blake (Hrsg.), 27–180
- Heine, Bernd (1997), *Possession – Cognitive Sources, Forces, and Grammaticalization*. Cambridge: University Press
- Hewitt, Brian George (1979), *Abkhaz*. Amsterdam: North-Holland
- Hinds, John (1986), *Japanese*. London usw.: Croom Helm
- Hinnebusch, Thomas J. & Kirsner, Robert S. (1980), "On the Inference of Inalienable Possession in Swahili". *Journal of African Languages and Linguistics* 2, 1–16
- Hyman, Larry M. (1977), "The Syntax of Body Parts". In: Byarushengo, Ernest R. & Duranti, Alessandro & Hyman, Larry M. (1977, Hrsg.), *Haya Grammatical Structure*. Los Angeles: University of Southern California (Southern California Occasional Papers in Linguistics 6), 99–117
- Isachenko, Alexander V. (1965), "Das syntaktische Verhältnis der Bezeichnungen von Körperteilen im Deutschen". *Studia grammatica* 5, 7–27
- Isachenko, Alexander V. (1974), "On 'have' and 'be' Languages (A Typological Sketch)". In: Flier, Michael S. (Hrsg.), *Slavic Forum: Essays in Linguistics and Literature*. The Hague, Paris: Mouton, 43–77
- Jelinek, Eloise & Escalante, Fernando (1988), "'Verbless' Possessive Sentences in Yaqui". In: Shipley (Hrsg.), 411–429
- Kastenholz, Raimund (1988), "Note sur l'expression enonciative de la possession en bambara". *Mandenkan* 14–15, 193–201
- Kay, Paul & Zimmer, Karl (1976), "On the Semantics of Compounds and Genitives in English". *California Linguistics Association: Conference Proceedings* 6, 29–35
- Koptjevskaja-Tamm, Maria (2002), "Adnominal Possession in the European Languages: Form and Function". *Sprachtypologie und Universalienforschung* 55.2, 141–172
- Kukuczka, Elena (1982), "Verwandtschaft, Körperteile und Besitz: Zur Possession im Tamil". *Arbeitspapiere* 39. Köln: Institut für Sprachwissenschaft, Universität zu Köln
- Langacker, Ronald W. (1977), *Studies in Uto-Aztecan Grammar, Vol. I: An Overview of Uto-Aztecan Grammar*. Arlington: SIL Publications
- Lévy-Bruhl, Lucien (1914), "L'expression de la possession dans les langues mélanésiennes". *Mémoires de la Société de Linguistique de Paris* 19.2, 96–104
- Lewis, Geoffrey L. (1985), *Turkish Grammar*. Oxford: Clarendon Press
- Lichtenberk, Frantisek (1985), *Possessive Constructions in Oceanic Languages and in Proto-Oceanic*. In: Pawley, Andrew & Carrington, Lois (Hrsg.), *Austronesian Linguistics at the 15th Pacific Science Conference*. Canberra: Pacific Linguistics (Pacific Linguistics Series C, 88), 93–140
- Masica, Colin P. (1991), *The Indo-Aryan Languages*. Cambridge: Cambridge University Press
- McCaughay, Terence P. (1976), "The Possessive Construction in Scottish Gaelic". *Celtica* XI, 141–149
- Meinhof, Carl (1937), *Die Sprache der Herero in Dt.-Südwestafrika*. Berlin: Reimer (Deutsche Kolonialsprachen, 1)
- Mey, Jacob (1969), "Possessive and Transitive in Eskimo". *Journal of Linguistics* 6, 47–56
- Miller, George A. & Johnson-Laird, Philip N. (1976), *Language and Perception*. Cambridge/MA: Harvard University Press
- Nichols, Johanna (1988), "On Alienable and Inalienable Possession". In: Shipley (Hrsg.), 557–609
- O'Leary, de Lacy (1955), *Colloquial Arabic*. London: Routledge & Keegan Paul Ltd.
- Payne, Doris L. & Barshi, Immanuel (1999, Hrsg.), *External Possession*. Amsterdam: Benjamins
- Pasch, Helma (1986), *Die Mba-Sprachen*. Hamburg: Buske (Sprache und Geschichte in Afrika, Beiheft 6)
- Porízka, Vincenc (1972), *Hindi Language Course, Part I*. Praha: Státní pedagogické nakladatelství

- Pustet, Regine (1985), "Possession in Dakota". *Arbeitspapiere 47*. Köln: Institut für Sprachwissenschaft, Universität zu Köln
- Reh, Mechthild (1985), *Die Krongo-Sprache (niino mó-dì): Beschreibung, Texte, Wörterverzeichnis*. Berlin: Reimer (Kölner Beiträge zur Afrikanistik 12)
- Santandrea, Stefano (1976), *The Kresh Group, Aja and Baka Languages (Sudan)*. Napoli: Istituto Universitario Orientale
- Schuh, Russel (1981), "Types of Genitive Constructions in Tschadic". *Studies in African Languages, Supplement 8*, 117–121
- Seiler, Hansjakob (1982 a), "Inherent versus Established Relation, Proximity versus Obviation, and Two Types of Cahuilla Kinship Expressions". *International Journal of American Linguistics* 48.2, 185–196
- Seiler, Hansjakob (1982 b), "Possessivity, Subject and Object". *akup (Arbeiten des Kölner Universalienprojekts) 43*. Köln: Institut für Sprachwissenschaft, Universität zu Köln
- Seiler, Hansjakob (1983), *Possession*. Tübingen: Gunter Narr Verlag
- Serzisko, Fritz (1983), *Der Ausdruck der Possessivität im Somali*. Tübingen: Gunter Narr Verlag (Continuum 1)
- Serzisko, Fritz (1992), *Sprechhandlungen und Pausen: Diskursorientierte Sprachbeschreibung am Beispiel des Ik*. Tübingen: Niemeyer
- Shipley, William (1988, Hrsg.), *In Honor of Mary Haas. From the Haas Festival Conference on Native American Linguistics*. Berlin usw.: Mouton de Gruyter
- Simons, Linda (1986), "The Pronouns of To'a-baita". In: Wiesemann (Hrsg.), 21–36
- Taylor, John R. (1989), *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press
- Thompson, E. David (1976), "Nera". In: Bender, M. Lionel (Hrsg.), *The Non-Semitic Languages of Ethiopia*. East Lansing/MI: African Studies Center, Michigan State University (Monograph No. 5, Occasional Papers Series, Committee of Ethiopian Studies), 484–494
- Tucker, Archibald N. (1940), *The Eastern Sudanic Languages, Vol. I*. London usw.: Oxford University Press
- Ultan, Russel (1978), "Towards a Typology of Substantival Possession". In: Greenberg (Hrsg.), *Vol. IV*, 11–49
- Voeltz, Erhard F. K. (1976), "Inalienable Possession in Sotho". *Studies in African Linguistics Supplement 8*, 255–266
- Westrum, Peter N. & Wiesemann, Ursula (1986), "Berik Pronouns". In: Wiesemann (Hrsg.), 37–46
- Wiesemann, Ursula (1986, Hrsg.), *Pronominal Systems*. Tübingen: Gunter Narr Verlag (Continuum 6)
- Wilson, William H. (1982), *Proto-Polynesian Possessive Marking*. Canberra: Australian National University (Pacific Linguistics Series B, 85)
- Wolfart, H. Christoph & Carroll, Janet F. (1981), *Meet Cree: Guide to the Cree Language*. Edmonton: University of Alberta Press
- Young, Robert W. & Morgan, William (1980), *The Navaho Language*. Albuquerque: University of New Mexico Press

*Fritz Serzisko, Köln (Deutschland)*

## XIV. Semantische Kategorien und Operationen in der Morphologie II: Sachverhalts-, Eigenschafts- und verwandte Begriffe

Semantic categories and operations in morphology II:  
State-of-affairs, property and related concepts

### 104. State-of-affairs concepts

1. Introduction
2. States-of-affairs
3. Predicates
4. Aktionsart
5. Conclusion
6. Uncommon abbreviations
7. References

#### 1. Introduction

This article takes the notion of state-of-affairs as its point of departure and studies the ways in which this concept is encoded in the structure of languages. In 2 a basic definition of states-of-affairs is given which sets them off from other types of entity, such as individuals and propositional contents, and stresses the fact that states-of-affairs are temporal entities. The subcomponents of a state-of-affairs are (i) a property or relation as it manifests itself in time and (ii) the participants for which this property or relation holds. The linguistic correlates of these components are the **predicate** and its **argument(s)**, respectively. The structure of predicate expressions is dealt with in 3. For the structure of argument expressions see Chapter XIII. The linguistic correlate of a state-of-affairs is a **predication**. The form a predication takes may depend on the nature of the state-of-affairs it describes. This issue is dealt with in 4.

Three preliminary remarks are in order with respect to the material presented in the following sections. First of all, many of the semantic notions dealt with in this article can be expressed by syntactic and by morphological means, but only the latter are dealt with here, while the former are only men-

tioned in passing. Secondly, the semantic notions dealt with in this article may either be reflected by morphological categories, in which case they do not themselves contribute to the meaning of a construction, or they may be expressed by morphological categories, in which case they do contribute to the meaning of a construction. These cases will be distinguished where relevant in what follows. Thirdly, most of the issues dealt with in this article receive a more detailed treatment in other articles. This article restricts itself to a general overview of the basic concepts involved in the formation of descriptions of states-of-affairs, and, where relevant, provides references to the more specialized articles for a more detailed treatment.

#### 2. States-of-affairs

##### 2.1. Introduction

This section presents a further characterization of the state-of-affairs and its components (2.2) and surveys their linguistic correlates (2.3).

##### 2.2. Characterization

States-of-affairs are best characterized in terms of the threefold classification of entity types presented in Lyons (1977: 442–447; cf. also Art. 94). Lyons distinguishes three different orders of entities. An **individual** is a first order entity. It can be located in space and can be evaluated in terms of its existence. A **state-of-affairs** is a second order entity. It can be located in space and time and can be evaluated in terms of its reality. A **propositional content** is a third order entity. Being a mental construct, it can neither be located in

space nor in time. It can be evaluated in terms of its truth.

To these three types of entity we may add one more. **Properties** and **relations** may be characterized as zero order entities (cf. Hengeveld 1992; Keizer 1992; Dik 1997). These have no independent existence and can only be evaluated in terms of their applicability to other types of entity. Thus, the property ‘green’ can be applied to first order entities only, the property ‘recent’ to second order entities only, and the property ‘undeniable’ to third order entities only. Table 104.1 lists the various types of entity.

order	description	evaluation
0	property/relation	applicability
1	individual	existence
2	state-of-affairs	reality
3	propositional content	truth

Tab. 104.1: Entity types

States-of-affairs can be set off from other types of entity by the fact that they can (i) be located in time, and (ii) be characterized in terms of their reality status. States-of-affairs can thus be said to ‘(not) occur’, ‘(not) happen’, or ‘(not) be the case’ at some point or interval in time.

The subcomponents of a simple state-of-affairs are (i) a property or relation as it manifests itself in time and (ii) the individuals for which this property or relation holds. Zero order and first order entities thus enter into the constitution of second order entities. States-of-affairs, in their turn, are the subject matter of propositional contents, i.e. they are thought about, known to be (un)real, presented in a speech act, etc. Thus, second order entities enter into the constitution of third order entities.

### 2.3. Linguistic correlates

There is no one-to-one relation between the various entity types distinguished in 2.2 and the ways in which these entities manifest themselves linguistically. This is mainly due to the fact that all entity types concerned may be described by means of lexical elements. Table 104.2 lists some nominal elements that are used to designate the different orders of entities (cf. Art. 94).

order	examples
0	<i>color, weight, manner</i>
1	<i>man, chair, house</i>
2	<i>meeting, wedding, war</i>
3	<i>idea, opinion, thought</i>

Tab. 104.2: Nominal expression of entity types

In English different derivational strategies are used to form nouns designating entities of the various orders, as shown in the examples (some of which are taken from Quirk et al. 1985: 1550 f.) in Table 104.3.

order	examples
0	<i>mean-ness, kind-ness, false-ness</i> <i>elastic-ity, rapid-ity, san-ity</i>
1	<i>writ-er, employ-er, sing-er</i> <i>inhabit-ant, contest-ant</i>
2	<i>explor-ation, starv-ation</i> <i>break-age, cover-age</i>
3	<i>hope-Ø, wish-Ø, belief-Ø</i>

Tab. 104.3: Derived nominal expression of entity types

Apart from lexical expressions such as the ones given above, syntactic expressions may be used, and are indeed used more frequently, to represent states-of-affairs linguistically. These syntactic units may be called **predications**. A predication is the product of the application of a predicate to a sufficient number of arguments. Predications may occur in various syntactic environments and may take different forms, as illustrated in the following examples:

- (1) *John left.*
- (2) *John having left, we decided to cancel the meeting.*
- (3) *John<sub>i</sub> decided to Ø<sub>i</sub> leave.*

In each of the examples (1)–(3) the predicate *leave* is applied to its single argument *John*, even when this argument is not overtly present, as in (3). The resulting predication forms part of a main clause in (1), an adverbial clause in (2), and a complement clause in (3). What all these predications have in common is that they describe an entity that may be interpreted in terms of its temporal setting and in terms of its actuality status, the two criterial features of states-of-affairs.

The temporal status of the entities described in the predication in (1) and (2) is evident from the fact that the clauses in which they appear are marked for absolute (1) and relative (2) past tense, respectively. But even in the absence of such marking, as in (3), the entity described can be given a temporal interpretation. Thus, in (3) the *leaving*-event is necessarily interpreted as posterior to the *deciding*-event. The temporal status of the entities described can furthermore be made explicit by means of temporal adverbs, as in (4)–(6):

- (4) *John left yesterday.*
  - (5) *John having left the day before, we decided to cancel the meeting.*
  - (6) *John<sub>i</sub> decided Ø<sub>i</sub> to leave the next day.*
- Similar things can be said about the interpretation of the entities described in the predication in (1)–(3) in terms of their actuality: in each case the *leaving*-event has positive polarity. This becomes evident if (1)–(3) are compared with their negative counterparts in (7)–(9):
- (7) *John didn't leave.*
  - (8) *John not having left, we decided to cancel the meeting.*
  - (9) *John<sub>i</sub> decided Ø<sub>i</sub> not to leave.*

Within predication the predicate, often but not always a verb or a verbal expression, designates a zero order entity, i.e. a property or relation that holds for or between the participants designated by the arguments of this predicate. The predicate occupies a central position within the predication for two reasons. Firstly, it is the only indispensable element of a predication, as is evident from the existence of argumentless predication, as in the following example from Spanish (10):

- (10) *Lluev-e.*  
rain-3.SG.PRES  
'(It) rains.'

Secondly, grammatical categories semantically relevant to the state-of-affairs as a whole are often encoded on the (verbal) predicate. Thus, in the following example from Quechua (Cole 1982: 142) the entire state-of-affairs 'Maria's living in Agato' is to be interpreted as situated in the past, but the past tense marking is attached to the verb, i.e. the relational part of the description of the state-of-affairs:

- (11) *Marya-ka Agatu-pi-mi kawsa-rka.*  
María-TOP Agatu-in-VAL live-PAST.3  
'María lived in Agato.'

Given the centrality of predicate expressions within predication I will concentrate on the structure of predicates in the next section. For the structure of argument expressions see Chapter XIII.

### 3. Predicates

#### 3.1. Introduction

A predicate is the core element of a predication. Whereas a predication designates a state of affairs as a whole, the predicate designates the relation or property structuring the internal constitution of a predication. A **predicate** is a syntactic unit, and may be realized by a variety of **lexemes**, i.e. lexical units. Predicates, being phrasal units, may be simple or complex. Lexemes may be basic, or derived by a lexical rule.

Complex predicates include serial verb constructions, auxiliary constructions and periphrastic constructions. In serial verb constructions two lexical verbs enter into the description of a single event. In auxiliary constructions a lexical verb is modified by a non-lexical verb. In periphrastic constructions a lexical verb is modified by a verb that retains some of its lexical properties. All these cases involve meaning extensions and modifications realized by syntactic means, and will thus not be treated here.

Derived lexemes are those that are created on the basis of other lexemes, which may themselves be basic or derived. Derivational processes are discussed in more detail in Art. 89. Basic and derived lexemes may belong to various categories, which is the issue of 3.2. The valency of basic and derived lexemes is discussed in 3.3.

#### 3.2. Categories

In many languages only verbs may be used as predicates directly, but in others non-verbal predicates have this possibility too, as the following examples from Turkish (Lewis 1967: 127; Ersen-Rasch 1980: 203, 188) illustrate:

- (12) *Gel-di-m.*  
come-PAST-1.SG  
'I came.'
- (13) *İssiz-di-m.*  
unemployed-PAST-1.SG  
'I was unemployed.'

- (14) *Eskiden öğretmen-di-m.*  
formerly teacher-PAST-1.SG  
'I used to be a teacher.'

Note that the past and personal endings of *gel-* 'come' in (12), *işsiz* 'unemployed' in (13) and *ögretmen-* 'teacher' in (14) are identical. The reason to call the latter two predicates non-verbal is that the lexemes occupying the predicate slot may also be used in the construction of noun phrases.

The extent to which languages allow the direct predicative use of various categories of lexemes can be described systematically in terms of the following hierarchy (Stassen 1992; 1997; Hengeveld 1992):

- (15) V > A > N

This hierarchy says that if a language allows the direct predicative use of nouns, it will also allow the direct predicative use of adjectives and verbs; if it does not allow the direct predicative use of adjectives, it will neither allow the direct predicative use of nouns; etc.

English is, of course, an example of a language which allows the direct predicative use of verbs, but not of adjectives and nouns. In the latter two cases a copula construction is used. Examples (12)–(14) illustrate that Turkish, on the other hand, allows the direct predicative use of all three categories in (15). Guarani exemplifies the third possibility, since it allows the direct predicative use of verbs and adjectives, but uses simple juxtaposition with nominal predicates, as shown in the following examples (Gregores & Suárez 1967: 138, 173, 158):

- (16) Še-manu?á.  
1.SG-remember  
I remember.'
- (17) Šé-yemiahíi.  
1.SG-hungry  
'I am hungry.'
- (18) Né soldádo.  
you soldier  
'You are a soldier.'

The lexeme *yemiahíi* 'hungry' in (17) can be considered an adjective rather than a verb since, as Gregores & Suárez (1967: 138) note, it belongs to a class of items that 'may also occur uninflected as attributes to a noun'.

For some languages it makes little sense to distinguish the lexical classes mentioned in the hierarchy in (15). A case in point is Samoan, a language in which lexemes are not

tied to a specific syntactic slot. Consider the following examples (Mosel & Hovdhaugen 1992: 80, 73, 74).

- (19) (a) 'Ua mālosi le lā.  
PF strong ART sun  
'The sun is strong.'
- (b) 'Ua lā le aso.  
PF sun ART day  
'The day suns. (The day is sunny.)'
- (20) (a) E alu le pasi i Apia.  
GENR go ART bus DIR Apia  
'The bus goes to Apia.'
- (b) le alu o le pasi i Apia.  
ART go CONN ART bus DIR Apia  
'the going of the bus to Apia'

In Samoan the translational equivalents of English nouns can not only be used as the head of a noun phrase (19 a) but also as a predicate (19 b), whereas the translational equivalents of English verbs can not only be used as a predicate (20 a) but also as the head of a noun phrase (20 b). This is a systematic feature of Samoan, and hence it makes little sense to distinguish lexeme classes in this language. As a result, every lexeme may be used as a predicate. This is quite the opposite of what happens in languages like English, in which lexeme classes are clearly distinguished and only verbs may be used as predicates directly, i.e. without the intervention of a copula. For further information on lexeme classes see Chapter X.

Many languages apply morphological means to enable a non-verbal lexeme to occur in predicative position. In descriptive grammars this process is normally called **verbalization**. In some cases the only function of verbalization is to allow the predicative use of the non-verbal lexeme. This is illustrated in the following example from West Greenlandic (Fortescue p.c.), in which the verbalizing suffix *-u* is functionally equivalent to a copula:

- (21) Uanga Tuumasi-u-vunga.  
I Tuumasi-VR-1.SG.IND  
'I am Tuumasi.'

A similar situation obtains in Krongo, witness the following example from Reh (1985: 242).

- (22) Aakù m-àa-nímyà.  
she F-IPFV:COP-woman  
'She is a woman.'

These are cases in which the verbalizing morpheme simply indicates that the lexeme to which it attaches occupies the predicate slot. In many other cases, verbalization adds a meaning component. Consider the following examples from Kayardild (Evans 1995):

- (23) *ngarrku*  
strong  
'strong'  
  
*ngarrku-watha*  
strong-INCH.VR  
'become strong'  
  
*ngarrku-rutha*  
strong-FACT.VR  
'strengthen'

The inchoative verbalizer *-watha* added to a non-verbal lexeme expresses ingress into the state described by that lexeme; the factitive verbalizer *-rutha* expresses the causation of the state described by the lexeme.

### 3.3. Valency

Predicatively used lexemes can not only be characterized in terms of their category, but also in terms of their valency, i.e. the number of arguments they require. Languages differ widely in the extent to which they encode differences in valency lexically or grammatically. Consider the following examples from English:

- (24) (a) *The water boiled.*  
 (b) *Peter boiled the water.* (= *Peter made the water boil*)  
  
 (25) (a) *The duckling died.*  
 (b) *Peter killed the duckling.* (= *Peter made the duckling die*)

In (24) the same lexeme is used to express both an intransitive and a transitive state of affairs. In (25) two different lexemes are used. In English this is a purely lexical issue. In other languages, this difference is expressed morphologically in two different ways: (i) by using markers which reflect the valency of a construction, or (ii) by applying derivational processes which change the valency of a lexeme and add a meaning component. These two processes will be illustrated separately.

Fijian (like Samoan, as shown above) does not distinguish between clearly delimited lexeme classes. It is therefore not surprising that in this language lexemes are hardly ever intrinsically intransitive or transitive either. Instead, the intransitive or transitive use of a

lexeme is explicitly marked, as shown in the following examples from Boumaa Fijian (Dixon 1988: 34):

- (26) (a) *Au la'o.*  
 I go  
 'I am going.'  
  
 (b) *Au la'o-va*  
 I go-TR  
 'I am going for it.'  
  
 (27) (a) *Au rai.*  
 I look  
 'I am looking.'  
  
 (b) *Au rai-ca*  
 I look-TR  
 'I see him/her/it.'

In Fijian the absence of the transitivity suffix indicates that a lexeme is used as a one-place predicate. The presence of the transitivity suffix indicates that it is used as a two-place predicate. This suffix does not add a specific meaning component, its function is simply to mark the transitive use of a lexeme.

A similar phenomenon may be observed in Wolof, as discussed in Comrie (1985: 316). In this language a suffix (*-al*) added to a verb reflects the presence of an additional argument which one would not expect on the basis of the basic valency of that verb. This additional argument may have a variety of semantic functions. The following examples illustrate:

- (28) (a) *Nga dem.*  
 AUX.2.SG go  
 'You went.'  
  
 (b) *Kan nga dem-al?*  
 who AUX.2.SG go-al  
 'Who did you go with?'  
  
 (29) (a) *Mungi dyàng téére bi.*  
 PRES.3.SG read book ART  
 'He is reading the book.'  
  
 (b) *Mungi dyàng-al eleew yi téére-ém.*  
 PRES.3.SG read-al pupil ART.PL book-his  
 'He is reading his book to the pupils.'  
  
 (30) (a) *Di naa toogal nenne bi.*  
 fut AUX.1.SG seat child ART  
 'I will seat the child.'  
  
 (b) *Di naa la toogal-al nenne bi.*  
 fut AUX.1.SG you seat-al child ART  
 'I will seat the child for you.'

Note that again the suffix does not add a specific meaning component to the construction.

Next to these morphological means, which simply function as signals of the quantitative valency of a lexeme, there are derivational operations which bring about both a change in quantitative valency and in meaning (see Art. 107). The following examples are from Hungarian (de Groot 1989: 138, 141):

- (31) (a) *Mari kimos-t-a a*  
Mary wash-PAST-3.SG ART  
*ruhák-at.*  
clothes-ACC  
'Mary washed the clothes.'
- (b) *Mari-val kimos-at-t-am*  
Mary-INSTN wash-CAUS-PAST-1.SG  
*a ruhák-at.*  
ART clothes-ACC  
'I had Mary wash the clothes.'
- (32) (a) *A borbély borotválja Feri-t.*  
ART barber shave Feri-ACC  
'The barber shaves Feri.'
- (b) *Feri borotválkozik.*  
Feri shave-REFL  
'Feri shaves himself.'

The quantitative valency of the basic lexeme *kimos* in (31 a) is extended with one argument slot in the causative construction in (31 b), in which the derived lexeme *kimosat* is the predicate. The quantitative valency of the basic lexeme *borotvál* in (32 a) is reduced with one argument slot in the reflexive construction in (32 b), in which the derived lexeme *borotválkozik* is the predicate.

#### 4. Aktionsart

The combination of a predicate with the appropriate number of arguments is a predication, which designates a state of affairs. The nature of the state of affairs may determine part of the form the predication takes. Four major subclasses of states-of-affairs which may be reflected in the form in which they are expressed can be distinguished on the basis of two basic parameters: (i) control, and (ii) dynamicity (cf. Dik 1978; 1997). For a more detailed treatment and classification of types of states of affairs see Art. 109.

A state-of-affairs is controlled if a participant has the power to determine whether or not the state-of-affairs obtains. A state-of-affairs is dynamic if it involves a change. The types of states-of-affairs that may be defined

in terms of these parameters are given in Table 104.4.

	+ control	- control
+ dynamic	action	process
- dynamic	position	state

Tab. 104.4: Types of states-of-affairs

Examples of these four types of states-of-affairs are given in (33)–(36):

- (33) *John opened the door.* (action)
- (34) *John kept the door open.* (position)
- (35) *John fell ill.* (process)
- (36) *John was ill.* (state)

A language in which the parameter of dynamicity is clearly reflected in the morphological system is Abkhaz. In this language dynamic and static stems enter into different tense systems. Consider the following examples (Spruit 1986: 95, 98):

- (37) *də-z-ba-wá-yt'*.  
3.SG.M-1.SG-see-PROG-DECL  
'I see him.'
- (38) *yə-s-taxá-w-p'*.  
3.SG.IRRAT-1.SG-want-PRES-DECL  
'I want it.'

The suffix *-wá* 'progressive/situational' in (37) is one of the 'Tense A' suffixes, which only combine with dynamic stems. The suffix *-w* 'present' in (38) is one of the 'Tense B' suffixes, which only combine with non-dynamic verbs or with a dynamic verb + Tense A suffix (Spruit 1986: 116–117). The suffix *-p* 'declarative' in (38) is furthermore only used with the present tense of non-dynamic verbs.

In Abkhaz many stative intransitive stems also occur as dynamic intransitive stems. In these cases dynamicity is signalled exclusively by the tense suffixes used (Spruit 1986: 95, 96):

- (39) *d-t'w a-wá-yt'*.  
3.SG.M-sit-PROG-DECL  
'He sits down.'
- (40) *d-t'w á-w-p'*.  
3.SG.M-sit-PRES-DECL  
'He is sitting.'

The parameter of control seems to be reflected in the verbal system less frequently. Comrie (1989: 54), referring to Munro &

Gordon (1982), mentions Chickasaw as an example. In this language one finds oppositions like the following:

- (41) *Sa-ttola.*  
1.SG.UGR-fall.down  
'I fell down (by accident).'
- (42) *Ittola-li.*  
fall.down-1.SG.ACR  
'I fell down (on purpose).'

The system is, however, not fully productive.

It is probably more common to find the parameter of control reflected in the ways arguments are realized. Foley (1986: 121–127) notes that the distinction between controlled and uncontrolled states-of-affairs is pervasive in many Papuan languages. In Barai (Olson 1981), for instance, with controlled predicates only Agent arguments take a special set of modal clitics, whereas with uncontrolled predicates these are only attached to Patient arguments. Consider the following examples (cited from Foley 1986: 124):

- (43) (a) *Fu-ka na kan-ie.*  
he-really I hit-1.SG.UGR  
'He really hit me.'
- (b) *\*Fu na-ka kan-ie.*  
he I-really hit-1.SG.UGR  
'He really hit me.'
- (44) (a) *Ije na-ka visi-nam-ie.*  
it I-really sick-TR-1.SG.UGR  
'It really sickened me.'
- (b) *\*Ije-ka na visi-nam-ie.*  
it-really I sick-TR-1.SG.UGR  
'It really sickened me.'

With the controlled predicate *kan* 'hit' in (43) the modal clitic must be attached to the Agent argument *fu* 'he', whereas with the uncontrolled predicate *visi* 'sicken' in (44) it must be attached to the Patient argument *na* 'I'.

The cases discussed so far concern morphological means which reflect Aktionsart differences. But there are also derivational processes which have the effect of changing the Aktionsart of the predication in which the derived lexeme is used. Consider the following examples from Hungarian (de Groot 1989: 138–139):

- (45) (a) *Mari szép.*  
Mary pretty  
'Mary is pretty.'

- (b) *A kozmetikus*  
ART beauty.specialist  
*szép-it-i Mari-t.*  
pretty-CAUS-3.SG.DEF Mary-ACC  
'The beauty specialist makes Mary pretty.'

The adjective *szép* 'pretty' in (45 a) forms the center of a stative, non-controlled state of affairs. After applying the causative formation rule to this lexeme, the resulting verb *szépit* forms the center of a dynamic, controlled state of affairs.

## 5. Conclusion

In this article I have reviewed the basic concepts involved in the formation of state-of-affairs expressions, concentrating on the properties of predicates, of basic and derived lexemes occupying predicate slots, and of predication. More detailed treatments of these issues can be found in other articles in this handbook, as has been indicated at the relevant places. Inflectional categories characteristic of verbs, such as Tense, Mood, and Aspect are dealt with in later articles (Art. 109–111).

## 6. Uncommon abbreviations

IRRAT	irrational
VAL	validator
GENR	general tense

## 7. References

- Cole, Peter (1982), *Imbabura Quechua*. Amsterdam: North Holland (Lingua Descriptive Studies 5)
- Comrie, Bernard (1985), "Causative Verb Formation and Other Verb-Deriving Morphology". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press, 309–348
- Comrie, Bernard (?1989), *Language Universals and Linguistic Typology*. Oxford: Blackwell
- Dik, Simon C. (1978), *Functional Grammar*. Amsterdam: North-Holland Publishing Company (North Holland Linguistic Series 37)
- Dik, Simon C. (1997), *The Theory of Functional Grammar*, 2 Vols. edited by Kees Hengeveld. Berlin: Mouton de Gruyter (Functional Grammar Series 20–21)

- Dixon, R[obert] M. W. (1988), *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press
- Ersen-Rasch, Margarete I. (1980), *Türkisch für Sie: Grammatik*. München: Max Hueber Verlag
- Evans, Nicholas (1995), *A grammar of Kayardild*. Berlin: Mouton de Gruyter (Moutin Grammar Library 15)
- Foley, William A. (1986), *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press
- Gregores, Emma & Suárez, Jorge A. (1967), *A Description of Colloquial Guarani*. The Hague: Mouton
- Groot, Casper de (1989), *Predicate Structure in a Functional Grammar of Hungarian*. Dordrecht: Foris
- Hengeveld, Kees (1992), *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter (Functional Grammar Series 15)
- Keizer, M. Eveline (1992), *Reference, Predication and (In)definiteness in Functional Grammar: A Functional Approach to English Copular Sentences*. Dissertation, Free University Amsterdam
- Lewis, Geoffrey L. (1967), *Turkish Grammar*. Oxford: Oxford University Press
- Lyons, John (1977), *Semantics*. Cambridge: Cambridge University Press
- Mosel, Ulrike & Hovdhaugen, Even (1992), *Samoaan Reference Grammar*. Oslo: Scandinavian University Press
- University Press (Instituttet for sammenlignende kulturforskning, Series B: Skrifter, LXXXV)
- Munro, Pamela & Gordon, Lynn (1982), “Syntactic Relations in Western Muskogean”. *Language* 58, 81–115
- Olson, Michael (1981), *Barai Clause Junctures: Towards a Functional Theory of Interclausal Relations*. Dissertation, Australian National University
- Quirk, Randolph & Greenbaum, Sydney & Leech, Geoffrey & Svartvik, Jan (1985), *A Comprehensive Grammar of the English Language*. London, New York: Longman
- Reh, Mechtilde (1985), *Die Krongo-Sprache (niño mó-di): Beschreibung, Texte, Wörterverzeichnis*. Berlin: Reimer (Kölner Beiträge zur Afrikanistik 12)
- Spruit, Arie (1986), *Abkhaz Studies*. Dissertation, University of Leiden
- Stassen, Leon (1992), “A Hierarchy of Main Predicate Encoding”. In: Kefer, Michel & Auwera, Johan van der (eds.), *Meaning and Grammar: Cross-linguistic Perspectives*. Berlin: Mouton de Gruyter (Empirical Approaches to Language Typology 10)
- Stassen, Leon (1997), *Intransitive Predication*. Oxford: Clarendon Press

Kees Hengeveld, Amsterdam  
(The Netherlands)

## 105. Property concepts

1. Introduction
2. Functions of property concept forms
3. How languages express property concepts
4. Characteristics of property concepts
5. References

### 1. Introduction

The idea of property concepts arises from a desire to identify a notional basis for the grammatical category **adjective**, which can be isolated for many, though by no means all, languages (cf. Art. 74). “Property concepts” are of interest, therefore, both because of this concern to establish the notional basis for the widespread lexical category of adjective, and because of an interest in exploring how languages with no category of “adjective” (or a

restricted such category) express the conceptual material expressed by adjectives in other languages.

We may assume that “it is possible to draw a distinction between the relatively simple perceptual properties which are distributed among individuals and the more complex conjunctions and disjunctions of properties in terms of which individuals are categorized as members of particular classes” (Lyons 1977: 447).

We may also assume that certain semantic types are probably linguistic universals, and each type has certain basic associations with a certain lexical class (Dixon 1977). These universal semantic types might include MOTION, AFFECT, GIVING, OBJECTS, DIMENSION, COLOR, VALUE, etc. Each language arranges

these types into a small number of groups, which are known as its “parts of speech”, or its **lexical classes**. MOTION, AFFECT, GIVING, and other types are generally classed together in the class called Verb; OBJECTS and other types are generally classed together in the group called Noun. And DIMENSION, COLOR, VALUE, and other types are typically expressed by members of the adjective class (Jespersen 1924: 74; Dixon 1977).

For languages which have the major class Adjective, the semantic content of the class is fairly consistent from language to language. Thus an adjective in English will normally be translated by an adjective in the Australian language Dyirbal (Dixon 1977: 20).

This set of ideas has been called **semantic types for adjectives** (Dixon 1977) or **property concepts** (Thompson 1988).

In order to establish a universal characterization for property concepts, Dixon (1977) proposes to examine a language with a class of adjectives, and determine what its basic members are. Basic members are those which are monomorphemic and not derived from another lexical item. Thus, *washable* would not be considered a basic member of the adjective class for English, but *limp* would be. Looking at English, then, which is a language with a well-defined adjective class, he arrives at seven universal semantic types for adjectives (Dixon 1977: 31):

- (a) DIMENSION: e.g., *big, large, little, wide, thin*
- (b) PHYSICAL PROPERTY: e.g., *hard, soft, rough, smooth*
- (c) COLOR
- (d) HUMAN PROPENSITY: e.g., *jealous, happy, kind, clever, wicked*
- (e) AGE: e.g., *new, young, old*
- (f) VALUE: e.g., *good, bad, pure, delicious, fine*
- (g) SPEED: e.g., *fast, slow, quick*

These seven types are found to be expressed by the class of adjectives when the language has such a class. In addition, however, such a class may express many other property concepts, including many non-basic ones. To take the example given above, *washable* could be said to express a property concept since it names a property or quality of an entity, though it is not one of the seven basic property concepts identified by Dixon.

## 2. Functions of property concept forms

Some writers have proposed “modification” as the primary function for property concept forms (Croft 1990; 1991; 2000; Bhat 1994). However, since this term is not defined by these scholars and since it is difficult to characterize it in communicative terms, it is preferable to describe property concept forms as having two primary communicative functions: as **predicates** and as **attributes** in referring (Thompson 1988). (Indeed, Croft (1990; 1991: 52) essentially defines “modification” in terms of referring and predicating functions.) English predicate adjectives illustrate the predicate function of property concept forms:

- (1) *we were hungry*

where the property concept *hungry* constitutes the semantic predicate. An attributive property concept form plays a role in the act of referring and is grammatically part of a noun phrase, as in the French:

- (2) *les beaux villages*  
ART beautiful villages  
'the beautiful villages'

## 3. How languages express property concepts

### 3.1. Property concepts as adjectives

Property concepts tend to be expressed by the class of adjectives for languages which have such a class. A language can be said to have a class of adjectives if it has a set of lexical items expressing property concepts which can be distinguished on morpho-syntactic grounds from the classes of noun and verb (Dixon 1977: 63).

For some languages, including most Indo-European languages, the class of adjectives is an **open class**, that is, one to which new members may be added by speakers of the language, either by morphological means or by borrowing or creating new forms. For others, including many languages of Africa, the class of adjectives is a **closed class**, meaning that its membership is restricted and cannot be augmented at will.

An example of a language with a closed class of adjectives is Igbo, a Kwa language of the Niger-Congo family spoken in Nigeria (Dixon 1977: 20 f.; Welmers 1973). Morpho-

syntactic criteria support the recognition of an adjective class with just eight members. A survey of twenty languages with small adjective classes showed that the same property concepts tend to recur; the most common are forms meaning ‘large’, ‘small’, ‘black’, and ‘white’. The survey can be summarized as follows (Dixon 1977):

property concept	no. of languages	property concept	no. of languages
‘large’	20	‘good’	13
‘small’	19	‘bad’	14
‘long’	14	‘black’	13
‘short’	15	‘white’	14
‘new’	15	‘red’	8
‘old’	14	‘raw’	7

Tab. 105.1: Property concepts in a survey of twenty languages (Dixon 1977: 23)

For languages with a class of adjectives, predicated property concepts may be expressed as either nouns or adjectives; in such cases, the choice may often rest on whether the speaker is creating a category, or adding a feature without creating a category. Thus, in English, for example, the property of having certain hair color is expressible as a predicate noun (*a blonde*, *a redhead*, but not *\*a brown* or *\*a brownhead*) on the grounds that socially and culturally, people have tended to be categorized according to these particular hair colors (Wierzbicka 1986).

An interesting question is how languages without a class of adjectives or with only a small closed adjective class express property concepts (cf. Dixon 1977; Schachter 1985; Thompson 1988; Wetzer 1996; Croft 2000). The rest of this section addresses this question.

### 3.2. Property concepts as a sub-class of verbs

In many languages with no separate category of adjective, property concepts are expressed as a sub-class of verb. In such languages the property concept form is sometimes called **stative verb** or **adjectival verb**. An example of such a language is Mandarin Chinese:

- (3) *ta hen gao*  
3.SG very tall  
's/he is very tall'

The form *gao* in this example can be identified as a verb according to the following three criteria:

- (a) there is no copular verb form distinguishing this form from other verb forms in the language, as in the English translation;
- (b) forms such as *gao* can occur with a range of auxiliaries, modal, and adverbial elements just as other verbs can. Compare (3) with (4) below; note that both can occur with *hen* ‘very’:

- (4) *ta hen ai wo*  
3.SG very love 1.SG  
's/he loves me very much'

Examples (5) and (6) below show that forms such as *gao* ‘tall’ can occur with the inchoative suffix just as verbs do:

- (5) *ta gao-qilai le*  
3.SG tall-INCH PAST  
's/he suddenly got tall'

- (6) *ta ku-qilai le*  
3.SG cry-INCH PAST  
's/he started to cry'

Examples (7) and (8) show that forms such as *gao* ‘tall’ are also negated exactly as verbs are:

- (7) *ta bu gao*  
3.SG NEG tall  
's/he is not tall'

- (8) *ta bu ku*  
3.SG NEG cry  
's/he isn't crying'

- (c) in **attributive** modification, there may be no distinction between verbs and forms such as *gao*; the grammar of **relative clauses** is used for attributive modification of both types:

- (9) *gao de ren*  
tall REL person  
'tall person'

- (10) *ku de ren*  
cry REL person  
'person who is crying'

These examples show that in Mandarin Chinese, as in many other languages of Asia, at least for many forms, property concept words can be considered as a sub-class of verbs. (Mandarin does have some forms which suggest a closed class of adjective; see Li & Thompson 1981 for discussion.)

### 3.3. Property concepts as a sub-class of nouns

For other languages with no class of adjective, property concepts are expressed by forms closely resembling nouns. That is, they are inflected for the same categories as are nouns in that language, particularly case, number, and gender. For such languages a cover term which includes both sub-classes is often used. The term **nominal** is thus sometimes used for forms which inflect similarly, including nouns, adjectives, pronouns, and numerals (Karlsson 1983); the term **noun** has also been used to include **substantives** and **adjectives** (Jespersen 1924). Current typological usage disfavors the use of such a cover term.

Finnish is an example of a language in which property concept forms may best be considered a sub-class of nouns; thus adjectives agree with the nouns they modify and take the same range of inflectional categories as nouns do.

- (11) *auto on sininen*  
car:NOM COP blue:NOM  
'the car is blue'
- (12) *isossa autossa*  
big-INNESS car-INNESS  
'in the big car'

Furthermore, property concepts functioning as nouns occur with the copula just as nouns do; compare (11) with (13):

- (13) *Pekka on mies*  
Pekka COP man  
'Pekka is a man'

In such languages, however, there are nearly always two important distinctions between such property concept forms and nouns:

- (a) for languages with gender, nouns generally have inherent gender, while property concept forms have variable gender, and show the gender of the noun they agree with (Jespersen 1924: 72);
- (b) property concept forms can typically take comparative and superlative morphology, while nouns cannot:

- (14) *minun auto-ni on iso-mpi*  
1.SG.POSS car-POSS COP big-CMPR  
*kuin sinun*  
than 2.SG.POSS  
'my car is bigger than yours'

Thus in such languages property concept forms cannot be said to be a sub-class of

nouns as clearly as property concept forms can be a sub-class of verbs for languages in which they pattern like verbs.

### 3.4. Mixed systems

Some languages express property concepts in a variety of ways, with no one mode of expression being dominant. In Bembe, for example, a language with a closed adjective class containing about twelve members, most property concepts not finding expression among these twelve are expressed by means of abstract nouns (of the type *a woman of beauty*), while color forms are classified as verbs. In Japanese property concepts are expressed by a sub-class of verbs, with different inflectional properties; while large (several hundred members), this sub-class is closed. In addition, there are several other classes of property concept forms (Uehara 1998; Croft 2000), including a class of "adjectival nouns" (Martin 1975) or "nominal adjectives" (Shibatani 1990) (e.g., *kirei* 'pretty'), which take the copula. However, most 'human propensity' property concepts are expressed by nouns (e.g., *shinsetu* 'kindness', *jiman* 'pride'), and many physical property concepts are expressed by ordinary verbs (e.g., *shimetta* 'wet') (Dixon 1977: 48 f.).

### 3.5. Derived property concepts

Many languages have grammatical devices for deriving property concept words from other lexical classes. These derivational options may be of two types: The first type involves the addition of a suffix expressing aspectual meanings of a verb; these are commonly called "participles", with meanings such as 'having the property of currently being engaged in' or 'having the property of resulting from' a given action. For example, in the English *the broken glass*, the participle *broken* attributes the property of resulting from breaking to the 'glass'. The second type of derivation of property concepts involves the addition of an affix with specific non-aspectual meaning to a noun or verb; in English, for example, the suffix *-able* derives property concept words (adjectives) from verbs, yielding the meaning 'having the property of being able to undergo, a given action, as in *erase* > *erasable*. In German, the suffix *-lich* derives property concept words (adjectives) from nouns, with the resulting meaning 'having the property of being like' a given object, entity, or concept, as in *friedlich* 'peaceful' < *Friede* 'peace'.

#### 4. Characteristics of property concepts

##### 4.1. Semantic properties

###### 4.1.1. Oppositions

Property concept forms show two kinds of semantic oppositions: **antonymy** and **complementarity** (Lyons 1968: 460–470; Dixon 1977: 31–34; Croft 1990). In the case of pairs of forms which are complementaries, such as *single/married*, *male/female*, the denial of one implies the assertion of the other. In the case of antonymy, however, as in the pairs *large/small*, *hot/cold*, the antonyms are always ‘relative’ to some implicit norm, or **gradable** (cf. Art. 114).

“... such words as *big* and *small*, or *good* or *bad*, do not refer to independent, ‘opposite’ qualities, but are merely lexical devices for grading as ‘more than’ or ‘less than’ some implicit norm.” (Lyons 1977: 466)

That is, most property concepts express **degree**. One grammatical correlate of the ability to express degree is the ability to be morpho-syntactically compared (e.g., *taller*) or superlativized (e.g., *tallest*).

The vast majority of property concepts are of this type and readily form antonym pairs. The set of property concepts denoting color, however, are among those which are not best described in terms of antonymy, but can instead be said to form a complement set: although denial of one does not imply assertion of any other, if we assert one, it implies denial of all the others. Human propensity property concepts, such as *jealous*, do not generally have monomorphemic antonyms (Dixon 1977: 34); it appears that such property concepts are used to describe certain human proclivities relative to the norm, which is the lack of such proclivities.

For certain pairs of antonymous pairs of property concept forms, one member is generally unmarked (Givón 1970; Clark 1973; Dixon 1977: 33). The unmarked term is the one which is used in neutral questions; thus (15 a), with the unmarked *long*, can be used without implying anything about the stick’s length, but (15 b) includes an expectation that it is relatively short:

- (15) (a) *How long is the stick?*  
 (b) *How short is the stick?*

The unmarked member of such a pair is also the form that appears in quantifying a property, as illustrated in (16):

- (16) (a) *The stick is 3 feet long.*  
 (b) *?The stick is 3 feet short.*

Property concepts of dimension best illustrate markedness value differences of the sort shown in (15) and (16). For property concepts of physical property or human propensity, markedness values may be less clear; thus, for example, both (17 a) and (17 b) are equally useful in ordinary speech, and neither of the property concepts expressed by these adjectives is typically quantified:

- (17) (a) *How soft is that pillow?*  
 (b) *How hard is that pillow?*

###### 4.1.2. Stative vs. active

Property concepts may also differ from one another along the dimension of **stativity** (Givón 1970). Property concepts with stative meanings generally cannot occur with morphology indicating durative or progressive meanings. For example, the property concept ‘rich’ has a stative meaning and in English, it cannot occur in the progressive, as illustrated in (18):

- (18) (a) *?She was being rich.*  
 (b) *She was being careful.*

Active, but not stative, property concepts easily occur in imperative utterances:

- (19) (a) *?Be rich!*  
 (b) *Be careful!*

Active property concepts include ‘noisy’, ‘naughty’, ‘obedient’, ‘co-operative’, ‘abusive’, ‘reasonable’, ‘vindictive’, etc. Stative property concepts include ‘stupid’, ‘tall’, ‘deaf’, ‘blonde’, ‘shallow’, ‘far’, etc.

#### 4.2. Property concepts in everyday usage

##### 4.2.1. Variation across languages in property concepts’ meanings

Whereas the nouns in a language can reveal something of the relevant objects in the speaker’s environment, and the verbs the important cultural actions, many property concepts appear to vary in their denotation relatively little across languages. At least some of Dixon’s seven main property concepts have few if any language-specific meanings: value, dimension, age, speed, and color. Property concepts indicating human propensity, however, often reveal something of the attitudes and social norms which speakers of certain cultures find it useful to talk about. No doubt every language has sets of property concepts

indicating human propensity exhibiting the kinds of subtleties described by Dixon for Dyirbal, where there are a number of different words for ‘sad’ or ‘offended’: *guynay* indicates ‘sad or offended because a favor was not repaid’, *guymbu* means ‘sad because of a broken relationship’, *yangur* is ‘sad or offended because one was not given anything’, and *mundu* denotes ‘sad or offended because one did not get something that one had a right to expect’ (Dixon 1977: 67).

An important contribution to the question of the meanings of property concept words is provided by Englebretson (to appear), whose survey of adjectives in English conversation reveals very little correspondence between Dixon’s ‘semantic types’ and the meanings exhibited by these adjectives. He further shows that the most frequent adjectives in English conversation are highly lexicalized forms, such as *little*, with social meanings which are as significant as their ‘denotations’.

#### 4.2.2. Property concepts in discourse

Recent discourse research has revealed a number of important characteristics of property concepts not immediately evident from considering decontextualized examples. Thompson (1988) shows that the lexical classes into which property concepts tend to fall across languages correlates with the two primary functions they have in conversation: attribution and predication. The attributing function predicts property concepts’ being classified as a sub-class of nouns, while the predicating function predicts their being classified as a sub-class of verbs. Englebretson (1997) shows that genre has a major effect on the ratio of attributing to predicating adjectives in English conversation.

## 5. References

- Bhat, D. N. S. (1994), *The Adjectival Category: Criteria for Differentiation and Identification*. Amsterdam: John Benjamins
- Clark, Herbert H. (1973), “Space, Time, Semantics, and the Child”. In: Moore, Timothy E. (ed.), *Cognitive Development and the Acquisition of Language*. New York: Academic Press, 27–63
- Croft, William (1990), “A Conceptual Framework for Grammatical Categories”. *Journal of Semantics* 7, 245–279
- Croft, William (1991), *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press
- Croft, William (2000), “Parts of Speech as Language Universals and as Language-Particular Categories”. In: Vogel, Petra M. & Comrie, Bernard (eds.), *Approaches to the Typology of Word Classes*. Berlin: Mouton de Gruyter, 65–102
- Dixon, R. M. W. (1977), “Where Have All the Adjectives Gone?”. *Studies in Language* 1, 19–80 [reprinted in: Dixon, R. M. W. (1982), *Where Have All the Adjectives Gone? and Other Essays in Syntax and Semantics*. Berlin: Mouton Publishers, 1–62]
- Englebretson, Robert, (1997), “Genre and Grammar: Predicative and Attributive Adjectives in Spoken English”. *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 23, 411–421
- Englebretson, Robert (to appear), *Why Don’t All the Adjectives Go There? Semantic Classification of Adjectives in Conversational English*. [Paper presented in CSDL 3]
- Givón, Talmy (1970), “Notes on the Semantic Structure of English Adjectives”. *Language* 46, 816–847
- Jespersen, Otto (1924), *The Philosophy of Grammar*. New York: W. W. Norton & Company
- Karlsson, Fred (1983), *Finnish Grammar*, translated by Andrew Chesterman. Helsinki: Werner Söderström Osakeyhtiö
- Li, Charles N. & Thompson, Sandra A. (1981), *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press
- Lyons, John (1968), *Theoretical Linguistics*. Cambridge: Cambridge University Press
- Lyons, John (1977), *Semantics*, Vol. 1–2. Cambridge: Cambridge University Press
- Martin, Samuel E. (1975), *A Reference Grammar of Japanese*. New Haven/CN: Yale University Press
- Schachter, Paul (1985), “Parts-of-speech Systems”. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. I*. Cambridge: Cambridge University Press, 3–61
- Shibatani, Masayoshi (1990), *The Languages of Japan*. Cambridge: Cambridge University Press
- Thompson, Sandra A. (1988), “A Discourse Approach to the Cross-Linguistic Category ‘Adjective’”. In: Hawkins, John (ed.), *Explanations for Language Universals*. Oxford: Basil Blackwell, 167–185 [reprinted in: Corrigan, Roberta & Eckman, Fred & Noonan, Michael (1989, eds.) *Lin-*

- guistic Categorization.* Amsterdam: Benjamins, 245–265]
- Uehara, Satoshi (1998), *Syntactic Categories in Japanese: A Typological and Cognitive Introduction.* Tokyo: Kuroshio
- Welmers, William E. (1973), *African Language Structures.* Berkeley, Los Angeles: University of California Press

Wetzer, Harrie (1996), *The Typology of Adjectival Predication.* Berlin: Mouton de Gruyter

Wierzbicka, Anna (1986), “What’s in a Noun?”. *Studies in Language* 10, 353–389 [reprinted in: Wierzbicka, Anna (1988), *The Semantics of Grammar.* Amsterdam: John Benjamins, 463–497]

*Sandra A. Thompson,  
Santa Barbara (U.S.A.)*

## 106. Circumstance concepts

1. Introduction
2. Types of circumstance concepts
3. Head-marking expression of circumstantial meanings
4. Dependent-marking expression of circumstance meanings
5. Uncommon abbreviations
6. References

### 1. Introduction

#### 1.1. Delimiting the topic

All grammatical theories have some way of recognizing two kinds of semantic and/or syntactic relationships that elements subordinated to a verb can have to their lexical head. Paired terms for expressing this opposition include center/core/nucleus vs. periphery; complements vs. modifiers of the verb; subjects-and-complements vs. adjuncts; or actant role/participant role vs. circumstantial role. This article deals with the kinds of semantic notions that are typically associated with the second members of such lists.

The tradition that gives us the term **circumstantial** goes back to the distinction made by Lucien Tesnière (1959: 102 f.) between *actants* and *circonstants*:

“Les actants sont les êtres ou les choses qui, à un titre quelconque et de quelque façon que ce soit, même au titre de simples figurants et de la façon la plus passive, participent au procès. (...) Les actants sont toujours des substantifs ou des équivalents de substantifs. Inversement les substantifs assument en principe toujours dans la phrase la fonction d’actants.” (Tesnière 1959: 102)

“les circonstants expriment les circonstances de temps, lieu, manière, etc ... dans lesquelles se déroule le procès. Ainsi dans la phrase fr. *Alfred fourre toujours son nez partout* (...), il y a deux circonstants, un de temps (*toujours*) et un de lieu (*par-*

*tout*). Les circonstants sont toujours des adverbes (de temps, de lieu, de manière, etc.) ou des équivalents d’adverbes. Inversement les adverbes assument en principe toujours dans la phrase la fonction de circonstants.” (Tesnière 1959: 102 f.)

Two characteristics typical of most later discussions of this topic are exhibited in this quotation: (i) a tendency to blend syntactic categories (actants are substantives) and semantic roles (actants participate in a process), and (ii) the incomplete list “time, place, manner, etc.” For our purposes it will be useful to begin the delimitation of our topic by reviewing four closely but imperfectly correlated ways of subdividing the components of a clause and/or of its meaning, especially in relation to a verb and its dependents.

The first is the extralinguistic distinction between “things” and “circumstances”. In the interpretation of a given sentence one distinguishes (i) those entities which are seen as participating in an event or state of affairs expressed by the sentence, and (ii) circumstances surrounding or accompanying the state of affairs such as its time of occurrence, its general spatial setting, aspects of manner, the participants’ intentions, and the like.

The second is a distinction based on the meaning or semantic frame associated with a clause’s head verb, a distinction between **frame-internal** and **frame-external** elements. Certain verbs or verb-stems require the presence of particular elements in their semantic valence by virtue of their meaning. Such elements are termed frame-internal. Other elements that may be expressed in the same clause are either optional embellishments of the scene (*through a megaphone, with his brother*) or indications of the scene’s spatial and temporal coordinates (*in the garden, yes-*

*terday morning); these are frame-external.* (Time of occurrence and spatial setting may be a necessary part of an event in the real world, but these features need not be specific and necessary to the meaning of any particular verb which expresses that event.)

This second division is orthogonal to the first. Specifications of space, time, or manner (“circumstances”) can in fact be inherent to the meanings of certain verbs and hence frame-internal (*the troops occupied the chapel [place], we approached the entrance [goal], the war lasted 100 days [temporal extent], I spent the week-end finishing the manuscript [temporal location], we worded our response carelessly [manner]*); and so on. Conversely, frame-external elements of a clause often include reference to extralinguistic objects (*on the table, with a friend*).

A third distinction is based on the grammatical relations which syntactic constituents can have in their clause. In particular, this involves distinguishing expressions which have ‘direct’ syntactic relations to the verb (e.g., subject or object), from others which have indirect or ‘oblique’ relations with the verb (e.g., in English, adverbs and prepositional adjuncts). In familiar languages, this has the effect of splitting constituent types into nominals vs. adverbials. The distinction corresponds generally to a traditional terminological distinction in descriptions of case languages separating grammatical cases from concrete cases. (There are no theory-neutral names for this contrast.)

Fourthly, it is possible to make a distinction between **obligatory** and **optional** elements in a sentence. An obligatory element is one which cannot be omitted without disturbing either the grammaticality of the sentence or the nature of the syntactic and semantic relations obtaining among the remaining constituents. An optional element is one which could be missing without affecting such properties of the remaining sentence. This distinction is partly orthogonal to the other three, and in particular to the second (frame internal/external), though of course elements which are semantically unnecessary will also be syntactically optional.

The Tesnièrean distinction between *actants* and *circonstants* represents a confusing amalgam of these four methods of sorting out the dependent constituents of a clause. All four can be problematic, and in those instances where they are not problematic, they frequently yield different results. Such confu-

sion is possible precisely because in the simplest cases the criteria do overlap: the elements of a clause which have circumstantial meaning are frame-external rather than frame-internal, have an oblique or adverbial relation to their verb, and are syntactically optional.

For this article we will take the first of the above criteria as definitional. Circumstantial elements will be understood in terms of real-world semantics, as specifying conditions of “time, place, manner, etc.” independently of questions of semantic valence, syntactic verb government, and optionality. This means, for example, that place notions that are conceptually and syntactically necessary dependents of verbs of motion will be considered as part of our topic; this is necessary because many morphological systems for expressing spatial notions apply equally well whether the place specification is a part of a verb’s valence or not. Further, it also means that semantically circumstantial morphemes which occur within the verb itself are to be included in our topic.

Even with this attempted semantic delimitation of circumstance concepts, there remain areas of uncertainty. It is not possible to divorce the specification of circumstances from the specification of things if only because many circumstances are most naturally stated with reference to specific things. Consider expressions which could be glossed as ‘for you’ or ‘with one’s foot’. Each of these can easily be interpreted in “thing”-oriented fashion, as semantically inserting a new entity into the scene represented by the verb, augmenting the cast of characters or the set of props. If the morphological means for introducing these elements involves a head-marking device, e.g. verbal affixation, it would be easy to see the process as an instance of **incorporation** (Art. 88) or **valence-augmenting derivation** (Art. 107). But suppose those same expressions were to be glossed respectively as ‘while having one’s interlocutor’s interests in mind’ or ‘by moving against [something] with one’s foot, in the prototypical way in which one moves or strikes something with one’s foot’. In that case they would appear to metamorphose into something much less “thing”-like and more immediately circumstantial, i.e., as directly expressive of a participant’s attitude or a manner of acting. Since distinctions among analysts’ glossing practices need not be accompanied by any real distinctions in the cat-

egories themselves, we have no recourse but to recognize inevitable areas of overlap between the two domains.

In the end it is necessary to resort to a list. This article, then, will be concerned with the inflectional or derivational means of expressing notions of the following types:

- the location in space of the participants in the event or of the event as a whole;
- locations with reference to which a movement is understood – source, path, goal;
- the region of time within which the event occurs; the temporal extent of an event;
- the manner in which a process takes place; the intentions of the major actor in an event and the effect on the event of such intentions;
- the means or instrumentality by which an actor carries out an act;
- etc.

No such list can avoid the “etc”. It is reasonable and natural to believe that the number of types of circumstantial elements is unlimited (cf. Tesnière 1959: 125).

## 1.2. Connections with other linguistic systems

The affixal expression of circumstantial concepts links up with other components of a linguistic system in three ways. Problems of determining the status of a given morpheme are touched on in 1.2.1; relations with deixis, tense, and aspect in 1.2.2; relations with valence phenomena in general in 1.2.3.

### 1.2.1. Affixes, clitics, particles or stems

Circumstantial information can be conveyed as part of the semantic structures of morphologically simple lexical heads, or through adpositions, clitics, particles, and affixes. Morphologically simple verbs in English which incorporate notions of place, time, or manner include *approach* and *surround* [place], *last* and *endure* [time], *slap* and *crawl* [manner]. Morphologically simple adverbs containing analogous semantic notions include *here*, *then*, *fast*, etc. Morphologically complex adverbs with circumstantial meanings include *slowly*, *downwards*, etc. Phrases with circumstantial meanings can be headed by simple prepositions, as in *toward the door*, *between her parents*, *with a pencil*, etc., or by prepositional complexes, as in *on behalf of the children*, *by means of trickery*, etc.

Even within a single grammar, and even when a language offers clear criteria for dis-

tinguishing affixes, clitics, and adpositions, a description of the purely affixal devices for expressing circumstance concepts cannot always be kept apart from other components of a grammar. There are several reasons for this. Multiple systems for expressing circumstance concepts can coexist in the same language, often using the same or clearly related morphological material; closely related languages can differ from each other in respect to the morphosyntactic status of cognate material; and, in many cases, analysts reach, or competing frameworks lead to, different conclusions on whether particular elements should count as affixes or clitics or lexical heads.

In Rama (Chibchan, Nicaragua), for example, verb prefixes with circumstantial meanings show clear relationships with postpositions used in the same language. The list of corresponding postpositions and prefixes is given in Tab. 106.1 (Craig & Hale 1988: 320). It would be clearly misleading to describe the two morpho-semantic systems separately.

post-positions	prefixes	meaning
<i>bang</i>	<i>bu-</i>	'goal, purpose'
<i>u</i>	<i>yu-</i>	'associative, instrumental'
<i>kang</i>	<i>k(a)-</i>	'ablative, source'
<i>su</i>	<i>su-</i>	'locative'
<i>aak</i>	<i>yaa-(y-a-)</i>	'dative'

Tab. 106.1: Rama postpositions and verb prefixes

In Stanley Newman's reconstruction of spatial prefixes in Salishan, he finds the daughter language cognates appearing variably as clitics, particles, stems, or true prefixes (Newman 1976: 237). Under the heading of headward migration, Johanna Nichols reviews cases of dependent elements and their markers migrating historically to their verbal heads, perhaps through stages as clitics, to become affixes on the verb (Nichols 1986: 84). Given such realities, there must be cases of systems in transition for which it is in principle impossible to tell whether something is or is not "really" an affix.

### 1.2.2. Other semantic systems: tense, aspect, deixis

Temporal locating expressions (e.g., *tomorrow*) can have specific co-occurrence relations with given tenses (e.g., future), just as specific

kinds of temporal extent expressions can have co-occurrence relations with specific categories of verbal aspect (cf. Art. 109 on aspect, Art. 110 on tense). When such information is incorporated into a verb, it comes to belong to the study of tense or aspect. Morphological markings of tenses can sometimes have quite specific meanings ('yesterday', 'since the last full moon'), in paradigmatic alternation with simpler or more abstract tense notions (cf. Comrie 1985: 83–101). Such phenomena are dealt with in Art. 110.

That part of **deixis** which appeals to knowledge of the location of the speaker or the addressee, or of the time of speaking, overlaps with the spatial and temporal aspects of circumstantial. In many cases such features are embedded in complex morphological structures, sometimes in lexical verbs, and sometimes in adverbs, in which deictic information is just one among many morphosemantic systems operating together, in tandem with such notions as recognition of the shape of a reference object, features of the local topography, etc. Further, many complex morphological systems in the structure of verbs contain directional information, including information about identification of the speaker with the starting point or the destination of the motional aspects of the verb's meaning. Description of such systems can be found in Art. 95, on deixis.

### 1.2.3. Valence

In the study of verb valence the question can arise as to whether given circumstantial elements are internal or external to the grammatical valence of a given verb, or to verbs of a particular class, or to verbs in general in the language. To the extent that such considerations, within a given theoretical framework, are equivalent to concerns about the difference between central or peripheral elements of the clause, certain concepts would appear to be at the borderline. The Papuan languages tend to register central constituents affixally on the verb but peripheral constituents on nominals, by means of case suffixes; but the concept 'recipient' is represented as central (with head-marking) in some languages and as peripheral (with dependent-marking) in others (Foley 1986: 96–99). For theory-specific reasons, to account for what they see as the incorporation of postpositions into the verb as prefixes, Craig & Hale (1988: 333–335) suggest that

instrumental, comitative, and benefactive relations are valence-internal in some languages (governed by the verb) but not in others – specifically not in languages for which a verb phrase constituent appears not to be justified. All of this variability is further reason for identifying circumstantial concepts independently of facts about their grammatical realization.

## 2. Types of circumstance concepts

In exploring the kinds of concepts that figure in the general area of circumstantial, we treat notions of place in 2.1, and notions of time – very sketchily – in 2.2. Circumstantial notions involving entities which are not specifically sentient beings (manner, means, instrument) are dealt with in 2.3, and those which do involve sentient beings are discussed in 2.4 (beneficiary, recipient). Several additional circumstantial notions are simply listed in 2.5.

### 2.1. Place categories

A locating expression can be frame-internal or frame-external; i.e., it can identify a location that is relevant to a participant or a sub-event in the frame signalled by the verb, or a spatial setting for the event as a whole. This distinction is reflected in the selection of Korean locative postpositions: when speaking of cooking eggs *in a pot* one chooses a different postposition than when speaking of cooking eggs *in the garden*. This distinction is shown in (1) (data from Jeong-Woon Park).

- (1) *Inho-ka cengwen-eyse*  
*Inho-NOM garden-LOC.EXL*  
*naympi-eytaka kyecklan-ul salm-ass-ta*  
*pot-LOC.INL egg-ACC boil-PAST-IND*  
 'Inho boiled eggs in a pot in the garden.'

Locating expressions in the sphere of movement can distinguish **source** (the region from which a movement originated), **goal** (the region toward which a movement is directed), and **path** (with reference to mid-points in a trajectory). In a common Bantu pattern for verbs of motion, the unadorned verb is seen as inherently, or by default, licensing a source complement, while a morphological extension – through the applicative suffix – is needed in order to refer to the destination. The following Luganda examples are from Ashton et al. (1954: 331). The 'APPL' suffix in (2 b) and (3 b) has the effect of replacing a source complement with a goal complement.

- (2) (a) *abaana ba-dduk-a mu nnyumba*  
 children SBJ-run-FV LOC house  
 'the children are running out of the house'  
 (b) *abaana ba-dduk-ir-a mu nnyumba*  
 children SBJ-run-APPL-FV LOC house  
 'the children are running into the house'
- (3) (a) *ente zino zigob-e*  
 those cattle drive.away-IMP  
 'drive away those cattle'  
 (b) *ente zino zibo-er-e mu*  
 those cattle drive.away-APPL-IMP LOC  
*kiraalo*  
 kraal  
 'drive those cattle into the kraal'

In some case systems, case markers alone, without accompanying adpositions, can distinguish location, origin, and destination. In the Uralic languages, such distinctions combine with another system of contrasts having to do with the topological character of the reference region, differences corresponding to the distinctions represented in English with *in/into/out of* (enclosure), *on/onto/off of* (surface), *at/to/from* (neutral or point) (see discussion of Finnish in 4.1 below).

The simplest semantic notions needed to account for the less exotic of such systems can be specified by referring to the Figure (locandum) and the Ground (the reference object or region with respect to which the Figure's location is specified). The Figure can be stationary or moving; when the Figure is stationary, we speak of the Ground as **place**; when the moving Figure moves toward the Ground, we speak of the Ground as **goal**; when it moves away from the Ground, we speak of the Ground as its **source**; and when it moves past the Ground we speak of the Ground as the **path**. The Ground can be an enclosure, and the place, source, goal, or path can be seen as either inside or outside the Ground; the Ground can be a physical object in space, and the place, source, goal or path can be specified as being above or below the Ground; a physical-object Ground can additionally have specific orientations in space, and the place, source, goal, or path can be specified as being in front of, behind, or beside the Ground, given such orientations. There can be two reference points, an explicit primary one (the Ground) and an implicit one establishing a point of view (most typically that of the speaker); given the re-

cognition of such a perspective on a spatial scene, the place, source, goal, or path can be specified as closer to the viewpoint center than to the Ground (cterior, 'this side of'), or farther from the reference object than the viewpoint center (ulterior, 'the other side of'). In any of the preceding configurations, the Figure might or might not be in contact with the Ground. And to all this, we can add various possibilities about the general shape of the reference object (*along the river*, etc.) or the relative orientations of the Figure and a Ground (*along* vs. *across*, etc.) (cf. Comrie & Smith 1977: 31f.).

In a number of highly elaborate morphological systems for registering spatial information in demonstratives, adverbs, and verbs – cases famous for their complexity – a wide range of rather specific semantic factors can complicate the preceding list of fairly global notions and contrasts. For Inuktitut (Eskimo) these include: a contrast between an extended or unextended reference region; contrasts represented by the paraphrases 'up-there', 'down-there', 'in-there', 'out-there', and 'over-there'; and a distinction between regarding the speaker or someone other than the speaker as the deictic reference-point (Denny 1982: 360–363). Many such systems have elaborate means for referring to local topographic features or reference objects relevant to the local culture (Casad & Langacker 1985; Hanks 1984).

## 2.2. Time

There are numerous conceptual distinctions to be made with respect to time as it is represented in languages. Verbal coding of the time of an event relative to speech time is dealt with in Art. 110 on tense, and indications of time-span or such temporal schema notions as continuous, interrupted, bounded, unbounded, etc. are taken up in Art. 109 on aspect and aktionsart. One finds occasional morphological structures for naming days relative to the day of speaking (structured ways of forming words meaning 'today', 'tomorrow', 'the day before yesterday', 'the day after tomorrow', etc.), but these are restricted to special lexical sets. Expression of temporal notions within case-marking systems is typically parasitic on spatial uses of the same case notions (see section 3.2 below).

## 2.3. Manner, means, instrument

The concepts that go by the names **manner**, **means**, and **instrument** overlap considerably. Indeed, the English question-word *how* is am-

biguous between a manner interpretation and a means interpretation. This ambiguity underlies such jokes as "How do you push a tiger into a telephone booth?" – "Very carefully". One understands the joke by recognizing that the first speaker had 'means' in mind as the intended interpretation of *how*, the second speaker 'manner'.

The three notions are naturally associated in language for fairly obvious reasons: one can use a hammer as a tool for working on a particular task [instrument]; use of this tool enables one to accomplish the task [means]; and one can use it in the standard way in which it was intended to be used – following the typical motor program, having the usual intentions, etc. [manner].

The category of manner is extremely broad, including: accompanying stances or postures (while sitting, standing, lying); cognitive-affective involvement of an actor in an act (willingly, knowingly, grudgingly); accompanying production of visual or auditory effects (loudly, brightly); various spatio-temporal schemata (speedily, in an arc-like motion), etc.

#### 2.4. Recipient, beneficiary, comitative

Common circumstantial role names involving co-participants in an activity include **beneficiary** (one who benefits from an event), **recipient** (one who receives something), and **comitative** (one who accompanies).

Recipient and beneficiary are frequently conflated with the spatial-motional notion of goal. Again, the reasons are easy to see: the movement of an object to a region [goal] can be understood as resulting in a person located in that region coming into possession of the object [recipient], and the person who receives this object might be seen as benefiting from having it [beneficiary]. These three notions are sometimes separated from each other linguistically, often combined into a single category, sometimes split into two categories.

Comitative is frequently associated with instrumental, as for example with the English preposition *with*.

#### 2.5. Other circumstantial notions

The following is a very partial list of circumstantial notions that appear in certain grammatical descriptions, with suggestive examples from English. In each case, the prepositional phrase represents an instance of the category named: **role** (I'm doing this *as a*

*friend*); **function** (he used his shoe *as a hammer*); **purpose** (he does that *in order to impress the neighbors*); **motivation** (you said that *out of jealousy*); **cause** (we acted *from fear*); **reason** (you did that *for a good reason*); **material source** (they made it *out of wood*); **concession** (she married him *in spite of your warning*); **reference** (they spoke *about the accident*); **negative instrument** (she got the right answer *without a calculator*); etc.

### 3. Head-marking expression of circumstantial meanings

Our classification of circumstantial morphemes will be based on two considerations. Using the term 'marker' for the morpheme or morphological process which expresses a circumstantial notion, we will first recognize a distinction between the situation in which the marker is a part of the verb and one in which it is (a part of) a constituent which is dependent on the verb. The distinction here is that between **head-marking** and **dependent-marking** structures (cf. Nichols 1986) with respect to verbal heads. The second distinction, orthogonal to the first, is between markers which express nothing more than the existence of a type of circumstantial (or in some cases merely the presence of a circumstantial element, uninformative as to the type), on the one hand, and markers which provide substantive information about the circumstance itself. This yields four sub-types. An applicative suffix in a verb signalling that a benefactive notion is present is an example of the first type. A verb prefix which indicates that a movement is outward is an example of the second type. A case suffix marking its nominal host as a location is an example of the third type. A morphologically complex adverb, with components indicating direction of movement and topographic features of the object with reference to which the movement is characterized, provides an example of the fourth type.

This section deals with head-marking structures. In section 3.1 we examine the flagging of circumstantial roles within the structure of the verb; in section 3.2 we consider substantive information marked in the verb. Dependent-marking structures will be examined in section 4.

#### 3.1. Marking circumstantial relations on the verbal head

Some verb-internal markers have the effect of licensing the presence in the clause of a particular circumstantial type (3.1.1). Some reg-

ister the information that a circumstantial meaning is expressed by a constituent which is syntactically nuclear in its clause (3.1.2). And some represent the anaphoric “record” of a circumstantial meaning, while the actual constituent that the marker stands in for is not found in the constituent directly built up around the verb; this antecedent may be missing altogether, or it may occupy a topic position at the beginning of the sentence (3.1.3).

### 3.1.1. Licensing

In some cases an affix on a verbal head has the effect of licensing the presence in the clause of a constituent which otherwise could not occur at all. This is the effect of the applied verb construction in some Bantu languages: certain circumstantial elements may become a part of the clause only if the applicative suffix is chosen; without this morpheme, a sentence containing the element is ungrammatical. The applicative suffix announces that a new element is licensed; the semantic role of the new element depends on the limited number of possibilities for applicative semantics available in the particular language and the meaning of the verb. A typical sample of the possibilities for Luganda follows (from Ashton et al. 1954: 330 f.): sentence (4) is plain; in (5) the ‘APPL’ suffix licenses a beneficiary; in (6) it licenses a place; and in (7) it licenses a reason.

- (4) *ba-sekul-a kasooli*  
SBJ-pound-FV maize  
'they are pounding maize'
- (5) *ba-sekul-ir-a omwami kasooli*  
SBJ-pound-APPL-FV chief maize  
'they are pounding maize for the chief'
- (6) *ba-sekul-ir-a wano kasooli*  
SBJ-pound-APPL-FV here maize  
'they are pounding maize here'
- (7) *ba-sekul-ir-a ki kasooli?*  
SBJ-pound-APPL-FV why maize  
'why are they pounding maize'

In some languages, and for some circumstantial meanings, the newly licensed constituent becomes a nuclear argument (e.g., direct object) of the clause; in other cases it does not. In the Runyambo examples in section 3.1.3, we see that the locative licensed by the ‘APPL’ morpheme is marked obliquely.

### 3.1.2. Nuclearizing

In some cases an affix on a verbal head has the effect of nuclearizing a particular clause element: something which otherwise could only be expressed obliquely loses its oblique marking, and the verb is now seen as directly governing that element. The case of the Winnebago relational preverbs discussed in Craig & Hale (1988: 314) is different from the Luganda situation just described. Without the *hoo-* prefix on the verb, the locative element is indeed possible, but it must be expressed as an oblique adjunct. Compare (8) and (9).

- (8) *kook-eja naanzhin-jee-nan*  
box-LOC stand-AUX-DECL  
'it is standing in the box'
- (9) *kook-eja ho-nanzhin-jee-nan*  
box-DEF INESS-stand-AUX-DECL  
'it is standing in the box'

Examples of the same phenomenon can be found in Ainu (Shibatani 1990: 65 f.). The *ko-* prefix in (10 b) renders the addressee of the story-telling nuclear; the *e-* prefix in (11 b) and (12 b) nuclearizes static and dynamic locative elements.

- (10) (a) *huci matkaci orun*  
grandmother girl to  
*upaskuma*  
tell.old.stories  
'grandmother told old stories to the girl'
- (b) *huci matkaci*  
grandmother girl  
*ko-paskuma*  
APPL-tell.old.stories  
'grandmother told old stories to the girl'
- (11) (a) *poro cise ta horari*  
big house in live  
'he lives in a big house'
- (b) *poro cise e-horari*  
big house APPL-live  
'he lives in a big house'
- (12) (a) *a-kor kotan ta sirepa-an*  
1.SG-have village to arrive-1.SG  
'I arrived at my village'
- (b) *a-kor kotan a-e-sirepa*  
1.SG-have village 1.SG-APPL-arrive  
'I arrived at my village'

The so-called focus system of Tagalog is of a different sort from the applicative markings

we have just seen. In Tagalog, as in other Austronesian languages, there are prepositions that mark various nominal roles; one of these is the preposition *ang*, which identifies one nominal per clause as syntactically privileged. This is called, variously and misleadingly, subject, topic, or focus. The semantic role of that privileged nominal (within a limited typology of semantic roles) is in turn expressed in the verb by a complex system of marking that uses prefixation, suffixation, infixation, and reduplication in patterns that combine information about tense and other categories with information about the semantic role of the focused constituent. Since there is always a focused element, this system is not limited to circumstantial roles; but circumstantial do figure in the roles which can be focused. In the examples (from Schachter 1987: 941) the category-label 'TRIG' stands for trigger, from the fact that it marks the element whose semantic role triggers the verb form (Schachter 1987: 939 f.). (The substructure of the verb stem is ignored. Focused constituents are italicized in the translation.)

- (13) (a) *mag-aalis ang tindero*  
acr.trig-take.out TRIG storekeeper  
*ng bigas sa sako para\_sa babae*  
PAT rice DIR sack BEN woman  
'the storekeeper will take some rice  
out of a/the sack for a/the woman'
- (b) *aalisi-n ng tindero*  
take.out-PAT.TRIG ACR storekeeper  
*ang bigas sa sako para\_sa babae*  
TRIG rice DIR sack BEN woman  
'a/the storekeeper will take the rice  
out of the sack to a/the woman'
- (c) *aali-san ng tindero*  
take.out-DIR.TRIG ACR storekeeper  
*ng bigas ang sako para\_sa babae*  
PAT rice TRIG sack BEN woman  
'a/the storekeeper will take some rice  
out of the sack for a/the woman'
- (d) *I-pagaalis ng tindero*  
BEN.TRIG-take.out ACR storekeeper  
*ng bigas sa sako ang babae*  
PAT rice DIR sack TRIG woman  
'a/the storekeeper will take some  
rice out of a/the sack for the woman'
- (e) *I-pangaalis ng*  
INSTR.TRIG-take.out ACR  
*tindero ng bigas sa sako ang*  
storekeeper PAT rice DIR sack TRIG  
*sandok*  
scoop  
'a/the storekeeper will take some  
rice out of a/the sack with the scoop'

The semantics and pragmatics of the focus system are quite complicated. We may note that the focused element is always translated with the definite article.

### 3.1.3. Anaphoric marking

A third type of verbal marking of a circumstantial relation is represented by the Runyambo (Bantu) locative verb-final suffix (or enclitic) which stands as an anaphoric element and requires that the locative element be missing from the clause proper, appearing, if at all, only as a topicalized element (cf. Rugemalira 1994). (The analogous morpheme in Luganda, as seen in examples (2) and (3), is not subject to this same restriction.)

Some verb stems in Runyambo have valence structures which independently license locative complements; for many verbs, though, a locative can only be licensed by the presence of the applicative suffix *-ir-*. In such cases, the suffix (or enclitic) *-mu* (for enclosures) or *-ha* (for surfaces) is appended to the verb if a locative complement (valence-internal or not) does not follow the verb. This same morpheme also occurs prefixed to a noun (after an "initial vowel") to mark it as having locative role.

A two-word sentence meaning 'she danced in the house', with the locative phrase occurring clause-internally in its normal post-verbal position and the verb containing the applicative suffix, is shown in (14 b) (data from Josephat Rugemalira). The expression meaning '(in) the house' bears the locative prefix *omu-* (*o-* = initial vowel, *mu-* = location marker), hence, *omunju*.

- (14) (a) *a-ka-zaan-a*  
SBJ-PAST-dance-FV  
'he danced'
- (b) *a-ka-zaan-ir-a*  
SBJ-PAST-dance-APPL-FV  
*o-mu-nju*  
IV-LOC-house  
'he danced in the house'

An ordinary noun phrase valence element, when topicalized and represented pre-verbally, requires the verb to contain an object-marking verbal concord prefix. By contrast, a topicalized or fronted oblique locative is reflected by a verb-final suffix, *-mu* or *-ha*, producing the sentence in (15).

- (15) *o-mu-nju*  
IV-LOC-house  
*a-ka-zaan-ir-a-mu*  
SBJ-PAST-dance-APPL-FV-LOC  
'he danced in the house'

The *-ir-* licenses the presence of a new valence element, a locative being one of the possibilities; the *-mu* indicates that the new element is locative and that it is not instantiated postverbally. But there is yet another possibility, and that is to recast the house as a topocalized object, requiring an agreement prefix to be located in the object slot within the verb stem, as in (16). Here, the verb-internal prefix is *ji-*; note that *mu-* is still present.

(16) *e-nju*

IV-house  
*a-ka-ji-zaan-ir-a-mu*  
 SBJ-PAST-OBJ-dance-APPL-FV-LOC  
 'he danced in the house'

In (16), the *-ir-* licenses the presence of a new valence element; the *-mu* tells us that that new element has the semantic role 'locative' and that it is not represented as a co-constituent with the verb; and the concord marker *-ji-* indicates that the locative has the grammatical role of direct object, allowing it to appear sentence-initially as a simple noun-phrase (*enju*) rather than as an oblique (*omunjū*). This sentence, thus, exhibits all three of the functions just considered: licensing, nuclearizing, and anaphoric.

### 3.2. Marking that conveys semantically substantive information

In the examples in 3.1.1 through 3.1.3, any substantive information about the circumstantial element was provided, or understood, outside of the verb. But the incorporation of substantive circumstantial information into the verb itself is common, especially in certain languages of North America.

In Atsugewi (Hokan), circumstantial information can be incorporated into a verb, often with sufficiently substantive information to require no further specification outside of the verb. See examples (17) and (18) (Talmy 1985: 108 f.; for these examples, standard morpheme-by-morpheme interlinear glossing is not particularly helpful).

(17) /m-w-phu-lup-mik.-<sup>a</sup>/ =  
 [mphol-úp<sup>h</sup>mik·a]

Analysis:

instrumental prefix: /phu-/ 'from the mouth, working egressively'

root: /-lup-/ 'for a small shiny object to move or be located'

directional suffix: /-mik-/ 'into the face or eyes of someone'

inflectional affix set: /m- w- -<sup>a</sup>/ 'thou subject, 3d person object (factual mood)'

Analytic Translation: 'you caused it that a small shiny spherical object move into his face by acting on it with your mouth working egressively'

Paraphrase for Context: 'you spat your candy ball into his face'

(18) /'-w-ma-ss̥taq'-ipsn<sup>u</sup>-ik.-<sup>a</sup>/ =  
 [ma·staqípsnuk·a]

Analysis:

instrumental prefix: /ma-/ 'from a person's foot/feet acting on patient'

root: /-ss̥taq-/ 'for runny icky material to move or be located'

deictic suffix: /-ik-/ 'hither'

directional suffix: /-ipsn<sup>u</sup>/ 'into a volume enclosure'

inflectional affix set: /'- w- -<sup>a</sup>/ '3d person subject (factual mood)'

Analytic Translation: 'He caused it that runny icky material move hither into a volume enclosure by acting on it with his feet'

Paraphrase for Context: 'he tracked up the house (coming in with muddy feet)'

In these examples the instrumental prefix (*phu-* in (17), *ma-* in (18)) is selected from a paradigmatic set expressing a variety of concepts of the type manner, means, cause, or instrumentality. A partial selection is given in Tab. 106.2 (cf. Talmy 1985: 112 f.).

---

#### natural forces

---

<i>ca-</i>	'from the wind blowing on patient'
<i>cu-</i>	'from flowing liquid acting on patient (a river on a bank)'
<i>ra-</i>	'from a substance exerting steady pressure on patient (gas in the stomach)'
<i>uh-</i>	'from the weight of a substance bearing down on patient (snow on a limb)'
<i>miw-</i>	'from heat/fire acting on patient'

---

#### objects in action

---

<i>uh-</i>	'from a linear object acting circumpivotally (swinging) on patient (in pounding, chopping, batting)'
<i>ta-</i>	'from a linear object acting within a liquid patient (in stirring, paddling)'
<i>mi-</i>	'from a knife cutting into patient'

---

body parts in action	
<i>tu-</i>	'from the hand(s) moving'
<i>ma-</i>	'from the foot/feet acting on patient'
<i>ti-</i>	'from the buttocks acting on patient'
<i>wi-</i>	'from the teeth acting on patient'
<i>pri-</i>	'from the mouth working ingressively on patient (sucking, swallowing)'
<i>phu-</i>	'from the mouth working egressively on patient (spitting, blowing)'

sensations	
<i>sa-</i>	'from the visual aspect of an object acting on experiencer'
<i>ka-</i>	'from the auditory aspect of an object acting on experiencer'
<i>tu-</i>	'from the feel of an object acting on experiencer'
<i>pri-</i>	'from the taste/smell of an object acting on experiencer'

Tab. 106.2: Atsugewi instrumental prefixes

The directional suffixes (-*mik* in (17) and -*ipsn"* in (18)) are chosen from a paradigm that identifies what Talmy calls path and ground, simultaneously specifying information about both the direction and the concrete nature of the destination of the motion or action. A sample is given in Tab. 106.3.

directional suffixes	
- <i>ict</i>	'into a liquid'
- <i>cis</i>	'into a fire'
- <i>wam</i>	'down into a gravitic container (a basket, a cupped hand, a pocket, a lake basin)'
- <i>ikn</i>	'over the rim into a volume enclosure (a gopher hole, a mouth)'
- <i>mik</i>	'into the face or eye, or onto the head of someone'

Tab. 106.3: Atsugewi directional suffixes

Additionally, the system requires mention of two other paradigmatic sets: (i) the root in the analyses of (17) and (18) – i.e., the system of 'patient' indicators, where the moving or located object itself is characterized in terms of various qualities involving size, shape, feel, etc., and (ii) a two-way system of deictic directional indicators, the ventive -*ik* 'hither' and the itive -*im* 'thither'.

#### 4. Dependent-marking expression of circumstance meanings

With dependent-marking morphologies for representing circumstantial notions, the primary conceptual split parallels that seen with head-marking morphologies. On the one hand, nominal elements may bear a mark, a case affix, indicating simply that the nominal fills a particular circumstantial role (4.1). On the other hand, we have morphologically complex adverbial constituents which provide detailed substantive information about the circumstance (4.2).

##### 4.1. Dependent marking of circumstantial relations

In general the categories in a case paradigm in a given language cannot always be cleanly divided into non-overlapping sets according to whether they have nuclear or circumstantial functions. Some case categories have primarily a nuclear function but in certain contexts can have circumstantial functions as well; e.g., the Latin accusative, centrally dedicated to expressing the direct object function, can express the meaning 'place whither' with nouns that name "towns, small islands, and home" (Hale & Buck 1903: 236): *domum redierunt* 'they went home again'. And many case categories which are capable of signalling circumstantial meanings without the help of adpositions can be found in other contexts where they appear governed by specific verbs or prepositions.

A typical set of circumstantial relationships recognized in a case-marking system is presented by William A. Foley for Papuan languages (New Guinea):

"instrumental (INSTR), the relation of an object manipulated by the actor used as a secondary cause in bringing about a change of state in the undergoer, as in *John cut the tree with an axe*; causal (CAUSE), the relation of an uncontrolled, unmanipulated, but also nonvolitional entity bringing about a change of state or experience in the undergoer, as in *I'm shivering from the cold*; locative (LOC), the relation of the place at which the action occurs, as in *I saw him at the store*; ablative (ABL), the relation of the place or entity from which the action proceeds, as in *light emanates from the sun*; allative (ALL), the relation of the place or entity toward which an action proceeds, as in *I walked to the house*." (Foley 1986: 98 f.)

Systems described as having very large numbers of contrasting categories in their case paradigms usually achieve this by elaboration

of spatial notions, thus combining relational with substantive information. In addition to distinguishing *static* ('at a place') from *dynamic* relations ('to a place', 'from a place'), they may encode contrasts within each of these that convey various spatial figure-ground schemata, expressing relations of the sort picked out by the English prepositions *at, on, in, to, onto, into, from, off of, out of*. Rasmus Rask in 1819 (discussed in Hjelmslev 1935: 64) described the local suffixes in Finnish nouns as a system of cases, and assigned them names that for the most part are current today: nominative, genitive, partitive, accusative, essive, absolute, allative, adessive, ablative, illative, inessive, elative. The last 6 case categories are given schematically (ignoring vowel harmony) in Tab. 106.4.

	exterior	interior
static	-lla	-ssa
	adessive	inessive
to	-lle	-Vn
	allative	illative
from	-lta	-sta
	ablative	elative

Tab. 106.4: Finnish local cases

This system in fact embodies a considerable amount of polysemy; but the basic sorting out and naming of the categories is based on what is considered their primary, locative functions, as illustrated with the following examples (ignoring vowel harmony):

illative	<i>talo-on</i>	'into the house'
elative	<i>talo-sta</i>	'out of the house'
ablative	<i>maa-lta</i>	'from the country'
allative	<i>tuoli-lle</i>	'onto the chair'
inessive	<i>talo-ssa</i>	'in the house'
adessive	<i>kato-lla</i>	'on the roof'

Tab. 106.5: Examples of Finnish local cases

As Tab. 106.4 shows, the Finnish system of local cases involves two intersecting sets of opposition, one case-like and one more topological or "configurational". The case-like notions are represented by the rows (motion to – allative and illative; location at – adessive and inessive; motion from – ablative and elative), and they cross-cut a second distinction, represented by the columns, involving topological spatial relations between Figure

and Ground, with Ground schematized as either surface (adessive, allative, ablative) or enclosure (inessive, illative, elative). These latter, configurational, distinctions frequently exhibit considerable idiomticity and polysemy of their own. Fred Karlsson's (1987: 100–104) description of the inessive and illative, for example, notes that in Finnish one speaks of wearing gloves, hats, and shoes 'in' one's hands, head, and feet, and similarly of putting gloves, hats, and shoes 'into' one's hands, head, and feet. Here the locative (more 'case-like') functions are preserved but the configurational schemata depart from those which motivated the case names.

Within the Finnish case system, dependency relations of nouns expressing time units or time periods are expressed with the same cases, parasitically on the locational notions. As can be seen in the following examples, from Karlsson (1987: 100–107), the spatial-metaphor motivation is sometimes easy to see, sometimes not.

- (19) *luin kirjan tunni-ssa*  
I.read book hour-INNESS  
'I read the book in an hour'
- (20) *tulen Norjaan ensi kuu-ssa*  
I.go to.Norway next month-INNESS  
'I'm coming to Norway next month'
- (21) *viiko-sta viikko-on*  
week-ELAT week-ILL  
'from week to week'
- (22) *en ole nähty häntä pari-in tunti-in*  
I not.be seen him couple-ILL  
hour-ILL  
'I haven't seen him for a couple of hours'

In laying out comparative descriptions of case systems, especially those of closely related languages, it is useful to start from a repertory of functions and to line up the language-specific morphological means for specifying or contrasting these functions. Common to the Papuan languages of New Guinea, Foley (1986: 99 f.) finds, are the five circumstantial notions listed in the quotation above: instrumental, causal, locative, ablative, and allative. Foley describes the ways in which these are represented in a number of related languages. In Yimas all are functions of a single suffix; Iatmul has two suffixes, one with allative function, the other covering the remaining four. Fore combines allative and

locative functions in one case category, putting the remaining three in another. Kewa combines instrumental and causal into one category, the remainder in another. Alambalak separates all of them except for ablative and causal, which are united in one category.

The New Guinea examples demonstrate clearly that the ability to make useful generalizations about case systems, in their typological and diachronic aspects, depends on an appeal to a language-independent repertory of circumstantial concepts. Foley's report of the Papua New Guinea situation points further to the importance of a distinction between the central and peripheral cases: the basic nuclear functions in these languages are handled by head-marking devices, while clearly circumstantial functions are expressed by case ('dependent-marking'). The 'recipient' function, however, is intermediate: it is handled by head-marking in some languages and by case in others. Foley also points out that when the recipient is identified by head-marking, its formal exponent also covers the function of beneficiary, but that when it is marked with a case it often includes the allative function. Such groupings and regroupings seem to fit recurring patterns of form-function mappings in the domain of grammatical functions – patterns characteristically linking beneficiary with recipient, recipient with allative, allative with locative, instrumental with ablative, ablative with cause, and many others. Such patterns of conflation would appear to be independent of whether the language's particular marking system involves cases, adpositions, head-marking, constituent order, or combinations of these.

#### 4.2. Morphologically complex circumstantial adverbs

Some languages have morphologically complex adverbs with distinct structural slots dedicated to different semantic domains and different semantic oppositions. In some cases such a morphological system is not dedicated purely to the expression of circumstantial or adverbial meanings but extends beyond such functions. For example, the morphological structure of deictic and interrogative words in Japanese, known as the *ko-so-a-do* system, combines in the first position deictic elements *ko-* (proximal), *so-* (medial), and *a-* (distal) with an interrogative marker *do-* (interrogative), while the second position is occupied by a head morpheme that determines the grammatical function of the word as a whole:

an inanimate demonstrative pronoun *-re*, determiner *-no*, a place adverb *-ko*, a direction adverb *-tira*, manner adverb (phonemic length), and a few others. For the circumstantial, this yields examples like *do-ko* 'where', *so-o* (*so-* plus length) 'in that manner', *ko-tira* 'in this direction'. Deictic and interrogative adverbs of place, direction-to-ward, and manner thus emerge from this cross-categorial morphological system.

A number of languages have a class of manner adverbs, typically with subtle but highly specific semantics, which appear to embody iconic or phonosymbolic components. Such forms can be morphologically interesting in numerous ways: they may be constrained to fit particular phonological templates, for instance, or involve reduplication or other semantically iconic patterns of vowel length, consonant gemination, etc. (Some aspects of such systems are taken up in Art. 30 on iconicity.) They also may involve segments not found in the language's regular inventory, or segments that violate the language's regular combinatory patterns. For the languages of West Africa a common term for this class of words is **ideophones**; for Austro-Asiatic languages (the Mon-Khmer family and Munda), **expressives**; for Japanese and Korean, **mimetic adverbs**. In some languages, like Japanese, they appear to be rigorously lexicalized; in others, they are sometimes described as independent of the lexicon. Igbo, for example, is said to be "rich in ideophones, which have been analyzed as being generated directly from meaning-related phonemes, or even from features slotted into templates, rather than being fixed items in the lexicon" (Williamson & Emenano 1992: 198). Regarding the Austro-Asiatic expressives, the following is said:

"Adverb-like, but without predicative force, these words are similar in some ways to African ideophones. They rely primarily on iconic means to evoke sensations of all kinds, especially in the domain of visual patterns, their rich morphology is made of iconic diagrams – e.g., partial reduplications, substitution, infix copying and systematic distortions – often akin to deliberate language games." (Diffloth & Zide 1992: 140)

A common templatic pattern for the Japanese mimetic adverbs is a  $C_1V_1C_2V_2$  structure, e.g., *pata*, which may be subject, inter alia, to the following modifications: reduplication of the whole (*patapata*), suffixation by *-ri* (*patri*), suffixation by *-ri* and gemination of  $C_2$  (*pattari*), suffixation by *-tto* (*patatto*), suffixa-

tion by moraic *-n* (*patan*), suffixation of the last-named pattern by *-to* (*patanto*), reduplication of the nasal-augmented form (*patan-patan*). Meanings conveyed by such words include the mood of the actor (cheerful and energetic vs. sad and listless), luminosity (gleaming, sparkling), producing an accompanying sound, various patterns of movement, a variety of ways of modifying specific kinds of activities such as eating, smiling, walking, and many others.

Japanese mimetic adverbs show many instances of morphophonemic alternations linking *h*, *p*, and *b*: ‘moving in a fluttering way’ – *hatahata*, *patapata*, *batabata*; ‘flowing, as of tears’ – *horohoro*, *poroporo*, *boroboro*, ‘twitching nervously’ – *hikuhiku*, *pikupiku*, *bikubiku*). The different forms apparently convey subtle semantic differences. There are frequent instances of sub-morphemic regularities, such as is suggested by the frequency of palatal stops and fricatives in words imitative of water sounds: *pityapitya*, *bityabitya*, *patyapatya*, *batyabatyaa*, *potyapotya*, *botyabotya*, *syaasyaa*, *zyabuzyabu*. Several adverbs of smiling begin with *ni-*: *nikoniko* (positively evaluated, ‘beamingly’), *nitanita*, *niyaniya* (negatively evaluated, ‘snidely’). Several adverbs indicating unsteady walking begin with *yo-*: *yotiyoti*, *yotayota*, *yoroyoro*. The structures of this family of words do not lend themselves to “ordinary” morphology: but, at least in Japanese, they involve more than language play.

The spatial semantic system for Cora (Uto-Aztecán, Mexico) includes a set of quite regularly constructed three-part adverbs which in their basic meanings locate objects in respect to landmarks in a mountainous terrain (Casad & Langacker 1985). Within these forms, the first morpheme indicates the “distance” of the reference object relative to the speaker’s real, or assumed, vantage point; the categories are proximal (‘PROX’), medial (‘MED’), and distal (‘DIST’). The second morpheme locates the Figure as being “inside” or “outside” the Ground, there being several metaphorical extensions of this contrast. The optional third morpheme situates the Figure with respect to the slope of the Ground object, as “top”, “face”, or “foot”. Thus *m-a-h* ‘away up there to the side in the face of the slope’ is glossed ‘MED-outside-face.of.slope’ and *y-ú-u* ‘right here at the foot of the slope’ is glossed as ‘PROX-inside-foot.of.slope’ (Casad & Langacker 1985: 258).

- (23) *u-h-t<sup>v</sup>ee*                                    *m-ú*  
     inside-vertical-be.long PROX-inside  
     *či?i-ta*  
     house-in

Example (23) shows the interaction and integration of three separate locative systems in a single three-word sentence. The verb has a complex prefix-set *u-h-*, with *u-* ‘on inside surface’ and *h-* ‘on vertical surface’, thus ‘on a vertical inside surface’. These prefixes combine with a verb root *t<sup>v</sup>ee* glossed as ‘it be long’ yielding *uht<sup>v</sup>ee* expressing the idea that something extends far upward on an inside surface. The locative adverb *m-u* is made up of the medial prefix *m-* in combination with the *-u-* element ‘inside’, thus ‘there inside’. The noun *či?i-ta* is made up of the noun *či?i* ‘house’ and the locative case suffix *-ta* ‘in’, thus ‘in the house’. The entire sentence, then, is *u-h-t<sup>v</sup>ee m-ú či?i-ta*, is translated as ‘it’s a long way up to the ceiling inside that house’ (Casad & Langacker 1985: 255).

We saw in the Runyambo example (16) in 3.1.3 a sentence which used the three means of exhibiting circumstantial roles discussed in that section: the licensing, nuclearizing, and anaphoric functions. In this Cora sentence too we see multiple means of conveying locative notions, this time distributed across elements of the clause: in head-marking prefixes, in a morphologically complex locative adverb, and in case-marking on the noun.

## 5. Uncommon abbreviations

FV	final vowel
IV	initial vowel
LOC.EXL	frame-external locative
LOC.INL	frame-internal locative
TRIG	trigger

## 6. References

- Ashton, Ethel & Mulira, E. M. K. & Ndawula, E. G. M. & Tucker, A. N. (1954), *A Luganda Grammar*. London etc.: Longmans Green
- Casad, Eugene H. & Langacker, Ronald W. (1985), “‘Inside’ and ‘Outside’ in Cora Grammar”. *International Journal of American Linguistics* 51.1, 247–281
- Comrie, Bernard (1985), *Tense*. Cambridge: Cambridge University Press
- Comrie, Bernard & Smith, Norval (1977), “Lingua Descriptive Studies: Questionnaire”. *Lingua* 42, 5–71

- Craig, Colette & Hale, Ken (1988), "Relational Preverbs in Some Languages of the Americas: Typological and Historical Perspectives". *Language* 64.2, 313–344
- Denny, J. Peter (1982), "Semantics of the Inuktitut (Eskimo) Spatial Deictics". *International Journal of American Linguistics* 48.4, 359–384
- Diffloth, Gérard & Zide, Norman (1992), "Austro-Asiatic Languages". In: Bright, William (ed.), *International Encyclopedia of Linguistics, Vol. I*. Cambridge: Cambridge University Press, 137–142
- Foley, William A. (1986), *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press
- Hale, William Gardner & Buck, Carl Darling (1903), *A Latin Grammar* [reprinted (1966): Tuscaloosa: University of Alabama Press (Alabama Linguistic and Philological Series 8)]
- Hanks, William F. (1984), *Referential Practice: Language and Lived Space Among the Maya*. Chicago: University of Chicago Press
- Hjelmslev, Louis (1935–1937), "La catégorie des cas: Etudes de grammaire générale". Part I (1935): *Acta Jutlandica* 7.1, i–xii, 1–184; Part II (1937): *Acta Jutlandica* 9.2, i–vii, 1–88
- Karlsson, Fred (1987), *Finnish Grammar*, translated by Andrew Chesterman. Helsinki: Werner Söderström [<sup>1</sup>1982]; translation of: *Suomen Peruskielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura]
- Newman, Stanley (1976), "Salish and Bella Coola Prefixes". *International Journal of American Linguistics* 42.3, 228–242
- Nichols, Johanna (1986), "Head-marking and Dependent-marking Grammar". *Language* 62, 56–119
- Rugemalira, Josephat Muhozi (1994), *Runyambo Verb Extensions and Constraints on Predicate Structures*. Ph.D. dissertation, University of California
- Schachter, Paul (1987), "Tagalog". In: Comrie, Bernard (ed.), *The World's Major Languages*. Oxford: Oxford University Press, 936–958
- Shibatani, Masayoshi (1990), *The Languages of Japan*. Cambridge: Cambridge University Press
- Talmy, Leonard (1985), "Lexicalization Patterns: Semantic Structure in Lexical Forms". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press, 57–149
- Tesnière, Lucien (1959), *Elements de syntaxe structurale*. Paris: C. Klincksieck
- Williamson, Kay & Emenanjo, E. Nolue (1992), "Igbo". In: Bright, William (ed.), *International Encyclopedia of Linguistics, Vol. II*, 195–199

*Charles J. Fillmore, Berkeley (U.S.A.)*

## 107. Valency change

1. Introduction
2. Valency-decreasing categories
3. Valency-increasing categories
4. General features of valency-changing morphology
5. Diachronic sources of valency-changing morphology
6. Uncommon abbreviations
7. References

### 1. Introduction

The **valency** of a lexical item is its inherent relationality that allows it to govern a particular number of **arguments** (or *actants*, Tesnière 1959) of a particular type. The grammatical meaning of certain morphological categories consists in changing the valency of a lexical item, and it is such categories that we will

deal with in the present article. (In this article, we use the term *category* in the sense of "grammatical morpheme" or "grammeme"; thus, notions like genitive or future are grammatical categories, while sets of categories like case or tense are termed *supercategories*.)

We will refer to the configuration of arguments that are governed by a particular lexical item as its **valency pattern** (in other terminologies: *argument structure* (Grimshaw 1990), *predicate frame* (Dik 1978: 15), *government pattern* (Russian linguistics, e.g. Mel'čuk 1988: 69)). Valency is characteristic of all the major word classes (verbs, nouns, adjectives) and of certain types of function words (in particular, adpositions and auxiliary verbs). However, it is verbs that show by far the most diverse and interesting valency patterns, as well as the most interesting valency-

changing operations. We will accordingly deal exclusively with verbal valency-changing categories in this article.

A verb's inherent relationality is obviously semantically motivated. The English verb *like* has two arguments in its valency pattern (as in *Sarah likes Farid*) because it describes a situation that involves two salient objects as participants. From a semantic point of view, participants are commonly characterized by the semantic roles they fulfill, e.g. experiencer and stimulus in the above example. However, a verb's valency pattern is not completely predictable on the basis of the semantic roles that its participants play in the situation in question. On the one hand, participants with identical semantic roles may show up as different types of arguments, as in *Sarah likes Farid* versus *Farid pleases Sarah*. On the other hand, participants with different semantic roles may show up as the same type of argument, as in *Sarah likes Farid* (experiencer – stimulus) versus *Sarah hits Farid* (agent – patient).

It is therefore common for grammarians to take valency as a syntactic notion and to characterize the verbal arguments by the **grammatical relations** they bear, such as subject, direct object, indirect object, etc. But most common is perhaps the characterization of valency both in semantic and in syntactic terms, reflecting both its semantic motivation and its partial conventionalization in terms of arbitrary linguistic rules.

There is no standard notation for valency patterns. In this article we use a very simple notation, consisting of two-line columns each of which corresponds to an argument. The upper line contains the semantic role of the arguments, and the lower line contains their grammatical relations, e.g.:

(1) (a) *like*

valency pattern:  
experiencer stimulus  
subject direct object

(b) *donate*

valency pattern:  
agent theme recipient  
subject direct object to-object

This notation is used here only for expository convenience. No claim is implied concerning the existence of levels of a semantic role structure or a grammatical relation structure. The notation is a crude oversimplification, but it provides a useful frame of reference for further discussion. Moreover, it can be easily

translated into various specific frameworks that posit such levels. Thus, our semantic roles and grammatical relations correspond, very roughly, to the initial stratum and later strata in Relational Grammar (Perlmutter & Postal 1977); to D-structure and S-structure in Government-Binding theory (Chomsky 1981); to semantic functions and syntactic functions in Functional Grammar (Dik 1978); to Semantic Role Structure and Syntactic Structure in work by the Leningrad/St. Petersburg typological group (e.g. Geniušienė 1987), etc.

Alternations in a verb's valency pattern are not necessarily the result of a morphological derivational process. Verbs or whole classes of verbs may have alternate valency patterns without any change in their formal makeup, as in (2) (the so-called "Dative shift").

(2) (a) *Sarah gave a mango to Farid.*

valency pattern:  
agent theme recipient  
subject direct object to-object

(b) *Sarah gave Farid a mango.*

valency pattern:  
agent theme recipient  
subject secondary primary  
object object

Since this article is concerned with valency-changing morphology, it does not deal with such cases.

## 2. Valency-decreasing categories

Valency-decreasing (or argument-removing) morphological categories can be further subdivided into patient-removing (or object-removing) categories and agent-removing (or subject-removing) categories.

### 2.1. Patient-removing categories

A derivational morpheme that removes the patient or direct object from the verb's valency pattern is here called **deobjective**. There is no standard term for this construction and for categories that signal it. Sometimes *indefinite object deletion* is used (Marantz 1984: 192–195), sometimes *absolut(iv)e* (Geniušienė 1987: 83, 314) or *absolutive antipassive* (e.g. Dayley 1989: 111).

In example (3) from Ainu, the prefix *i-* has deobjective function (Shibatani 1990: 46).

(3) (a) *Sake a-ku.*

sake 1.SG.TR-drink  
'I drink sake.'

- valency pattern:  
 agent patient  
 subject object
- (b) *I-ku-an.*  
 DEOBJ-drink-1.SG.INTR  
 'I drink.'
- valency pattern:  
 agent patient  
 subject Ø

Deobjective morphemes are not very common, and they seem to be applicable only to a restricted set of verb stems, mostly those verbs that denote actions that affect the agent as well as the patient (cf. Marantz 1984: 93).

A related construction is what could be called the **potential deobjective**, as in (4) and (5).

- (4) Lithuanian (Geniušienė 1987: 83–84)
- (a) *Berniuk-as muša vaik-us.*  
 boy-NOM.SG beat(3.SG) child-ACC.PL  
 'The boy beats children.'
  - (b) *Berniuk-as muša-si.*  
 boy-NOM.SG beat(3.SG)-DEOBJ  
 'The boy fights (is pugnacious).'
- (5) Udmurt (Geniušienė 1987: 315)
- (a) *Puny vanz-es kurtčyl-e.*  
 dog(NOM) all-ACC bite-3.SG  
 'The dog bites everybody.'
  - (b) *Puny kurtčyl-išk-e.*  
 dog(NOM) bite-DEOBJ-3.SG  
 'The dog bites.'

Whereas ordinary deobjectives express a real action without mentioning the patient, potential deobjectives express a disposition of an agent to perform an action. Potential deobjectives therefore occur only in irrealis or generic sentences, never in specific realis sentences.

A third type of object-removing category is labeled **deaccusative** in Geniušienė (1987: 94). (It could also be called “antiapplicative”.) In deaccusatives, the patient is not removed entirely, it only loses its direct-object status and is expressed as an oblique phrase. Thus, the deaccusative is strictly only a patient-backgrounding category, not a patient-removing category.

- (6) Hungarian (Károly 1982: 187)
- (a) *Az orvos szán-ja a beteg-et.*  
 the doctor pity-3.SG the patient-ACC  
 'The doctor pities the patient.'

- (b) *Az orvos szán-akoz-ik a beteg-en.*  
 the doctor pity-DEACC-3.SG the patient-SUPESS  
 'The doctor feels pity for the patient.'
- derived valency pattern:  
 agent patient  
 subject oblique

Deaccusatives are not common cross-linguistically and are heavily restricted lexically even in those languages where they occur.

## 2.2. Agent-removing categories

In many languages there is a strong requirement for all sentences to have subjects. When in such languages a valency-changing category removes the agent argument from the subject position, the patient argument must take up the subject position instead. In the following examples we will not pay special attention to this automatic promotion to subject.

The most radical agent-removing category is the **anticausative** (Haspelmath 1987; 1993 b). An anticausative affix eliminates the agent argument completely, as illustrated in (7).

- (7) Gothic
- (a) ... *aiphau distair-id bata niujo wein pans balgins.*  
 lest burst-3.SG the new wine bags  
 '... lest the new wine burst the skin bags (Lk 5,37).'
  - valency pattern: agent patient  
 subject direct object
  - (b) ... *aiphau distaur-n-and balgeis.*  
 lest burst-ACAU-3.PL bags  
 '... lest the skin bags burst (Mt 9,17).'
  - derived valency pattern:  
 patient subject

The notion of anticausative was unknown in traditional grammar. Among the terms that have been used for this category are *inchoative*, *pseudopassive*, *neutral passive*, *spontaneous* (Shibatani 1985), and others. (The term *anticausative* is due to Nedjalkov & Sil'nickij 1969.)

The anticausative is similar to the passive in that the agent argument loses its subject status and the patient becomes the new subject, but in the passive, the agent is not entirely eliminated. Semantically, a passive such as Russian *Dver' byla otkryta* 'The door was

opened' crucially differs from the anticausative *Dver' otkrylas'* 'The door opened' in that an agent is implied (though not mentioned explicitly) in the passive, whereas the anticausative action is thought of a happening spontaneously. Furthermore, the anticausative is subject to severe lexical restrictions, unlike the passive, which can generally be formed from the great majority of transitive verbs. Anticausatives can be formed only from verbs expressing actions that are performed without any specific instruments or methods, so that they can be thought of as happening spontaneously, without a (human) agent's intervention. Thus, the verbs 'break', 'tear', 'split', 'divide' often have anticausatives, while 'cut' and 'saw' do not; the verbs 'open', 'close', 'raise', 'cover' often form anticausatives, while 'write', 'wash', 'construct' do not; 'loosen' forms anticausatives, while 'tie' does not.

In (8) and (9) there are two more examples of transitive verbs and corresponding anticausatives.

(8) Hungarian

- (a) *András-t három tárgy-ból elvág-t-ák.*  
András-ACC three subject-ELAT fail-PAST-3.PL  
'They failed András in three subjects.'
- (b) *András három tárgy-ból elvág-ód-ott.*  
András three subject-ELAT fail-ACaus-Past(3.SG)  
'András failed in three subjects.'

(9) Turkish

- (a) *Anne-m kapi-yı aç-tı.*  
mother-1.SG door-ACC open-Past(3.SG)  
'My mother opened the door.'
- (b) *Kapı aç-il-di.*  
door open-ACaus-Past(3.SG)  
'The door opened.'

A category that is somewhat similar to the anticausative is the **resultative** (cf. Nedjalkov 1988, ed.). The resultative of stative verbs is essentially a stative variant of the anticausative, as illustrated in (10).

(10) Russian

- (a) non-derived:  
*Mira zakryvaet dver'.*  
Mira closes door  
'Mira is closing the door.'

valency pattern:

agent	patient
subject	direct object

(b) anticausative:

*Dver' zakryvaet-sja.*  
door closes-REFL  
'The door is closing.'

valency pattern:

patient
subject

(c) resultative:

*Dver' zakry-ta.*  
door close-PART  
'The door is closed.'

valency pattern:

patient
subject

The resultative is defined as a category that turns a verb that refers to an event into a verb referring to a state that results from that event.

Although the anticausative and the resultative have the same derived valency pattern, the two categories are very different in nature. The anticausative has the removal of the agent as its primary function, whereas the primary function of the resultative is the expression of a state by means of an event word. The removal of the agent is a secondary effect of this primary function: Since all actions are dynamic and states cannot be actional/agentive, states cannot have agents. Since the expression of a state is the primary function of the resultative, it can also be formed from intransitive verbs in many languages, e.g. in Homeric Greek (where the Perfect functions as a resultative).

- (11) (a) *pēg-nu-mi*  
stick-Pres-1.SG  
'I am sticking (tr.)'

- (b) *pé-pēg-a*  
RDP-stick(RES)-1.SG  
'I am stuck'

- (12) (a) *thnē-isk-ō*  
die-Pres-1.SG  
'I am dying'

- (b) *té-thnē-k-a*  
RDP-die-Res-1.SG  
'I am dead'

In resultatives of intransitive verbs (as in (12 b)), there is no valency change at all.

Another verbal category that is not uncommon cross-linguistically is the **reflexive**. Two examples are found in (13)–(14).

- (13) Modern Greek
- (a) *O Axmed ksíri-s-e ton*  
ART Ahmed shave-AOR-3.SG ART  
*Péðro.*  
Pedro  
'Ahmed shaved Pedro.'
- (b) *O Péðro ksíri-s-tik-e.*  
ART Pedro shave-AOR-REFL-3.SG  
'Pedro shaved (himself).'
- (14) Armenian (Kozinceva 1981: 83)
- (a) *Mayr-ə lvan-um e*  
mother-ART wash-PRES AUX  
*Seda-yi-n.*  
Seda-DAT-ART  
'Mother is washing Seda.'
- (b) *Seda-n lva-cv-um e.*  
Seda-ART wash-REFL-PRES AUX  
'Seda is washing (herself).'

In reflexive verbs, the number of semantic participants remains strictly speaking the same, but since subject and object participants are referentially identical, only one participant (the subject) is expressed. The change in the valency pattern may be represented as follows (following Krakovskij 1981):

transitive valency pattern:	
referents:	A            B
roles:	agent        patient
functions:	subject      object
derived valency pattern:	
referents:	A
roles:	agent        patient
functions:	subject

Reflexive verbs are most common with naturally reflexive actions such as body care or grooming verbs ('wash', 'shave', etc.) or non-translational motion ('turn', 'bow', 'lie down'). Often they cannot be used with actions that are carried out reflexively (e.g., in Modern Greek it is impossible to say *\*singrīθike* 'compared herself'; a reflexive pronoun must be used here) (cf. Kemmer 1994).

Some languages also have a special **reciprocal** category, e.g. Turkish:

- (15) (a) *Sara Farid-i sev-iyor.*  
Sara Farid-ACC love-IPFV  
'Sarah loves Farid.'
- (b) *Dost-lar sev-iş-iyor-lar.*  
friend-PL love-REC-IPFV-PL  
'The friends love each other.'

### 3. Valency-increasing categories

Let us turn to valency-changing categories that either establish a syntactic relation or redefine an already existing one. The most im-

portant cases fall into two types: With the first, the verb is supplied with a direct object, with the second, a new agent/subject is introduced. The former is called applicative, the latter, causative.

#### 3.1. Object-adding categories: the applicative

##### 3.1.1. Types of applicatives

**Applicatives** assign the status of a direct object to oblique roles of different kinds. There are three main types of applicatives in the languages of the world.

The first and probably most widespread type is the **benefactive applicative**, as shown in (16) and (17) (cf. also Shibatani 1996).

- (16) Indonesian (Chung 1976: 58)

- (a) *Orang itu me-masak ikan.*  
man ART TR-cook fish  
'The man cooked fish.'  
valency pattern:  
agent      patient  
subject     object
- (b) *Orang itu me-masak-kan*  
man ART TR-cook-APPL  
*perempuan itu ikan.*  
woman ART fish  
'The man cooked fish for the woman.'
- derived valency pattern:  
agent      patient      benefactive  
subject     (object)     object

That the argument corresponding to the benefactive of the English gloss has acquired the status of a direct object is testified by the fact that it is a possible passive subject in (17 b).

- (17) Fula (Atlantic; Arnott 1970: 355)

- (a) *ße-kirs-an-ii-min*  
3.PL.SBJ-PL:slaughter-APPL.BEN-  
*ŋaari*  
PAST-1.PL.OBJ bull  
'A bull has been slaughtered for us.'
- (b) *min-kirs-an-aama*  
1.PL.SBJ-PL:slaughter-APPL.BEN-  
*ŋaari*  
PASS.PAST bull  
'We have had a bull slaughtered for us.'

Various related meanings are centered around the benefactive applicative. The action may also proceed to the detriment of the object participant, or this may be the possessor of the original patient, as in Tzotzil

(Mayan; Aissen 1987: 126 ff.). Other notions are purpose or cause ('come-for this book' would be an applicative in Swahili, cf. Port 1981: 78), and the addressee, as with 'pray' in Kanuri (Saharan; Hutchison 1981: 143).

The second main type of applicative is the **goal** or **directive applicative**. In many languages it is formally identical to the benefactive applicative, and thus notionally related to it (for example, in Swahili). Example (18) demonstrates a genuine species:

- (18) Ika (Chibchan; Frank 1990: 68)  
*mi-ka-wa'ka*  
 2.OBJ-APPL.DIR-look  
 'He looked at you.'

The directive applicative is open to the same semantic extensions as the benefactive one. The Ika sentence in (19) demonstrates a possessive variant (often called "possessor ascension"):

- (19) *peri-kin-di mi-ka-ga*  
 dog-LIM-TOP 2.OBJ-APPL.DIR-eat  
 '(The jaguar) eats your two dogs.'

It may be necessary to set up another nuclear type, related to the aforementioned, in which the affectedness of the object participant is more important (and which might be called the **comprehensive applicative**). This would account for, e.g., German *be-* in *be-schmeißt ihn mit Eiern*, 'APPL.COMPR-throws him with eggs (bombards him with eggs)'. This enables the applicative to occur in cases where no direction can properly be attested, e.g. Fiji *dabe-ca*, 'sit-APPL.COMPR (sit on)'.

The third main type of applicative is the **instrumental** (or **comitative**) **applicative**, of which Grebo supplies an example:

- (20) Grebo (Kru; Innes 1966: 57)  
 $\begin{array}{ll} o & du-di-da \end{array}$  *bla*  
 3.PL pound-APPL.INSTR-REMPST rice  
*sū*  
 pestle  
 'They pounded rice with a pestle.'

Instrumental applicatives exist in Australian languages (e.g. Yidiny, cf. Dixon 1977: 302–304), the Arawakan languages of South America (cf. Wise 1990), North American languages (cf. Mithun 1989), Oceanic languages (cf. Harrison 1982), Nilo-Saharan languages (e.g. Nandi, cf. Creider & Creider 1989: 126 f.), Cushitic languages (e.g. Somali, cf. Saeed 1987: 185 f.), and throughout the Niger-Congo superphylum. The instrumental applicative exhibits a certain tendency to

acquire a causal (reason) and a stimulus function (like the English preposition *through*), the latter in turn having affinities to the directive applicative (e.g. in Oceanic languages, cf. Ross 1988: 375–377). Also, the instrumental applicative may develop a comprehensive variant via meanings like Fiji *voce-taka* 'row-APPL.INSTR (row it)', where an instrumental notion might be backgrounded.

There are many other, less frequent sorts of applicative that are too numerous to pass in review here. But we should add that Latin (and in part German) has a wealth of so-called "**preverbs**" that govern the accusative, e.g. (21).

- (21) *fama urbem per-vasit*  
 rumor town:ACC through-roamed  
 'The rumor went around the town.'

They are of limited productivity and mirror the inventory of (primary) adpositions. (They are therefore often conceived of as instantiating a compounding process.)

The applicative transitivizes an intransitive verb, providing it with a direct object. If a transitive verb is extended by an applicative, the original direct object (e.g. 'bull' in our example (17)) will normally give up its status, often becoming an instrumental. But there are instances (cf. the discussion in Baker 1988: 245–247, Bresnan & Moshi 1990), where the original patient retains its ability to become the subject of a passive even after applicative formation has taken place. This is so in Fula.

### 3.1.2. Applicative vs. adposition

The construction consisting of the verb base, the applicative formative and the object must not be regarded as isofunctional with that consisting of the verb, an adposition (or case form), and its noun phrase (cf. Craig & Hale 1988; Blake 1987: 69–71). The verb and the applicative formative jointly govern the direct object. That presupposes that the latter be subject to some sort of affectedness exerted by the extended verbal notion. This is particularly obvious in the case of **locative applicatives**, as in (22).

- (22) Kinyarwanda (Bantu; Kimenyi 1980: 92)  
*ábá-ana*  
 DEF:HUM.PL-child  
*b-iica-yé-ho*  
 HUM.PL.SBJ-sit-PF-APPL.LOC  
*ubu-riri*  
 DEF:CL.14-table  
 'The children are sitting on the table.'

With this construction it would not be possible to replace 'table' by 'mountain', because the subject participant is required to somehow dominate the object (or influence its condition; for a similar case in the Australian language Kalkatungu, cf. Blake 1987: 69 f.). Static verbs become dynamic in the applicative form. This need for this sort of affectedness lies at the base of many semantic specializations with applicatives. The Latin preverbs exhibit an applicative-like behavior specifically when the enlarged verb acquires a figurative meaning that induces a patient. Compare example (21) with the more literal, non-applicative *incendium per agros per-vasit* 'conflagration through fields through-roamed (the conflagration roamed through the fields)'.

The applicative is a verbal category, case forms and adpositions are nominal categories. Accordingly, applicative formatives usually cannot be traced back to adpositions, both diachronically (although there may be analogical assimilations) and synchronically (in the sense of a semantic mapping). The applicative is nothing like "case or adposition with the verb". But here certain specifications are appropriate. Under the conditions of topicalization, the complement of an adposition might be removed from the adposition, that is, get stranded. This is a well-known property of the English relative and interrogative clause, but also occurs in the passive (the so-called prepositional passive, cf. Couper-Kuhlen 1979). In this latter case, the adposition is becoming a verbal category, transferring its object to the then compound verb. Even then, affectedness comes into play. So one would not say, *This country has been [emigrated from] by John*, but *This country has been [emigrated from] by thousands of people* is fine. Similar phenomena of reanalysis of prepositions under topicalization of their object (also involving coordination reduction) have been recorded for Tongan (Churchward 1953: 148), Rama (Chibchan; Craig & Hale 1988) and Nadeb (Maku; Weir 1986).

As regards the relation of adpositions and applicatives, it must be admitted that there are languages that compensate for the lack of adpositions (or cases) by having a certain amount of applicatives at their disposal (e.g. the Arawakan languages, cf. Wise 1971 for Nomatsiguenga). Similarly, there are languages, a famous example being Oceanic, whose basic verbs are usually intransitive and which consequently use an applicative (usually of the directive kind) for each transitive

clause (hence the applicative is often called transitivizer here; cf., e.g., Schütz 1985 for Fijian).

With some languages, e.g. Athapaskan and North-West Caucasian, we find what could be called verbal prepositions. These do not form a unit with the verb that governs an object; instead, the object (in the form of a possessive personal affix) depends entirely on the internal adposition, while the respective noun may be supplied in the context. As a consequence, there are no restrictions as to which notions are expressed, and more than one verbal adposition may occur in one verb. An example of a verbal preposition is Navajo *-taa* 'among' in (23) (cf. Kibrik 1990 for details on the grammaticalization of such elements).

- (23) Navajo (Young & Morgan 1980: 93)
- |   |             |
|---|-------------|
| nihi-siláo                                    | náá+ts'ózí  |
| 1.PL.POSS-soldier                             | narrow.eyes |
| yi-taa-da-'a-sdon                             |             |
| 3.POSS-MED-DISTR-OBJ.INDEF-IPFV:shoot         |             |
| 'Our soldiers fired at (among) the Japanese.' |             |

There are other phenomena that should be kept apart from applicatives. Some notions evoke an optional *accusativus spatii*, as tradition has it. They exhibit at the same time certain adpositional properties and certain properties of a verbal operator (cf. Bolinger 1971: 23–25). To these notions typically belong 'up', 'down', and 'along'. Cf. *He strolled drowsily [along the paths]* with *He [strolled along] drowsily the paths*.

As a final point, there are valency-changing categories that introduce cases other than the accusative. The most significant is the dative (*dativus commodi*) in cases like German *zu* in *es ihr zu-werfen* 'it her:DAT to-throw (throw it to her)'. In general, the dative (in the sense of a *dativus commodi*) is due to a general valency regularity. It becomes possible in connection with (certain) directionals, as in *es ihr in den Wagen werfen* 'it her:DAT into the car throw (throw it in her car)', and *es ihr hinein-werfen* 'it her:DAT in-throw'. What is peculiar about *zu* is the fact that with it, the dative is obligatory. A further case of a dative-adding applicative seems to be the Georgian objective and superessive version (Boeder 1968).

### 3.2. Agent-adding categories: the causative

Many languages have a morphological category, called **causative**, which conveys the meaning of causation and adds a new agent

argument (the causer) to the valency pattern. According to Bybee (1985:29), the causative is the most common valency-changing category in her world-wide sample of 50 languages. The syntax and semantics of causatives has been studied extensively (cf. Xolodovič 1969, ed.; Shibatani 1976, ed.; Comrie 1985; Baker 1988; Song 1996; Dixon 2000). Example (24) illustrates the causative in Babungo, marked by the suffix *-s*.

- (24) Babungo (Niger-Congo; Schaub 1982: 211)

(a) *ŋwá ŋjì ták ŋjì*  
he enter in house  
'He entered the house.'

valency pattern:

agent	place
subject	adverbial

(b) *má nji-só ŋwá ták ŋjì*  
I enter-CAUS him in house  
'I made him enter the house.'

derived valency pattern:

causer	agent	place
subject	object	adverbial

### 3.2.1. Valency changes in the causative

The addition of the causer to the valency pattern leads to quite drastic changes in it because the causer always usurps the subject function, and the old subject (the lower subject, or causee) must occupy a different grammatical relation in the derived valency pattern. When the base verb is intransitive, no major problem arises, because the causee can take up the object position, as in (24 b). This happens almost universally, and it is semantically very plausible: The causee is the patient of the causation and as such would be expected to occupy the direct object position. In this respect, the representation of the derived valency pattern in (24 b) is somewhat misleading. A fuller representation would have to include the semantic roles in the causing event and the semantic roles in the caused event, as in (25).

- (25) derived valency pattern:

(caused event roles)  
agent place  
(causing event roles)  
causer patient/causee place  
(grammatical relations)  
subject object adverbial

Many languages have only causatives from intransitive verbs. For instance, in Lezgian (Nakh-Daghestanian, Haspelmath 1993 a), there are causatives in -(a)r- such as *aqwaz-*

*ar-* 'stop (tr.)' (from *aqwaz-* 'stop (intr.)'), *ksu-r-* 'put to bed' (from *ksu-* 'fall asleep'), etc., but there are no causatives from transitive verbs like *at'u-* 'cut' (\**at'u-r-*).

In those languages that do have causatives derived from transitive verbs, the causee may be treated in three different ways: (a) the causee becomes an indirect object; (b) the causee is expressed as an instrumental phrase; (c) the causee becomes a direct object, and the causative valency pattern contains two direct objects.

The first option is exemplified by Georgian (26). In such cases, the resulting valency pattern is identical to the valency pattern of non-derived ditransitive verbs such as 'give'.

- (26) Georgian (Kartvelian; Harris 1981: 75)

*Mama-m Mzia-s*  
father-ERG Mzia-DAT  
*daanteb-in-a cecxli.*  
light-CAUS-AOR:3.SG fire(ABS)  
'Father made Mzia light the fire.'

derived valency pattern:  
causer causee patient  
subject indirect direct  
object object

The second option is exemplified by Kannada:

- (27) Kannada (Dravidian; Cole & Sridhar 1977: 707)

*Raamanu manga-gal-inda*  
Rama(NOM) monkey-PL-INSTR  
*Sii-te-yannu huduki-si-danu.*  
Sita-ACC search-CAUS-3.SG  
'Rama had the monkeys search for  
Sita.'

derived valency pattern:  
causer causee patient  
subject instrumental direct  
object

The third option is chosen, for instance, by Quechua (28):

- (28) Imbabura Quechua (Cole 1982: 135)

*Juzi-ka Juan-ta ruwana-ta*  
José-TOP Juan-ACC poncho-ACC  
*awa-chi-rka.*  
weave-CAUS-3.SG  
'José made Juan weave a poncho.'

derived valency pattern:  
causer causee patient  
subject direct 2nd direct  
object object

In cases like (28), it appears that the causative valency pattern has two direct objects. How-

ever, a more detailed syntactic analysis usually shows that the patient of the caused event is not a “true” direct object in that it does not have all the properties of direct objects in the language (e.g., *ruwana* in (28) may not become the subject when the verb is passivized; see Cole 1982: 136–141 for discussion of Imbabura Quechua).

### 3.2.2. Semantic subtypes of causatives

One may distinguish two main semantic types of causatives. With the first one, the causer actively participates in the action, acting on the causee (in order to get the content of the base verb realized), which will imply some sort of coercion in case the causee is animate. This type of causative is often called the **direct causative**. In contradistinction to it, the **indirect causative** ('have someone do something') implies that the causer is conceived of as a mere instigator or distant cause of the realization of the verb content. This allows in some sort of mediation in the relation towards the causee and, since the primary interest of an animate causer consists in the realization of the base verb content, even renders the causee optional. For example, in Tamil, the verb ‘place’ has been grammaticalized as a direct causative, and the verb ‘make’ as an indirect causative:

- (29) Tamil (Dravidian; Fedson 1985: 15)
- (a) *piLLaiyait tuunka vai-tt-een*  
child:ACC sleep:INF place-PAST-1.SG  
'I made the child sleep.'
  - (b) *avaru jepam taan*  
3.HON(GEN) prayer:NOM indeed  
*noNTiyai naTakka*  
cripple:ACC walk:INF  
*cey-t-atu*  
make-PAST-3.SG.N  
'His (someone else's) prayer really  
made the cripple walk.'

The absence of a causee with an indirect causative is demonstrated in (30).

- (30) Nomatsiguenga (Arawakan; Wise 1990: 94)
- i-p-agant-ë-ri*  
3.M.ERG-give-CAUS.IND-RLS-3.M.ABS  
*kireki*  
money  
'He had (someone) give the money to  
him', i.e. 'He sent him the money.'

Often the indirect causative subsumes an **assistive** meaning ('help causee to V'), as in

Telugu, or a **permissive** meaning, as in Georgian (Comrie 1985: 334). A corollary of the act of causation and the base verb content being less integrated in the case of the indirect causative is the fact that the base verb may be negated (or qualified) independently.

The direct causative has its bias on intransitive verbs, the indirect causative, on transitive verbs. The independence of both types of causatives is evidenced by the fact that with a certain verb one causative type can be negated while the other is asserted. Apart from that it is normal for an indirect causative to take the direct one as its base.

The direct causative has a special subtype that may be termed the **immediate causative**. With it, the action represented by the base verb and the causation constitute an integral unit. This is the realm of the non-productive causative, where often the causative morpheme is fused with the base. For example, Hindi-Urdu exhibits ablauting pairs as in (31).

- |                      |                          |
|----------------------|--------------------------|
| <i>khull/khol</i>    | 'open (intr.)/(tr.)'     |
| <i>marl/maar</i>     | 'die/kill'               |
| <i>ghirl/gher</i>    | 'be surrounded/surround' |
| <i>nikall/nikaal</i> | 'emerge/extract'         |

Here the causative appears as the exogenous variant of the same state of affairs that the base represents as endogenous. (Patterns like the one from Hindi-Urdu might alternatively be regarded as anticausative variations, although historically they are causative.) Apart from so-called unaccusative (passive) intransitive verbs this pattern also applies to transitive verbs like ‘learn’, yielding ‘teach’. Immediate causatives are typically reflected by simplex verbs in languages like English.

Many languages unite the indirect and the direct causative in one formative, which allows it to occur iteratively, the inner token representing the direct causative, the outer one the indirect causative, e.g. Turkish *öldür-t* 'die-CAUS-CAUS (have killed)' (where the two suffixes are allomorphs of the same morpheme).

The place of the causee in the derived valency pattern (cf. 3.2.1) may depend on the degree of potential control that the causee participant has over the realization of the base verb content, or, vice versa, the degree of affectedness by the causer. A completely dependent causee will be realized as a direct object (or accusative), a constellation found with many intransitive bases. A less affected or more active (or simply animate) causee

will tend to be expressed as an indirect object (or dative). The causee as the executor of an action often surfaces as an instrumental. This is illustrated in (32).

- (32) Bolivian Quechua (Cole 1983)
- (a) *Nuqa wawa-ta waqa-či-ni.*  
I child-ACC cry-CAUS-1.SG  
'I made the child cry.'
  - (b) *Nuqa kurandero-man*  
I medicine.man-DAT  
*yuya-či-ni.*  
remember-CAUS-1.SG  
'I reminded the medicine man of it.'
  - (c) *Nuqa Fan-wan rumi-ta*  
I Juan-INSTR rock-ACC  
*apa-či-ni.*  
carry-CAUS-1.SG  
'I had Juan carry the rock.'

The categories of the applicative and the causative are not unrelated to each other. The directive and benefactive applicative are in various languages homophonous with a causative (cf. Song 1990, who attributes this to the utilization of allative morphemes in both cases). The instrumental applicative can turn into a causative with meanings like 'get something done by someone' providing the link between the two categories (for example in the Arawakan languages (Wise 1990), and in Kinyarwanda (Kimenyi 1980: 164–166)).

#### 4. General features of valency-changing morphology

##### 4.1. Derivation and inflection

Valency-changing categories generally have many of the properties that are considered as characteristic of derivation as opposed to inflection (cf. Art. 38). In particular, they often exhibit formal and semantic idiosyncrasies and arbitrary restrictions on productivity. Furthermore, combinations of the verbal root and the valency-changing morpheme show a tendency toward lexicalization, i.e. loss of semantic segmentability.

An example of a formal idiosyncrasy is the formation of comitative applicatives in Fijian (Schütz 1985: 132–139). There are a number of different applicative suffixes (*-tak*, *-vak*, *-mak*, etc.). Which verb takes which suffix has to be lexically specified (*-a* cross-references a 3rd person singular object):

- (33)
- |                    |                 |
|--------------------|-----------------|
| <i>lako</i>        | 'go'            |
| <i>lako-vak-a</i>  | 'go with it'    |
| <i>kaba</i>        | 'climb'         |
| <i>kaba-tak-a</i>  | 'climb with it' |
| <i>dromu</i>       | 'sink'          |
| <i>dromu-cak-a</i> | 'sink with it'  |

Semantic idiosyncrasy is found, e.g., in some German verbs with the applicative prefix *be-* (*schreiben* 'write' vs. *beschreiben* 'describe'; *nehmen* 'take' vs. *benehmen* 'behave').

A good example for lexicalization of valency-changing categories is provided by **deponents**. *Deponent* is the traditional term for Latin Passive forms (and Greek Middle forms) that lack corresponding Active forms, e.g. Latin *loquor* 'I speak' (\**loquo*), Greek *gignomai* 'I become' (\**gignō*). Since no base form exists, the Passive/Middle forms cannot be segmented into two meaning components, and the root and the valency-changing morphology are lexicalized together.

Deponent-like situations occur with valency-changing categories in many languages (cf. Croft et al. 1987). Thus, the Russian reflexive verbs in (34) could be called *reflexive deponents*.

- (34)
- |                               |                   |
|-------------------------------|-------------------|
| <i>smejat'-sja</i> 'laugh'    | * <i>smejat'</i>  |
| <i>nadejat'-sja</i> 'hope'    | * <i>nadejat'</i> |
| <i>bojat'-sja</i> 'be afraid' | * <i>bojat'</i>   |

And the Hebrew causatives in (35 b) do not have corresponding non-causative forms (they might be called *causative deponents*).

- (35)
- |                                |                  |
|--------------------------------|------------------|
| (a) <i>hi-zkir</i> 'remind'    |                  |
| <i>zaxar</i> 'remember'        |                  |
| <i>hi-lbiš</i> 'dress, clothe' |                  |
| <i>lavaš</i> 'wear'            |                  |
| (b) <i>hi-šlix</i> 'throw'     | * <i>šalax</i>   |
| <i>he-žel</i> 'begin'          | * <i>hal(al)</i> |
| <i>hi-š?ir</i> 'leave'         | * <i>ša?ar</i>   |

Finally, valency-changing can sometimes be combined with other word classes, especially adjectives and nouns. For example, the Gothic anticausative suffix *-n* (cf. (7 b)) also forms deadjectival verbs (e.g. *full-n-an* 'become full'), and the Kannada causative suffix *-s* (cf. (27)) also forms denominal verbs (e.g. *prayatn-is-* 'try' from *prayatna* 'attempt').

That valency-changing categories are located more toward the derivational end on the derivational-inflectional continuum is closely related to their function. Changing the number of participants or the nature of their semantic roles generally entails a substantial change in the situation described. In

Bybee's (1985: 13) terms, valency-changing categories are highly **relevant** to the verb's meaning, i.e., their semantic content directly affects the semantic content of the verb stem. As Bybee (1985: 81–110) demonstrates, a category's position on the derivational-inflectional continuum depends to a considerable extent on the degree of relevance of its meaning. Highly relevant meanings like those that change valency are most likely to be derivational.

#### 4.2. Valency change and voice

The verbal super-category of voice was originally used to describe the distinction between Active and Passive in Latin (as well as the analogous distinction between Active and Middle in Greek). Similar verbal categories and constructions were found in a large number of other languages, and the passive has been established as a cross-linguistic syntactic and morphological category (cf. Perlmutter & Postal 1977; Siewierska 1984; Shibatani 1985; Haspelmath 1990; Art. 108).

However, the question of how the notion of voice could or should be generalized beyond the active-passive distinction is not clear. In the Russian and Soviet linguistic tradition, the causative and the reflexive are commonly referred to as voices as well. And a proposal by Xolodovič (1970) explicitly defines voice as a regular formal expression of a **diathesis**, where diathesis is defined as a particular pairing of semantic roles and grammatical relations. Thus, "regular" valency-changing categories would be an instance of voice, according to this definition.

But more often voice is taken to be a less inclusive supercategory, comprising mainly the passive and the antipassive (in addition to the unmarked or active form). What these two categories have in common is that they do not make a dramatic change in the semantic content of a verb; rather, they present the event expressed by the verb in a different perspective. A passive clause presents the event from the perspective of the underlying direct object, but the agent is still present in the background (either omitted from the surface or expressed by an oblique phrase). This contrasts with the anticausative, where the agent is completely eliminated (cf. 4.1).

The ability to change the perspective of an event is often quite useful in discourse in order to maintain topicality. For this reason voice categories may have a high text frequency, whereas valency-changing categories

are rarer in texts. Changing the perspective from which an event is viewed is not as radical a change as changing the participants, so voice categories are not as relevant to the verb's meaning as valency-changing categories (cf. Bybee 1985: 20). It follows from this that voice categories are located further toward the inflectional end on the derivational-inflectional continuum.

#### 4.3. Regularities of morpheme order

The ordering of valency-changing morphemes with respect to the other morphemes in the word is by no means random, but is largely predictable on the basis of the principle of iconicity.

As has been demonstrated by Bybee (1985: 33–35), the order of affixes correlates strongly with the degree of their semantic relevance to the verb's stem: the more relevant categories (such as voice and aspect) tend to occur closer to the stem, while the less relevant categories (such as mood and person/number agreement) occur at a greater distance from the stem. Valency-changing categories are highly relevant to the verbal event because a change in the number and/or arrangement of the participants is invariably connected with a change in the way the event is viewed. Thus, we find that valency-changing morphemes tend to occur very close to the stem. In some languages, valency-changing categories are even expressed by changing the stem itself, e.g.:

- (36) Lithuanian
 

<i>laužti</i>	'break (tr.)'
<i>lūžti</i>	'break (intr.)'
- (37) German
 

<i>sinken</i>	'sink (intr.)'
<i>senken</i>	'sink (tr.)'

The order of valency-changing morphemes with respect to each other is also generally iconic. The valency changes expressed by a morpheme that stands closer to the stem are applied before the valency changes are expressed further away from the stem. In this way, formal changes directly reflect changes in the valency of the verb. This is the essence of Baker's (1985) Mirror Principle.

This iconic ordering is best illustrated with cases in which both orders are possible. For instance, Quechua allows both orders of causative and reflexive, with radically different resulting meanings (Muysken 1988; Baker 1985: 392).

- (38) (a) *Maqa-ku-ya-chi-n.*  
          [beat-REFL]-DUR-CAUS-3.SG  
          ‘He<sub>i</sub> is causing [him<sub>j</sub> to beat him-self<sub>j</sub>.]’  
 (b) *Maqa-chi-ku-n.*  
       beat-CAUS-REFL-3.SG  
       ‘He<sub>i</sub> lets someone<sub>j</sub> beat him<sub>j</sub>.’

Meaning differences that are iconically reflected in ordering differences can also be observed when valency-changing affixes interact with other affixes, e.g. a desiderative affix in (39).

- (39) Capanahua (Payne 1990: 228; data from Eugene Loos)  
 (a) *pi-catsih-ma-hue*  
       eat-DES-CAUS-IMP  
       ‘Make him hungry. (lit. Make him want to eat)’  
 (b) *pi-ma-catsihqu-i*  
       eat-CAUS-DES-PRES  
       ‘He wants to feed it. (lit. He wants to make it eat.)’

#### 4.4. Restrictions on possible valency changes

It seems clear that only a subset of logically possible valency changes are actually realized and expressed by morphological categories in the world's languages. For example, the following valency changes are unattested: subject-object switch (and other switches), changes from subject or direct object to indirect object, removal of an indirect object or an oblique argument. The most important constraint on valency changes is that derived valency patterns must be identical to valency patterns that occur with at least some non-derived verbs. Although it is perhaps violated in a few marked cases (e.g. doubly transitive causatives, cf. (28) above), it excludes a vast number of logically possible but unattested changes. Furthermore, valency changes most often affect subjects and direct objects, the most grammaticalized argument types, while oblique or adverbial arguments tend to resist changes (presumably because they carry their own function marking and are not as dependent on the properties of the verb).

There is an extensive literature that tries to formulate principles governing these restrictions from a perspective of autonomous syntax (cf., in particular, Williams 1981; Marantz 1984; Baker 1988). Williams (1981) argues that valency-changing morphemes may only affect the external argument (roughly,

the subject), either by externalizing a previously internal argument and thereby internalizing the previous subject (as, for instance, in the resultative), or by internalizing the previous external argument and adding a new external argument (as, for instance, in the causative). However, these restrictions are too tight, not allowing, e.g., for applicatives.

Marantz (1984) claims that valency changes that involve a change in the number or semantic roles of the participants (such as the anticausative or the deobjective) cannot be expressed solely by affixes, and cannot be wholly productive. While it is true that lexical generality is lower with more radical valency changes and higher with less radical ones (cf. the remarks in 4.2), uniquely anticausative and deobjective affixes certainly exist (e.g. Gothic *-na* and Swahili *-i/e-* are unambiguous anticausative suffixes).

In contrast to autonomous-syntactic approaches, functionalist approaches to valency-changing morphology take into account the verbal semantics and the semantic and pragmatic functions of these categories (e.g. Croft 1991). From this perspective, valency-changing morphemes serve to express an unusual or marked view (or construal) of the event expressed by the predication. The causative widens the scope of the verb's meaning to include a prior causer, while the anticausative narrows the scope of the verb's meaning by excluding the normally present agent (and, in the case of the resultative, even the dynamic part of the event, focusing entirely on the resulting state). The deobjective narrows the scope of the verb's meaning in the opposite direction, excluding the object from the construal of the event. Since the direct object is generally the most affected participant, the applicative allows an event view where an unusual participant is the most affected participant.

#### 5. Diachronic sources of valency-changing morphology

Not much is known about the diachronic sources of valency-changing categories. In part, this may be due to the fact that the languages whose history is best attested (i.e., Indo-European languages) are not rich in valency-changing morphology. But another reason is probably that valency-changing categories, like other derivational categories, do

not change as fast as some inflectional categories (e.g. tense and case categories), so in most cases the historical depth that we have is insufficient. Also, derivational categories often arise by analogy (cf. Art. 148), rather than from the grammaticalization of periphrastic constructions (cf. Art. 146), which makes their origins more difficult to identify. Nevertheless, some things can be said.

### 5.1. Sources of valency-decreasing morphology

A common source for valency-decreasing categories are reflexive pronouns which are first grammaticalized as reflexive markers. For example, the Russian reflexive suffix *-sja* (e.g. *moet-sja* ‘washes (itself)’) was grammaticalized from the early Slavic reflexive pronoun *se*. As a next step, reflexive markers may be extended to verbs with an inanimate object, whose reflexive form is naturally understood as anticausative (e.g. Russian *ot-krylsja* ‘opened (intr.)’, originally ‘opened itself’). Finally, reflexive forms may be extended to the deobjective and deaccusative functions (cf. the Lithuanian example (4 b)). The semantic mechanisms for these extensions are obscure.

The resultative differs from the other valency-decreasing categories in that it never derives from reflexive pronouns. Rather, resultatives are commonly based on resultative participles (cf. Haspelmath 1994) (plus the copula), as in the Russian example (10 c), or they are based on a perfective converb plus the copula (cf. Nedjalkov 1988, ed.: 19).

### 5.2. Sources of valency-increasing morphology

Applicatives seem to arise from the grammaticalization of an adverb or an oblique case marker and its attachment to the verb. For example, the directive applicative in German (*be-*) and Oceanic (Pawley 1973) can be traced back to a dynamic local adverb ‘at’. And according to Reh (1986), the Lango (Western Nilotic) benefactive applicative suffix *-i* (e.g. *ò-kèl-li dákô* ‘3.SG-bring-APPL woman (she brought it for the woman)’) derives from an earlier benefactive/dative preposition *ni* (\**ò-kèl-ni* > *ò-kèlli*) which is still found in the cognate languages Alur and Dholuo. Another source of applicatives is verb serialization or compounding, where one verb grammaticalizes to an applicative marker. The benefactive applicative has as its most obvious source the verb ‘give’ (cf. Foley

1986: 141 f. for Papuan languages). The directive applicative goes back to ‘see’ in the Papuan (Irian Jaya) language Dani (Bromley 1981: 107–109). The instrumental applicative comes from ‘take’ in Chickasaw (Muskocean; Munro 1983).

Morphological causatives may arise from the grammaticalization of periphrastic causatives (which typically include a causative verb such as ‘make’, ‘let’, ‘give’, ‘put’ plus a non-finite form of the caused verb). For example, the Avar (Northeast Caucasian) causative suffix *-zabi* (e.g. *t'eha-zabi* ‘make blossom’, from *t'eha-* ‘blossom’) clearly derives from a periphrastic construction with the infinitive in *-ze* plus the verb *habi* ‘make’ (*t'eha-ze habi* > *t'eha-z-abi*). Two kinds of verbal sources of the causative have been exemplified for Tamil above (29 a–b). See Song (1990) for further cross-linguistic data and speculation.

An example for the analogical origin of a causative affix is Hindi-Urdu *-vaa* (e.g. *nigal* ‘swallow’, *nigal-vaa* ‘cause to swallow’). This is the regular phonological reflex of Sanskrit *-apaya*, which in turn is an analogical extension of the older *-aya* (Proto-Indo-European \**-ejole-*). This example shows that valency-changing affixes may be so old that even the best historical-comparative information does not point to an ultimate lexical source.

## 6. Uncommon abbreviations

APPL.BEN	benefactive applicative
APPL.COMPR	comprehensive applicative
APPL.DIR	directive applicative
APPL.INSTR	instrumental applicative
APPL.LOC	locative applicative
DEACC	deaccusative
DEOBJ	deobjective
LIM	limitative

## 7. References

- Aissen, Judith L. (1987), *Tzotzil Clause Structure*. Dordrecht: Reidel
- Arnott, Donald W. (1970), *The Nominal and Verbal Systems of Fula*. Oxford: Clarendon
- Baker, Mark C. (1985), “The Mirror Principle and Morphosyntactic Explanation”. *Linguistic Inquiry* 16.3, 373–415
- Baker, Mark C. (1988), *Incorporation: A Theory of Grammatical Function Changing*. Chicago: The University of Chicago Press

- Blake, Barry (1987), *Australian Aboriginal Grammar*. London: Croom Helm
- Boeder, Winfried (1968), "Über die Versionen des georgischen Verbs". *Folia Linguistica* 11.1/2, 31–60
- Bolinger, Dwight (1971), *The Phrasal Verb in English*. Cambridge/MA: Harvard University Press
- Bresnan, Joan & Moshi, Lioba (1990), "Object Asymmetries in Comparative Bantu Syntax". *Linguistic Inquiry* 21.2, 147–85
- Bromley, Myron H. (1981), *A Grammar of Lower Grand Valley Dani*. Canberra: Australian National University (Pacific Linguistics C 63)
- Bybee, Joan L. (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam, Philadelphia: Benjamins (Typological Studies in Language 9)
- Chomsky, Noam A. (1981), *Lectures on Government and Binding*. Dordrecht, Providence/RI: Foris
- Chung, Sandra (1976), "An Object-creating Rule in Bahasa Indonesia". *Linguistic Inquiry* 7, 41–87
- Churchward, C. Maxwell (1953), *Tongan Grammar*. London: Oxford University Press
- Cole, Peter (1982), *Imbabura Quechua*. London etc.: Croom Helm
- Cole, Peter (1983), "The Grammatical Role of the Causee in Universal Grammar". *International Journal of American Linguistics* 49.2, 115–133
- Cole, Peter, & Sridhar, S. N. (1977), "Clause Union and Relational Grammar: Evidence from Hebrew and Kannada". *Linguistic Inquiry* 8, 700–713
- Comrie, Bernard (1985), "Causative Verb Formation and Other Verb-deriving Morphology". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III*. Cambridge: Cambridge University Press, 309–348
- Couper-Kuhlen, Elisabeth (1979), *The Prepositional Passive in English*. Tübingen: Niemeyer (Linguistische Arbeiten 81)
- Craig, Colette & Hale, Ken (1988), "Relational Preverbs in some Languages of the Americas: Typological and Historical Perspectives". *Language* 64, 312–344
- Creider, Chet A. & Creider, Jane Tapubei (1989), *A Grammar of Nandi*. Hamburg: Buske (Nilo-Saharan 4)
- Croft, William (1991), *Syntactic Categories and Grammatical Relations*. Chicago, London: The University of Chicago Press
- Croft, William & Shydkroft, Hava Bat-Zeev & Kemmer, Suzanne (1987), "Diachronic Semantic Processes in the Middle Voice". In: Giacalone Ramat, Anna & Carruba, Onofrio & Bernini, Giuliano (eds.), *Papers from the 7th International Conference on Historical Linguistics* [Pavia 1985]. Amsterdam, Philadelphia: Benjamins (Current Issues in Linguistic Theory 48), 179–192
- Dayley, Jon P. (1989), *Tümpisa (Panamint) Shoshone Grammar*. Berkeley etc.: University of California Press (University of California Publications in Linguistics 115)
- Dik, Simon C. (1978), *Functional Grammar*. Amsterdam etc.: North Holland (North Holland Linguistic Series 37)
- Dixon, Robert M. W. (1977), *A Grammar of Yidiñy*. Cambridge: Cambridge University Press (Cambridge Studies in Linguistics 19)
- Dixon, Robert M. W. (2000), "A Typology of Causatives: Form, Syntax and Meaning". In: Dixon, Robert M. W. & Aikhenvald, Alexandra Y. (eds.), *Changing Valency*. Cambridge: Cambridge University Press, 30–83
- Fedson, Vijayarani Jotimuttu (1985), "A Note on Periphrastic Serial or Compound Causative Constructions in Tamil". *Chicago Linguistic Society* 21.2, 13–20
- Foley, William A. (1986), *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press
- Fox, Barbara & Hopper, Paul (1994, eds.), *Voice: Form and Function*. Amsterdam, Philadelphia: Benjamins
- Frank, Paul (1990), *Ika Syntax*. Arlington: Summer Institute of Linguistics, University of Texas at Arlington (Publ. Ling. 93)
- Geniušienė, Emma (1987), *The Typology of Reflexives*. Berlin, New York: Mouton de Gruyter (Empirical Approaches to Language Typology 2)
- Grimshaw, Jane A. (1990), *Argument Structure*. Cambridge/MA: MIT Press
- Harris, Alice C. (1981), *Georgian Syntax: A Study in Relational Grammar*. Cambridge: Cambridge University Press
- Harrison, Sean P. (1982), "Proto-Oceanic \*aki(ni) and the Proto-Oceanic Periphrastic Causatives". In: Halim, Amran & Carrington, Louis & Wurm, Stephen A. (eds.), *Papers from the Third International Conference on Austronesian Linguistics, Vol. I: Currents in Oceanic*. Canberra: Australian National University (Pacific Linguistics C 74), 179–230
- Haspelmath, Martin (1987), *Transitivity Alternations of the Anticausative Type*. Köln: Institut für Sprachwissenschaft der Universität zu Köln (Arbeitspapiere, N. F. 4)

- Haspelmath, Martin (1990), "The Grammaticization of Passive Morphology". *Studies in Language* 14.1, 25–71
- Haspelmath, Martin (1993a), *A Grammar of Lezgian*. Berlin: Mouton de Gruyter (Mouton Grammar Library 9)
- Haspelmath, Martin (1993b), "More on the Typology of Inchoative/Causative Verb Alternations". In: Comrie, Bernard & Polinsky, Maria (eds.), *Causatives and Transitivity*. Amsterdam: Benjamins, 87–120 (Studies in Language Companion Series 23)
- Haspelmath, Martin (1994), "Passive Participles across Languages". In: Fox & Hopper (eds.), 151–177
- Hutchison, John P. (1981), *The Kanuri language*. Madison: University of Wisconsin, African Studies Program
- Innes, Gordon (1966), *An Introduction to Grebo*. London: School of Oriental and African Studies
- Károly, Sándor (1982), "Intransitive-transitive Derivational Suffixes in Hungarian". In: Kiefer, Ferenc (ed.), *Hungarian Linguistics*. Amsterdam, Philadelphia: Benjamins (Linguistic & Literary Studies in Eastern Europe 5)
- Kemmer, Suzanne (1994), "Middle Voice, Transitivity, and the Elaboration of Events". In: Fox & Hopper (eds.), 179–230
- Kibrik, A[ndrej] A. (1990), "‘Poslelogi’ v navaxo: poterja statusa časti reči". In: Alpatov, V[ladimir] M. (ed.), *Časti reči: teorija i tipologija*. Moskva: Nauka, 234–253
- Kimenyi, Alexandre (1980), *A Relational Grammar of Kinyarwanda*. Berkeley etc.: University of California Press (University of California Publications in Linguistics 91)
- Kozinceva, Natalija A. (1981), "Refleksivnye glagoly v armjanskom jazyke". In: Krakovskij (ed.), 81–98
- Marantz, Alec (1984), *On the Nature of Grammatical Relations*. Cambridge/MA: MIT Press (Linguistic Inquiry Monographs 10)
- Mel'čuk, Igor A. (1988), *Dependency Syntax*. Albany: State University of New York Press
- Mithun, Marianne (1989), "The Subtle Significance of the Locus of Morphologization". *International Journal of American Linguistics* 53.3, 265–82
- Munro, Pamela (1983), "When ‘same’ is not ‘not different’". In: Haiman, John & Munro, Pamela (eds.), *Switch Reference and Universal Grammar*. Amsterdam, Philadelphia: Benjamins (Typological Studies in Language 2), 223–243
- Muysken, Pieter (1988), "Affix Order and Interpretation in Quechua". In: Everaert, Martin & Evers, Arnold & Huybregts, Riny & Trommelen, Mieke (eds.), *Morphology and Modularity*. Dordrecht, Providence/RI: Foris (Publications in Language Sciences 29), 259–279
- Nedjalkov, V[ladimir] P. & Sil'nikič, Georgij G. (1969), "Tipologija morfoložičeskogo i leksičeskogo kauzativov". In: Xolodovič (ed.), 20–50
- Nedjalkov, Vladimir P. (1988, ed.), *Typology of Resultative Constructions*. Amsterdam, Philadelphia: Benjamins (Typological Studies in Language 12)
- Pawley, Andrew (1973), "Some Problems in the Proto Oceanic Grammar". *Oceanic Linguistics* 12, 103–188
- Payne, Doris L. (1990), "Morphological Characteristics of Lowland South American Languages". In: Payne (ed.), 213–241
- Payne, Doris L. (1990, ed.), *Amazonian Linguistics*. Austin: University of Texas Press
- Perlmutter, David M. & Postal, Paul M. (1977), "Toward a Universal Characterization of Passivization". *Berkeley Linguistics Society* 3, 394–417
- Port, Robert F. (1981), "The Applied Suffix in Swahili". *Studies in African Linguistics* 12.1, 71–82
- Reh, Mechthild (1986), "Where Have all the Case Prefixes Gone?". *Afrikanistische Arbeitspapiere* 5, 121–134. (Köln: Institut für Afrikanistik der Universität zu Köln)
- Ross, Malcom D. (1988), *Proto Oceanic and the Austronesian Languages of Western Melanesia*. Canberra: Australian National University (Pacific Linguistics C 98)
- Saeed, John Ibrahim (1987), *Somali Reference Grammar*. Wheaton/MD: Dunwoody
- Schaub, Willi (1982), *Babungo*. London, etc.: Croom Helm
- Schütz, Albert (1985), *The Fijian Language*. Honolulu: The University of Hawaii Press
- Shibatani, Masayoshi (1976, ed.), *The Grammar of Causative Constructions*. New York: Academic Press (Syntax and Semantics 6)
- Shibatani, Masayoshi (1985), "Passives and Related Constructions: A Prototype Analysis". *Language* 61, 821–848
- Shibatani, Masayoshi (1990), *The Languages of Japan*. Cambridge: Cambridge University Press
- Shibatani, Masayoshi (1996), "Applicatives and Benefactives: A Cognitive Account". In: Shibatani, Masayoshi & Thompson, Sandra A. (eds.), *Gram-*

- matical Constructions: Their Form and Meaning.* Oxford: Oxford University Press, 157–194
- Siewierska, Anna (1984), *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm
- Song, Jae Jung (1990), “On the Rise of Causative Affixes: A Universal-Typological Perspective”. *Lingua* 82, 151–200
- Song, Jae Jung (1996), *Causatives and Causation: A Universal-Typological Perspective*. London: Longman
- Tesnière, Lucien (1959), *Éléments de syntaxe structurale*. Paris: Klincksieck
- Weir, E. M. Helen (1986), “Footprints of Yesterday’s Syntax: Diachronic Development of Certain Verb Prefixes in an OSV Language (Nadeb)”. *Lingua* 68.4, 291–316
- Williams, Edwin (1981), “Argument Structure and Morphology”. *The Linguistic Review* 1, 81–114
- Wise, Mary Ruth (1971), *Identification of Participants in Discourse: A Study of Aspects of Form and Meaning in Nomatsiguenga*. Summer Institute of Linguistics Publ. 28
- Wise, Mary Ruth (1990), “Valence-changing Affixes in Maipuran Arawakan Languages”. In: Payne (ed.), 89–116
- Xolodovič, A[leksandr] A. (1969, ed.), *Tipologija kauzativnyx konstrukcij. Morfoložičeskij kauzativ*. Leningrad: Nauka
- Xolodovič, A[leksandr] A. (1970), “Zalog. 1: Opredelenie. Isčislenie”. In: *Kategorija zaloga. Materialy konferencii*. Leningrad: Institut jazykoznanija AN SSSR, 2–26
- Xrakovskij, V[iktor] S. (1981), “Diateza i referentnost’ (k voprosu o sootnošenii aktivnyx, passivnyx, refleksivnyx i reciproknyx konstrukcij)”. In: Xrakovskij (ed.), 5–38
- Xrakovskij, V[iktor] S. (1981, ed.), *Zalogovye konstrukcii v raznostrukturnyx jazykax*. Leningrad: Nauka
- Young, Robert W. & Morgan, William (1980), *The Navajo Language: A Grammar and Colloquial Dictionary*. Albuquerque: University of New Mexico

Martin Haspelmath, Leipzig (Germany)  
 Thomas Müller-Bardey, Mainz (Germany)

## 108. Voice

1. Introduction
2. Definition: The active and the passive
3. Voice as a family of constructions
4. The middle
5. Valency-increasing passives
6. The inverse system
7. Voice in ergative languages
8. Voice in Philippine languages
9. Two dimensions of voice
10. Voice morphology
11. Concluding remarks
12. Uncommon abbreviations
13. References

### 1. Introduction

Voice (Greek *diáthesis*, Latin *vox*), among morphological categories, is perhaps the most elusive. For one thing, unlike categories such as person and tense, it is not clear what its semantic basis is. For another, there is wide disagreement among linguists as to what is to be included in the category of voice. Furthermore, there is terminological complexity emanating from different applications of terms such as passive, ergative, im-

personal, and middle to different constructions by different linguists. This last problem points to the difficulty of doing cross-linguistic comparison of voice systems, which is essentially caused by the first two problems.

There are two major factors contributing to these difficulties. One has to do with the relationship between morphology and syntax in the recognition of a grammatical category. If one restricts the notion of voice to the domain of morphology, some languages, especially isolating-type languages, will not have certain voice categories, despite the fact they may have comparable constructions that are marked morphologically in other languages. However, once category labels are freed from morphological manifestations, they may be used in reference to phenomena widely different from the conventional use. Such is the case, for instance, in the use of the terms ‘middle’ and ‘ergative’ in some syntactic work.

The other factor has to do with the multiplicity of function that a voice-related morphology tends to have. For example, an ele-

ment recognizable as a passive morpheme may often have other functions such as turning clauses into reflexive, spontaneous, or resultative constructions. Therefore, if one further narrows down the scope of search for morphological elements that uniquely mark voice categories, then the actual number of languages that come under scrutiny becomes extremely small.

Thus, a fruitful study of voice must be negotiated between the two extremes of unconstrained use of category labels and too narrow morphological restrictions. In this article we endeavor to clarify various voice-related terms and identify points of disagreement among linguists and their causes.

## 2. Definition: The active and the passive

The narrowest definition of voice – and the one most widely maintained – identifies it as an inflectional category whose members exhibit different relationships of subject and object to the verb, as illustrated in the following pair of Japanese sentences:

- (1) (a) *Taroo ga Hanako o korosi-ta.*  
Taro NOM Hanako ACC kill-PAST  
(active)  
'Taro killed Hanako.'
- (b) *Hanako ga Taroo ni korosa-re-ta.*  
Hanako NOM Taro DAT kill-PASS-PAST  
(passive)  
'Hanako was killed by Taro.'

As the labels ‘active’ and ‘passive’ above indicate, the active voice refers to a subcategory whose typical manifestation involves (i) a basic form of a verb, and (ii) the alignment of an agentive argument of a transitive clause with the grammatical function of subject, and (iii) the alignment of the patient with the object function. The passive voice transposes these alignments such that now the morphologically marked verb co-occurs with the patient subject and with the agent argument either in an oblique function or unrealized syntactically.

The contemporary, popular understanding of voice thus points to the alternations in the correspondence of semantic (participant) roles and grammatical functions. The relationship between voice categories and this semantico-syntactic role correspondence has

been made clear by separating out the notions of voice and diathesis as in the practice of Russian linguists, especially those associated with the Leningrad school tradition (cf. Xolodovič 1974; Khrakovskiy 1979; Mel'čuk 1994). In this tradition, the Greek grammatical term **diathesis** ‘state, disposition’ (see 4) has been given a special significance representing patterns of correspondence between syntactic level nominal expressions (**actants**) and semantic-level participant roles (**participants**). In essence a verbal lexeme is associated with a particular set of participant roles. These roles are associated with syntactic noun phrases according to unique patterns with specific lexical forms of the lexeme. For example, the Japanese lexeme KOROSU ‘kill’ has associated with it participant roles **agent** and **patient**. In its basic, underived lexical form *koros-*, the agent role is associated with a subject noun phrase and the patient role with an object noun phrase. However, when the verb occurs with the suffix *-(ra)re*, the agent role is linked with an oblique noun phrase, and the patient role with a subject noun phrase. The pattern of correspondence exhibited by the basic form of a given lexeme (e.g. Japanese *koros-* ‘kill’) is considered to be the basic diathesis, whereas the patterns of correspondence displayed by other forms are derived diatheses whose derived status is typically indicated by morphological means.

In the Leningrad practice, the term **voice** is used in reference to the morphological categories reflecting diathetic patterns. Thus **active voice** refers to the category of underived verb forms associated with the basic diathesis, while **passive voice** refers to the category of verb forms that encode the derived diathesis in which the agent role is not linked with a subject noun phrase. The relationships between verb forms, diatheses, and voices are shown below, where Geniušienė's (1987) system of diathetic representations is adopted:

- (2) verb form: *koros-* ‘kill’  

diathesis:	person 1	person 2
	agent	patient
	subject	direct object
voice:	active	
- (3) verb form: *korosa-re-* ‘kill-PASS’  

diathesis:	person 1	person 2
	agent	patient
	$\emptyset$ /oblique	subject
	object	
voice:	passive	

A possible definition of voice along the lines of the conception outlined above would be the following one:

"The category of voice is an inflectional category such that its grammemes specify such modifications of the basic diathesis of a lexical unit that do not affect its propositional meaning." (Mel'čuk 1994: 11).

Under this definition, voice is not associated with any semantic category, and its function lies in signaling changes in the syntactic alignment of arguments associated with a particular lexeme.

While the definition above renders those working definitions that are tacitly held by practicing linguists in identifying voice and its subcategories, one encounters a number of problems in actually applying such a definition. Prominent among them are: (i) the grammatical expression of inflectional categories, (ii) the semantic and valency effects of voice conversion, (iii) the notion of grammatical functions, and (iv) the notion of the basic diathesis. In what follows these and other related topics will be taken up separately.

### 3. Voice as a family of constructions

The idea of voice as an inflectional category derives from the voice systems of classical languages of the Indo-European family, in which voice categories manifest themselves in combination with inflectional endings for person and number. Observe, for example, Latin 2nd person singular active *-s*, passive *-ris*; 2nd person plural active *-tis*, passive *-mini*, where the component parts for the categories of person, number, and voice are not clearly separable.

However, in many languages voice morphology is not directly connected with inflectional paradigms. Indeed, while voice categories are typically associated with affixing morphology, whether a morphological expression itself is to be taken as an integral component of the definition of voice can be called into question. Those working within the Leningrad tradition take a narrower view:

"[Voice is] a regular marking in the verb of the correspondences between units at the syntactic level and units at the semantic level. In short, voice is a diathesis grammatically marked in the verb." (Xolodovič 1970 as quoted in Geniušienė 1987: 52 f.)

"... diathesis is a semantico-syntactic and universal category: any verbal lexeme of any language has at

least one diathesis; voice, however, is only a morphological and therefore, non-universal category ..." (Khrakovskiy 1979: 291)

As an example of language in which there is no voice, Khrakovskiy points out Old Chinese, which, in keeping with its isolating character, shows no change in verbal form. See the following Mencius phrase (Hashimoto 1988:340):

- (4) *Láo xīn zhě zhì rén láo*  
trouble mind person rule people trouble  
*lì zhě zhì yú rén.*  
strength person be.rules at/by people  
'Those who work with their brain rule [people], and those who work with their brawn are ruled [by people].'

Certain verbs of Central Yupik Eskimo provide another example of passive that involves no morphological marking in the verb (Miayoka 1984).

- (5) (a) *Arna-m neqa iir-aa.*  
woman-ERG fish.ABS hide-3.SG&3.SG  
'The woman hides the fish.'  
(b) *Neqa iir-tuq.*  
fish.ABS hide-3.SG  
'The fish is hidden (by someone).'

Notice that (5 b) involves more than simple agent suppression of a transitive clause, as the clause has been turned into intransitive; notice the change in the verbal endings, from the transitive *-aa* to the intransitive *-tuq*.

While these languages show extreme cases where diathetic alternation is not signaled by a change in form of the main verbs, familiar European languages also present some problem for a view that voice requires a morphological manifestation localized in the main verb. On the other hand, one may maintain that an inflectional category need not be expressed locally and that its manifestation can be dispersed over several grammatical elements as in the so-called periphrastic passives such as the English *be*-passive and the German *werden* (become)-passive shown in (6):

- (6) (a) *John was hit by Bill.*  
(b) *Hans wurde von seinem Vater*  
Hans became by his father  
*bestraft.*  
punished  
'Hans was punished by his father.'

Once such a move is made, it is a minor step to widen the concept of voice as something definable in terms of grammatical construction attaching a minor significance to the

morphological aspect. Relational Grammar, thus, offers a universal characterization of the passive voice in terms of a construction which embodies a change in grammatical functions of direct object to subject (Perlmutter & Postal 1983). Morphological repercussions of this change vary from one language to another and thus play a minor role in the universal definition of voice within this framework.

Relational Grammar's disregard for the morphological aspect in its definition of voice is matched by a more recent functional framework, in which not only morphology but also syntax plays a considerably minor role, as voice is now defined in terms of communicative functions. The fundamental pragmatic domain of voice is taken to be the distribution of topicality between the agent and patient arguments of a transitive clause, according to which major voice categories are defined as follows (Givón 1990: 566; 1994, ed.):

**Active voice:** The agent is more topical than the patient, but the patient retains considerable topicality.

**Passive voice:** The patient is more topical than the agent, but the agent is extremely non-topical ('suppressed', 'demoted').

**Antipassive voice:** The agent is more topical than the patient, and the patient is extremely non-topical ('suppressed', 'demoted').

**Inverse voice:** The patient is more topical than the agent, but the agent retains considerable topicality.

In the extreme functionalist framework, functional unity alone determines the membership of a particular voice category such that even word order change without any alterations in morphology or grammatical function is considered as a voice mechanism that converts an expression from, for example, the active to the inverse (see Givón 1994, ed.).

A more tempered approach avoiding all these extreme views defines voice categories in terms of prototypes associated with features ranging over pragmatic functions and overall grammatical properties that characterize the central members of each category. Shibatani (1985: 837) offers the following characterization of the passive prototype:

- (7) Prototype characterization of the passive
  - (a) Primary pragmatic function: Defocusing of agent.

- (b) Semantic properties:
  - (i) Semantic valence: Predicate (agent, patient).
  - (ii) Subject is affected.
- (c) Syntactic properties:
  - (i) Syntactic encoding:
    - agent  $\Rightarrow \emptyset$  (not encoded)
    - patient  $\Rightarrow$  subject.
  - (ii) Valence of P[predicate]:
    - Active = P/n
    - Passive = P/n-1.
- (d) Morphological property:
  - Active = P;
  - Passive = P[+passive].

There are several motivations for this kind of prototype approach. One is a consideration that all the extreme views reviewed above capture only partial aspects of voice constructions ignoring features that typically covary with voice conversion. For example, the diathetic alternation between the active and the passive is typically associated with morphological alteration. And so is the functional difference.

Another motivation for the prototype approach is the fact that there are various constructions that show partial resemblances to what can be recognized as a prototypical passive construction such that it is difficult to demarcate constructions in a discrete manner – various constructions are similar to a certain extent; what needs to be clarified are the respects in which these constructions are similar and dissimilar. For example, although Old Chinese has no morphological indication in the verb, it exhibits a diathetic alternation similar to the one involved in passive conversion, and the derived structure shares pragmatic, semantic, and syntactic properties with a prototypical passive construction. More interesting is the fact that, as discussed below, both synchronically and diachronically voice constructions show similarities both among themselves and to other types of construction motivating the view that voice categories are inhabited by a family of grammatical constructions which share resemblances of one kind or another. The prototype characterization in (7) leaves the nature of the passive marking unspecified, but from a morphological perspective on voice phenomena, the nature of passive morphology is of prime interest. Indeed, it is the morphological unity, which will be dealt with below, that overtly indicates the nature of voice as something comprising of a family of constructions. This

concept of voice as a family of constructions is in fact a useful one in understanding the voice opposition observed in the earliest forms of Indo-European languages (e.g. Sanskrit and Classical Greek) and elsewhere, namely the opposition of the active and the middle.

#### 4. The middle

The **middle** (or **medial**) voice is considered to be the most heterogeneous voice category. But this heterogeneity can be only apparent because the variety of expressions encoded as middle are expressed by separate constructions in other languages. A typical middle form and its ‘ambiguity’ can be observed in the following Classical Greek middle expression from Homer (Andersen 1991: 51):

- (8) *és hr' asamínthous*  
 to then bathing.tub (ACC.PL)  
*bá-nt-es*  
*go* (AOR)-PART.AOR-NOM.PL  
*eūxéstas*                    *loúsa-nto*  
 well.polished (ACC.PL) wash-AOR-3.PL.MID  
 ‘Then having climbed into the well polished bathing tubs they (a) washed themselves, (b) washed each other, (c) were (automatically) cleansed, (d) were washed.’

The middle inflection in Classical Greek expresses a range of concepts that are expressible, in English and some other languages, by four (and more) separate constructions; (a) the reflexive, (b) the reciprocal, (c) the spontaneous, and (d) the passive construction. Because of this, modern grammarians tend to view the middle voice as having a number of functions or senses each of which is uniquely expressed by a separate construction in some other languages. But such a view fails to capture the essence of the voice mechanisms of early Indo-European languages, from which the modern notions of voice as well as voice-related grammatical terms are inherited – some via Latin translations.

Unfortunately there is some terminological inconsistency in the Greek grammatical tradition handed down from the classical grammarians such as Dionysius Thrax, whose *Téchnē grammatikē* (ca. 100 B.C.) is considered a standard work on Classical Greek (cf. Art. 6). Against the two distinct sets of inflectional endings for person and number, which are identified as ‘active’ and

‘middle’, Dionysius distinguishes three diatheses ‘state, dispositions’ or voices: *enérgeia* ‘activity’ (from which the modern term ‘active’ evolved via Latin *activum*), *páthos* ‘affection’ (> Latin *passivum* > ‘passive’), and *mesótēs* (> ‘medium’, ‘middle’). Among these, the first two voice categories are primary ones that correlate directly with the inflectional categories of ‘active’ and ‘middle’. Dionysius exemplifies the diathesis *energeia* with *tuptō* (hit. 1. SG.ACT) ‘I hit’ and *pathos* with *tuptomai* (hit. 1.SG.MID) ‘I undergo hitting’. That is, as a first approximation, the major voice categories of Classical Greek of *energeia* and *pathos* can be said to be correlated with the semantic distinction of whether the subject of the verb affects others (*energeia*) or it itself is affected (*pathos*). The *mesótes* category, as its name implies, combines the features of the two major voice categories; i.e. the active inflection with the *pathos* meaning or the middle inflection with the *energeia* meaning. Thus, the form identified as the middle inflection does not correlate directly with Dionysius’ *mesótes* voice. Rather the middle inflection represents the *pathos* voice, and this is what is normally recognized as the middle voice by modern grammarians. (One needs to exercise discretion in reading Kemp’s (1987) translation of *Téchnē*, in which ‘*páthos*’ is straightforwardly translated as ‘passive’.) Notice that the four ‘modern’ interpretations of the middle form in (8) all express the *pathos* meaning, i.e. the subject represents an affected entity or ‘the locus of the principal effects of the verbally denoted action’ (Klaiman 1991: 106).

In some modern syntactic work, the term ‘middle constructions’ is used to designate the construction type represented by the English sentence *This book sells well* and its analogs in other languages (e.g. German *Dieses Buch verkauft sich gut*) (Fagan 1992). To the extent that these constructions combine the features of active inflection and the *pathos* meaning, the designation of them as ‘middle constructions’ is in line with Dionysius’ original *mesótes* voice. But one must take caution in not identifying them with the classical middle-inflecting forms expressing the *pathos* voice, whose members show greater resemblances to actively inflected reflexive and reciprocal constructions, on the one hand, and to passives and the anticausative, spontaneous construction, on the other (see below).

In the Sanskrit grammatical tradition, the active-middle opposition has been discussed

in the following two terms: *parasmai padam* ‘a word for another’ characterizing the active and *ātmane padam* ‘a word for oneself’ the middle (Burrow 1973: 293). The distinction is illustrated by *kaṭam karoti* ‘he makes a mat (for another)’ and *kaṭam kurute* ‘he makes a mat (for himself)’.

The term ‘medio-passive’ is often employed in designating the passive use/sense of the middle voice. Languages, however, may later develop distinct passive constructions as in both Sanskrit and Greek, where, in the former, they are accommodated within the middle inflectional category. The passive in these languages has the function of isolating the medio-passive sense of the middle voice, namely that the subject is uniquely patientive, with the implication of an agent distinct from it. Notice that the middle subject can be simultaneously agentive and patientive as in those forms conveying reflexive and reciprocal senses.

The active-middle voice system is not confined to these classical Indo-European languages and is seen in various modern languages of different genetic affiliations, and the range of the middle functions is strikingly similar from one system to another (see Klaiman 1991; Kemmer 1993). This fairly widespread voice system poses a number of problems for the definition of voice set forth earlier in 2. Languages with a middle voice seem generally to split up the verbal lexicon into three types of inflecting forms: (i) those that show alternations between active and middle, (ii) active-only forms (**activa tantum**, e.g. Classical Greek *bainō* ‘go’, Sanskrit *ad-* ‘eat’, and (iii) middle-only forms (**media tantum**), e.g. Classical Greek *keimai* ‘lie’, Sanskrit *labh-* ‘receive’ – middle-only forms with an active meaning (e.g. Greek *kharizomai* ‘gratify’), are known as deponent verbs (Smyth 1956: 107). The latter two categories contain a large number of intransitive verbs. Now, the active-only and the middle-only verbs obviously do not show alternations in their basic diathesis, and thus the active/middle distinction here is not correlated with ‘modifications of the basic diathesis of a lexical unit’.

One of the important roles of the middle inflection for the alternating forms is the process known as anticausative or decausative that yields intransitive verbs of the spontaneous meaning.

E.g. Classical Greek, *tēk-ein* ‘melt-INF.ACT (melt (something))’ vs. *tēk-esthai* ‘melt-INF.MID ((something) melts); Sanskrit *vardhati*

‘increase:3.SG.ACT (increases, makes bigger)’ vs. *vardhate* ‘increase:2.SG.MID ((something) increases); Fula *bill-ii* ‘squash-GENERAL.PAST.ACT (squashed (something))’ vs. *bill-ake* ‘squash-GENERAL.PAST.MID ((something) squashed)’.

Here, not only the diathesis but also the meaning is changed from causative to spontaneous. In other words, the middle inflection for alternating verbs may change the propositional meaning of the lexical item, again conflicting with the definition of voice given in 2.

The active-middle voice system cuts across transitive/intransitive valence division, as its basis is semantic rather than diathetic alternations upon which modern conceptions of voice are founded. Indeed, a general understanding of voice in modern times is typically based on only one specific voice opposition, namely that of active and passive, an alternation of which is generally held not to affect the propositional meaning of the basic diathesis. Yet, the middle voice systems of classical Indo-European languages and others countenance alternations that change the semantic valence of the basic diathesis and hence its propositional meaning. As we shall see in later sections, this situation turns out to be rather normal across languages, indicating the difficulty of separating voice phenomena clearly apart from valence-changing, and hence derivational, processes.

## 5. Valency-increasing passives

The middle voice expresses not only those situations where the diathetic arrangement is altered at the syntactic level, as in a typical passive situation, but also situations where the number of arguments is affected, as in the anticausative derivation that has the effect of deleting the agent role altogether. Some passive constructions show an opposite effect, namely an increase in valence. The constructions in question are known as indirect or adversative passives, and Japanese is well-known for this type of passive exemplified below:

- (9) (a) *Ame ga hut-ta.*  
rain NOM fall-PAST  
‘Rain fell/It rained.’
- (b) *Taroo ga ame ni huru-re-ta.*  
Taro NOM rain DAT fall-PASS-PAST  
‘It rained on Taro.’
- (10) (a) *Hanako ga sin-da.*  
Hanako NOM die-PAST  
‘Hanako died.’

- (b) *Taroo ga Hanako ni*  
 Taro NOM Hanako DAT  
*sina-re-ta.*  
 die-PASS-PAST  
 'Hanako died on Taro.'

Here the passive morpheme introduces a new argument, the passive subject, and as such the resulting passive form changes a propositional meaning in an opposite direction from that of the middle, anticausative derivation. Notice that the active (a) sentences above do not admit a nominal being adversely affected either as a direct object or in the dative or other oblique case. The problem with the Japanese adversative passive is that, unlike derivational processes, it is completely regular applying to both intransitive and transitive verbal bases. Again, if we were to maintain that voice phenomena do not alter the basic propositional meaning, then we must divide Japanese passives into two types – one being voice and the other something else – disregarding the morphological unity. Tungusic languages such as Even and Evenki (see Nedyalkov 1991; Malchukov 1993) as well as Korean and Chinese also pose the same problem with their valence-increasing passives.

## 6. The inverse system

The Algonquian Amerindian family is famous for its system of verbal marking that is often confused with the active-passive opposition. In Plains Cree, for example, a topicality difference between a third person agent and a third person patient has the effect of triggering different inflectional forms. When the agent is more topical than the patient such that it is 'typically the most central to the discourse' (Dahlstrom 1991: 31), a **direct** form of the verb is used, whereas in the reverse situation an **inverse** form is chosen. The topical third person is called 'proximal' and the peripheral one 'obviative'.

- (11) (a) /wa.pam-e.-w/  
 see-DR-3  
 'He [prox] sees him [obv].'  
 (b) /wa.pam-ekw-w/  
 see-INV-3  
 'He [obv] sees him [prox].'

The status of the (11 b) form has been controversial. The relationship between the direct and the inverse is in some sense similar to that between the active and the passive. In

the latter when the agent is the subject, the active occurs, and when the patient is subject, the passive obtains. In the direct-inverse system, when the agent is proximal, the direct form occurs, and when the patient is proximal, the inverse form obtains. Indeed, if we had agent John and patient Bill in (11), switching of the proximal status causes no meaning difference; i.e. John [prox] sees-direct Bill [obv] vs. John [obv] sees-inverse Bill [prox], the latter of which can be rendered as 'Bill was seen by John.'

There are, however, two considerations that differentiate the two systems. The first concerns situations in which a non-third person is involved, where the obviation hierarchy is fixed as 2 > 1 > 3 (abbreviated) and discourse-conditioned switching of the proximal status is not permitted. Thus when a first person is acting on a third person, the direct form of a verb is consistently used. The use of the inverse form in this case does not express an equivalent situation in which only the proximal status of a third person patient is affected. Instead it expresses a new situation where a third person acts on a first person.

- (12) (a) /ni-wa.pam-a.-w/  
 1-see-DR-3  
 'I see him.'  
 (b) /ni-wa.pam-ekw-w/  
 1-see-INV-3  
 'He sees me.'

Notice that the only difference between (12 a) and (12 b) is the direct/inverse choice, but the meanings of the two forms differ. Certainly one may maintain that (12 b) is to be rendered as the passive form 'I was seen by him' deriving from the hypothetical underlying direct form 'he sees me' via obligatory passivization (see Jolley 1982 for such an analysis). However, Dahlstrom (1991) offers convincing evidence that the inverse forms (11 b) and (12 b) are active sentences in which the agents are not only arguments but also subjects, in contradistinction to the passive agent implied in real passive clauses of the following type, where the agent is neither an argument nor a subject.

- (13) *ni-sa.kih-ikawi-n*  
 1-love-PASS-SG  
 'I am loved.'

The inverse system here is part of the inflectional paradigm. The system can be consid-

ered as a voice system as it plays the role of determining grammatically prominent arguments, subjects and proximal nominals (see 9).

## 7. Voice in ergative languages

The definition of voice set out in section 2 is stated in terms of the notion of diathesis, which in turn hinges on the notion of grammatical functions of subject and object. There are at least two groups of languages in which the notion of grammatical voice as we understand it in terms of subject and object is unclear. The so-called ergative languages is one group and Philippine languages constitute the other group.

Understandings of the nature of ergative languages among linguists are not uniform and the question of which constituent of a transitive, ergative clause is the subject is among those topics hotly debated. The traditional view regards the patient nominal marked by an absolute (or nominative) case (often zero marking) as the subject of a transitive clause, whereas a more recent view holds that the subject is the ergative-marked agent nominal. The former considers the morphological case-marking property, which is often reinforced by an agreement phenomenon, to be a decisive factor in the determination of the grammatical function, whereas the latter emphasizes the syntactic roles played by the agentive nominal such as the control of co-referential deletion and of reflexive-binding.

These two views maintain that the grammatical functions primarily defined on the basis of accusative-type languages are applicable to ergative constructions. There is yet a third view that posits grammatical relations different from subject and object, namely ergative and absolute (or nominative) that are indicated by the ergative marking and the absolute (zero) marking. We assume the correctness of this approach, which necessitates positing of a neutral relational system that embraces the subject-object system and the absolute-ergative system. (Such a system is needed for accounting for the Philippine situation too; see 8.) A system embodying the relational hierarchy in terms of the notions of the primary term (or argument) and the secondary term serves the purpose (Shaumyan 1985). The primary arguments are the most indispensable arguments of

clauses, being instantiated by the subject of a clause encoded in the nominative-accusative pattern, and by the absolute of a clause encoded in the absolute-ergative pattern. The secondary terms are the object and the ergative nominals.

A once-popular view that the ergative construction is a passive no longer seems to be maintained, partly in view of the fact that some ergative languages such as Eskimo and Mayan languages have what appear to be passive constructions; e.g.

- (14) Greenlandic Eskimo (Woodbury 1977: 323 f.)  
 (a) *Ajut-ip arnaq*  
 man-ERG woman.ABS  
*taku-vaa.*  
 see-IND.3.SG&3.SG  
 'The man saw the woman.'  
 (b) *Arnaq (ajuti-mit)*  
 woman.ABS man-ABL.SG  
*taku-tau-puq.*  
 see-PASS-IND.SG  
 'The woman was seen (by the man).'

But since a language may possess more than one type of passive forms (see 10), simple possession of a passive construction itself does not preclude the possibility of an ergative construction being a passive construction. More telling evidence for considering an ergative construction as active comes from the facts that: (i) the verb of an ergative clause is a basic, rather than marked form (cf. (14 a) vis-à-vis (14 b)), (ii) the ergative agentive nominal is an integral constituent rather than being an optional adjunct like the passive agent, and (iii) the ergative construction is a textually unmarked form of transitive expression in that its text frequency is as high as that of an active clause in accusative-type languages, whereas the marked passive construction typically has a very low text frequency (Comrie 1988).

Languages with ergative constructions generally lack the passive construction of the type shown above. In its place they may possess a construction-type known as **antipas-sive**, illustrated by the following Greenlandic Eskimo example (Woodbury 1977: 323; see also 10.2):

- (15) (a) *piniartu-p qijmiq*  
 hunter-ERG dog.ABS  
*unatar-aa*  
 beat-TP.3.SG&3.SG  
 '(The doctor saw that) the hunter  
 beat the dog.'

- (b) *piniartuq qijmi-mik*  
 hunter.ABS dog-INSTR  
*unata-i-šuq*  
 beat-APASS-IP.3.SG  
 '(The doctor saw that) the hunter  
 beat a dog.'

The term ‘antipassive’, based on ‘passive’, was invented by Silverstein (1976) as a way of pointing up the ‘mirror image’ effects the two conversions bring about. (When you wave the right hand, the mirror image waves the left hand.) The passive is an agent-defocusing voice as opposed to the patient-defocusing effect brought about by the antipassive; their syntactic consequences being demotion of an agent nominal into an oblique position or its total suppression in the former, and demotion of a patient nominal into an oblique position or its total suppression in the latter. The passive typically promotes a patient nominal into the primary term position of subject, and the antipassive converts an agent nominal to the primary absolute term.

We noted in 3 that a construction that syntactically qualifies as a passive may have no morphological marking in the verb, casting some doubt on a narrow morphology-bound definition of voice. Antipassives, too, raise this issue, for a fair number of accusative-type languages allow deletion of indefinite objects without any morphological marking in the verb, as witnessed in the following English and Japanese expressions.

- (16) (a) *John drinks.*  
 (b) *Dogs bite.*
- (17) (a) *Kono ko wa yoku tabe-ru.*  
 this child TOP lot eat-PRES  
 ‘This child eats a lot.’
- (b) *Uma wa ker-u.*  
 horse TOP kick-PRES  
 ‘Horses kick.’

These constructions (also known as ‘neuter’ or ‘absolute’ constructions) do share the patient-defocusing function of antipassives, but they do not partake of another important function of true antipassives. That is, antipassivization promotes a secondary ergative term into a primary absolute term, whereby the derived primary term is made susceptible to absolute-based phenomena such as relativization and cross-sentential co-referential deletion. Notice that passives in ergative-type languages achieve a similar half-effect of a

true passive as they do not affect the primarity of the absolutely coded patient argument (cf. (5 a), (14 a) vis-à-vis (5 b), (14 b)). Again, the passive in an ergative language and the antipassive in an accusative language are mirror images with respect to the terms failing to undergo the promotional effect of these voice conversions.

## 8. Voice in Philippine languages

Philippine languages constitute another group of languages in which a definition of voice like the one set out in 2 encounters problems. First, Philippine languages are often given as examples of those languages that permit passives of various constituent types. While many languages, e.g. the Korean *ci*-passive, the German *werden*-passive, restrict passive conversion to a patient role or something assimilated to it, Philippine languages are said to allow not only patient but all kinds of roles to be converted to a passive subject (Keenan 1985). Observe the following Cebuano sentences (Shibatani 1988: 88 f.):

- (18) (a) *Ni-hatag si Juan sa libro*  
 AF-give TOP (ACR) GOAL book  
*sa bata.*  
 RECIP child  
 ‘Juan gave the book to the child.’
- (b) *Gi-hatag ni Juan ang libro sa bata.*  
 GF-give ACR Juan TOP (GOAL) RECIP child  
 ‘Juan gave the book to the child.’
- (c) *Gi-hatag-an ang bata ni Juan sa libro.*  
 D-give-F TOP (RECIP) ACR Juan GOAL book  
 ‘Juan gave the book to the child.’
- (d) *Gi-palit-an ni Juan si Maria ug libro.*  
 D-buy-F ACR Juan TOP(BEN) Maria GOAL book  
 ‘Juan bought a book for Maria.’
- (e) *Gi-adtu ni Juan ang estasionan.*  
 DF-go ACR Juan TOP station  
 ‘Juan went to the station.’
- (f) *Gi-pritu-han ni Juan ug isda ang kaha.*  
 D-fry-F ACR Juan GOAL fish TOP pan  
 ‘Juan fried a fish in the pan.’
- (g) *Gi-duan ni Juan ang bata.*  
 DF-play ACR Juan TOP child  
 ‘Juan played with the child.’

- (h) *I-hiwa ang kutsilyo sa mangga*  
 IF-cut TOP knife GOAL mango  
*ni Maria.*  
 ACR Maria  
 'Maria will cut the mango with the  
 knife.'

The Philippine data present two major problems in relation to the definition of voice as given in 2. First, there is a problem of the grammatical function of subject. As the above data indicate, there is an instance, (18 a), in which the agentive nominal converges with what in Philippine linguistics is called the topic. In this so-called actor-topic construction all the known subject properties concentrate on this constituent. The actor-topic construction can thus be considered as representing the basic, active diathesis. However, a problem arises in other non-actor topic constructions, (18 b–h), especially in the so-called goal-topic construction (18 b), in which the patient nominal is made topic.

If Philippine 'topicalization' is the same as the subjectivization accompanying the passive conversion in English and other languages, we except a diathetic change such that the derived topic nominal acquires a great majority of the subject properties. But this does not happen in Philippine languages. The non-topic agentive nominal, e.g. Actor *ni Juan* in (18 b), retains a considerable degree of subjecthood. First of all, unlike the oblique agentive phrase of a prototypical passive, the actor-nominal is an integral constituent of a goal-topic sentence, and as such it is not normally suppressed. In other words, a goal-topic sentence is both semantically and syntactically transitive. Secondly, non-topic actors actively participate in syntactic phenomena exhibiting subject properties such as reflexive binding, controlling of an actor gap in the second coordinated clause, deletion in imperatives, deletion in the second coordinated clause, controlling of a gap in a subordinated clause, and topicalization into sentence initial position. The topic nominal, on the other hand, acquires such subject properties as extractability in relativization, focus constructions, quantifier float, controlling of a gap in the *samtang* 'while'-clause, raising out of the *nga* 'that' subordinate clause as well as some of those that are controlled by the non-topic actor (see Shiba-tani 1991). Thus, it is not at all clear whether a diathetic alternation is as discrete in the case of the Philippine voice system as in the

more familiar active-passive alternation. In the Philippine system, only a partial transfer of the total subject properties is achieved in non-actor topic constructions.

The Philippine topic is perhaps another case that calls for a neutral term for a grammatically salient argument. While non-topic agent retains a great many properties that a subject in other languages may possess, there is no question that it is the topic nominal that is most indispensable and hence primary, as it must occur in major clause types. The notion of a primary term is useful here too; it allows us to subsume the subject, the absolute, and the Philippine topic under the rubric of primary term without committing us to the total equivalence among these terms.

The other difficulty the Philippine voice system poses has to do with the determination of the basic diathesis. In the above discussion we assumed that the actor-topic construction represents the basic, active diathesis, as the agent aligns with the topic and all the subject properties accrue on the actor-topic nominal. However, this assumption can be challenged. The opposition between the basic unmarked voice and the marked voice is normally correlated with morphological marking in the verb form and with the text frequency, such that active voice (as well as ergative voice) sentences involve a basic verb form, whereas the passive and the antipassive sentences employ marked verb forms, indicating their marked status. Indeed, marked voice sentences are marginal in that their text frequency is quite low compared to the basic voice sentences; the ratio being somewhere around 20 : 80.

Now, in the case of Philippine languages it is typically the case that the verb forms in both actor-topic and non-actor topic constructions involve marking, as shown in the Cebuano examples in (18), and therefore the morphological markedness does not point out the basic diathesis. Furthermore, text frequency is not a strong indicator of the basic sentence type either, as the text frequencies of the actor-topic and the goal-topic transitive constructions are highly comparable (see Shiba-tani 1988; Payne 1994). Thus, in the case of Philippine languages, the notion of the basic diathesis and the conception of voice as a mechanism for modifying the basic diathesis do not apply as straightforwardly as in some other languages. One may indeed question whether Philippine topicalization is to be

understood as a voice system. This of course depends on one's definition of voice, which, as we have seen, is difficult to come by.

### 9. Two dimensions of voice

We have seen that cross-linguistic comparison of voice systems reveals some significant differences among them such that imposing a view based on one system onto another may result in a distorted characterization of a given voice system. Such has been the case in applying the term passive to the goal-topic construction of Philippine languages. The goal-topic construction is very different from the prototypical passive construction, as briefly surveyed (see Shibatani 1988a for further detail). On the other hand, there is some significant similarity that runs through all the systems surveyed above. In order to capture that essence, we need to countenance a broader view of voice.

There are two dimensions in a voice system. One has to do with the grammatical status of event participants, and hence with the notion of transitivity. Here a voice system regulates what is chosen as a primary grammatical term as well as what status other non-primary participants should have. The second task of a voice system is to represent semantic contrasts between sentences of contrasting voice categories or more narrowly to indicate the semantic relation that a primary term holds with respect to the predicate.

Let us consider the first role of voice systems. In the active-passive and the ergative-antipassive systems, the basic transitive clauses receive two formal expressions depending on whether they are coded according to the nominative-accusative format or the ergative-absolutive format. In the nominative-accusative construction the primary, nominative constituent encodes the agent role. Under this system the passive voice indicates that the primary term does not encode the expected agent role; instead it encodes the patient, affected entity, if there is such a constituent. (Many languages allow impersonal passives, which do not have a primary term (see 10.3.1).)

On the other hand, the unmarked voice couched in the ergative-absolutive pattern indicates that the primary, absolute nominal encodes the patient role. The marked, anti-passive voice indicates that the primary term does not encode the expected patient role; instead it encodes the agent role.

In addition, the passive in the absolutive-ergative system has the effect of signaling the lowered status of the secondary, ergative term, as the antipassive in the nominative-accusative system has the effect of signaling the lowered status of the secondary, object term.

The notion of the lowering of the status of an event participant is to be viewed as applicable at both semantic and syntactic levels, so that the passive-anticausative continuum can be accounted for by a single principle. That is, the passive, for example, involves a syntactic downgrading of an agentive participant. The anticausative, on the other hand, goes one step further and obliterates the agentive participant at the semantic level, turning the clause into intransitive both semantically and syntactically. Historically passives may develop into anticausatives or vice versa, and thus it is desirable to maintain a neutral treatment of the grammatical status of an agentive participant – the development of a passive from an anticausative involves the introduction of an agentive participant at the semantic level and possibly (at a later stage of development) at the syntactic level as well.

The second perspective that voice oppositions reflect semantic oppositions requires some radical reconsideration of some familiar voice systems. First, the active-middle system represents the binary system of *energeia* and *pathos*, the active voice indicating that a primary constituent stands in an energetic relation to the predicate and the middle voice indicating that the same stands in a pathetic relation to the predicate. This contrast is not directly correlated with diathetic alternations, as both intransitive and transitive verbs enter into the opposition and as direct objects may be left intact.

#### (19) Classical Greek

- (a) *politeú-ō*  
‘I am a citizen.’ (active)
- (a') *politeú-omai*  
‘I act as a citizen.’ (middle)
- (b) *loú-ō khitōna*  
‘I am washing a shirt.’ (active)
- (b') *loú-omai khitōna*  
‘I am washing a shirt for myself.’  
(middle)

#### (20) Sanskrit

- (a) *yáj-a-ti*  
‘performs a sacrifice for someone else’ (active)

- (a') *yáj-a-te*  
‘performs a sacrifice for oneself’  
(middle)
- (b) *devadattáh = yajnadattásya kródham ví-nay-a-ti* (active)  
‘D. pacifies the anger of Y.’
- (b') *kródham ví-nay-a-te* ‘subdues one’s own anger’ (middle)

We started our discussion of voice with Japanese examples that appeared to illustrate the prototypes of the voice forms based on the active-passive opposition along the generally held understanding of the concept of voice. But we also noticed in 5 that the passive voice in Japanese includes intransitive based forms that increase valence, contrary to the general tendency where marked voices are associated with valency decrease. As it turns out, the Japanese voice system is much more similar to the active-middle system than the active-passive system of the English type; and as such the kinds of opposition we posited for the Japanese examples in 2 and 5 are not in fact appropriate. Instead, the patterns of opposition for Japanese must be as illustrated below, which parallels the Greek and the Sanskrit patterns in (19) and (20) – especially take notice of the proposed opposition involving simple transitive forms as seen between (21 b) and (21 b'):

- (21) Japanese
  - (a) *Taroo ga nai-ta.*  
Taro NOM cry-PAST  
(active)  
‘Taro cried.’
  - (a') *Taroo ga Hanako ni naka-re-ta.*  
Taro NOM Hanako DAT cry-PASS-PAST  
(passive)  
‘Taro was adversely affected by Hanako’s crying.’
  - (b) *Taroo ga Hanako o korosi-ta.*  
Taro NOM Hanako ACC kill-PAST  
(active)  
‘Taro killed Hanako.’
  - (b') *Taroo ga Hanako ni korosa-re-ta.*  
Taro NOM Hanako DAT kill-PASS-PAST  
(passive)  
‘Taro was killed by Hanako.’
  - (c) *Taroo ga piano o hii-ta.*  
Taro NOM piano ACC play-PAST  
(active)  
‘Taro played the piano.’

- (c') *Taroo ga Hanako ni piano o hika-re-ta.*  
Taro NOM Hanako DAT piano ACC play-PASS-PAST  
(passive)  
‘Taro was adversely affected by Hanako’s playing the piano.’

The voice opposition in Japanese, in other words, reflects the meaning opposition of whether the referent of the subject nominal induces action under his own control (active) or it is affected by an externally induced action (passive). Indeed, the close historical – and morphological – affinity of the Japanese passive and its spontaneous construction points to the likeliness, as believed by many Japanese grammarians, that the early Japanese voice system was much similar to the active-middle system of the classical Indo-European languages.

The Philippine voice system too seems to require a semantically based characterization as not only a direct object but also various other constituents enter into the opposition. In this system thematic relations are divided into several macro relations – four in the case of Cebuano, as reflected in the four types of focus marking in the verb, the actor-focus (AF), the goal-focus (GF), the directional-focus (DF), and the instrumental focus (IF) (see (18) and Shibatani 1988: 87). Each focus marking in the verb indicates that a primary constituent, the topic nominal, holds one of these macro relations.

Viewed in the perspective laid out above, the English-type active-passive opposition that revolves just around the direct object is in fact a highly specialized, narrow system that disguises the semantically-based nature of voice systems. However, as seen below, even such a highly constrained system is not an isolated grammatical category that can be clearly set apart from other categories or constructions.

We have above offered a broader, semantically-based understanding of voice, which, contrary to the prevailing modern views, lays less stress on the notions of basic diathesis, diathetic alternations, and the preservation of propositional meanings across voice distinctions, all of which can be challenged when a large body of voice data is examined. Particularly compelling evidence for such a broad view is found in the domain of voice morphology, especially its historical sources, to which we now turn.

## 10. Voice morphology

Earlier we mentioned the view that considers voice categories as comprising of families of constructions. There are at least three senses in which this simile is used. A first sense pertains to a cross-linguistic perspective. As we saw in 4, a voice category in one language may subsume a family of constructions whose members are distinct constructions in another language. The middle category of Classical Greek is a case in point, as it includes various senses or functions that are expressed by distinct constructions such as the spontaneous, the reflexive, the reciprocal, and the passive construction in English and other languages.

A second sense in which a voice category is said to consist of a family of constructions has to do with the fact that often a single language has more than one construction instantiating a particular voice category. Thirdly, a voice construction may show a family resemblance to other constructions. These two points are illustrated by the Russian passive constructions below (Babby & Brecht 1975):

- (22) (a) *Kalitka otkryvalas' Olegom.*  
          gate open:REFL Oleg:INSTR  
          'The gate was being opened by Oleg.'  
       (b) *Kalitka byla otkryta Olegom.*  
          gate be open Oleg:INSTR  
          'The gate was opened by Oleg.'

Sentence (22 a), with the verbal suffix -sja (realized as -s' or -sja) instantiates the imperfective passive construction, whereas (22 b) represents the periphrastic perfective version that, like the English *be-passive*, involves *byt* 'to be' and the participle forming suffix (realized as -n-, -en-, or as -t-). These two constructions, which inhabit the passive voice category of Russian, are not isolated constructions in that there are other constructions, which do not qualify as passive, that show resemblances to them. For example, (22 a) has affinity, among others, to the following constructions in (23), most of which can be subsumed under the middle voice category (Channon 1974; Babby & Brecht 1975; Geniušienė 1987) and (22 b) is related to the sentences in (24):

- (23) (a) *Kalitka otkryvaetsja.*  
          gate open:REFL  
          'The gate is opening.'

- (b) *Ivan moetsja mylom.*  
     Ivan wash:REFL soap:INSTR  
     'Ivan washed himself with soap.'  
      (c) *My s nim billis'*  
     we with him hit:PL.REFL  
     'He and I hit each other.'  
      (d) *Jemu slyšitsja muzyk-a.*  
     he.DAT hears:REFL music-NOM  
     'He can hear music.'  
      (e) *Mne mečtajetsia.*  
     I.DAT dreams:REFL  
     'I am dreaming.'  
      (f) *Babuška rugajetsja.*  
     granny:NOM scolds:REFL  
     'Granny is scolding.'

- (24) (a) *Ego ruki byli smorščeny.*  
         his hands be wrinkled  
         'His hands were wrinkled.'  
       (b) *Knigi byli postavleny na stol.*  
         books be placed on table  
         'The books were placed on the table.'

What points us to these cross-constructional similarities is the morphology, like the Russian reflexive -sja suffix in (23 a), employed in voice constructions.

### 10.1. Reflexives, middles, and passives

Though both Sanskrit, in the form of suffix -ya, and Classical Greek, in the form of suffix -(th)ē in the aorist and future tenses, had passive constructions formally distinct from the middle constructions, they generally subsumed passives in the middle category; i. e. the passive reading was one of those interpretations permitted by the middle construction. In fact, the Russian passive form in (22 a) and its affinity to the middle constructions in (23) shows that a passive construction can arise from the constructions forming a middle category. The Russian examples in 10 also indicate that the constructions forming a middle category can sprout up around the reflexive construction. Indeed this is a very general pattern of development of the middle and the passive categories in European languages, especially among Slavic and Romance languages. But the inclusion of the **reflexive** in the middle category is by no means confined to the Indo-European family. For example, the Cochabamba Quechua reflexive suffix -ku also makes up a middle category with a meaning range comparable to the middle categories in other languages, as witnessed below (Bills et al. 1969):

- (25) (a) *Maqa-ku-nqa.*  
hit-REFL-3.SG.FUT  
'He will hit himself.'
- (b) *Alqu atuq-wan riku-na-ku-nku.*  
dog fox-with look-REC-REFL-3.PL  
'The dog and the fox looked at each other.'
- (c) *Runtu-kuna p'aki-ku-san.*  
egg-PL break-REFL-PROG  
'Eggs are breaking/Eggs are being broken.'
- (d) *Ropas t'aqsa-ku-n mayu-pi.*  
clothes wash-REFL-3.SG river-in  
'Clothes are washed in the river.'

The rise of the middle and the passive morphology from the reflexive source seems to be motivated by at least the following two characteristics of the reflexive construction. First, the basic direct reflexive construction (e.g. *John hit himself*) is notionally intransitive in the sense that the action performed by the referent of the subject nominal is not transferred to another entity as in regular transitive constructions with two distinct arguments. In this regard, the reflexive is similar to intransitive constructions such as *John walked* and *John died* in which the actions and processes expressed are confined to the sphere of the subject entity. Secondly, unlike regular transitive expressions, the subject of the reflexive construction is affected (by its own action). These two properties of the reflexive construction, namely the notional intransitivity and the subject affectedness, drive the grammaticalization process that turns a transitive reflexive clause into a syntactically intransitive one with no object nominal via cliticization and eventual affixation of the erstwhile reflexive noun (see Haspelmath 1990 and Kemmer 1993 on the grammaticalization of reflexive nouns and pronouns).

The meaning of the reflexive construction that the affectedness of the subject referent is caused by no external agency other than by itself gives rise to the spontaneous reading that the process affecting the subject entity occurs of its own accord. The extension of the construction to an inanimate subject yields a true spontaneous construction such as (23a) and (25c), where no agency may be entertained. This spontaneous pattern provides the basis for the passive interpretation of the middle, yielding a situation where such reading is predominant as in (25d), and which eventually extends to unambiguous

passives in which external agents are consistently conceptualized and may be overtly expressed, as in (22a).

### 10.2. Reflexives, middles, and antipassives

Curiously enough, reflexives not only give rise to passives but also to what we earlier considered to be in the mirror image relationship to them, namely antipassives. A fair number of Australian languages show this correlation illustrated by the use of the suffix *-:di-n* in the following Yidiñ examples in (26b) and (27b) (Dixon 1977: 277, 280):

- (26) (a) *wagudangu yanay bambi:l*  
man.ERG I.ACC covered  
'The man covered me.'
- (b) *wagu:da bambi-dinu*  
man.ABS covered-REFL  
'The man covered himself.'
- (27) (a) *wagudangu guda:ga wawa:l*  
man.ERG dog.ABS saw  
'The man saw the dog.'
- (b) *wagu:da wawa:dijnu gudaganda*  
man.ABS saw.APASS dog.DAT  
'The man saw the dog.'

Whether the antipassive arises from the reflexive via the middle is not entirely clear as the status of the middle category itself is not clear in these Australian languages, although available descriptions of the reflexive or the antipassive such as Dixon (1977), Haviland (1979), Austin (1981), and Tsunoda (1985) point out a number of additional uses of the relevant affixes, which may include clear spontaneous constructions such as the following:

- (28) (a) *yuri watyu-kali-n*  
kangaroo.ABS cook-REFL-PRES/PAST  
'The kangaroo is cooking/cooked.'  
(Warrungu; Tsunoda 1985: 638)
- (b) *tana kupa tinta-tadi-yi*  
3.PL.SBJ child.ABS lose-REFL-PRES  
(*wila-pi*)  
woman-LOC  
'The children got lost (on the woman).'  
(Diyari; Austin 1981: 156)

The development of the passive and the anti-passive from the reflexive are due to the Janus-faced nature of the reflexive. The subject of a prototypical reflexive construction is both agent and patient at the same time. It can thus give rise to both patient-oriented constructions such as the spontaneous (or the

anticausative) and the passive, on the one hand, and agent-oriented constructions such as the antipassive, on the other. This contrast in the perspective between the passive and the antipassive correlates with the opposition in the aspectual meanings associated with these two constructions (Tsunoda 1985: 636):

- (29) passives: perfect/perfective/resultative/stative  
 antipassives: imperfective/durative/continuative/repetitive/habitual

If the reflexive can give rise to both passive and antipassive in the manner described above, it stands to reason that the reflexive should be able to give rise to these constructions within single languages. Indeed, some languages bear out this prediction. For example, the Diyari reflexive suffix *-tadi* occurs in the reflexive (30 b), in the antipassive (31 b), and in the passive (32 b) (as well as the spontaneous (28 b) – but not the reciprocal) (Austin 1981: 152–155):

- (30) (a) *ŋat̪u yinapa muduwa-yi*  
          1.SG.AGT 2.SG.OBJ scratch-PRES  
          'I scratch you.'  
 (b) *ŋapi muduwa-tadi-yi*  
          1.SG.SBJ scratch-REFL-PRES  
          'I scratch myself.'
- (31) (a) *ŋat̪u napa wila*  
          1.SG.AGT 3.SG.F.OBJ woman.ABS  
          *kalka-yi*  
          wait.for-PRES  
          'I wait for the woman.'  
 (b) *ŋani kalka-tadi-yi*  
          1.SG.SBJ wait.for-APASS-PRES  
          *ŋanjaku wila-ŋi*  
          3.SG.F.LOC woman-LOC  
          'I wait for the woman.'
- (32) (a) *wat̪ara-li ŋipa mana*  
          wind-ERG 3.SG.NF.OBJ door.ABS  
          *ŋndawalka-ŋa wara-yi*  
          close-PART AUX-PRES  
          'The wind closed the door.'  
 (b) *ŋawu mana*  
          3.SG.NF.SBJ door.ABS  
          *ŋandawalka-tadi-ŋa wara-yi*  
          close-PASS-PART AUX-PRES  
          *waṭara-ŋi*  
          wind-LOC  
          'The door got closed by the wind.'

Russian also shows the double use of the reflexive in the passive and the antipassive, as

revealed by the comparison of (22 a) and the following antipassive forms – (33 b) and (34 b) – (Geniušienė 1987: 9, 56):

- (33) (a) *Babušk-a rugajet devočk-u.*  
          granny-NOM scolds girl-ACC  
          'Granny is scolding the girl.'  
 (b) *Babušk-a rugajet-sja.*  
          granny-NOM scolds-REFL  
          'Granny is scolding.'
- (34) (a) *On rojet zeml-iu.*  
          he.NOM digs ground-ACC  
          'He is digging the ground.'  
 (b) *On rojet-sja v zeml-e.*  
          he.NOM digs-REFL in ground-LOC  
          'He is digging in the ground.'

### 10.3. Other sources of voice morphology

The morphological affinity of the reflexive and the middle (or its subconstructions; e.g. the spontaneous) is widely attested, and the reflexive certainly seems to be a favorite historical source for the middle voice forms. But this affinity is not absolute either semantically or historically. Whereas the Classical Greek middle allows a direct reflexive reading, the Sanskrit counterpart does not, where the direct reflexive is expressed by means of the accusative *ātmānam* 'self' (Burrow<sup>3</sup> 1973: 194). Neither is there evidence that these classical middle systems were historically connected with reflexives, as they are said to have developed from a contamination of the active and perfect sets of endings of Proto-Indo-European (see Andersen 1991: 49). Indeed, clear cases of middle and/or passive morphology due to non-reflexive sources are attested.

#### 10.3.1. The indefinite person constructions

In Ainu the first person inclusive pronominal affixes (*-an* in intransitives and *a-* in transitives) developed into the middle marking as well as the passive marking.

- (35) basic person marking pattern

- (a) *Itak-an.*  
          speak-1.PI  
          'We speak.'  
 (b) *Tampe a-e-kore.*  
          this 1.PI-2.SG-give  
          'We give you this.'

- (36) indefinite-person construction

*Tepeka paye-an yak Sat ta paye-an.*  
          here go-1.PI if Saru to go-1.PI  
          'If one goes here, one goes to Saru.'

## (37) middle

- (a) *Chip a-nukar.*  
ship 1.PI-see  
'A ship is seen/visible.'
- (b) *Pirka hawe a-nu.*  
beautiful voice 1.PI-hear  
'A beautiful voice is heard/audible.'

## (38) passive

- (a) *Kamui umma raike.*  
bear horse kill  
'A bear killed a horse.'
- (b) *Umma kamui orowa a-raike.*  
horse bear from 1.PI-kill  
'A horse was killed by a bear.'

(36) is analogous to impersonal constructions in other languages, which subsume both active constructions with indefinite pronominal subjects, as in (39), which may be rendered as passives in translations into other languages, and a variety of impersonal passives (also known as 'pseudo-passives'), in which passive morphology may be employed but no regular personal subjects occur, as in (40).

## (39) (a) French

- On vendit la maison.*  
one sold the house  
'The house was sold.'

## (b) German

- Man hat mir gesagt.*  
one have me told  
'I have been told.'

## (40) (a) Danish

- I dag danse-s der ikke.*  
today dance-PASS there no  
'Today there is no dancing.'

## (b) German

- Es wurde gestern*  
it became yesterday  
*getanzt.*  
dance.PASS.PART  
'There was dancing yesterday.'

## (c) Finnish

- Syöttiin sen.*  
east.PASS.PART it.ACC/GEN  
'It was eaten.'  
(Siewierska 1984)

(d) *Polja pobilo gradom.*  
fields.ACC crushed.3.SG.N hail.INSTR  
'The crops were destroyed by hail.'  
(Mel'čuk 1979)

The Ainu indefinite person construction exemplified in (36) seems to be directly con-

nected to the spontaneous use in (37). However, other than verbs such as *nukar* 'to see' and *nu* 'hear,' there do not seem to be other detransitivized verbs with the use of the affixes in question. The passive form in (38) is a further development from the middle use and in this construction, the personal affix *a-* is best considered as having been reanalyzed as a passive marker. In addition, the indefinite person construction gave rise to a honorific construction, in which the replacement of a regular personal affix with *a-* signals the speaker's deference toward the subject referent. Similar developments of passives from indefinite person constructions are observed in Kimbundu (Bantu) and Trukic (Micronesian) (see Shibatani 1985).

## 10.3.2. The causative source

Another major non-reflexive source for voice morphology is causative formatives. The causative-passive correlation, mediated specifically by the permissive causative reading, is particularly conspicuous among Tungusic languages but it is also seen elsewhere: Korean, Sonrai, etc. (see Nedyalkov 1991). This can be illustrated by the Evenki suffix *-v* in (41 a) and (41 b), but the causative-to-passive development here, as in Korean, has reached the stage where the passive morphology is used as an anticausative (or middle) marker, as seen in the forms in (41 c) (Nedyalkov 1991):

(41) (a) *beye mo-l-va*

- man wood-PL-ACC  
*l-v-re-n*  
enter-CAUS-PAST-3.SG  
'The man brought fire-wood into the house.'

(b) *Oron (beyumimni-du)*

- deer (hunter-DAT)  
*va-v-cha-n*  
kill-PASS-PAST-3.SG  
'The deer was killed (by the hunter).'

(c) *mana-* 'finish (something)':

- mana-v-* '(something) finishes'  
*ula-* 'soak (something)':  
*ula-v-* '(something) soaks'  
*n'i-* 'open (something)':  
*n'i-v-* '(something) opens'  
*sukcha-* 'break (something)':  
*sukcha-v-* '(something) breaks'

## 10.3.3. The auxiliary-verb source

Whereas the voice morphologies surveyed above vary in their origins and in the degree of etymological transparency, they fall square-

ly in the narrow definition of voice as morphological categories involving marking in the verb. Many languages, however, have (in addition) the so-called periphrastic passives that go beyond this narrow definition in that they typically involve independent auxiliary-like verbs in addition to certain morphological alterations in main verbs, as in the case of the English *be*-passive, the German *werden*-passive, and the Russian *byt'*-passive seen above. Whether voice morphology is realized in the affixal form or occurs as free forms largely depends on the overall morphological typology of the language and on the degree of grammaticalization that the source form has undergone. Therefore, the distinction of periphrastic versus morphological manifestations of the voice morpheme is not as interesting as the etymological origins and the semantic scope that typical periphrastic voice constructions share with certain affixal forms. For example, the Japanese passive morpheme *-(ra)re* is clearly suffixal, as it cannot stand alone. The Korean passive form *-ci*, too, is suffixal, but it still shows high transparency to the independent verb *ci-ta*. What interests us about these and some other affixal voice morphology is the fact that their etymology and their semantic scope are highly similar to those of independently occurring auxiliary voice formatives of the so-called periphrastic passives. This similarity is implicit in the slogan "today's morphology is yesterday's syntax" of the grammaticalization theory (Givón 1971; Heine et al. 1991).

Though not definitively known, the Japanese *-(ra)re* is said to come from the verb *aru* 'to be', 'to have' or from the combination of *uru* 'to get' and *aru* 'to be'. The Korean verb *ci-ta* means 'to become'. Indeed, the verbs meaning 'to be,' 'to become,' 'to get,' 'to go,' 'to suffer,' etc. are favorite sources from which passive auxiliaries arise (see Siewierska 1984; Keenan 1985). As for the semantic scope, the use of the Japanese *-(ra)re* suffix includes, in addition to passive, the spontaneous, the potential, and honorific constructions (see Shibatani 1985).

Many of the so-called periphrastic passives are not isolated constructions either. In the case of English *be*-passive and the Russian *byt'*-passive, they are close kin of the resultative and adjectival stative constructions of the following type (see (22)–(24) for Russian data):

- (42) (a) *The fish was all eaten (when I reached the table).*

- (b) *He was tired.*  
 (c) *The door was open.*

Particularly noteworthy is the fact that many periphrastic passives also have a strong affinity to, or are not distinguishable from, the spontaneous constructions. Thus, there is a great deal of controversy over whether the Persian construction in (43 c), which makes use of the auxiliary verb *šodan* 'to become', is to be understood as a passive or still as a spontaneous like (43 b) (see Dabir-Moghadam 1982).

- (43) (a) *āb sard ast.*  
 water cool is  
 'The water is cool.'  
 (b) *āb sard šod*  
 water cool become  
 'The water cooled/became cool.'  
 (c) *āb (tavassot-E mahmud) sard*  
 water (through Mahmud) cool  
*šod*  
 become  
 'The water was cooled (by Mahmud).'

The Indo-Aryan language Maithili, like Hindi, makes use of the verb *jæb* 'to go' in its passive formation. But again the passive is not an isolated construction as seen below, where the perfect participle form, *gel*, of the verb *jæb* is used in the spontaneous constructions as well (Williams 1973).

- (44) (a) *həm sinema gellu:*  
 I cinema went  
 'I went to the cinema.'  
 (b) *Jibach blirh gel*  
 Jibach grown went  
 'Jibach grew.'  
 (c) *am pɔ'k gel*  
 mango ripe went  
 'The mango ripened.'  
 (d) *Jibach mlir gel*  
 Jibach died went  
 'Jibach died.'  
 (e) *gach suikh gel*  
 tree dried went  
 'The tree withered.'  
 (Cf. *həm gach sukh-lu-lu*: 'I dried the tree.')  
 (f) *Jibach marəl gel*  
 Jibach murdered went  
 'Jibach was murdered.'  
 (Cf. *Jibach hunka marləi* 'Jibach murdered him.'')

In fact, we don't have to search far in finding the spontaneous-passive continuum, as it is also observed in the English *get*-passives (see Arce-Arenales et al. 1994; Givón & Yang 1994):

- (45) (a) *John got lost/bogged down.*  
 (b) *John got killed in the accident.*  
 (c) *You got influenced by some people.*

#### 10.3.4. Other sources of the antipassive morphology

Finally, the available data on the affinity or the etymological relationship of the antipassive morpheme to other forms than reflexives (see 10.2) are meager. But two comparatively clear cases point to first person object markers as a possible non-reflexive source for antipassive markers. First, among the three sets of antipassive affixes in the ergative language Chukchee, the prefixes *ena-/lina-* are related to the first person singular object marking (Kozinsky et al. 1985: 652, 658):

- (46) (a) *Gənan gəm*  
 thou.ERG I.ABS  
*ena-pela-gʔe.*  
 1.SG.OBJ-leave-2.SG.SBJ  
 'Thou left me.'  
 (b) *ətləg-e qora-ŋə*  
 father-ERG deer-ABS  
*qərir-nim.*  
 look.for-3.SG&3.SG.AOR  
 'The father looked for the deer for some time.'  
 (c) *ətləg-ən ena-rer-gʔe.*  
 father-ABS APASS-look.for-3.SG.AOR  
 'The father did some searching.'

The other less frequently attested antipassive suffix sets *-tukul-iko* and *-etl-at* indicate in addition the iterative meaning and the reflexive meaning respectively. Both also derive denominal verbs (see Kozinsky et al. 1985) like passive forming affixes in some languages.

In a manner parallel to Chukchee, Ainu makes use of the first person object prefix *i-* in its detransitivizing process that can be construed as a case of antipassivization (Shibatani 1990):

- (47) (a) *Akor achapo, i-oira na!*  
 my uncle 1.OBJ-forget PARTICLE  
 'My uncle, don't forget me!'  
 (b) *Sake a-ku.*  
 sake 1.SG.SBJ-drink  
 'I drink sake.'

- (b') *I-ku-an.*  
 1.OBJ-drink-1.SG.SBJ  
 'I drink (typically sake).'  
 (c) *Ya a-oshke.*  
 net 1.SG.SBJ-knit  
 'I weave a net.'  
 (c') *I-yoshke-an.*  
 1.OBJ-knit-1.SG.SBJ  
 'I weave (typically a net).'

Notice that the prefix *i-* converts a transitive clause into an intransitive one, as evidenced by the change in the person markers from the transitive first person subject marker prefix *a-* to the intransitive subject marker suffix *-an*. (See Greenberg 1991 for a related discussion.)

The use of object person markers for the antipassive is the converse in its function to the use of subject person markers for the passive seen in 10.3.1. Both cases appear to first permit the use of person markers as generalized subject or object constructions, which then give rise to agent-defocusing passive constructions or patient-defocusing antipassive constructions depending on whether the person markers are subject markers or object markers.

Finally, the Eskimo antipassive marker *-gil-i* is said to have developed from the valence-increasing suffix that introduces an adversely affected experiencer into the clause. Observe the following Central Yupik forms (Miyaoka 1984):

- (48) (a) *Nayiq kit'-uq.*  
 seal.ABS.SG sink-3.SG  
 'The seal sank.'  
 (b) *Angute-m nayiq*  
 man-ERG.SG seal.ABS.SG  
*kic-i-a.*  
 sing-gi-3.SG&3.SG  
 'The seal sank on the man.'  
 (c) *Angun kic-i-uq nayir-mek.*  
 man.ABS.SG sink-gi-3.SG seal-ABL.SG  
 'A seal sank on the man.'
- (49) (a) *Arna-m neqa*  
 woman-ERG.SG fish.ABS.SG  
*iir-aa.*  
 hide-3.SG&3.SG  
 'A/the woman is hiding the fish.'  
 (b) *Arnaq neq-mek*  
 woman.ABS.SG fish-ABL.SG  
*iir-i-uq.*  
 hide-gi-3.SG  
 'The woman is hiding a fish.'

How the adversely affected experiencer (e.g. the ‘man’ in (48 c)) got reinterpreted as an agent of an antipassive clause (e.g. the ‘woman’ in (49 b)) is an interesting question to be pursued (see Miyaoka 1984 for an attempt). Perhaps not to be ignored is the syntactic parallelism between (49 a)–(49 b) and the other co-existing ergative-antipassive pairs of the following type, in which no antipassive morpheme is involved.

- (50) (a) *Arna-m neqa*  
woman-ERG.SG fish.ABS  
*ner-aa.*  
eat-3SG&3.SG  
'A/The woman is eating the fish.'  
(b) *Arnaq neq-mek ner'-uq.*  
woman.ABS fish-ABL.SG eat-3.SG  
'The woman is eating a fish.'

## 11. Concluding remarks

This article examined several major voice constructions together with their morphological characteristics. Whereas the currently popular definitions of voice delimit voice phenomena to those cases in which diathetic alternations do not affect propositional meanings of the relevant constructions, morphological facts indicate that voice constructions are not easily delimitable as such. Particularly noteworthy is their affinity with spontaneous constructions that may be derived via anticausativization or denominal/deadjectival processes. Indeed, both synchronic and diachronic facts indicate that voice phenomena straddle the inflectional/derivational division, calling for a framework both more comprehensive and more dynamic than those currently available.

## 12. Uncommon abbreviations

AF	actor focus
D(-)F	directional focus
GF	goal focus
IF	instrumental focus
IP	intransitive participle
NF	non-feminine
RECIP	recipient
TP	transitive participle

## 13. References

Andersen, Paul (1991), *A New Look at the Passive*. Frankfurt am Main: Peter Lang

Arce-Arenales, Manuel & Axelrod, Melissa & Fox, Barbara (1994), “Active Voice and Middle Diathesis: A Cross-Linguistic Perspective”. In: Fox & Hopper (1994, eds.), 1–22

Austin, Peter (1981), *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press

Babby, Leonard H. & Brecht, Richard (1975), “The Syntax of Voice in Russia”. *Language* 51, 342–367

Bills, Garland & Vallejo, Bernardo & Troike, Rudolph (1969), *An Introduction to Spoken Bolivian Quechua*. Austin: University of Texas Press

Burrow, Thomas (³1973), *The Sanskrit Language*. London: Faber & Faber

Channon, Rachel (1974), “Pseudo-Reflexive Verbs in Russian”. In: Brecht, Richard & Chvany, Catherine V. (eds.), *Slavic Transformational Syntax*. Ann Arbor: Dept. Of Slavic Languages and Literatures, University of Michigan, 66–77

Comrie, Bernard (1988), “Passive and Voice”. In: Shibatani (ed.), 9–23

Dabir-Moghaddam, Mohammad (1982), “Passive in Persian”. *Studies in Linguistic Sciences* 12, 63–90

Dahlstrom, Amy (1991), *Plains Cree Morphosyntax*. New York: Garland

Dixon, Robert M. W. (1977), *A Grammar of Yidiš*. Cambridge: Cambridge University Press

Fagan, Sarah (1992), *The Syntax and Semantics of Middle Constructions: A Study with Special Reference to German*. Cambridge: Cambridge University Press

Fox, Barbara & Hopper, Paul (1994, eds.), *Voice: Form and Function*. Amsterdam: John Benjamins

Geniušienė, Emma S. (1987), *The Typology of Reflexives*. Berlin: Mouton de Gruyter

Givón, Talmy (1971), “Historical Syntax and Synchronic: An Archeologist’s Field Trip”. *Papers from the upperuse seventeenth Regional Meeting*, Chicago Linguistic Society, 394–415

Givón, Talmy (1990), *Syntax: A Functional-Typological Introduction, Vol. II*. Amsterdam: John Benjamins

Givón, Talmy (1994, ed.), *Voice and Inversion*. Amsterdam: John Benjamins

Givón, Talmy & Yang, L. (1994), “The Rise of the English GET-Passive”. In: Fox & Hopper (1994, eds.), 119–150

Greenberg, Joseph H. (1991), “The Last Stage of Grammatical Elements: Contractive and Expan-

- sive Desemanticization". In: Traugott, Elisabeth & Heine, Bernd (eds.), *Approaches to Grammaticalisation*, Vol. I. Amsterdam: John Benjamins, 301–314
- Hashimoto, Mantaro J. (1988), "The Structure and Typology of the Chinese Passive Construction". In: Shibatani (ed.) 329–354
- Haspelmath, Martin (1990), "The Grammaticalization of Passive Morphology". *Studies in Language* 14, 25–71
- Haviland, John B. (1979), "Guugu Yimidhirr". In: Dixon, Robert M. W. & Blake, Barry (eds.), *Handbook of Australian Languages*, Vol. I. Amsterdam: John Benjamins, 27–180
- Heine, Bernd & Claudi, Ulrike & Hünnemeyer, Frederike (1991), *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press
- Jolley, Catherine (1982), "On the Plains Cree Passive: An Analysis of Syntactic and Lexical Rules". In: Joseph, Brian (ed.), *Grammatical Relations and Relational Grammar*. Working Papers in Linguistics No. 26. Department of Linguistics, The Ohio State University, 1–33
- Keenan, Edward (1985), "Passive in the World's Languages". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description: Clause Structure*. Cambridge: Cambridge University Press
- Kemmer, Suzanne (1993), *The Middle Voice*. Amsterdam: John Benjamins
- Kemp, Alan (1987), "The Tékhnē Grammatikē of Dyonymus Thrax: Translated into English". In: Taylor, Daniel (ed.), *The History of Linguistics in the Classical Period*. Amsterdam: John Benjamins, 169–189
- Khrakovs'kyj, Viktor (1979), "Diathesis". *Acta Linguistica Academiae Scientiarum Hungaricae* 29, 289–307
- Klaiman, Miriam (1991), *Grammatical Voice*. Cambridge: Cambridge University Press
- Kozinsky, Ivan S. & Nedjalkov, Vladimir & Polinskaja, Maria (1988), "Antipassive in Chukchee: Oblique Object, Object Incorporation, Zero Object". In: Shibatani (ed.), 651–706
- Malchukov, Andrej (1993), "The Syntax and Semantics of Adversative Constructions in Even". *Gengo Kenkyū* 103, 1–36
- Mel'čuk, Igor (1979), "Syntactic or Lexical Zero in Natural Language". *Berkeley Linguistics Society: Annual Meeting* 5, 224–260
- Mel'čuk, Igor (1994), "The Inflectional Category of Voice: Towards a More Rigorous Definition". In: Comrie, Bernard & Polinsky, Maria (eds.), *Causatives and Transitivity*. Amsterdam: John Benjamins, 1–46
- Miyaoka, Osahito (1984), "On the So-called Half-Transitive Verbs in Eskimo". *Études/Inuit/Studies, Vol VIII: Supplement Issue: The Central Yupik Eskimos*, 193–218
- Nedyalkov, Igor (1991), "Recessiv-Accessiv Polysemy of Verbal Suffixes". *Languages of the World* 1, 4–31
- Perlmutter, David & Postal, Paul M. (1983), "Towards a Universal Characterization of Passivization". In: Perlmutter, David (ed.), *Studies in Relational Grammar 1*. Chicago: University of Chicago Press
- Payne, Thomas (1994), "The Pragmatics of Voice in a Phillipine Language: Actor-Focus and Goal-Focus in Cebuano Narrative". In: Givón (ed.), 317–364
- Shaumyan, Sebastian (1985), "Ergativity and Universal Grammar". In: Plank, Frans (ed.), *Relational Typology*. Berlin: Mouton (Trends in Linguistics, Studies and Monographs 28), 311–338
- Shibatani, Masayoshi (1985), "Passives and Related Constructions: A Prototype Analysis". *Language* 61, 821–848
- Shibatani, Masayoshi (1988), "Voice in Philippine Language". In: Shibatani (ed.), 85–142
- Shibatani, Masayoshi (1988, ed.), *Passive and Voice*. Amsterdam: John Benjamins
- Shibatani, Masayoshi (1990), *The Languages of Japan*. Cambridge: Cambridge University Press
- Shibatani, Masayoshi (1991), "Grammaticalization of Topic into Subject". In: Traugott, Elisabeth & Heine, Bernd (eds.), *Approaches to Grammaticalisation*, Vol. II. Amsterdam: John Benjamins, 93–134
- Siewierska, Anna (1984), *The Passive: A Comparative Linguistic Analysis*. London: Croom Helm
- Silverstein, Michael (1976), "Hierarchy of Features and Ergativity". In: Dixon, Robert M. W. (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112–171
- Smyth, Herbert (1956), *Greek Grammar*. Cambridge: Harvard University Press
- Tsunoda, Tasaku (1988), "Antipassives in Warunru and Other Australian Languages". In: Shibatani (ed.), 595–650

Williams, Jennifer (1973), "Clause Patterns in Mai-thili". In: Trail, Ronald (ed.), *Patterns in Clause, Sentence and Discourse in Selected Languages of India and Nepal II, Clause*. Summer Institute of Linguistics: University of Oklahoma, 345–453

Woodbury, Anthony (1977), "Greenlandic Eskimo, Ergativity and Relational Grammar". In: Cole, Peter & Sadock, Jerrold (eds.), *Grammatical Rela-*

*tions*. New York: Academic Press (Syntax and Semantics 8), 307–336

Xolodovič, Alexandr (1974, ed.), *Tipologija Passivnyx Konstrukcij: Diatezy i Zagoli*. Leningrad: Nauka

Masayoshi Shibatani,  
Kobe (Japan)/Houston (USA)

## 109. Aspect and Aktionsart

1. Aspectuality
2. Aspect versus Aktionsart
3. Aktionsart
4. Aspect
5. Interaction of aspect and Aktionsart
6. References

### 1. Aspectuality

This article treats the domain of **aspectuality**, which is used as a cover-term for **aspect** and **Aktionsart**. The two categories are treated together since both of them concern temporal characteristics of states of affairs as expressed in language. They should, however, be distinguished from the category of tense, to be treated in article 110. Tense serves to situate a state of affairs with respect to a deictic reference point. For instance, the past tense in (1) and (2) situates both states of affairs before the point of speech.

(1) *He died.*

(2) *He was dead.*

Since tense concerns the external relationship between a state of affairs and the deictic center of the discourse, it is considered to be a deictic category. Aspectuality is not deictic but rather concerns the internal temporal structure of states of affairs (Comrie 1976). Both (1) and (2) contain a past tense, but aspectually, they are different: (1) describes a change of state, (2) describes a durative state of affairs that does not have clearly definable boundaries. (Interactions between tense and aspectuality are addressed in 4.3.) Thus, (1) and (2) differ with respect to the kind of action, the Aktionsart, they express; (1) presents an **event**, (2) presents a **state**. The term **state of affairs** is used here, as elsewhere in this volume, as a cover-term for states, events

and other Aktionsarten. Now compare (1) and (3).

(3) *He was dying.*

(1) and (3) do not differ in Aktionsart; both refer to the same kind of event, as expressed by the English verb *to die*. However, (1) and (3) offer a different perspective on the same state of affairs; they focus on different parts. Whereas (1) describes the dying in its totality, including its endpoint, (3) presents the dying as a process going on at a certain point in time, to be provided by the context, and actually leaves open the possibility of miraculous recovery. We will say that the difference manifested in (1) and (3) is a difference of **aspect** rather than Aktionsart.

The distinction between aspect and Aktionsart, as it is made here, is neither unproblematic nor undisputed. It will be further discussed in 2. 3 and 4 will deal with the categories of aspect and Aktionsart separately. The main question to be answered here is how languages formally express these different notions and whether there is, in this respect, a systematic difference between the two categories.

### 2. Aspect versus Aktionsart

#### 2.1. A semantic distinction

The distinction between aspect and Aktionsart has been stated both in notional and in formal terms. The characterization which was given above is primarily a notional, or semantic one. In this view, Aktionsart is used to refer to a typology of states of affairs (event, state, etc.); by far the most influential one was proposed by Vendler (1957) (see 3.2). Aspect is seen as a linguistic means to offer a

certain perspective on any given state of affairs. More specifically, aspect indicates whether a state of affairs is seen either from an external viewpoint, as completed (perfective aspect), or from an internal viewpoint, as ongoing (imperfective aspect) (see 4.2). The latter distinction, in principle, applies to states of affairs of all kinds – all Aktionsarten – but the categories of Aktionsart and aspect do interact (see 5). The semantic distinction between Aktionsart and aspect is well captured terminologically when talking about them as **situation aspect** and **viewpoint aspect**, respectively (Smith 1991).

Less felicitously, the semantic difference is sometimes described in terms of objective (Aktionsart) versus subjective (aspect) information. This suggests that Aktionsart is given more or less automatically as an inherent property of states of affairs in the world, while aspect is more dependent on the choice of the speaker. However, it is, of course, the speaker who presents, by his choice of words, a state of affairs as either a state or an event. Indeed, the Aktionsart distinction between (4 a) and (4 b) is no less a matter of speaker choice than is the aspect distinction between (5 a) and (5 b).

- (4) (a) *He ate.*
- (b) *He ate an apple.*
- (5) (a) *He ate an apple.*
- (b) *He was eating an apple.*

The importance of making a semantic distinction between aspect and Aktionsart can be demonstrated most clearly by means of the distinction between **telicity**, a semantic feature at the level of Aktionsart, and **perfectivity**, a semantic notion at the level of aspect. Both concepts include reference to the endpoint of a state of affairs but they do so in different ways. As for Aktionsart, the predicate *eat an apple* in (4 b), (5 a) and (5 b) presents a telic situation: it has a well-defined endpoint, namely the moment at which the apple will be eaten. This is true irrespective of the question whether the situation is presented by means of a simple past tense, as in (5 a), or by means of a progressive, as in (5 b).

The simple past tense in (5 a) is interpreted as expressing **perfective** aspect: it is asserted that the endpoint of the situation was actually reached, i.e. the apple was eaten. The progressive in (5 b) expresses **imperfective** aspect: for (5 b) to be true, it is not necessary that the whole apple was actually eaten.

However, (5 b) may still be called telic, at the level of Aktionsart, since it refers to a situation that has a potential endpoint. Thus, aspect (perfectivity) and Aktionsart (telicity) are independent notions (cf. Declerck 1979 and Depraetere 1995, who refer to the semantic notion of perfectivity as **boundedness**). This was shown for French by Garey (1957) with the examples in Table 109.1.

	imperfective	perfective
telic	<i>Pierre arrivait</i>	<i>Pierre est arrivé</i>
atelic	<i>Pierre jouait</i>	<i>Pierre a joué</i>

Tab. 109.1: Telicity and perfectivity

The situation expressed by the French verb *arriver* ‘to arrive’ is telic, but it may be presented by either the imperfective past *imparfait* (*arrivait*) or the perfective past *passé composé* (*est arrivé*). Likewise, the verb *jouer* ‘to play’ refers to an atelic situation, but such a situation may be presented either imperfectively (*jouait*) or perfectly (*a joué*). Atelic situation types do not, at the level of Aktionsart, have a natural endpoint, but when such a state of affairs is presented by means of a perfective form it is understood to have ended at some, relatively arbitrary, point in time (cf. 5 on the interaction of aspect and Aktionsart).

## 2.2. A formal distinction

Another way of looking at the distinction between aspect and Aktionsart takes into account the way in which these aspectual notions are expressed in language. In this formal approach, one can find Aktionsart and aspect being referred to as **lexical aspect** and **grammatical aspect**, respectively. The intuition behind this terminology seems to be that Aktionsart is a lexical property of, most notably, verbs (cf. the distinction of telicity in Table 109.1), whereas aspect is expressed grammatically – usually by means of inflection, but sometimes by means of aspectual grammatical constructions, such as the English progressive in (5 b).

This approach to the aspect/Aktionsart distinction is not incompatible with the semantic approach outlined above. In fact, the formal approach seems to presuppose some semantic distinction. Indeed, given the difference between, in particular, telicity and perfectivity, discussed above, it would be hard

to maintain that these constitute one and the same semantic notion which some languages happen to express lexically (Aktionsart), while other languages express it by grammatical means (aspect). Rather, the way in which aspect and Aktionsart are realized linguistically should be seen as a possible additional distinction between the two.

There are, however, quite a few problems attached to distinguishing between Aktionsart and aspect as pertaining either to the lexicon or to the grammar.

As for the lexical nature of Aktionsart, it should be noted from the start that it is difficult to regard Aktionsart as a lexical property of verbs. This should be evident from the sentences in (4). In (4 a), the verb *eat* is used to refer to a state of affairs that has no well-defined end-point (atelic). The same verb is used in (4 b), however, as part of the predicate *eat an apple*, which refers to a state of affairs that does have a well-defined endpoint (telic). Thus, other elements in the clause co-determine Aktionsart and this makes the computation of “lexical” aspect quite a complicated undertaking (Verkuyl 1993). Moreover, even independent of other elements in the clause, verbs turn out to be surprisingly flexible when it comes to presenting various Aktionsarten (see 3.3). A final problem for the lexical approach to Aktionsart is constituted by the fact that some distinctions that do seem to indicate Aktionsart, in particular **phasal Aktionsart**, are, in many languages, expressed by grammatical means (see 3.4).

The grammatical view of aspect is problematic mainly because it is not obvious what should count as “grammatical”, or for that matter as “lexical”, in this approach. It has been argued, especially in the older literature, that use of the term “aspect” should be restricted to those languages in which the notion is obligatorily expressed, independently of tense, by morphological means. Under this definition, it can be, and has been argued, that the Germanic languages “do not have aspect”. The definition excludes, for instance, a periphrastic construction such as the English progressive, exemplified in (5 b), which clearly belongs in the domain of aspect from a semantic perspective. Not covered by the strictly formal definition of aspect either are the past tenses of the Romance languages, such as the French *passé simple/passé composé* (cf. 4.4.3) and *imparfait*, which were exemplified in Table 109.1. These tenses arguably express the semantic notions of perfec-

tive and imperfective past. There is, however, no single one morpheme in French expressing aspect *per se*, independently of tense.

As “true” aspect languages in the formal sense, the Slavic languages are often mentioned. In Russian, for instance, the perfective counterpart of an imperfective verb is formed through prefixation; the distinction is obligatory and it is independent from tense. However, it is a matter of some debate to what extent the “perfectivizing” prefixes of Russian in fact derive perfective verbs with a different lexical content, i.e. with a different Aktionsart (see 3.4 and 4.4.1). More generally, studies on grammaticalization have shown convincingly that grammatical markers often develop out of lexical material. As a consequence, it is not always possible, or useful, to classify aspectual forms at a given point of their diachronic development as belonging either to the lexicon or to the grammar (Bybee et al. 1994).

In the following sections, I will offer a separate treatment of Aktionsart (see 3) and aspect (see 4). In both sections, I will discuss in more detail the semantics of the category under scrutiny, and then describe how these semantic notions are expressed in language. Each section starts out with some terminological remarks.

### 3. Aktionsart

#### 3.1. Terminology

In the English linguistic literature, it was common to refer to all aspectual phenomena discussed thus far as *aspect*. Use of the German word *Aktionsart* (‘kind of action’), which is becoming more and more widespread in linguistics, arose out of a need to distinguish between qualitatively different kinds of “aspect”. The German term, which was introduced by Curtius in the late 19<sup>th</sup> century, was itself originally used as the German cover term for what we now distinguish as aspect and Aktionsart.

The practice of using the term “Aktionsart” specifically to refer to typologies of state of affairs, or rather the linguistic manifestation thereof, is even more recent. Originally, the term was used in the Slavicist tradition to refer to those morphological means that single out a particular phase of a situation – its beginning, middle, or end – as is the case, for instance, with inchoatives and resultatives (see 3.4). Although one can still find authors,

like Binnick (1991), who argue in favour of restricting the notion of *Aktionsart* in this way, it is becoming more common to use the term for typologies of the Vendler-type (see 3.2).

As English alternatives for “Aktionsart” have been used “verb aspect”, “verb character”, “intrinsic verb meaning”, “Aristotelian aspect”, “actionality”, “situation aspect” (see 2.1), and “lexical aspect” (see 2.2), but one can also still find the term “aspect” being used to cover both Aktionsart and aspect in the strict sense.

### 3.2. Semantics of Aktionsart: Vendler classes

In contemporary linguistics, the notion of Aktionsart is used most commonly to refer to a typology of states of affairs, or, more often, to the linguistic manifestation of such a typology. Zeno Vendler introduced the most influential of these typologies as four distinct “time schemata implied by the use of English verbs” (Vendler 1957: 144). The **Vendler classes** are, therefore, sometimes referred to as **verb classes**, but it should be clear, also from Vendler’s own examples given in (6), that entire predicates are involved in the determination of Aktionsart. (In this section, I will therefore use the term “Aktionsart” to denote a property of states of affairs as presented by predicates; I will return to the linguistic manifestation of these classes in 3.3.)

- (6) (a) state: *have, love, hate, know, believe*  
(b) activity: *walk, swim, push, pull*  
(c) accomplishment: *paint a picture, make a chair, read a novel, grow up, recover*  
(d) achievement: *recognize, realize, spot, lose, find, reach the summit, stop/start*

Semantically, Vendler's quadripartition is based on three underlying binary distinctions, namely **telic/atelic**, **dynamic/stative** and **durative/punctual** (cf. Smith 1991). I will briefly discuss each of these.

In the preceding sections, we already saw that the states of affairs as presented by the predicates in (7a) and (7b) differ from those in (7c) and (7d) in terms of telicity: the former sentences do not include explicit reference to the endpoint of the state of affairs (atelic), whereas the latter do (telic) (cf. 2.1).

- (7) (a) *He was dead.* (atelic)  
 (b) *He ate.* (atelic)  
 (c) *He died.* (telic)  
 (d) *He ate an apple.* (telic)

Telicity has also been called **boundedness** or **terminativity**. (Sometimes, however, the latter terms refer to what is called **perfectivity** here, i.e. to aspect rather than Aktionsart.) Some approaches to Aktionsart focus on the fact that telic predicates refer to a **change of state**: they present states of affairs as having a natural endpoint (or culmination), followed by a specific outcome (or result state) (Dowty 1979; Moens & Steedman 1988).

In order to determine which class a given predicate belongs to, "diagnostic tests" are used. (An elaborate survey of such tests is given by Dowty 1979.) A well-known test for telicity makes use of a distinction between two classes of temporal adverbials: *in*-adverbials and *for*-adverbials. The compatibility of telic and atelic predicates with these types of adverbials is shown in (8) and (9).

- (8) (a) *He walked for two hours.*  
     (b) *?He walked a mile for two hours.*

(9) (a) *?He walked in two hours.*  
     (b) *He walked a mile in two hours.*

Atelic predicates such as *walk* in (8 a) can be combined with *for*-adverbials, whereas telic predicates such as *walk a mile* in (8 b) cannot or receive an iterative reading. The reverse is true for adverbials of the *in-kind*: they are fine with telic predicates (see (9 b)), but marked with atelic predicates, such as in (9 a) which can only get a kind of inceptive or inchoative reading – meaning it took him two hours to start walking.

The states of affairs presented by the predicates in (7a) and (7b) are both atelic, but in another respect they are different: (10a) does not involve change, whereas (10b) does, and so do (10c) and (10d). The predicate in (10a) is therefore called **stative**, whereas the other sentences present **dynamic** states of affairs. This distinction corresponds to the one between **states** and **events**.

- (10) (a) *He was dead.* (stative)  
(b) *He ate.* (dynamic)  
(c) *He died.* (dynamic)  
(d) *He ate an apple.* (dynamic)

Stative states of affairs may begin or end at some point in time, but as long as they are holding, they remain the same throughout, at every moment of their duration. Since dynamic states of affairs involve change from one moment to the next, one moment in time does not suffice to determine whether or not a dynamic state of affairs is holding.

A diagnostic test that, in English, is often used in order to distinguish between stative and dynamic predicates is their compatibility with the progressive verb form. The predicate in (11a) can not be presented with a progressive form; the predicates in (11b)–(11d) are compatible with the progressive.

- (11) (a) *?He was being dead.*  
       (b) *He was eating.*  
       (c) *He was dying.*  
       (d) *He was eating an apple.*

It should be noted, however, that some states are fine with the progressive. This applies, in particular, to stative clauses presenting **temporary** as opposed to **permanent** properties (*Mary was living in London*, *Mary is missing her mother very much*).

Combining the features of telicity and dynamicity still does not enable us to distinguish between the states of affairs expressed in (7c)/(10c) and (7d)/(10d): both are telic and dynamic. A clear difference between them can be seen in (11c) and (11d). The progressive in (11d) presents the eating of the apple, from the inside, as going on at a particular point in time. The progressive in (11c), however, rather refers to the process preceding the actual dying. This is a consequence of the fact that the predicate in (12d) presents a **durative** state of affairs, whereas the predicate in (12c) presents a **punctual** one.

- (12) (a) *He was dead.*      (durative)  
       (b) *He ate.*                  (durative)  
       (c) *He died.*                (punctual)  
       (d) *He ate an apple.*      (durative)

A punctual state of affairs lacks temporal duration – its begin point and endpoint coincide, as it were – so there is no “internal” portion to be focused upon by the progressive in (11c). While the use of a progressive to present a punctual state of affairs does not result in an ungrammatical utterance, such utterances do receive a marked interpretation. They either refer to the process leading up to the actual state of affairs, as was the case in (11c), or they receive an **iterative** reading, as is the case in (13).

- (13) *He was knocking at the door.*

In (13), the temporal duration which is a necessary condition for the use of the progressive arises from interpreting the punctual state of affairs as occurring several times in a row.

In Table 109.2, the four Vendler classes first presented in (6) are characterized in terms of the three binary distinctions discussed so far.

Vendler class	dynamic	telic	durative
state	–	–	+
activity	+	–	+
accomplishment	+	+	+
achievement	+	+	–

Table 109.2: Aspectual parameters underlying the Vendler classification

The three distinctions, of course, allow for more than four possible combinations. Some of these have been argued to constitute Aktionsart classes of their own. Smith (1991), for instance, distinguishes five rather than four classes. She introduces the category of **semelfactives** for states of affairs that are dynamic, punctual and atelic; such states of affairs are called **points** by Moens & Steedman (1988). This class of atelic achievements would be constituted by predicates, such as *to knock* in (13), that receive an iterative reading when combined with the progressive. Other examples include *to cough* and the flickering of a light. Unlike telic achievements, such as exemplified by *to die* in (7c), semelfactives have no clear result or outcome associated with them. It is, therefore, hard to present a semelfactive by means of a present perfect, which is supposed to refer to a result state holding at the present moment (see also 4.4.3 and Art. 110); this is evidenced in (14).

- (14) (a) *He has died.*      (achievement)  
       (b) *?He has hiccupped.* (semelfactive)

Another addition to the four Vendler classes has been proposed by Croft (1999), who introduces the term **point state** for a state of affairs that is stative, and atelic, but not durative. This class would be exemplified by such sentences as *the train is on time* or *it is eight o'clock*.

Exactly how many, and which, Aktionsart classes one distinguishes is determined partly by the particular language one is dealing with, as well as by the specific linguistic phenomenon one is investigating. Some authors regard Aktionsart as representing a universal, cognitive distinction, independent of its linguistic manifestation in a given language. For the basic distinction between states of affairs that involve change (events) and those that

do not (states), this may be valid. However, as could be seen in the foregoing discussion, the Vendler classes themselves have been motivated primarily by means of diagnostic tests involving language-specific constraints of English. Not all Aktionsart features used in the analysis of one language necessarily have grammatical ramifications in every other language. It is likewise to be expected that other, or more, Aktionsart distinctions than the ones discussed thus far may be needed to account for other data or other languages (see 3.4).

### 3.3. The form of Aktionsart

Even though Vendler, being a philosopher, seems to have considered his classes to represent ontological rather than linguistic categories, the classification was well received in linguistics and has been used extensively in the linguistic literature to date. As for the relationship between the Vendler classes and linguistics, two important questions should be raised. The first question concerns the role of the Vendler classes in actual linguistic analyses: When and how are the four classes relevant? The second question concerns the expression of the Vendler classes in language. Phrased from the hearer's, or the linguist's, perspective: How does one decide which of the four classes a state of affairs as expressed by any given expression belongs to?

As for the first question, it can not be denied that Aktionsart has to be invoked in order to explain certain language-specific constraints on the use of adverbials or tense/aspect-marking (see, for instance, the findings on the progressive and the present perfect in 3.2). Moreover, the Vendler classes have been used in discourse-oriented approaches to explain the temporal interpretation of narrative discourse. It is, to some extent, determined by Aktionsart whether or not a clause "moves narrative time forward" with respect to the time of a preceding clause (Hinrichs 1986; Kamp & Reyle 1993). More specifically, events are usually understood to have happened in sequence, whereas states temporally overlap with surrounding events. The different behaviour of states and events in discourse is exemplified in (15).

- (15) (a) *John opened the door. He walked to the bookcase.*
- (b) *John opened the door. The room was brightly lit.*

Thus, the accomplishment of walking to the bookcase in (15a) is to be situated after the

opening of the door. The state of the room being lit in (15b), however, temporally overlaps the event of opening the door: the room was presumably lit both before and after John opened the door. The distinction between sequence and simultaneity in discourse has also been phrased in terms of **foreground** and **background** (Hopper 1979) (see also 4.3). Such findings on the significance of the Vendler classes for discourse issues have added to the popularity of the classification in recent years.

However, in all of the instances mentioned thus far, it is one of the underlying aspectual parameters presented in 3.2 that is explanatory, rather than the Vendler class as such. For instance, the specific, repetitive reading of (13), *He was knocking at the door*, arises out of the interaction of, on the one hand, the semantics of the progressive (aspect) and, on the other hand, the fact that we normally conceptualize the state of affairs presented by the predicate *knock at the door* as punctual (Aktionsart).

We clearly need the aspectual parameter of durativity to explain such cases, but this is in fact obscured by using the Vendlerian label *achievement* for the class of predicates involved, since the label includes information on such things as telicity and dynamicity, which are irrelevant to this particular issue.

Likewise, the different behaviour of the Vendler classes in discourse, illustrated in (15), seems to be a matter of telicity rather than anything else. This is clear from the fact that both accomplishments and achievements move narrative time forward. Apparently, the parameter of durativity, which may be used to distinguish between accomplishments and achievements (see 3.2), does not have a role to play here. Even though the names of the Vendler classes may be used as convenient, and well known, cover terms for certain clusters of aspectual features, it has been remarked that, in actual analyses, "linguists do not use his classes if they express linguistically relevant generalizations" (Verkuy 1993: 33).

The obvious relevance of Aktionsart for linguistic issues – irrespective of the labels one prefers to use, raises another important question, namely how, in a given case, to determine Aktionsart. In 2.2, it was already noted that Aktionsart cannot be regarded as a lexical property of verbs alone. For instance, the verb *to run*, in and by itself, does

not refer to the endpoint of the running activity, and (16a), is, therefore, atelic.

- (16) (a) *John was running.*  
       (b) *John was running a mile.*

Adding a “countable” object to such verbs, such as *a mile* in (16b), means adding an endpoint to the situational concept: at one point, John will have run a mile. The concept expressed by (16b) is, therefore, telic. (The progressive in (16b) expresses imperfective aspect, but this does not affect the Aktionsart of the predicate, as explained in 2.1.)

It is, however, not just the *presence* of an object phrase that is crucial in such cases, but rather the *nature* of the object. This can be seen in (17), which contains an object but still refers to an atelic state of affairs which may go on indefinitely.

- (17) *John was running miles.*

In contrast to *a mile* in (16b), *miles* in (17) does not refer to a countable object; it does not, in Verkuyl's (1993) terms, express a **specified quantity**. As a result, the predicate is atelic. In fact, both the subject and the object should express a specified quantity in order for the sentence as a whole to refer to a telic situation. Thus, even though (18), like (16b), contains a countable object, the sentence does not necessarily refer to a telic situation. This is a consequence of the fact that in (18) the subject does not present a specified quantity of people.

- (18) *Many people were running a mile.*

If, in addition to the object, the subject should be taken into account in order to determine Aktionsart, this means that Aktionsart is a property to be assigned at clause-level rather than at the level of the verb or the predicate. Verkuyl (1993) has shown how, in languages such as English and Dutch, Aktionsart is built up compositionally out of the temporal information provided by the verb, and the atemporal information ([ $\pm$  specified quantity]) provided by its arguments.

It is clear that the nature of the arguments has a crucial role to play when determining whether a clause presents a state of affairs as either telic or atelic. (Verkuyl, incidentally, refers to these properties as **terminative** and **durative**, respectively.) However, telicity is just one of the three aspectual features we distinguished in 3.2 and the information provided by the arguments is arguably of less significance when deciding if a state of affairs

is presented as durative/punctual or stative/dynamic. The verb *to run* in (16a), for instance, refers to a durative state of affairs. This is true irrespective of the nature of the subject or object: (16b)–(18) all remain durative.

However, this does not mean that the aspectual features of durativity and dynamicity differ from telicity to the extent that the former can be determined by looking solely at the lexical information provided by the verb. The determination of all Aktionsart features is partly dependent on other elements in the clause, context, and, ultimately, world knowledge. For instance, the verb *to love* is one of Vendler's own examples of the category of states (see (6)) and, in accordance with this, it does not easily appear in the progressive form, as can be seen in (19a).

- (19) (a) *?I am loving her.*  
       (b) *I am loving her more and more, the better I get to know her.*

As discussed in 3.2, compatibility with the progressive is often used as a criterion to distinguish between stative predicates, that do not express change, and dynamic predicates, that do involve change. In (19b), however, other elements in the sentence make clear that in this case the concept of loving does involve change and, therefore, use of the progressive is possible here. Even a sentence such as (11a), *?He was being dead*, may be used felicitously when talking about an actor who was doing his best to portray a dead man on stage.

The sentences in (19b) and (11a) confirm the idea that the progressive can only combine with dynamic predicates; both *to love* and *to be dead* get a kind of dynamic interpretation in these cases. However, such examples obviously cast serious doubts on attempts to divide verbs or predicates into fixed Aktionsart classes on the basis of the constructions in which they occur (Croft 1999).

### 3.4. “Grammatical Aktionsart”

The considerations on the linguistic manifestation of Aktionsart in 3.3 do not invalidate the claim, put forward in 2.2, that Aktionsart is a *lexical* category. Problems posed by the compositionality of meaning (see examples (16)–(18)) and the role of context (see examples (19) and (11a) as discussed in 3.3) are hardly restricted to the domain of Aktionsart and they confirm rather than contradict the

idea that we are dealing with a lexical category. In any case, overt morphological marking of Aktionsart was lacking in all examples discussed thus far.

However, many languages have derivational morphological means to modify the verbal action expressed by simple verbs and these means are sometimes referred to as expressing “Aktionsart”, or even “aspect”. In fact, this use of the term “Aktionsart” is older than its use for Vendler-type distinctions; the latter use has been called “an unfortunate terminological innovation” by Binnick (1991: 400), who introduces the term **Aristotelian aspect** for classifications such as Vendler’s.

Use of the term “Aktionsart” for morphological categories can be found particularly in the Slavic linguistic literature. This is not surprising since the Slavic languages abundantly use derivational morphological means to modify the action expressed by simple verb stems. Some of the prefixes used in this domain primarily serve to indicate the distinction between perfective and imperfective aspect (see, however, 4.4); other prefixes are said to indicate different Aktionsarten. Take, for instance, the Russian verb stem *pisat'* ‘to write’ in (20) (cf. Smith 1991: 242; Binnick 1991: 137). Used as a simple verb, in (20 a), it expresses imperfective aspect (‘to be writing’), like simple verbs in Russian typically do.

- (20) (a) *pisat'* ‘to write’ (imperfective)
- (b) *na-pisat'* ‘to write’ (perfective)
- (c) *nad-pisat'* ‘to write above’
- (d) *pere-pisat'* ‘to copy’
- (e) *po-pisat'* ‘to write for a little while’
- (f) *za-pisat'* ‘to start writing’
- (g) *do-pisat'* ‘to finish writing’

Adding a prefix to such imperfective verb stems, as in (20 b–g), renders the verb perfective (aspect), but, in addition, may also change the lexical meaning of the verb (Aktionsart). Whereas the distinction between (20 a) and (20 b) is said to be strictly a matter of aspect (see, however, 4.3), a change in lexical meaning can clearly be observed in (20 c–e), as should be evident from the English translations of these verbs. The sentences in (20 f) and (20 g) exemplify cases of what is sometimes called **phase**, **phasal aspect** or **phasal Aktionsart**. In such cases, a particular phase of a situation is singled out or focused upon. This may be the beginning of the situation (**inchoative**, sometimes called **ingressive**

or **inceptive**), as in (20 f), or its end (**completive**), as in (20 g). Phasal aspect is in many languages expressed by **aspectual verbs** such as *to start* and *to finish* in the English rendering of (20 f) and (20 g).

In form and meaning, the “perfectivizing” prefixes of Russian, only a few of which were exemplified in (20), show great resemblance to the verbal particles and prefixes in Germanic languages, such as, for instance, in English (21) and Dutch (22).

- (21) (a) *to write* (atelic)
- (b) *to write down* (telic)
- (c) *to write up* (telic)
- 
- (22) (a) *schrijven* ‘to write’ (atelic)
- (b) *op-schrijven* ‘up-write (to write down)’ (telic)
- (c) *over-schrijven* ‘over-write (to copy)’ (telic)
- (d) *in-schrijven* ‘in-write (to subscribe)’ (telic)

It is clear that grammatical means to modify a verb’s meaning may, among other things, change the Vendler-type Aktionsart of a verb: The bare verbs in (20 a), (21 a), and (22 a) are atelic, whereas the prefixed verbs in (20 b–g) and (22 b–d) as well as the verb particle combinations in (21 b–c) all present a telic state of affairs. It is equally clear, however, that such modifications of verb meaning should be distinguished from Vendler-type Aktionsart since none of the prefixes or particles mentioned marks telicity *per se*; the semantic difference between the verbs in (20), (21) and (22) cannot be described exclusively in terms of Aktionsart. Moreover, as discussed in 3.3, Vendler-type Aktionsart can only be determined at clause-level and it does not make much sense to refer to the verbs in (20), (21) and (22) as telic, given that such “telic” verbs may be used to present an atelic state of affairs. For instance, the verb-particle combination *to write down* in (21 b) can still be combined with an argument referring to an uncountable object, as in (23 b), resulting in an atelic sentence.

- (23) (a) *He wrote down the number.* (telic)
- (b) *He wrote down numbers.* (atelic)

Within the domain of “grammatical Aktionsart” one can, furthermore, find means to indicate, for instance, **iterativity** (the same action repeated on one occasion) or **habituality** (occurrences of the same action on different occasions). Such meanings may be expressed

by periphrastic constructions such as English *used to*, but are in many languages expressed by inflectional means or, not surprisingly, by means of reduplication. The latter is illustrated in the example from Mwera in (24) (Bybee et al. 1994: 160).

- (24) *lyā* ‘to eat’  
*lyalyalyā* ‘to eat and eat and eat’

Reduplication may also be used to signal **diminutive** meaning (to do something “a little”), as exemplified in the example from Tongan in (25) (Bybee 1985: 152).

- (25) *kata* ‘to laugh’  
*katakata* ‘to laugh slightly, to smile’

At this point, however, it can be questioned if there is anything to gain by using the term “Aktionsart” as a cover term for all derivational morphological means that serve to modify the action expressed by a simple verb. Some of these, such as (25), do not really concern temporal characteristics of states of affairs at all and may, therefore, fall outside of the domain of aspectuality altogether. Other grammatical means may in fact change the temporal structure of a state of affairs as expressed by the simple verb, but such means should still be distinguished from Vendler-type Aktionsart, especially since the latter concerns the aspectual meaning of entire clauses and the semantics of the verb, whether simple or derived, is just one of the contributing factors (see 3.3).

## 4. Aspect

### 4.1. Terminology

As was remarked in 3.1, the English term “aspect” is sometimes used as a cover term for both aspect and Aktionsart. In this contribution, as in most recent work on the subject, a distinction is made between the two categories (see especially 2). Aspect is thus specifically used to indicate the interpretation of a language-specific grammatical category of the verb, as expressing either imperfectivity or perfectivity. This “grammatical category of the verb” expressing aspect is in the grammars of many languages referred to as a tense, but most tenses express both temporal and aspectual information; tense and aspect are different, but interrelated phenomena (see 4.3).

The term “aspect” was used for the first time, at the beginning of the nineteenth cen-

tury, as a translation of Russian *vid* ‘view’ (cf. Latin *videre*). The root of *aspect* is *spect-*, which likewise means ‘to see’ (cf. Latin *spectare*). Even though the term originates from Russian, and this language has often been considered to represent an aspect language par excellence, it is not uncommon nowadays to find claims that Russian aspect is in fact an instance of Aktionsart (see 4.4.1).

### 4.2. Semantics of aspect

The sentences in (26) both contain a past tense, and both present a state of affairs that is telic, dynamic and durative (a Vendlerian accomplishment).

- (26) (a) *He was writing a letter.*  
(b) *He wrote a letter.*

The difference between (26a) and (26b) is, thus, not one of tense or Aktionsart; it is a difference of aspect. By his choice of aspect, the speaker indicates whether the state of affairs presented is either ongoing (imperfective aspect), or completed (perfective aspect). Thus, in (26a), the writing of the letter is presented as going on at a definite point in time, to be provided by the context or the situation; (26a) does not make any claims about whether or not the letter ever got finished. By using (26b), however, the speaker indicates that the writing of the letter got executed from beginning to end. (The distinction between simple and progressive forms in English does not in all cases coincide with the perfective-imperfective distinction, but it does in the case of past tense clauses presenting events rather than states, see 4.4.2 and 4.4.4.)

Presenting a state of affairs as either ongoing or completed is a typical function of the perfective-imperfective distinction in the languages of the world, but it is not the only possible one. In (26), we are dealing with a state of affairs that is telic at the level of Aktionsart; it has a well-defined endpoint, namely the moment at which the letter is finished. Therefore, completion/incompletion is well suited to characterize the effect of using either a perfective or an imperfective form to present such a state of affairs. However, the specific interpretation of aspect in a given case differs depending on lexical content (Aktionsart). For instance, in many languages states are typically presented by means of imperfective forms, such as the *imparfait* in French (27a).

- (27) (a) *Il était malade.*  
 (b) *Il fut malade.*

Since atelic states of affairs do not have a natural endpoint associated with them, it does not make much sense to say that the being ill in (27 a) is presented as in any way incomplete. Also, using the perfective past *passé simple* in such cases, as in (27b), does not necessarily present the state of affairs as completed. In fact, the sentence in (27b) prefers an inchoative reading, as in English *he fell ill* or *he got ill*, and does not even exclude the possibility that the illness is still holding at the present moment. Therefore, the following, more general, characterization of the perfective/imperfective distinction is arguably needed: perfective forms make reference to the boundaries of a state of affairs (which, depending on *Aktionsart* and context, may be either the starting point, the endpoint, or both), while imperfective forms present a state of affairs without regard to temporal boundaries.

In more recent approaches to aspect, the focus has shifted from defining the category in terms of (in)completion or boundaries to emphasizing the function of aspect at discourse level. In particular, aspect is said to indicate **perspective**, and to play an important role in establishing relations across clauses. I will briefly discuss both views.

The “perspectival” approach to aspect has a long tradition. Indeed, the term itself arguably means perspective (see 4.1) and it may well be considered somewhat of a tautology when Smith (1991) refers to (grammatical) aspect as “viewpoint aspect”. Imperfective forms are said to indicate an internal viewpoint: they look at a state of affairs from the inside. Perfective forms, on the other hand, present a state of affairs from the outside, as a “single, undivided whole” and thus indicate an external viewpoint (Comrie 1976).

The perspectival approach to aspect is not incompatible with the temporal approach. As for perfective aspect, one can only claim that a state of affairs is completed if one looks at it from a point in time after the completion happened, more specifically, from the point of speech, the “external” perspective of the speaker. In order to be able to say that a state of affairs is going on, one needs to refer to some point in time that is “internal” to the state of affairs. This may be the point of speech, as in the imperfective present in (28 a), but it may also be a point in time pre-

ceding the point of speech, as in the imperfective past in (28 b).

- (28) (a) *He is writing a letter.*  
 (b) *He was writing a letter.*

In the latter case, the “internal” point in time required for the interpretation of the imperfective past, may be identified with the perspective of someone other than the speaker, such as, in narrative discourse, a story character. This may explain why many languages predominantly use imperfective forms when representing the speech or thought of someone other than the speaker, such as in (free) indirect speech or reported perception, even in cases where there is no sense of “incompletion” at all (Ehrlich 1987). In such contexts, perfective forms may in fact be infelicitous altogether, such as in French (29).

- (29) ?*Il a dit qu'il fut malade.*

The French perfective past *passé simple* hardly ever occurs in the embedded clause of indirect speech. Apparently, the external viewpoint required for the interpretation of the embedded perfective past is incompatible with the internal viewpoint (of the reported speaker) suggested by the indirect speech context (Landeweerd & Vet 1996).

A further discourse-level notion which has been related to aspect is the distinction between **foreground** and **background** in narrative discourse (Hopper 1979). This distinction is exemplified in (30).

- (30) (a) *When John came into the room, Pete was writing a letter.*  
 (b) *When John came into the room, Pete wrote a letter.*

In (30 a), the writing of the letter, presented by means of a progressive, is interpreted as going on at the time of John’s entrance. The latter state of affairs, presented by means of a (perfective) simple tense, is part of the foreground, or the story-line; it “moves narrative time forward”. The writing of the letter is merely going on in the background, it temporally overlaps with states of affairs of the story-line. The sequence in (30 b), containing a simple (perfective) tense both in the sub-clause and in the main clause, is preferably interpreted such that Pete started writing the letter after John’s entrance. In this example, both states of affairs are part of the temporal-causal chain of events constituting the foreground, or the backbone, of the story.

On the basis of examples such as (30), equivalents of which can be found in nearly every language that marks perfective and imperfective aspect, it has been claimed that it is in fact the meaning of aspect to indicate temporal sequence or foreground versus simultaneity or background. (In languages that do not systematically mark grammatical aspect, Aktionsart is used as a clue for the temporal interpretation of discourse, as was noted in 3.3.) A problem of reducing the meaning of aspect to its discourse function is that the latter may vary depending on genre, or discourse type; the distinction between foreground and background as it is defined here is suited exclusively for narrative discourse. Furthermore, the linguistic information provided by aspect is hardly ever sufficient to determine the temporal interpretation of discourse, which is dependent to a large extent on extralinguistic information, such as provided by world knowledge.

#### 4.3. Aspect and tense

In 1, the categories of aspect and Aktionsart were separated from the category of tense (to be treated in Art. 110). Tense is a deictic category that serves to situate a state of affairs with respect to a reference point, usually constituted by the point of speech. Aspectuality, on the other hand, concerns the internal temporal structure of a state of affairs and does not present deictic information. The distinction between tense and aspect as it is made here is purely a semantic distinction; it does not exclude the possibility that, formally, tense and aspect are expressed by one and the same morpheme. The French *imparfait*, for instance, expresses both past tense and imperfective aspect.

The semantic independence of aspect and tense is clear for the category of imperfective aspect: a state of affairs may be ongoing at a point in time preceding the point of speech, as in (31 a), but it may also be ongoing at the present moment, as in (31 b), or at a point in time following the point of speech, as in (31 c).

- (31) (a) *He was writing a letter.*
- (b) *He is writing a letter.*
- (c) *He will be writing a letter.*

Thus, one can find instances of imperfective past, imperfective present, and imperfective future. This does not mean that every language that has means to indicate imperfective aspect, necessarily marks it in every tense,

but at least there is no conceptual clash between the notion of imperfective aspect and the notion of tense.

However, the semantics of tense and aspect are not always compatible and the situation is somewhat more complex for the category of perfective aspect. This is understandable since a state of affairs can not be at the same time completed, or bounded (perfective aspect), and holding at the present moment (present tense). In some sense, the present tense is inherently imperfective since it presents states of affairs that are ongoing at the point of speech. (In fact, the category of imperfective past, which presents a state of affairs as ongoing at a point in time preceding the point of speech, has been characterized as a **present in the past**.)

The linguistic consequences of the conceptual clash between perfective aspect and present tense are not the same in every language. In English, imperfective marking by means of the progressive is obligatory in clauses presenting events such as (31 a); the sentence in (32), containing a simple present tense, cannot be used to present the state of affairs as going on at the time of speech. (It may be used as a historical present or to present a habit.)

#### (32) *He writes a letter.*

In languages that mark aspect but not tense there is no real problem; in such languages, perfective forms are typically interpreted as referring to the past. In languages that do mark tense, the expression of perfective aspect may be restricted to the past. French, as most Romance languages, has two past tenses, a perfective and an imperfective one, but it makes no similar distinction in the present tense. Yet another possibility is manifested in Russian. As stated earlier, Russian marks aspect in every tense, and independently from tense (but see 4.4.1). However, forms that are both present (tense) and perfective (aspect) are usually interpreted as referring to the future (Binnick 1991: 138).

#### 4.4. Forms of aspect

Formal expression of the semantic distinction between perfective and imperfective aspect ranges from inflectional and derivational to periphrastic and zero-expression. (Large-scale typological research into the area has been reported on by Bybee 1985, Dahl 1985, and Bybee et al. 1994). I will discuss three issues which have generated quite a bit of

confusion in the aspect literature. First, I will briefly return to the topic of derivational “aspect”, as illustrated for Russian in 3.4. Then I will discuss two types of constructions that arguably express aspect, namely locative/progressive constructions and perfect constructions. Finally, I will address the aspectual interpretation of aspectually unmarked or “zero” forms, such as the simple tenses of English.

#### 4.4.1. Derivational “aspect”

As noted in 4.1, the terminology of perfective and imperfective aspect originated from the Slavic linguistic literature. These days, however, some forms that were called “aspect” in the Slavic languages are considered to represent Aktionsart rather than aspect (Bertinetto & Delfitto 2000).

In 3.4, it was shown how, in Russian, “perfective verbs” are formed through prefixation. Some of these prefixes have a clear lexical meaning of their own, thus changing the lexical content of the simple, imperfective verb. It was argued that, both formally and semantically, these prefixes belong in the domain of Aktionsart (even though they cannot be identified completely with Vendler-type Aktionsart). The same thing may in fact be argued for all “perfectivizing” prefixes of Russian, including those that are considered devoid of lexical content and merely grammatical markers of perfectivity, such as *na-* in (33 b).

- (33) (a) *pisa* ‘to write’
- (b) *na-pisa* ‘to write (up)’

The latter form may be compared to English *to write up* – all the more so since most of the aspectual prefixes in Russian derive from prepositions or adverbs (*na* ‘on’). Now, even though the semantic difference between (33 a) and (33 b), as the one between *to write* and *to write up*, may be difficult to pin down precisely, it is clear that the distinction between these lexical entries is different from the distinction between, for instance, simple and progressive forms in English or the one between the different past tenses of the Romance languages. If we consider the latter grammatical distinction one of aspect, we need a different term for lexical distinctions such as exemplified in (33). Since the distinction between lexical and grammatical “aspect” has come to be known as the distinction between Aktionsart and aspect (see 2), the Russian forms in (33) are now often la-

beled as differing in Aktionsart rather than aspect; semantically, they express telicity rather than perfectivity. In this view, Russian forms such as (33 b) are not overtly marked for the category of aspect at all; instead, Russian may be argued to have a systematic, grammatical way of indicating Aktionsart distinctions. (Dahl 1985: 89 refers to them as “grammaticalized lexical categories”.) Support for this analysis can be found in other Slavic languages, such as Bulgarian and Macedonian, that know both the Russian-type lexical means to derive telic verbs and the Romance-like inflectional means to indicate perfective and imperfective aspect; the latter distinction is expressed by the different past tenses **Aorist** (perfective past) and **Imperfect** (imperfective past), respectively (Bertinetto & Delfitto 2000).

In fact, Russian itself arguably has means to express grammatical aspect. In addition to “perfectivizing” prefixation, Russian knows imperfectivizing suffixation. Both in form and in meaning, imperfectivization by means of a suffix is more grammatical in nature than the kind of prefixation discussed above: there is only one productive imperfective suffix in Russian (-*yva-*), which does not change the lexical meaning of the verb. When attached to a prefixed perfective verb, the suffix produces a **derived imperfective** verb, as can be seen in (34). (Not all prefixed verbs in Russian have a secondary imperfective counterpart; (33 b), for instance, does not.)

- (34) (a) *pisa* ‘to write’ (atelic, imperfective)
- (b) *pere-pisa-* ‘to copy’ telic, perfective)
- (c) *pere-pis-yva* ‘to copy’ (telic, imperfective)

The arguments given in favour of the claim that the distinction between (34 a) and (34 b) is a matter of lexical Aktionsart rather than aspect, do not apply to the distinction between (34 b) and (34 c).

#### 4.4.2. Progressive/locative constructions

A further debated issue when it comes to the formal expression of aspect concerns the question to what extent constructions such as the progressive and the perfect express imperfective and perfective aspect.

In the preceding sections, I often used the English progressive to illustrate the semantics of imperfective aspect. It should be clear, however, that the English progressive covers only a subdomain of the category imperfec-

tive aspect and the same is true for many progressive-like constructions in the languages of the world. Such progressive constructions are often derived from a locative source. They either contain prepositions (or postpositions) indicating location, or they express a locative notion in the auxiliary used (Bybee et al. 1994: 129). Both types can be found in Dutch, see (35).

- (35) (a) *Hij is aan het lezen.*  
he is on the read
- (b) *Hij zit te lezen.*  
he sits to read

The verb formation in (35a) contains the preposition *aan* ('on'); in (35b), the auxiliary derives from a postural verb (other verbs that may be used in this construction are equivalents of *to lie*, *to stand*, and *to walk*).

The original meaning of locative/progressive constructions is something like "subject is in the midst of doing something" and therefore they give rise to an imperfective reading: the state of affairs is presented without reference to its boundaries, as going on at a definite point in time. The reverse, however, does not hold: not all imperfective readings can be marked by progressive/locative constructions. In particular, the use of such constructions is usually restricted to dynamic predicates (cf. 3.2), which arguably can be related to their locative origin (Bybee et al. 1994: 136). The difference between imperfective and progressive can be illustrated by means of (36) and (37).

- (36) *John était fâché.*
- (37) (a) *John was angry.*
- (b) *John was being angry.*

The French *imparfait* in (36) presents the state of being angry as holding at a definite moment in time. Since states are inherently atelic (or unbounded), they are typically presented by means of imperfective forms. The English progressive, however, can hardly ever be used to present a stative predicate. In order to obtain the "unmarked", imperfective reading of French (36), English must use a simple tense, as in (37a). If English uses a progressive, as in (37b), the predicate, in addition to being imperfective, necessarily receive a dynamic, activity-like, interpretation.

Thus, even though the English progressive expresses imperfective aspect, it does so for only a restricted set of predicates. Highly similar restrictions hold for most locative/

progressive constructions, even though some of them may be developing into more general imperfectives (Bybee et al. 1994: 140).

#### 4.4.3. Perfect versus perfective

Another construction which has been claimed to express aspect is the perfect. However, **perfect** should be carefully distinguished from perfective. The distinction may be clarified by the English sentences in (38).

- (38) (a) *He was reading a book.* (imperfective)
- (b) *He read a book.* (perfective)
- (c) *He has read a book.* (perfect)

The term "perfective" is used to characterize the interpretation of (38b) as opposed to (38a): (38a) is imperfective, it presents the state of affairs without reference to its boundaries; (38b) is perfective, it presents the reading of the book as bounded, and, in this case, completed. Clearly, we need a different term to distinguish between the interpretation of (38b) and (38c). The latter pair of sentences exemplifies the difference between perfective and perfect (the latter is sometimes referred to as **anterior**).

Contrary to the perfective past in (38b), the perfect in (38c) is a present tense: in addition to referring to a situation in the past (as expressed by the past participle), the present perfect explicitly refers to the present moment, the moment of speech. In fact, the finite form of the present perfect arguably presents a state of affairs in its own right: it presents a state holding at the present moment. The exact nature of this state is not always equally clear. In the case of telic predicates, it can often be identified with the result state of a past event. Indeed, most perfect constructions derive from **resultative** constructions (Nedjalkov & Jaxontov 1988), presenting the result rather than the event which caused it. In other instances, the link between the past situation and the present moment expressed by the present perfect will be more abstract; this includes all cases which have been labeled "current relevance" in the literature. In any case, the semantic notion "perfect" differs from the semantic notion "perfective", mainly since the latter lacks the additional dimension of explicit linking to the speech situation (cf. Art. 110).

However, the formal category labeled "present perfect" can in some languages be used to express the semantic notion of perfective past. This is understandable if we look at

it from a diachronic perspective. It has been argued for perfects in many, genetically unrelated, languages that the semantics of this category systematically develops from resultative to perfect to perfective. Thus, the present perfect would start out as a present tense, presenting a (result) state holding at the present moment, and develop into a past tense, expressing a situation that precedes the point of speech. The English present perfect does not really allow for such perfective readings, but in French, for instance, the *passé composé* (present perfect) has taken over the functions of the earlier perfective past *passé simple*. Also, Southern German is often cited as a variety of German in which the perfect expresses simply past time rather than present perfect. The Dutch present perfect takes up a middle position between the English and the French/German ones (Boogaart 1999; many more instances of the development can be found in Bybee et al. 1994: Chapter 3).

To sum up: even though the semantic notions of perfect and perfective are clearly different, the formal category perfect can in many languages be used as a perfective past tense.

#### 4.4.4. Invisible aspect

To conclude this section on the formal manifestation of aspect, I want to mention those cases in which aspect is not formally marked at all. There are of course tenses that are neither perfective or imperfective, but simply present or past. The aspectual interpretation of such forms will depend on tense (see 4.3), lexical content (*Aktionsart*), and context. But this cannot be the whole story.

For instance, I have been using sentences like English (39 a) to illustrate the category of perfective past.

- (39) (a) *He wrote a letter.*
- (b) *He was sick.*

The English simple tenses are, however, not marked for perfectivity. Rather, given sentences such as (39 b), which are clearly imperfective (cf. discussion of (36) and (37)), we have to say that the simple past tense of English can express both perfective and imperfective aspect. The disambiguation of such forms can, however, not just be a matter of tense and lexical content (*Aktionsart*): there is no intrinsic reason why a state of affairs that is durative, dynamic and telic, such as in (39 a), could not be interpreted as going on at a definite moment in the past. Indeed, the Dutch

unmarked past tense in (40) does allow for such an imperfective reading.

- (40) *Hij schreef een brief.*  
he wrote a letter

In fact, Dutch (40) allows for both perfective and imperfective readings. The difference between English (39 a) and Dutch (40) is a consequence of the fact that in English, but not in Dutch, the marking of imperfective aspect by means of a progressive construction has grammaticalized to the extent that it has become obligatory in clauses presenting events. As a result, English (39 a) can only be used to express perfective aspect. For stative clauses such as (39 b), on the other hand, the expression of imperfective aspect by means of a progressive is not possible (see 4.4.2), and, therefore, (39 b) does allow for an imperfective reading. In Dutch, the expression of imperfective aspect by means of the locative constructions in (35) is often possible, but never obligatory, and, therefore, sentences such as (40) still allow for an imperfective reading (Boogaart 1999).

As English (39 a) shows, the absence of aspectual marking may be as meaningful as the presence of such marking. This development of aspectual meaning in unmarked forms has been described as **grammaticalization of zero** by Bybee (1994); the pragmatic, inferential mechanisms involved are described by Bickel (1996) for the Himalayan language Belhare.

### 5. Interaction of aspect and *Aktionsart*

At the outset of this article, especially in 2, some effort was made to distinguish between *Aktionsart* and aspect. *Aktionsart* concerns the type of state of affairs as presented by the lexical information in a clause (in terms of telicity, durativity, and dynamicity) (see 3.2). Aspect concerns the interpretation of language-specific grammatical categories, as either perfective or imperfective (see 4.2). Even though the distinction, like the general distinction between lexicon and grammar, is not always clear-cut (see especially 3.4 and 4.4.1), it is important to be able to make it, especially since the two categories interact in interesting ways. This was apparent at various points in the preceding sections.

For instance, we may say that an atelic state of affairs (*Aktionsart*) is typically imperfective (aspect). Thus, in languages that

mark imperfective aspect, statives will typically be presented by means of imperfective forms, as in (41).

(41) *Il était malade.*

This is understandable since such a state of affairs does not have an inherent, natural endpoint, and an imperfective form presents a state of affairs without regard to its endpoint. Similarly, a non-durative state of affairs (Aktionsart), such as in (42), will standardly be perfective (aspect).

(42) *He died.*

Imperfective aspect presents a state of affairs from an “internal” point in time, but non-durative state of affairs lack internal structure (their starting point and endpoint coincide). On the basis of such examples, it will be clear that in languages in which aspect has not been grammaticalized, lexical information will be an important clue when deciding between a perfective or an imperfective reading.

However, grammatical aspect may overrule the standard interpretation on the basis of lexical information. Thus, the atelic state of being ill may be presented by means of an explicitly perfective form, such as the *passé simple* in (43).

(43) *Il fut malade.*

Likewise, the punctual situation of dying may be presented by means of the explicitly imperfective progressive form, as in (44).

(44) *He was dying.*

In order for the interpretations of (43) and (44) to be compatible with the semantics of aspect, these sentences receive a non-standard reading. Thus, the state of being ill in (43) is interpreted such that it does have a bound; depending on the context of use, (43) may receive an inchoative reading. The sentence in (44) no longer presents the punctual event of dying, but rather the durative process preceding the actual dying. In order to obtain such “marked” readings, languages that lack grammatical aspect, will have to rely on other lexical means and context. (A discussion of marked versus unmarked aspect choice can be found in Smith 1991.)

We might say that the French sentence in (43) no longer refers to an atelic state of affairs at all. Also, the sentence in (44) arguably

presents a durative rather than a punctual state of affairs. In other words, grammatical aspect can establish **coercions** from one Aktionsart category to the next (Moens & Steedman 1988). In a similar vein, one can find claims that, irrespective of the Aktionsart of the verb or the predicate, the French perfective past *passé simple* always presents an event, whereas the imperfective past *imparfait* always presents a state (Kamp & Reyle 1993). In such proposals, aspect is actually defined in terms of Aktionsart. Even though there is something intuitively correct about such an approach, it should be clear that it blurs the distinction between aspect (perfectivity) and Aktionsart (telicity). A distinction about which, finally, some consensus seems to have been arrived at.

## 6. References

- Bertinetto, Pier Marco & Delfitto, Denis (2000), “Aspect vs. Actionality: Why They Should Be Kept Apart”. In: Dahl, Östen (ed.), *Tense and Aspect in the Languages of Europe*. Berlin, New York: Mouton de Gruyter, 189–226
- Bickel, Balthasar (1996), *Aspect, Mood, and Time in Belhare: Studies in the Semantics – Pragmatics Interface of a Himalayan Language*. Zürich: Universität Zürich (Arbeiten des Seminars für Allgemeine Sprachwissenschaft 15)
- Binnick, Robert I. (1991), *Time and the Verb: A Guide to Tense and Aspect*. New York, Oxford: Oxford University Press
- Boogaart, Ronny (1999), *Aspect and Temporal Ordering: A Contrastive Analysis of Dutch and English*. The Hague: Holland Academic Graphics
- Bybee, Joan L. (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam, Philadelphia: John Benjamins
- Bybee, Joan L. (1994), “The Grammaticization of Zero: Asymmetries in Tense and Aspect Systems”. In: Pagliuca, William (ed.), *Perspectives on Grammaticalization*. Amsterdam: John Benjamins, 235–254
- Bybee, Joan L. & Perkins, Revere & Pagliuca, William (1994), *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago, London: The University of Chicago Press
- Comrie, Bernard (1976), *Aspect*. Cambridge: Cambridge University Press

- Croft, William (1999), "The Structure of Events and the Structure of Language". In: Tomasello, Michael (ed.), *The New Psychology of Language*. Mahwah/NJ, etc.: Lawrence Erlbaum, 67–92
- Dahl, Östen (1985), *Tense and Aspect Systems*. Oxford: Basil Blackwell
- Declerck, Renaat (1979), "Aspect and the Bounded/Unbounded (Telic/Atelic) distinction". *Linguistics* 17, 761–794
- Depraetere, Ilse (1995), "On the Necessity of Distinguishing between (Un)boundedness and (A)telicity". *Linguistics and Philosophy* 18, 1–19
- Dowty, David R. (1979), *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel
- Ehrlich, Susan (1987), "Aspect, Foregrounding and Point of View". *Text* 7, 363–376
- Garey, Howard B. (1957), "Verbal Aspect in French". *Language* 33, 91–110
- Hinrichs, Erhard (1986), "Temporal Anaphora in Discourses of English". *Linguistics and Philosophy* 9, 63–82
- Hopper, Paul J. (1979), "Aspect and Foregrounding in Discourse". In: Givón, Talmi (ed.), *Discourse and Syntax*. New York: Academic Press, 213–241
- Kamp, Hans & Reyle, Uwe (1993), *From Discourse to Logic: An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht etc.: Kluwer Academic Publishers
- Landeweerd, Rita & Vet, Co (1996), "Tense in (Free) Indirect Discourse in French". In: Janssen, Theo A.J.M. & Wurff, Wim van der (eds.), *Reported Speech: Forms and Functions of the Verb*. Amsterdam: John Benjamins, 141–164
- Moens, Marc & Steedman, Mark (1988), "Temporal Ontology and Temporal Reference". *Computational Linguistics* 14, 15–28
- Nedjalkov, Vladimir P. & Jaxontov, Sergej Je. (1988), "The Typology of Resultative Constructions". In: Nedjalkov, Vladimir P. (ed.), *Typology of Resultative Constructions*. Amsterdam: Benjamins, 3–62
- Smith, Carlota S. (1991), *The Parameter of Aspect*. Dordrecht etc.: Kluwer Academic Publishers
- Vendler, Zeno (1957), "Verbs and Times". *Philosophical Review* 66, 143–160
- Verkuyl, Henk J. (1993), *A Theory of Aspectuality: The Interaction between Temporal and Atemporal Structure*. Cambridge: Cambridge University Press

Ronny Boogaart, Amsterdam  
(The Netherlands)

## 110. Tense

1. General
2. The notion of "tense" and its relation to "mood" and "aspect"
3. Formal models of the semantics of tense
4. Past, present, and future
5. Past time reference and past tense
6. Future time reference and future tense
7. Present tenses
8. Perfect
9. Past and future perfects
10. Experiential
11. Narrative
12. Remoteness
13. Tense and subordination
14. Interaction between tense, negation, and expectation markings
15. References

### 1. General

This article treats tense, viewed as grammaticalized time-reference. The closely related phenomenon of aspect is treated in Art. 109.

The structure of the article is as follows: First, the notion of tense and its relations to mood and aspect are treated in general terms, and two influential ways of formalizing tense (Reichenbach and Prior) are summarized. Then, the traditional tenses past, future, and present are discussed (in that order) in the context of grammatical marking of reference to the respective segment of time. Then, other tenses are treated individually under headings such as perfect, narrative tense, and remoteness distinctions. Finally, the question of tense in subordinate constructions is considered.

The literature on tense is immense. For a recent survey of theories on tense and aspect from the ancient Greeks to modern times, see Binnick (1991). Some other recent works of a general nature are Comrie (1985), Dahl (1985), Bybee et al. (1994), and Klein (1994).

## 2. The notion of “tense” and its relation to “mood” and “aspect”

**Tense** is commonly seen as a grammatical category that expresses the notional category of time. The problem is that it is usually difficult to find anything in the grammatical systems of natural languages that fits this view very well. The semantics of the grammatical phenomena traditionally treated under the heading “tense”, such as the distinction between English verb forms and periphrastic constructions such as *knows*, *knew*, and *will know*, tends to involve a variety of factors, both temporal factors and ones that would rather be called aspectual or modal. Conversely, temporal considerations may show up also in other parts of the grammar than those normally thought of as “tense”. For instance, even in languages where there is seemingly a clear distinction in grammar between tense and aspect, such as Russian, it is not possible to describe adequately the semantics of the latter without invoking temporal notions. Furthermore, the “tenses” of traditional grammar tend to be heterogeneous as to their mode of expression: for instance, in English, the past tense is inflectional, whereas what is commonly referred to as the future tense is an auxiliary construction. Rather than speaking of the grammatical categories of tense, mood, and aspect, it might be more fruitful to talk of “tense-mood-aspect systems”, comprising all the phenomena usually subsumed under the three traditional notions. Tense, mood, and aspect would then be understood rather as indications of the main semantic domains providing the distinctions relevant to tense-mood-aspect systems, although not necessarily separated formally. However, when using the term “tense” by itself, as we shall in this article, it still seems wise to follow the tradition in modern linguistics that restricts it to what Comrie (1985: 9) calls “grammaticalised expression of location in time” rather than to time reference in general. The term “tense”, obviously, goes back to Latin *tempus*, which of course just meant ‘time’, like the corresponding Greek *chrónos*. The terminological distinction between “tense” and “time” came later and is still not made in e.g. French and Russian, which translate both English words in the same way (*temps* and *vremja*, respectively).

The central cases of tense in the sense just defined operate on the sentence or verb phrase level, realized as inflectional markings

on verbs, auxiliaries, or other words in predicate positions, or as periphrastic constructions. Sometimes, grammars also talk of tense on noun phrases (with interpretations similar to that of *former* in *my former husband*), but since no systematic treatment of such cases has been made they will be neglected here, like other distinctions such as the one between the German temporal subordinators *als* and *wenn*, which among other things involves temporal considerations and could accordingly be seen as a manifestation of tense.

It is important to understand the consequences of the distinction between grammatical and other means of referring to time. One often sees statements like the following: “Since language X has no tenses, time reference is indicated by time adverbials.” Such claims presuppose a basic functional equivalence between tenses and non-grammaticalized means of temporal reference. An empirical consequence ought to be a noticeable frequency difference with regard to the occurrence of time adverbials between languages with well-developed tense systems and those without. However, if such a difference exists, it is rather marginal. It follows from the grammaticalized nature of tense marking that tenses are used in addition to rather than instead of time adverbials. That is, the distribution of tense markers is fairly independent of that of time adverbials. Being grammaticalized, tense markers tend to be used even if they are redundant from an informational point of view. Indeed, there may even be a tendency for tense choice to be more constrained (or for tense markers to be more obligatory) when there is a time adverbial in the same clause than when there is not, especially if the latter is deictic. Thus, one may often in English choose a “historical” present rather than the simple past, but this option is normally blocked if the clause contains an adverb such as *yesterday* (*John arrives yesterday* is felt to be rather deviant). Such facts suggest that tense by nature comes close to other grammaticalized phenomena like agreement.

## 3. Formal models of the semantics of tense

On the most simplistic view, tenses relate the time when something happens to the time when the utterance is made. Such a two-point theory would yield the traditional three basic

tenses, past, present, and future. Many contemporary analyses of tenses are based on the model of the German philosopher Reichenbach (1947), in which there is in addition to the **event point** and the **speech point** a third point, the **reference point**. In most cases, the reference point coincides with either the event point or the speech point. An example where all three points are distinct would be the second clause of the sentence *I called John, but he had left*. Here, the point in time when I phoned serves as reference point whereas the event point is the time of John's leaving. Reichenbach actually never gave a very explicit definition of the notion of reference point; this has probably increased the popularity of his analysis, since everyone may find an interpretation of it to one's own liking.

One limitation of Reichenbach's theory is that it only accounts for what is traditionally called **absolute tense**, i.e. tense anchored in the point of speech. For tense based on other anchoring points, see 13. Reichenbach also has little to say about reference to time intervals (rather than to points).

Another formal tradition – usually referred to as **tense logic** – emanates from the logician Arthur Prior (e.g. Prior 1967), and was later integrated into Montague semantics (Montague 1974). Prior postulates a number of tense operators such as “it has at least once been the case that *p*”, “it will always be the case that *p*”, etc. While this system captures some interesting features of temporal reference in natural language, it has to be modified in rather fundamental ways before being usable for modelling tense systems in those languages.

#### 4. Past, present, and future

The ternary division of the time-line into past, present, and future has obtained cliché status; the assumption that there are three basic tenses with the same names is usually either taken for granted or regarded as the traditional view which is there to be criticized. It is of some interest to note that alternatives to the three-tense theory have been around from the very start; thus, Aristotle acknowledged only the past and the future as tenses (or rather, “inflections” (*ptōseis*)) of the verb, regarding the present simply as the absence of tense marking. Later, many scholars have noted the asymmetries in expression

and content between the past and the future, and the situation is further complicated by the fact that languages often have several tense-aspect forms for expressing each segment of the time-line. In spite of this, it may be most practical to take the traditional notions as a starting-point for the further discussion.

#### 5. Past time reference and past tense

Intuitively, a **past tense** should be a grammatical form used to refer to the past, that is, to points in time that precede the temporal reference point, whether identical to the point of speech or not. However, marking past time reference is relatively seldom a completely straightforward matter. In most languages, there are a number of alternative ways of referring to a past state-of-affairs, which makes the identification of past tenses problematic.

It is thus extremely common for there to be an interaction between the perfectivity-imperfectivity distinction and tense marking. As is noted in Art. 109, the perfective aspect is often restricted to past time reference. Such a restriction is combinable with several different kinds of tense marking. To start with, the language may have no past tense at all, which means that the use of the perfective aspect alone carries the information that the sentence refers to the past. An example of such a system is Yoruba, where the (unmarked) perfective aspect in a sentence such as (1 a) is understood to refer to the past, whereas a sentence in the imperfective aspect such as (1 b) does not distinguish between present and past time reference.

- (1) (a) *mo pàdè re*  
I meet him  
'I met him'
- (b) *ó-n ko letà*  
he-IPFV write letter  
'he is/was writing a letter'

Other systems complement a perfective/imperfective distinction with a past tense marking that is restricted to the imperfective aspect, like the “imperfect” tenses of many Indo-European languages. In a variant of such a system, the past marker may also be added to perfective forms, but then with an interpretation analogous to pluperfects (past perfects) in other languages. Such markers are claimed to be common especially in Cre-

ole languages and are then often labelled ‘anterior’ (Bickerton 1981), but are also found in other languages, notably a number of West African ones. As an example, consider the following sentences from Karaboro, a Gur (Niger-Congo) language.

- (2) (a) *kē ká?a gbɔɔ?*  
the house big  
'the house is big'
- (b) *kē ká?a ba gbɔɔ?*  
the house PAST big  
'the house had been big'
- (c) *ù kū*  
he die  
'he died/has died'
- (d) *ù bá kū*  
he PAST die  
'he had died'

Finally, there may be a general past which occurs independently of the perfectivity-imperfectivity marking, like the Russian past tense:

- (3) (a) *otkry-l*  
open-PAST  
'opened (perfective)'
- (b) *otkry-va-l*  
open-IPFV-PAST  
'opened (imperfective)'

A very frequent situation is for there to be a distinction analogous to the one between the “simple past” and the “perfect” in English. In the Reichenbachian model, ‘past’ time could in principle mean two things: that the reference point precedes the speech point or that the event point precedes the reference point. One of the major claims made by Reichenbach is that these two possibilities account for the Simple Past:Perfect in English. We shall discuss perfects in more detail in 8. At this point, it may be noted that there is a common diachronic path from perfects to “simple” or “general” pasts. One can also show in various ways that the former synchronically show a lower degree of grammaticalization than the latter, being for instance more often expressed analytically (peripherastically).

The marking of past time reference is also influenced by discourse type. Narrative texts, that is texts which relate sequences of events in the order they occur, often follow special rules. In many languages, the verbs in such texts tend not to be marked for tense (except for the very first sentence of a narrative).

Since the time reference in narratives is more or less wholly determined by the preceding context, tense marking carry no or little information. Only highly grammaticalized past time markings can therefore be expected to show up in narratives. Another possibility is the use of special narrative tenses (see 11).

Finally, past time marking may interact with evidentiality. We may know about past states-of-affairs from different sources: we may have experienced them ourselves, we may have heard about them from someone else or we may have made an inference from other states-of-affairs. Many languages have different markers for these different types of evidence, or morphemes that simultaneously mark pastness and type of evidence, as in Komi (Finno-Ugric, Leinonen & Vilkuna 2000):

- (4) [What do you know about this book?]
  - (a) *Sijös gižis Grem Grin.*  
it:ACC write:PAST<sub>1</sub> Graham Greene  
'Graham Greene wrote it (I am convinced of it.)'
  - (b) *Sijös Grem Grin gižöma.*  
it:ACC Graham Greene write:PAST<sub>2</sub>  
'Graham Greene wrote it (according to what I heard.)'

Past tenses in a narrow sense, then, will be those tense-mood-aspect devices one of the main functions of which is to mark the pastness of the reference point relative to the speech point, without at the same time indicating perfectivity or evidentiality. An obvious example is the Simple Past in English. An example of a past tense that does not contrast with a perfect is the one found in Russian. Pasts seem to be highly grammaticalized and predominantly marked inflectionally, although some caution may be called for here, given that free morphemes marking past time reference appear to be more common in some linguistically less well mapped parts of the world. Europe happens to be an area where pasts are more frequent than elsewhere. In the languages having a grammaticalized past/non-past distinction, there also tend to be copulas in nominal and adjectival predictions.

Pasts (in the narrow sense) often have uses which are not definable strictly in terms of time reference. For instance, in English counterfactual constructions such as *if I got the Nobel Prize, I would become famous* the form *got* does not have to refer to a time in

the past but rather indicates that the condition is not fulfilled in the real world or is highly improbable. Some rather spectacular uses of imperfective pasts of a parallel kind can be found in the Romance languages, where the *imparfait* (or the corresponding form) may be used e.g. to render the content of dreams. Similarly, in child language, past tense marking is often used to represent what is pretended in playing. Observations of this kind have been taken as a basis for claiming that the meaning of past tenses is something like "distance", where "distance" may pertain both to time and reality. An alternative view is to see the non-temporal uses rather as extensions of the primary or prototypical temporal meaning of pasts. An argument in favor of the latter treatment is that it is relatively uncommon cross-linguistically for past tense morphemes to serve as exclusive markers of counterfactuality.

## 6. Future time reference and future tense

The existence of **future** tenses in the proper sense, that is, a grammatical form that marks future time reference, has been an issue of debate for quite some time. It has been repeatedly pointed out that the interpretation of e.g. *shall/will* in English involves modal notions in various ways (such as the use of *will* to express a supposition with non-future time reference, as in *he will be about fifty years old now*), and this has been taken as an argument for treating them as modal auxiliaries rather than as tense markers. On the other hand, most if not all tense-mood-aspect devices have a more or less mixed semantics, and there is no doubt that temporal factors do play a role also for English *shall/will*, as well as for e.g. the inflectional future tenses in Romance languages. From the point of view of grammaticalization processes, an important consideration is that the temporal element in originally modal constructions like *shall/will* tends to become more pronounced over time (Fleischman 1983), as they come to be more systematically used in sentences with future time reference. The claim that the future is less grammaticalized than the past may be "Eurocentrically" biased: Europe happens to be a region where highly grammaticalized future tenses are less common than elsewhere. Futures tend to develop from the following sources:

- (a) constructions expressing intentions, desires, wishes, etc. (e.g. English *will* with the earlier interpretation 'want');
- (b) constructions expressing obligation, necessity, etc. (e.g. English *shall*);
- (c) constructions involving verbs of motion with original meanings such as 'go' and 'come' (as in English *be going to*);
- (d) temporal adverbs.

Developing futures usually go through preliminary stages when their use is more restricted. For instance, in such a grammaticalization process a construction like 'he wants to write a letter' starts out by expressing pure desire, then it begins to be used to say that he has an intention that the agent will carry through, and finally the construction may eventually come to be used also for cases when the action is not planned by the agent, and thus be extended to cases like 'it will rain'.

A further way that futures may arise is when a form or construction which has earlier been used for both present and future time reference loses its present time uses, as a result of the expansion of an earlier progressive construction. Such constructions often have other uses such as habitual ones. In some languages, such as Russian, this development is restricted to verbs in the perfective aspect.

Typically, futures first show up in main clauses. In particular, temporal and conditional clauses exhibit resistance against future marking and usually only highly grammaticalized futures are systematically used in those contexts. Also in main clauses there are some kinds of future time reference that are not usually marked as such. One type that tends to be unmarked cross-linguistically is what has been called "scheduling", as in *the library opens at nine tomorrow*.

Futures are expressed both inflectionally and periphrastically, with roughly equal frequency (Dahl 1985; Bybee et al. 1994). However, inflectional futures differ from periphrastic ones in being less advanced from the point of view of grammaticalization. Thus, inflectional futures have greater overall frequency and are more often used e.g. in temporal and conditional clauses.

Tense-mood-aspect devices often combine future uses with others, such as habitual ones, which were mentioned above.

## 7. Present tenses

A **present** tense would, reasonably, be a tense that is used to refer to the present, that is, to the time of speech. One problem that we have already touched upon (cf. 4) is the question whether the present is a tense in its own right or just something that can be defined negatively, as the absence of marking for past or future – the default or unmarked alternative.

Indeed, there is seldom an entirely symmetric (or in structuralist terms, equipollent) relation between different forms in a tense-mood-aspect system, and it is perhaps not too astonishing to find that forms referring to the present tend to be less marked in various respects, both with respect to expression and content, than forms referring to the future and the past. However, this general statement has to be somewhat modified. The simplest forms in paradigms are not always those used for present time reference. In particular, it is not uncommon that completed events in the past, i.e. such states-of-affairs for which perfective forms are typically used, are zero-marked (cf. the Yoruba examples in (1)). If this is combined with obligatory marking of progressive aspect for on-going states-of-affairs in the present, we obtain a system in which at least for non-stative verbs, past time reference is in practice overtly marked to a lesser extent than present time reference. Still, of course, there would not be overt marking of present tense in any strict sense. Such marking is not infrequent, however, and diachronically often arises through the expansion of earlier progressives, as in (Eastern) Armenian, where a periphrastic construction (*gr-um em* ‘write-PART.IPFV be: PRES:1.SG (I write/am writing)’) is used as the present tense of virtually all verbs (with some exceptions such as *gite-m* ‘know-PRES.1.SG (I know)').

The generally unmarked character of presents also shows up in the fact that the range of uses is largely determined by what other members of the tense-mood-aspect system are. Thus, forms labelled “present tense” in traditional grammars are more often than not regularly used both for past and future time reference, especially if past or future tenses are lacking or not fully grammaticalized. We thus find that in European languages the present tense is used for future time reference to varying degrees, being the standard way of talking about the future e.g. in languages such as Finnish and Estonian.

Similarly, “historical presents” are a common although not a universal phenomenon. (In some languages, such as Persian or Japanese, it is impossible to tell a story in the present tense.) In addition, present tenses are commonly used with generic and habitual interpretations and other cases with less specific time reference, such as descriptions of pictures (*in this picture, St. George slays the dragon*).

The most important objections against treating presents simply as the absence of tense marking are of a practical nature. One needs a convenient label for the forms in question, and since grammatical tradition provides such a label, there is no reason not to use it.

## 8. Perfect

The term **perfect** is used here to refer to constructions with a semantics and a distribution similar to that of the (present) perfect in English. It is argued in Dahl (1985) that it is possible to establish the perfect, identified in this way, as a cross-linguistically valid notion. (In Bybee et al. 1994, the label “anterior” is used for what is called perfect here. Notice that this is a different use of this term from the one referred to above.)

The perfect is often regarded as an aspect, in spite of the fact that it is rather hard to make it fit with traditional definitions of that term. The issue, it should be noted, is of importance only in a model where a sharp delimitation is made between tense and aspect as grammatical categories.

The perfect is most often expressed by periphrastic means, either by an auxiliary construction, as in English (*I have sung*), or by a particle, as in the Yoruba example (5) (where *ti* is a perfect marker with the original meaning ‘already’).

- (5) ó *ti ka iwe na*  
he PF read book this  
‘he has read this book’

Historically, perfects develop out of a number of sources, among which the most well-known are resultatives, that is, constructions expressing that the result of some event is valid (see Art. 109). Relevant resultatives are of two types, intransitive and transitive.

Intransitive resultatives (like the English *it is gone*) give rise to so-called ‘be’-perfects, as in Finnish:

- (6) *olen luke-nut kirja-n*  
 be.PRES.1.SG read-PAST.PART book-ACC  
 'I have read the book'

The traditional label *be* or *esse* perfect is suitable for languages like Finnish where these constructions do contain a copula but are less adequate for the analogous constructions in other languages which consist of a bare participle or the like. Cf. (7) from non-standard Russian (Tommola 2000).

- (7) *My davno ne spa-mši, ne e-mši.*  
 we long NEG sleep-GER.PAST NEG eat-GER.PAST  
 'We have neither slept nor eaten for a long time.'

An alternative term would be **intransitive de-resultative perfect**.

Possessive resultatives, i.e. constructions like *I have two letters written* (that is, originally 'I have two letters and they are written'), develop into *have* or *habere* perfects (**possessive de-resultative perfects**) like the one found in English. This type of perfect is common in Europe, where it often co-exists with a *be* perfect. It is less frequent elsewhere and is rare in languages in which possessive constructions are not expressed by a transitive verb like *have*. Cf. however the dialectal Russian construction with the possessive-marking preposition *u* (Tommola 2000):

- (8) *U syn-a institut za-konč-eno.*  
 at son-GEN institute PF-finish-PART  
 'My son has higher education [has finished the institute].'

At least two other sources can be noted: morphemes meaning 'already' as in the Yoruba example (5) and verbs or adverbs meaning 'finish' or 'finished', as in Thai where *laew* is a perfect marker originating in a verb meaning 'finish':

- (9) *khaw taay laew*  
 he die finish  
 'he has died'

Opinions are divided about the proper semantic characterization of the perfect. According to one relatively widespread view, the perfect is ambiguous between a number of readings, which would be the following using the terminology of Comrie 1976:

- (a) the perfect of result (as in *I have finished my Ph.D.*),  
 (b) the experiential perfect (as in *have you ever eaten oysters?*),

- (c) the perfect of recent past (as in *Malcolm X has just been assassinated*), and  
 (d) the perfect of persistent situation (as in *I have lived here for five years*).

Other researchers prefer to treat the perfect as basically unambiguous, but opinions are again divided as to what the general meaning of the perfect would be. Surveying earlier theories of the meaning of the perfect, McCoard (1978) lists the following main alternatives: the "indefinite past" theory, the "current relevance" theory, the "extended now" theory, and the "embedded past" theory. In Reichenbach's analysis, the perfect is a tense where the event time precedes the point of reference, which coincides with the point of speech (whereas in the simple past, the point of reference coincides with the time of the event).

The category of perfect should be distinguished from perfective aspect. There is considerable terminological confusion in this area, which has its historical explanation in the non-distinctness of perfect and perfective in classical Latin.

The perfect plays an important role in a number of different grammaticalization processes. Above, the sources from which perfects may arise were mentioned; in addition, the perfect is also the starting-point or sometimes an intermediate point in different "grammaticalization paths". Most well-known is probably the development by which an original perfect develops into a general past tense, as in Southern German, Yiddish, and Afrikaans, or sometimes rather a past perfective, as the *passé composé* (*j'ai mangé* 'I ate/I have eaten') in spoken French. Intermediary stages of this development are common, as in many Germanic dialects. By another path, a perfect may develop into a **recent** or **hodiernal past** (e.g. in some Romance languages (Spanish and Occitan) and Bantu languages (Zulu, Sotho)). Last but not least, perfects often acquire inferential uses. It is not always clear if these uses develop directly from the resultatives which are the historical sources of the perfects in question or if they should be seen as extensions from the perfects themselves. Whichever is the case, such developments tend to cluster areally, the most well-known 'hotbed' being the Black Sea region (e.g. Persian, Kurdish, Georgian).

## 9. Past and future perfects

In many languages, the perfect partakes in the formation of other tense-aspect forms with a more complex marking. Thus, in Eng-

lish, we find the **pluperfect** or **past perfect** (*I had sung*) and the **future perfect** (*I shall have sung*). Whether such forms should be seen as tense-aspect categories in their own right or merely as “tenses” of the perfect is an open question. While the processes giving rise to forms like the pluperfect seem quite regular, there is a clear tendency for them to develop non-compositional semantic properties. Thus, according to Reichenbachian semantics, a pluperfect ought to be used when the event point precedes the reference point and the reference point precedes the speech point. Although this certainly accounts for the central cases of pluperfects in most languages, one may also find cases where pluperfects are used e.g. as remote pasts, or to refer ‘absolutely’, to events that took place within a closed past interval, as in Modern Greek:

- (10) *O papús mu ixe pandrefí*  
 DEF grandfather my had married  
*téseris forés.*  
 four times  
 ‘My grandfather married four times.’  
 (Hedin 1987: 22)

It should also be added that there are cases of pluperfects that do not synchronically form a paradigm with a (present) perfect, such as the Romanian *dormise* ‘had slept’, an inflectional form which goes back to the Latin pluperfect, whereas the perfect in Romanian is formed peripherastically (*a dormit* ‘has slept’).

## 10. Experiential

**Experiential constructions**, which are usually periphrastic, indicate that a state-of-affairs of a certain type took place at least once during a certain period, usually ending in the reference point. An example would be the Japanese *koto ga aru* construction:

- (11) *watashi no ani ni atta*  
 I GEN brother with meet:PAST  
*koto ga aru no?*  
 thing SUBJ exist INT  
 ‘Have you (ever) met my brother?’

The meaning of experentials thus overlaps with that of perfects. As formally differentiated constructions they are probably less common than the latter. They seem to be more frequent in certain geographical areas (such as East Asia).

## 11. Narrative

Some languages have special verb forms or markers for the non-initial elements in a sequence of events. The term **narrative** for these is motivated by the fact that they typically show up in narrative texts. In the following Swahili example the first verb is in an ordinary past tense, the second is narrative, indicating that the second event took place after the first:

- (12) *Tu-li-kwenda mji-ni,*  
 we-PAST-go town-LOC  
*tu-ka-mw-oná rafiki y-etu.*  
 we-NARR-him-see friend CL-our  
 ‘We went to town and saw our friend.’

Narrative verb forms are characteristic of but by no means restricted to Niger-Congo languages.

## 12. Remoteness

The relevance for tense-aspect systems of **remoteness** distinctions, that is, how far two points in time are from each other, has only lately been the object of systematic study. Yet, such distinctions show up in unrelated languages all over the world.

In many cases, the connection between a certain degree of remoteness and the use of a certain tense-aspect form is rather vague and implicature-like. The perfect in English, for instance, may sometimes carry an implicature of recency, as when one says *I have been ill*. In some languages, on the other hand, the degree of remoteness implied by a tense-aspect form may be quite precise. In systems of the latter kind, there are one or more ‘cut-off points’ determining the choice between different forms.

In most if not all systems with “objective” remoteness distinctions, the major distinction is between what happened or will happen today (that is, the day when the utterance is made) and everything else. There may also be further cut-off points, often less exact than the first one. Remoteness distinctions are more common in the past than in the future. They also show up less often in narrative texts.

A typical system of past time marking might involve a ‘today’ past (**hodiernal past**), a ‘yesterday’ past (**hesternal past**), and a **remote past**. Examples from Nkora-Kiga (also called Runyankore; Morris & Kirwan 1972):

- (13) (a) hodiernal (“very near past”)  
*n-aa-teera*  
 1.SG-HODPST-beat  
 ‘I beat (today)’
- (b) hesternal (“near past”)  
*n-teire*  
 1.SG-beat(MODIFIED.STEM).HEST  
 ‘I beat (yesterday)’
- (c) remote (“far past”)  
*n-ka-teera*  
 1.SG-REMPST-beat  
 ‘I beat (last week)’

There is sometimes a difference in markedness in such a way that more remote forms are more marked. In the simplest case, there is just one type of past marking which is restricted to less recent events (e.g. Luo (Nilo-Saharan)).

Diachronically, marking of remoteness distinctions may arise from the grammaticalization of time adverbs. Recent or hodiernal pasts may develop out of perfects; various auxiliary constructions are also reported, as when verbs meaning *spend the night* give rise to hesternal pasts.

### 13. Tense and subordination

The question of tense marking in subordinate constructions is a complex matter, which is hard to generalize about in view of the syntactic heterogeneity of such constructions cross-linguistically. By ‘subordinate construction’, we here mean any predicative phrase which is not the main predicate of a main clause. This would thus include both e.g. finite subordinate clauses and participial, infinitival, and converbal constructions of various kinds.

Tense in main clauses is normally anchored in the point of speech – in traditional terminology, **absolute tense**. The temporal interpretation of subordinate constructions may also relate to time points provided in the superordinate syntactic context – **relative tense**. In the simplest case, relative tense relates the reference point of a subordinate predicate to the event point of the matrix predicate. Cf. the following two English sentences:

- (14) (a) *I believed him to be honest.*  
 (b) *I believed him to have been honest.*

In (14 a), the reference point of the embedded predication is equal to the event point of the

main clause, that is, I believed that he was honest at the time of the belief. In (14 b), the embedded reference point is seen as preceding the event point of the main clause, that is, I believed that he was honest at some earlier time. Here, then, the simple infinitive *to be honest* expresses simultaneity or if we like, relative present, whereas the perfect infinitive *to have been honest* expresses anteriority or relative past. A relative future is of course logically possible but seems to be only marginally expressible in this construction in English (*I believed him to be going to be honest*).

Relative tense, as illustrated here, seems cross-linguistically to be the normal case for non-finite subordinate constructions, although cases of absolute tense do occur (e.g. imperfective past participles in Russian).

In finite subordinate clauses, where the morphology often is the same as in main clauses, absolute tense is more common. Thus, tense in English relative clauses, like in many other languages, is straightforwardly absolute:

- (15) *Yesterday, I met the man who is here*  
 (= who is here at the moment of speech).

Tense in adverbial clauses often deviates from that of main clauses (and consequently, that of relative clauses). Thus, in English and many other languages, future time reference is normally not marked in conditional and temporal clauses – rather, the present is used:

- (16) *If he comes, we will have dinner early.*

A widespread phenomenon is for perfective forms, used in main clauses with past time reference, to be used with future time reference, as in Arabic.

Finite clauses expressing indirect speech and the content of propositional attitudes (*know*, *believe*, etc.) constitute a special case. In many languages, such clauses normally have relative tense. Another way of putting this is to say that indirect speech does not differ from direct quotations in tense interpretation. Cf. the following Russian examples:

- (17) (a) *On skaza-l: “Ja bolen”.*  
 he say-PAST I ill  
 ‘He said: I am ill.’
- (b) *On skaza-l, čto on bolen.*  
 he say-PAST that he ill  
 ‘lit. He said that he is ill.’

These two sentences are two ways of reporting the same utterance, *ja bolen* ‘I am ill’. In both cases, the present tense is used, relating the time of the illness to the time when the reported utterance was made.

However, some languages, notably most West and North European ones, display a phenomenon traditionally referred to as **sequence-of-tenses**. In the idiomatic English translation of (18) the past tense is used in the subordinate clause:

- (18) *He said that he was ill.*

The treatment of sequence-of-tenses is not uncontroversial. Two basic explanations of the past tense in (18) have been proposed: i) the subordinate clause has absolute tense, i.e. the past tense is interpreted relative to the moment of speech; ii) there is a syntactic rule switching an underlying present tense into a past tense. The first hypothesis has the advantage of being the simplest one but cannot account for cases like *he said that he would come tomorrow* where a verb in the past (or rather, future in the past) is used although the time of reference is not past relative to the moment of speech. The second hypothesis, then, would appear to be the correct one; it has to be modified, however, to account for the use of non-shifted present tenses in sentences such as (19).

- (19) *The Egyptians did not realize that the earth is round.*

In this and many similar cases, the motivation for not applying the sequence-of-tense principle is that the truth-value of the propositional content of the that-clause does not depend on whether we evaluate it from the time-perspective of the Egyptians or from our own time-perspective.

#### 14. Interaction between tense, negation, and expectation markings

There is sometimes interaction between tense and negation, notably when the latter is marked inflectionally. Thus, tense distinctions may be neutralized in negated clauses, as in at least one of the negation constructions in Tamil. We also find the inverse situation, viz. that there are negated tenses that do not have affirmative counterparts. Consider e.g. the following examples from Swedish as spoken in the province of Västerbotten (in standard orthography):

- (20) (a) *jag har läs-t bok-en*  
I have read-SUP book-DEF  
'I have read the book'  
(b) *jag har inte läs-t bok-en*  
I have not read-SUP book-DEF  
'I have not read the book'  
(c) *jag har o-läs-t bok-en*  
I have un-read-SUP book-DEF  
'I have not yet read the book (lit. I have unread the book)'

The affirmative perfect in (20 a) has two negated counterparts, (20 b) and (20 c). In (20 b), we find the standard negation *inte*. In (20 c), the negation marker is the prefix *o-* (corresponding to English *un-*) on the participle. Semantically, (20 b) and (20 c) differ in that (20 c) suggests that I may yet read the book, or that I have to read the book sooner or later – nuances expressed in English by particles such as *yet* and *still*. Similar phenomena are found e.g. in many Bantu languages (see Contini-Morava 1989 for a discussion of the Swahili system). Apparently, however, also affirmative tenses may carry similar nuances of “expectation”, as the so-called ‘already’ tense (*n-aa-teire* ‘1.SG-HODPST-beat(MODIFIED.STEM) (I have already beaten)’) and ‘still’ tense (*a-ki-rwaire* ‘3.SG-still-ill(MODIFIED.STEM) (he is still ill)’) in Nkore-Kiga (cf. (13)). Exactly how such forms relate to perfects and other tenses is a matter that demands further study.

#### 15. References

- Bickerton, Derek (1981), *Roots of Language*. Ann Arbor: Karoma
- Binnick, Robert I. (1991), *Time and the Verb: A Guide to Tense and Aspect*. Chicago: The University of Chicago Press
- Bybee, Joan L. & Perkins, Revere & Pagliuca, William (1994), *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago Press
- Comrie, Bernard (1976), *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press
- Comrie, Bernard (1985), *Tense*. Cambridge: Cambridge University Press
- Contini-Morava, Ellen (1989), *Discourse Pragmatics and Semantic Categorization: The Case of Negation and Tense-Aspect with Special Reference to Swahili*. Berlin: Mouton de Gruyter

- Dahl, Östen (1985), *Tense and Aspect Systems*. Oxford: Blackwell
- Dahl, Östen (2000, ed.), *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter
- Fleischman, Suzanne (1983), "From Pragmatics to Grammar: Diachronic Reflections on Complex Pasts and Futures in Romance". *Lingua* 60, 183–214
- Hedin, Eva (1987), *On the Use of the Perfect and the Pluperfect in Modern Greek*. Stockholm: Almqvist & Wiksell (Acta Universitatis Stockholmiensis: *Studia Graeca Stockholmiensia VI*)
- Klein, Wolfgang (1994), *Time in Language*. London, New York: Routledge
- Leinonen, Marja & Vilkuna, Maria (2000), "Past Tenses in Permian Languages". In: Dahl (ed.), 495–514
- McCoard, Robert W. (1978), *The English Perfect: Tense-Choice and Pragmatic Inferences*. Amsterdam: North-Holland
- Montague, Richard (1974), *Formal Philosophy*, ed. by Richmond H. Thomason. New Haven/CN: Yale University Press
- Morris, Henry Francis & Kirwan, Brian Edmond Renshaw (1972), *A Runyankore Grammar*. Nairobi: East African Literature Bureau
- Prior, Arthur (1967), *Past, Present, and Future*. Oxford: Clarendon Press
- Reichenbach, Hans (1947), *Elements of Symbolic Logic*. New York: Collier-Macmillan
- Tommola, Hannu (2000), "On the Perfect in North Slavic". In: Dahl (ed.), 441–478
- Östen Dahl, Stockholm (Sweden)*

## 111. Illocution, mood, and modality

1. Introduction
2. Illocution
3. Modality
4. Mood
5. Uncommon abbreviations
6. References

### 1. Introduction

The term **mood** is used in language descriptions for the morphological category that covers the grammatical reflections of a large semantic area. Although the term is applied surprisingly consistently across language descriptions, attempts at defining this semantic area in positive terms have never been entirely successful, in the sense that all definitions proposed leave certain distinctions unaccounted for. While a positive definition thus seems to require a disjunctive formulation, in negative terms the morphological category of mood may be said to comprise all grammatical elements operating on a situation/proposition that are not directly concerned with situating an event in the actual world, as conceived by the speaker. In this respect mood differs crucially from tense, aspect, and negation, which do have this situating function (Art. 110, 109, 113, respectively).

The large semantic area covered by this negative definition can be subdivided into

two smaller ones: the first concerns the area of **illocution**, the second the area of **modality**. This subdivision is warranted on semantic grounds: the category of illocution is concerned with identifying sentences as instances of specific types of speech act, whereas the category of modality is concerned with the modification of the content of speech acts. But apart from these semantic differences, there are also formal reasons to distinguish between the two areas. As will be shown in 4.1, in the expression of illocution the morphological category of mood has to compete with word order and intonation as markers of particular subdistinctions, whereas modality is expressed by mood markers only.

In 2 and 3 below the semantic categories of illocution and modality will be discussed separately. In 4 an inventory of the different ways of expressing illocutionary and modal distinctions is given, and the distribution of these expression formats across the various subdistinctions is specified.

### 2. Illocution

#### 2.1. Basic illocution

The **basic illocution** of a sentence can be defined as the conversational use conventionally associated with the formal properties of

that sentence (cf. Sadock & Zwicky 1985: 155), which together constitute a **sentence type**. Apart from word order and intonation, these formal properties may include specific mood morphemes, which may in these cases be interpreted as the morphological markers of basic illocutions. By their very nature, basic illocutions are restricted to independent sentences and quotations. This feature will be of help in distinguishing illocutionary from modal categories.

The most frequently attested basic illocutions are **declarative**, **interrogative**, and **imperative**. These are illustrated in the following examples from Tauya (MacDonald 1990: 209–212):

- (1) *Ya-ni tei-mene-amu-?a.*  
1.SG-ERG catch-STAT-1.SG.FUT-DECL  
'I will have it.'
- (2) *Nen-ni sen-yau-i-nae?*  
3.PL-ERG 1.PL-see-3.PL-INT  
'Did they see us?'
- (3) *Ni-a-e!*  
eat-2.SG.FUT-IMP  
'Eat!'

The declarative sentence in (1) is conventionally associated with an assertion, the interrogative in (2) with a question, and the imperative in (3) with a command.

Apart from these most frequently attested basic illocutions there are several others that occur with some frequency (cf. Sadock & Zwicky 1985). Among these are **prohibitive**, **hortative**, and **optative**, conventionally associated with prohibitions, exhortations, and wishes. All three may be found in Tauya (MacDonald 1990: 212 f.):

- (4) *Yate-?atene!*  
go-PROH.SG  
'Don't go!'
- (5) *Saniya te-amu-ne.*  
work get-1.SG.FUT-HORT  
'I must work.'
- (6) *?ei mene-?e-no!*  
there stay-3.SG.FUT-OPT  
'Let her be there!'

Two other basic illocutions that are worth mentioning are **imprecative** and **admonitive**, conventionally associated with curses and warnings, respectively. The following examples from Turkish (Lewis 1967: 115) and Mandarin Chinese (Li & Thompson 1981: 311) illustrate these two types:

- (7) *Geber-esi!*  
die.like.a.dog-IMPR.3.SG  
'May he die like a dog!'
- (8) *Xiāoxīn ou!*  
careful ADM  
'Be careful, OK?'

In languages not making all the distinctions listed here various groupings of basic illocutions may occur. The prohibitive may simply be a negative imperative; imperative, optative, and hortative may be combined on the basis of their shared impositive nature; optative, imprecative, and admonitive may be combined on the basis of their expressive nature; etc.

In 1 a distinction was made between illocution and modality as two basic categories that may be expressed through mood markers. The importance of this distinction can now be illustrated by comparing some basic illocutions with corresponding modalities. First note that declarative, as defined above, is not the same as **indicative**. The latter is a mood category with a wide range of applications, whereas the former is an illocutionary category. This difference is reflected in the restriction that declarative forms are used in main clauses and quotations only, whereas indicative forms may be used in varying sets of subordinate clauses (cf. Bybee 1985: 170). Declarative markers may furthermore freely combine with modal markers that would be in conflict with an indicative marker, as in the following example, again from Tauya (MacDonald 1990: 209):

- (9) *?ei-ra mene-a-rafo-?a.*  
there-TOP stay-3.SG-DUB-DECL  
'Maybe he's there.'

MacDonald (1990: 209) adduces this example as problematic for her analysis of -?a as an indicative marker. It is, however, unproblematic to have a combination of a modal and illocutionary marker: the former indicates the propositional attitude of the speaker (in (9) his less than full commitment to the truth of the proposition), the latter his communicative intention (in (9) his intention to provide the addressee with a certain piece of information).

Similarly, basic illocutions such as interrogative, imperative, and optative should be distinguished from modalities such as dubitative, necessitive, and volitive, respectively. To give one more illustration, consider the difference between interrogative and dubitative.

The former has been defined above as a basic illocution, whereas the latter will be presented below as a modality. The basic difference between the two is that sentences with interrogative basic illocution constitute questions, whereas sentences which contain a dubitative modality report doubt. Thus, a speaker may execute an assertive speech act using a declarative sentence, within which he presents his doubts, rather than execute a question as such. This difference is illustrated by the Tauya examples (2) and (9) given earlier. (2) is an interrogative sentence, (9) a declarative sentence which contains a dubitative marker.

## 2.2. Illocutionary modification

Basic illocutions may be further modified by markers of what I here call **illocutionary modification**. Like basic illocution, illocutionary modification should be interpreted in terms of the conversational use of sentences. But unlike basic illocution, markers of illocutionary modification do not identify sentences as speech acts of certain types, but rather mark much more general communicative strategies on the part of the speaker: they reinforce or mitigate the force of the speech act (cf. Haverkate 1979: 81–87; Hengeveld 1989: 140 f.). Strategies of illocutionary modification typically apply to sentences with different basic illocutions, and it is this property that makes it necessary to distinguish them from basic illocutions. Consider the following examples from Babungo (Schaub 1985: 119):

- (10) *Mà yé Làmbí mɔɔ!*  
1.SG see:PF:IND Lambi EMPH  
'I have seen Lambi!'
- (11) *Jwí mɔɔ!*  
come:IMP EMPH  
'Come now!'

The emphatic particle *mɔɔ* is used in Babungo to turn both assertions (10) and commands (11) into more insistent speech acts. It thus represents a more general communicative strategy than the one embodied by the indicative and imperative verb forms. This strategy may be called **reinforcement**.

The reinforcing strategy illustrated for Babungo in (10)–(11) may be contrasted with the mitigating strategy illustrated for Mandarin Chinese (Li & Thompson 1981: 316, 313, 315) in (12)–(14):

- (12) *Wō bìng méi zuò-cuò a.*  
1.SG on.the.contrary NEG do-wrong MIT  
'On the contrary, I didn't do wrong.'
- (13) *Nǐ xiǎng bu xiǎng tā a?*  
2.SG think NEG think 3.SG MIT  
'Do you miss her/him?'
- (14) *Chī-fàn a!*  
eat-food MIT  
'Eat food, OK?!'

The particle *a* (*ya* in some dialects) may be added to sentences representing assertions (12), questions (13), and orders (14). In each case it has the same function of reducing the forcefulness of the utterance. Thus again it embodies a more general communicative strategy than that of basic illocutions, which in this case may be called **mitigation**.

Illocutionary modification is not a category that has acquired an established position in language descriptions. Reinforcing means may often be found under sentence emphasis, mitigating means under a variety of labels, many of which will contain some reference to their polite nature. A frequently used term is *downtoner*.

## 3. Modality

### 3.1. Classifying parameters

In classifying modal categories two parameters have to be distinguished. The first concerns the target of evaluation of a modal distinction. It is on the basis of this parameter that a distinction can be drawn between e.g. objective and subjective modality. The second concerns the domain of evaluation of a modal distinction. It is on the basis of this parameter that a distinction is drawn between e.g. epistemic, deontic, and volitive modality. Although distinctions pertaining to both of these parameters are present in most treatments of modality, they are often not strictly kept apart.

#### 3.1.1. Target of evaluation

By the **target of evaluation** of a modal distinction is meant the part of the utterance that is modalized. Along this parameter the following types of modality can be distinguished (see Jakobson 1957; Lyons 1977; Foley & Van Valin 1984; Hengeveld 1988; 1989):

- (a) **Participant-oriented modality.** This type of modality affects the relational part of the utterance as expressed by a predicate

- and concerns the relation between (properties of) a participant in an event and the potential realization of that event (cf. Foley & Van Valin 1984: 215).
- (b) **Event-oriented modality.** This type of modality affects the event description contained within the utterance, i.e. the descriptive part of an utterance, and concerns the objective assessment of the actuality status of the event.
- (c) **Proposition-oriented modality.** This type of modality affects the propositional content of an utterance, i.e. the part of the utterance representing the speaker's views and beliefs, and concerns the specification of the degree of commitment of the speaker towards the proposition he is presenting.

The following example from Turkish (Lewis 1967: 151), containing all three types of modality, may serve as a first illustration of the differences between them:

- (15) *Anlı-y-abil-ecek-mış-im.*  
 understand-Ø-ABIL-IRR-INFR-1.SG  
 'I gather that I will be able to understand.'

In this example the ability suffix *-abil* (preceded by an obligatory intervocalic *-y-*) expresses a participant-oriented modality. The first singular subject is said to have the capacity of engaging in the relation expressed by the predicate. The irrealis suffix *-ecek* expresses an event-oriented modality. The event described by the sentence is characterized as non-actual, which is in this case, but not necessarily, reflected in the translation by means of a future tense. The inferential suffix *-mış* expresses a proposition-oriented modality. It signals that the speaker does not fully commit himself to the propositional content of his assertion.

The term "modality" has been restricted in various ways to cover only part of the categories of modality recognized here. Thus, Foley & Van Valin (1984: 213–220) restrict the term to participant-oriented modality, reserving the labels "status" and "evidentiality" for event-oriented and proposition-oriented modality, respectively. Halliday (1970: 336) takes the opposite position, restricting the term "modality" to proposition-oriented modality and using "modulation" for the remaining categories.

### 3.1.2. Domain of evaluation

By the **domain of evaluation** of a modal distinction is meant the perspective from which the evaluation is executed. By varying this perspective the following types of modality may be distinguished:

- (a) **Facultative modality** is concerned with intrinsic or acquired capacities.
- (b) **Deontic modality** is concerned with what is (legally, socially, morally) permissible.
- (c) **Volitive modality** is concerned with what is desirable.
- (d) **Epistemic modality** is concerned with what is known about the actual world.
- (e) **Evidential modality** is concerned with the source of the information contained in a sentence.

Extensive exemplification of all these types will be given below. A first illustration is given in the following English examples:

- (16) *John is able to swim.*  
 (Ability: Facultative)
- (17) *John has to swim.*  
 (Obligation: Deontic)
- (18) *John would rather not swim.*  
 (Wanting: Volitive)
- (19) *John may be swimming.*  
 (Possibility: Epistemic)
- (20) *John will be swimming.*  
 (Inference: Evidential)

### 3.1.3. Synthesis

The combination of three targets of evaluation with five domains of evaluation leads to 15 possible combinations of features of modality types. Some of these are logically excluded, however. To give a simple example, it is impossible to evaluate propositions in terms of their intrinsic or acquired capacities. The logically permitted combinations are listed in Table 111.1:

	Target	Participant	Event	Proposition
Domain \				
Facultative	+	+	–	
Deontic	+	+	–	
Volitive	+	+	+	
Epistemic	–	+	+	
Evidential	–	–	+	

Tab. 111.1: Cross-classification of modality types

The following sections describe the ten remaining subcategories of modality identified in Table 111.1, using the target of evaluation as the primary classificatory parameter and the domain of evaluation as the secondary.

### 3.2. Participant-oriented modality

Participant-oriented modalities are better known from the literature as agent-oriented modalities. Although widely used, this term is not too felicitous in that it suggests that only controlling participants in dynamic events may be subject to this type of modalization. That this is not the case is apparent from such examples as:

- (21) *John wants to be young again.*

The term *participant-oriented* modality is neutral as to the event type in which this class of modal expressions occurs. Three main subcategories of participant-oriented modality may be distinguished on the basis of the domain of evaluation they are concerned with.

#### 3.2.1. Facultative

Facultative participant-oriented modality describes the ability of a participant to engage in the event type designated by the predicate. In some languages a distinction is made between **intrinsic** ('be able to') and **acquired** ('know how to') **ability**, as shown in the following examples from Mapuche, which has separate auxiliaries for these two types of ability (Smeets 1989: 219):

- (22) *Pepí kuθaw-la-n.*  
INTR.ABIL work-NEG-DECL.1.SG  
'I am not able to work.'
- (23) *Kim tuku-fi-n.*  
ACQ.ABIL put.at-OBJ-DECL.1.SG  
'I know how to put it.'

Spanish makes the same distinction. Intrinsic ability is expressed by the modal verb *poder* 'be able to', acquired ability by the verb *saber* 'know (how to)' in its modal use.

Inability may also acquire the status of a separate category, as in the Turkish Impotential (25), which may be compared with its Potential (24), used for ability (Lewis 1967: 151):

- (24) *Gel-ebil-di-o.*  
come-ABIL-PAST-3  
'He was able to come.'
- (25) *Gel-eme-di-o.*  
come-INAB-PAST-3  
'He was unable to come.'

#### 3.2.2. Deontic

Deontic participant-oriented modalities describe a participant's being under the obligation or having permission to engage in the event type designated by the predicate. Obligation seems to be encoded by grammatical means more often than permission. Terms used in different grammatical traditions for verb forms expressing obligation are "obligative" and "necessitative". The following example is from Quechua (Cole 1982: 151):

- (26) *Miku-na ka-rka-ni.*  
eat-OBLG COP-PAST-1  
'I must eat.' (lit. 'I am to eat.')

#### 3.2.3. Volitive

Volitive participant-oriented modality describes a participant's desire to engage in the event-type designated by the predicate. The following example is from Guajajara (Bendor-Samuel 1972: 95):

- (27) *Za-hem rəm.*  
1.PI-leave VOL  
'We want to leave.'

#### 3.3. Event-oriented modality

Event-oriented modalities occupy a position in between participant-oriented modalities and proposition-oriented modalities. They are like participant-oriented modalities in that they form part of the descriptive content of the sentence. They are like proposition-oriented modalities in that the source of modalization is not a participant in the event described within the sentence.

Event-oriented modalities describe the existence of possibilities, general obligations, and the like, without the speaker taking responsibility for these judgements. This is best illustrated by means of the following sentence, which contains both a proposition-oriented and an event-oriented modal expression (Lyons 1977: 808):

- (28) *Certainly he may have forgotten.*

Through the epistemic proposition-oriented modal adverb *certainly* the speaker commits himself to the truth of the proposition *he may have forgotten*, which contains the epistemic event-oriented modal verb *may* that describes the existence of the possibility of the occurrence of the event *he has forgotten*. Although the two epistemic judgements contained in (28) are non-harmonic (Lyons 1977; Coates 1983; Bybee et al. 1994), no contradiction

arises, since the two judgements pertain to different levels: the speaker expresses his certainty about the existence of an objective possibility. For this reason epistemic proposition-oriented modality has been called “subjective” and event-oriented modality “objective” (Lyons 1977: 797–804; cf. also Halliday 1970; Coates 1983). Objective status may, apart from epistemic modality, also be assigned to facultative, deontic, and volitive varieties.

### 3.3.1. Facultative

Facultative event-oriented modality characterizes events in terms of the physical or circumstantial enabling conditions on their occurrence (Bybee et al. 1994; Olbertz 1998). This type of modality is often referred to as root modality (Coates 1983). Examples are the following:

- (29) *It can take three hours to get there.*
- (30) *I couldn't finish reading the book because it got too dark.*

In contrast to facultative participant-oriented modality, the possibility of occurrence of the event does not depend on the intrinsic capacities of a participant, but follows from the circumstances in which the event takes place. This sense can most easily be detected in impersonal constructions such as (29).

### 3.3.2. Deontic

Deontic event-oriented modality characterizes events in terms of what is obligatory or permitted within some system of moral or legal conventions (cf. Allwood et al. 1977: 111). In contrast to deontic participant-oriented modality, the obligations expressed by means of deontic event-oriented modality do not rest upon a particular participant, but represent general rules of conduct. This sense of general applicability can most clearly be identified in impersonal expressions such as the Turkish modal periphrases illustrated in (31) and (32) (van Schaik 1985):

- (31) *Bura-da ayakkabıları çıkar-mak var.*  
DEM-LOC shoes take.off-INF EXIST  
'One has to take off his shoes here.' (lit.  
'There is taking off of shoes here.)
- (32) *Avuç aç-mak yok.*  
hand open-INF EXIST.NEG  
'Begging prohibited.' (lit. 'There isn't  
begging.)

But the sense of general obligation may be present in personal constructions as well, as in (33) (cf. Coates 1983: 73):

- (33) *We ought to have a right to intervene.*

### 3.3.3. Volitive

Volitive event-oriented modality characterizes events in terms of what is generally desirable or undesirable. This category seems hardly ever to be encoded by specialized markers, but rather to group with deontic modality. An exception to this, however, is the Tauya avolitional, which “[...] implies that the action or state specified by the verb would be undesirable” (MacDonald 1990: 202 f.):

- (34) *Tepau-fe-?ate-e-?a.*  
break-TR-AVOL-1-DECL  
'It would be bad if I broke it.'

### 3.3.4. Epistemic

Epistemic event-oriented modality characterizes events in terms of the (im)possibility of their occurrence in view of what is known about the world. Although many different shades of meaning could be defined within this domain, grammatical encoding of this type of modality is generally restricted to a realis versus irrealis (or potentialis) opposition. An example of this type of opposition may be found in Mapuche (Smeets 1989: 307):

- (35) *Trür amu-a-y-u*  
together go-IRR-DECL-1.DU.SBJ  
*üyüw.*  
over.there  
'Together we will go over there.'

- (36) *Trür amu-o-y-u*  
together go-RLS-DECL-1.DU.SBJ  
*üyüw.*  
over.there  
'Together we went over there.'

In spite of the translation the Mapuche irrealis cannot be interpreted as a future tense morpheme, since it has a whole range of additional shades of modal meaning, including probability.

The opposition between realis and irrealis is sometimes further obscured by the fact that the realis domain is occupied by certain tenses, as a result of which the modal category irrealis stands in opposition to the temporal categories past and present. This is, for instance, the case in Ngiyambaa, where there

is “a three-way tense system, involving two contrasts, one of actuality (actualis versus irrealis) and, within the actualis category, one of time (past versus present)” (Donaldson 1980: 160). Again, the category of irrealis cannot be interpreted as a simple future tense, since it is also used for stating (objective) probabilities, as in:

- (37) *Yuruj-gu yidjal-agá.*  
rain-ERG rain-IRR  
'It may rain.' or 'It will rain.'

In order to avoid such ambiguities some languages make a distinction between a “certain future” and an “uncertain future”, where the latter might perhaps better be interpreted as an irrealis form. The following examples are from Garo (Burling 1961: 27 f.):

- (38) *Aya re'-ay-gen.*  
1.SG move-DIR-FUT  
'I will go.'
- (39) *Re'-ba-nabadoja.*  
move-DIR-IRR  
'He may come.'

Garo furthermore has an intentional future and two negative futures.

### 3.4. Proposition-oriented modality

As stated and illustrated above, proposition-oriented modalities specify the subjective attitude of the speaker towards the proposition he is presenting. The speaker may characterize the proposition as his personal wish (volitive modality), express several degrees of commitment with respect to the proposition (epistemic modality), or specify the source of the proposition (evidential modality).

#### 3.4.1. Volitive

Volitive proposition-oriented modality differs from its participant-oriented counterpart in that the source of the volitional attitude is the speaker, and not a participant in the event described within the sentence. In Pawnee (Parks 1976: 162) a special formation, in which the verb inflected passively is provided with ‘perfect intentive aspect’ suffixes, expresses volitive proposition-oriented modality:

- (40) *Ti-ku-itka-is-ta.*  
IND-1.SG.OBJ-sleep-PF-INTV  
'I want to sleep.' (lit. 'It is going to sleep on me.')

Note that the indicative mood morpheme *ti-* shows that this sentence cannot be interpre-

ted as having optative basic illocution (see 2.1), i.e. it is not a wish but an assertion concerning the speaker's wishes.

#### 3.4.2. Epistemic

In the introduction to 3.3 the distinction between objective and subjective epistemic modality was explained and it was shown that objective epistemic modality is event-oriented, whereas subjective epistemic modality is proposition-oriented. Ngiyambaa provides a further illustration of this distinction (Donaldson 1980: 256):

- (41) *Gali:-yjinda-gila nyianu baluy-agá.*  
water-PRIV-DUB 1.PL.NOM die-IRR  
'We'll probably die for lack of water.'

Apart from the irrealis marker discussed in 3.3.4, Ngiyambaa has a special marker for dubitative modality. Both may occur in a single sentence, as illustrated in (41), which may be paraphrased as 'I guess (DUB) the unrealized (IRR) event of our dying for lack of water will take place'. Thus, the dubitative gives the speaker's subjective assessment of a proposition containing an objective specification of the unrealized status of an event.

Just as objective epistemic modality groups with tense (3.3), so does subjective epistemic modality group with evidentials (3.4.3). This can be explained as a result of the fact that both tense and objective modality are event-oriented, while subjective modality and evidentiality are proposition-oriented. Ngiyambaa illustrates this grouping. In this language irrealis modality is expressed by means of a verb-suffix that is mutually exclusive with tense suffixes (see 3.3.4). Dubitative modality, on the other hand, is expressed by means of a particle that cliticizes to the first constituent in the sentence, in exactly the same way as evidential modalities. In Garo (Burling 1961; see also Bybee 1985: 180 f.) the uncertain future illustrated in 3.3.4 is a verb suffix that may immediately follow the verb stem, just like true tense suffixes, whereas the dubitative occupies the final position of the suffix string, just like evidential suffixes.

The most important subdistinctions to be made within the category of epistemic proposition-oriented modality are **doxastic**, **dubitative**, and **hypothetical**. Through a doxastic modality the speaker indicates that he believes the proposition he is presenting to be true. Since this is the usual assumption underlying assertions, this modality type is least frequently expressed by grammatical means.

The following example is from Hidatsa (Matthews 1964), where the sentence final particle *c* indicates that the speaker has reasonable grounds to believe that the proposition he is presenting is true:

- (42) *Wio i hirawe ki ksa c.*  
 woman 3.SG sleep INGR ITER DOX  
 'The woman fell asleep again and again.'

Through the much more frequently marked dubitative modality the speaker indicates that he has some doubts about the truth of the proposition he is presenting. The Ngiyambaa example (41) above illustrates this case. A second example comes from Mapuche (Smeets 1989: 431):

- (43) *Amu-y chi.*  
 go-DECL.3 DUB  
 'Maybe he went away.'

Through a hypothetical modality the speaker indicates absence of commitment (either positive or negative) with respect to the proposition he is presenting. In the following English examples this modality type is expressed by means of a particle that at the same time functions as a conjunction:

- (44) *if he comes, (I'll leave)*  
 (45) *if he came, (I would leave)*

Note, incidentally, that the distinction between realis and irrealis conditions, as illustrated in (44)–(45), is not a subdivision that obtains at the level of proposition-oriented modality, but at the level of event-oriented modality. Thus, in (44)–(45) the speaker indicates absence of commitment to the proposition introduced by *if*, and within that proposition he characterizes an event as real (44) or unreal (45) within the hypothesized world.

### 3.4.3. Evidential

**Evidential** proposition-oriented modality is concerned with the way the proposition the speaker is presenting came to his knowledge, i.e. it specifies the source on which the speaker relies for the information contained within his utterance. For this reason the term "epistemological modality" (Chung & Timberlake 1985) has been used for what is more generally referred to as "evidentiality" (Jakobson 1957; Chafe & Nichols 1986, eds.; Willett 1988; Aikhenvald & Dixon 2003, eds.).

The most basic grammatically encoded distinction within the domain of evidentiality is that between **sensory evidence** and **non-sensory evidence** (see Willett 1988: 57, who uses the terms "direct" and "indirect evidence", respectively). Markers of sensory evidence indicate that the speaker acquired the information he is presenting through perception, those of non-sensory evidence that he acquired it from any other source. A language making just this binary distinction (in the past tense only) is Turkish. Compare the following sentences (Lewis 1967: 128):

- (46) *Bir turist vapuru*  
 INDEF tourist ship  
*gel-di-o.*  
 come-PAST.SENS.EV-3.SG  
 'A tourist-ship arrived (I witnessed it.)'
- (47) *Bir turist vapuru*  
 INDEF tourist ship  
*gel-miṣ-o.*  
 come-PAST.NONSENS.EV-3.SG  
 'A tourist-ship arrived (I did not witness it.)'

A verb with the suffix *-miṣ*, as in (47), "conveys that the information it gives is based either on hearsay or on inference from observed facts, but not on the speaker having seen the action take place" (Lewis 1967: 122), whereas a verb with the suffix *-di*, as in (46), is used "when relating past events positively known to the speaker" (Lewis 1967: 128).

Markers of sensory evidence may be further subdivided according to the particular sensory mode through which the information was acquired (Palmer 1986: 67; Willett 1988: 57). A more fundamental subdivision, however, obtains within the domain of non-sensory evidence (Willett 1988). Within this class a distinction should be made between **reportative modality**, through which the speaker characterizes the information he is presenting as obtained through hearsay, and **inferential modality**, through which the speaker indicates that he has inferred the information he is presenting from other pieces of (non-sensory) information. The following examples from Inga (Levinsohn 1975: 15, 24; see also Palmer 1986: 52) show that this language has markers for these two types of modality next to its marker of sensory evidence:

- (48) *Iujpataca Pasto-ma-si ri.*  
 long.ago Pasto-DIR-RPRT go:3.SG  
 'Long ago someone went to Pasto (it is said).'

- (49) *Chipica diablo-char ca.*  
there devil-INFR COP:3.SG  
'A devil was presumably there.'
- (50) *Nispaca Santiago-ma-mi rini.*  
after.that Santiago-DIR-SENS.EV go.1.SG  
'After that I went to Santiago.'

Within the domain of reportative modality further distinctions may be made as to the particular source of the report, whereas in the case of inferential modality the type of information on which the inference is based may trigger further subdivisions. See Willett (1988) for an overview.

#### 4. Mood

The illocutionary and modal distinctions listed in sections 2 and 3 may be expressed by a variety of morphological markers, for which the term **mood** is commonly used. Besides mood there are non-morphological markers of illocution, such as word order and intonation, which will be included here for the sake of completeness.

The various semantic categories that have been distinguished above are often expressed differently in main and subordinate clauses. For this reason these two syntactic contexts are discussed separately below. Note that illocution is only expressed in main clauses, and therefore irrelevant to the analysis of subordinated clauses.

##### 4.1. The expression of modality and illocution in main clauses

The following strategies for the expression of modality and illocution in main clauses may be distinguished:

(a) **Word order.** The basic illocution of a sentence may be signalled by word order, as in the following Dutch examples of a declarative and interrogative sentence, respectively:

- (51) *Peter kom-t.*  
Peter come-PRES.3.SG  
'Peter comes.'
- (52) *Kom-t Peter?*  
come-PRES.3.SG Peter  
'Does Peter come?'

Note that apart from word order differences the intonation patterns of (51) and (52) are different as well.

(b) **Intonation.** As illustrated in (51)–(52), intonation may play a role in the expression

of basic illocution. Intonational strategies may furthermore be exploited to express illocutionary modification. Thus, Halliday (1970: 331) shows that in English the speaker may mitigate his statement through intonational means.

(c) **Particle.** Particles are widely used for the expression of a variety of illocutionary and modal distinctions. In most cases these particles either occupy the sentence-final position, or cliticize to the first constituent of the clause. In some cases they occupy the preverbal position. These three types are illustrated in the following examples from Hidatsa (Matthews 1964), Ngiyambaa (Donaldson 1980: 276), and Dutch, respectively:

- (53) *Wio a riiti rahe.*  
woman 3.SG hungry RPRT  
'I have been told that the woman is hungry.'
- (54) *Ijindu-dhan girambiyi.*  
2.SG-RPRT sick:PAST  
'You are said to have been sick.'
- (55) *Doe de deur even dicht!*  
do.IMP.SG the door MIT closed  
'Close the door, will you?'

(d) **Inflection.** Inflection of main predicates, mostly of verbs, is widely used to mark many of the illocutionary and modal distinctions described above. Terminology is abundant in this particular area, and there is little chance that terms are used in the same way across language descriptions. Some of the more frequent names of inflections are "indicative" (for verb forms used in clauses with declarative basic illocution), "imperative" (for the verb forms used in clauses with imperative basic illocution), "conditional" (for verb forms expressing epistemic possibility), and "counterfactual" (for verb forms expressing irrealis modality). Note that very often inflections express more than one illocutionary and/or modal value, and that the set of meanings associated with a given form may vary from language to language. The value given between brackets is the meaning which is generally included in the set of meanings of the forms mentioned.

The Turkish example (56) illustrates (Lewis 1967: 126) the inflectional strategy, here used to express event-oriented deontic modality:

- (56) *Gel-meli-ymış-im.*  
come-OBLG-NON.SENS.EV-1.SG  
'It seems I ought to come.'

category \ marker	marker	word order	intonation	particle	inflection	auxiliary	periphrasis	derivation
basic illocution	+	+	+	+	—	—	—	—
illocutionary modification	—	+	+	+	—	—	—	—
proposition oriented modality	—	—	+	+	+	—	—	—
event oriented modality	—	—	—	+	+	+	—	—
participant oriented modality	—	—	—	—	+	+	+	+

Tab. 111.2: Encoding of illocution and modality

(e) **Auxiliary.** Modal distinctions may be expressed by means of auxiliaries, as illustrated in the following examples of the auxiliaries *dū* in Babungo (Schaub 1985: 228) and *must* in English:

- (57) *Iwə nyii dī-dū.*  
3.SG run:PRES DUR-INAB  
'He is unable to run.'

- (58) *He must be home.*

(f) **Periphrastic construction.** In periphrastic constructions the modal meaning is the result of a particular configuration of elements, rather than that it can be attributed to a single element of the clause. Many modal periphrastic constructions involve some form of the verbs *have* or *be*, the latter either in its existential, locative, or copular sense (Hengeveld 1992: 257–290). The following examples are from Basque (Lafitte 1944: 221) and Quechua (Cole 1982: 151):

- (59) *Etche huntan ez da*  
house DEM-LOC NEG 3.SG.ABS-COP.PRES  
*bizitzer-ik.*  
live.INF-PRTV  
'It is impossible to live in this house.'  
(lit. 'There is no living in this house.')
- (60) *Miku-na ka-rka-ni.*  
eat-IRR COP-PAST-1  
'I must eat.' (lit. 'I was characterized by unrealized eating.')

(g) **Derivation.** Derivational means are used to a limited extent, and probably for the expression of participant-oriented modalities only. The ability and inability suffixes of Turkish, illustrated in 3.2.1 are of a derivational nature, and so is the volitional suffix in Ngiyambaa (Donaldson 1980: 115, 281). Note that the derivational strategy is here combined with a periphrastic one.

- (61) *yadhu dhingga: dhal-i-qinda ga-ra.*  
1:NOM meat:ABS eat-PURP-VOL COP-PRES  
'I want to eat meat.' (lit. 'I am in want of eating meat.')

A preliminary investigation of a sample of 20 languages (Hengeveld 1996, ed.) suggests that the various markers of modality and illocution presented above are not randomly distributed. The general tendency emerging from the data is as in Table 111.2.

As Table 111.2 shows, there appears to be a clearcut correlation between illocutionary and modal categories on the one hand, and expression type on the other. The particular ordering of illocutionary and modal categories given in Table 111.2 may be determined by the fact that the lower the category is in the table, the more directly relevant to the predicate and the less general in meaning it is (Bybee 1985). Alternatively, one might say that the higher the position in Table 111.2, the more personal (Traugott 1982) the category is.

#### 4.2. The expression of modality in subordinate clauses

In subordinate clauses the non-morphological markers listed above, word order and intonation, are not used as expressions of modality. This coincides with the fact that in subordinate clauses basic illocution and illocutionary modification can not be expressed. The remaining strategies can all be found in subordinate clauses, but in the case of particles and inflections there may be forms that are specific to subordinate clauses. These two strategies are discussed below.

Particles with a modal value may simultaneously act as conjunctions in subordinate clauses. Thus, the complementizer *if* in English signals absence of commitment on the part of the speaker, whereas *that* may signal

positive commitment, as in the following examples:

(62) *Peter didn't know if John was ill.*

(63) *Peter didn't know that John was ill.*

Note that through the complementizers the propositional attitude of the speaker, not that of the subject of the matrix clause, is expressed, i.e. the modality type involved is epistemic proposition-oriented modality (see 3.4.2). On the other hand, the complementizer *to* in English could be interpreted as an irrealis marker, i.e. the expression of epistemic event-oriented modality.

In many languages there are inflectional means, usually called ‘subordinate moods’, that are only or mainly used in subordinate clauses. These forms generally cover a wider range of modal values than inflectional forms in main clauses. The best known case of a subordinate mood is the **subjunctive** or **conjunctive** mood, which is generally used in opposition with the indicative mood, the latter also being used in main clauses. The range of modal values covered by a subjunctive mood varies from language to language. From the data in Noonan (1985: 91–103) it can be inferred that, at least for complement clauses, this variation can be described systematically. The determining factor is the modality type in terms of its target of evaluation. The distribution would be as follows:

	A	B	C
proposition oriented	—	—	SUBJ
modality			
event oriented	—	SUBJ	SUBJ
modality			
participant oriented	SUBJ	SUBJ	SUBJ
modality			

Tab. 111.3: Uses of subjunctives

In systems of type A the use of the subjunctive is restricted to complement clauses with dependent time reference, including the complement clauses of verbs lexically expressing participant-oriented modalities, such as the equivalents of *want* and *be able to*. In systems of type B the subjunctive has the additional use of expressing irrealis modality, i.e. it occurs in complement clauses which contain the description of an event the factuality of which has not been determined. In languages of type C the subjunctive is used in all cases

in which the speaker does not assert the proposition contained in the complement clause.

These facts suggest that the same parameter that may be used to describe the crosslinguistic distribution of mood markers in main clauses, may be used to describe the crosslinguistic distribution of subjunctive verb forms. Tables 111.2 and 111.3 thus lend further support to the classification of modal categories as to their target of evaluation, apart from the classification as to their domain of evaluation.

## 5. Uncommon abbreviations

ABIL	ability
ACQ.ABIL	acquired ability
AVOL	avolitional
DOX	doxastic
IMPR	imperative
INAB	impotential
INTR.ABIL	intrinsic ability
INTV	intitative
MIT	mitigation
NON.SENS.EV	non-sensory evidence
SENS.EV	sensory evidence

## 6. References

- Aikhenvald, Alexandra & Dixon, R[obert] M. W. (2003, eds.), *Studies in Evidentiality*. Amsterdam: Benjamins (Typological Studies in Language 54)
- Allwood, Jens & Andersson, Lars-Gunnar & Dahl, Östen (1977), *Logic in Linguistics*. Cambridge: Cambridge University Press
- Bendor-Samuel, David (1972), *Hierarchical Structure in Guajajara*. Norman: Summer Institute of Linguistics
- Burling, Robbins (1961), *A Garo Grammar*. Poona: Deccan College Postgraduate and Research Institute
- Bybee, Joan L. (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins (Typological Studies in Language 9)
- Bybee, Joan L. & Perkins, Revere D. & Pagliuca, William (1994), *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press
- Chafe, Wallace & Nichols, Johanna (1986, eds.), *Evidentiality: the Linguistic Coding of Epistemology*. Norwood: Ablex
- Chung, Sandra & Timberlake, Alan (1985), “Tense, Aspect, and Mood”. In: Shopen, Timothy (ed.),

- Language Typology and Syntactic Description, Vol. III.* Cambridge: Cambridge University Press, 202–258
- Coates, Jennifer (1983), *The Semantics of the Modal Auxiliaries*. London: Croom Helm
- Cole, Peter (1982), *Imbabura Quechua*. Amsterdam: North-Holland (Lingua Descriptive Studies 5)
- Donaldson, Tamsin (1980), *Ngiyambaa*. Cambridge: Cambridge University Press
- Foley, William A. & Van Valin, Robert D. (1984), *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press
- Halliday, M[ichael] A. K. (1970), “Functional Diversity in Language, as Seen from a Consideration of Modality and Mood in English”. *Foundations of Language* 6, 322–361
- Haverkate, Henk (1979), *Impositive Sentences in Spanish: Theory and Description in Linguistic Pragmatics*. Amsterdam: North-Holland
- Hengeveld, Kees (1988), “Illocution, Mood and Modality in a Functional Grammar of Spanish”. *Journal of Semantics* 6.3–4, 227–269
- Hengeveld, Kees (1989), “Layers and Operators”. *Journal of Linguistics* 25.1, 127–157
- Hengeveld, Kees (1992), *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter (Functional Grammar Series 15)
- Hengeveld, Kees (1996, ed.), *Verslag werkgroep functionale grammatica II*. Unpublished paper, University of Amsterdam
- Jakobson, Roman (1957), “Shifters, Verbal Categories, and the Russian Verb”. In: Jakobson, Roman, *Selected Writings, Vol. II*. The Hague: Mouton, 130–147
- Lafitte, Pierre (1944), *Grammaire Basque: Navarro-labourdin Littéraire*. Bayonne: Librairie “Le Livre”
- Levinsohn, Stephen H. (1975), “Functional Sentence Perspective in Inga”. *Journal of Linguistics* 11, 13–37
- Lewis, G[eoffrey] L. (1967), *Turkish Grammar*. Oxford: Oxford University Press
- Li, Charles N. & Thompson, Sandra A. (1981), *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press
- Lyons, John (1977), *Semantics, Vol. 1–2*. Cambridge: Cambridge University Press
- MacDonald, Lorna (1990), *A Grammar of Tauya*. Berlin: Mouton de Gruyter (Mouton Grammar Library 6)
- Matthews, G[eorge] H[ubert] (1964), *Hidatsa Syntax*. The Hague: Mouton
- Noonan, M[ichael] (1985), “Complementation”. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. II*. Cambridge: Cambridge University Press, 42–140
- Olbertz, H[ella] G. (1998), *Verbal Periphrases in a Functional Grammar of Spanish*. Berlin: Mouton de Gruyter (Functional Grammar Series 22)
- Palmer, F[rank] R. (1986), *Mood and Modality*. Cambridge: Cambridge University Press
- Parks, Douglas R. (1976), *A Grammar of Pawnee*. New York: Garland
- Sadock, Jerrold M. & Zwicky, Arnold M. (1985), “Speech Act Distinctions in Syntax”. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. I*. Cambridge: Cambridge University Press, 155–196
- Schaaijk, Gerjan van (1985), “Verb Based Terms and Modality in Turkish”. Unpublished paper, University of Amsterdam
- Schaub, Willi (1985), *Babungo*. London: Croom Helm
- Smeets, Ineke (1989), *A Mapuche Grammar*. Unpublished Ph.D. dissertation, University of Leiden
- Traugott, Elizabeth C. (1982), “From Propositional to Textual and Expressive Meanings: Some Semantic-Pragmatic Aspects of Grammaticalization”. In: Lehmann, W[infred] P. & Malkiel, Y[akov] (eds.), *Perspectives on Historical Linguistics*. Amsterdam: Benjamins, 245–271
- Willett, Thomas (1988), “A Crosslinguistic Survey of the Grammaticalization of Evidentiality”. *Studies in Language* 12, 51–97

*Kees Hengeveld, Amsterdam  
(The Netherlands)*

## 112. Interclausal relations

1. Introduction
2. Temporal relations
3. Switch-reference
4. Conditional
5. Cause/reason
6. Purpose
7. References

### 1. Introduction

**Interclausal relations** within a sentence are often characterized in terms of **coordination** versus **subordination**. Subordination is often taken as a hypernym for embedding (relativization, nominalization, complement clauses), for example in the Lingua Descriptive Series. In order to differentiate coordination of nominal elements from that of verbal or clausal material the term **parataxis** is sometimes used for the latter (Halliday & Hasan 1976; Lehmann 1988: 182), which is opposed to **hypotaxis**. However, for Lehmann hypotaxis is a subcategory of subordination, whereas Halliday & Hasan make a distinction between hypotaxis and embedding, avoiding the term ‘subordination’. Matthiessen & Thompson (1988: 283) contrast parataxis and hypotaxis: Parataxis includes coordination, apposition, and quoting; Hypotaxis includes clause combining, involving non-restrictive relative clauses, clauses of reported speech and clause combinations defined by various semantic relations, such as temporal, conditional, reason, concessive, purpose, means, and manner.

The label “subordination” for a grammatical category has been challenged by Haiman & Thompson (1984), who suggest that a number of factors underlie subordination as composite notion: identity of subject, tense, mood; reduction of one clause; one clause being in the scope of the other, etc. This is not the place to attempt to solve terminological confusion, since there do not seem to be morphological correlates for coordinate versus subordinate interclausal relations. For example, switch-reference systems in American Indian and Australian languages are said to operate on subordinate clauses (Oswalt 1976; 1977; Jacobsen 1983; Austin 1981), whereas for Papuan languages it is claimed that they are used on coordinate relations (Haiman 1980; Reesink 1983; Roberts 1988). However, Winter (1976: 166) suggests paratactic relations for switch-reference in Yuman lan-

guages, and Davies (1981) labels clauses with switch-reference affixation in Kobon (a Papuan language) as subordinate.

Cross-cutting the syntactic concepts are semantic notions that are used to characterize interclausal relations: temporal sequence and simultaneity, conditional, cause, reason, antithetical, contrafactual, concessive are some of the more perspicuous labels besides more idiosyncratic ones, such as paraphrase, attribution, elaboration, explanation (cf. Grimes 1975; Longacre 1976; 1985; Matthiessen & Thompson 1988). The morphological expressions for these semantic relations are by no means unequivocal in any language. Also, these concepts refer as much to intersentential relations as to interclausal ones, as illustrated by the plethora of conjunctions that can be listed for English (Halliday & Hasan 1976) and presumably for many more, if not all, languages.

As a generalization it seems safe to say that the more specific a semantic relation between clauses needs to be made, the more morphological material is used to make up such a conjunction. Simple **juxtaposition** of finite verbs (under one sentence intonation, cf. Mithun 1984: 502; Derbyshire 1985: 66; Houis 1968: 15 f.) or **asyndetic linkage** between a non-finite verb (participle, gerund, medial, etc.) and finite verb are in general unspecified with regard to the semantic relation. Two propositions are simply connected. Grimes’ term “collection” (1975: 221) may be appropriate for such a neutral relation between two clauses. When two events are expressed in close connection there is ipso facto a temporal relation implied, which may be either sequential or simultaneous. It is these two temporal relations that interact with switch-reference systems.

### 2. Temporal relations

**Temporal relations** between clauses can be divided into two main categories (cf. Art. 110): sequence and simultaneity, each of which can be subcategorized by finer distinctions, as for example in Givón (1990: 828). Juxtaposition of independent clauses can in most languages be used to express at least sequence in time, with Latin *veni vidi vici* ‘I came, I saw, I conquered’ as perhaps the most famous instance.

Languages belonging to very different families exhibit the same possibility: Austronesian Tagalog (Schachter & Otanes 1972: 541), Australian Gooniyandi (McGregor 1990: 411), Nilotic Lango (Noonan 1981: 55), Carib Hixkaryana (Derbyshire 1985: 66), to name but a sample from around the world.

Often, juxtaposition allows many other semantic relations, as Noonan (1981) discusses for “parataxis” as a subcategory of juxtaposition in Lango. Most semantic relations mentioned in the introduction can be expressed in this morphologically minimal way, but often coreferentiality of subject, sometimes of object as well, is required. Morphologically slightly more interesting are cases in which a more unequivocal indication of the temporal relation is expressed. Both sequence and simultaneity may be marked by non-finite verb forms, as in Indo-European participial constructions, or nominalized forms (“masdar”) in Caucasian (cf. Hewitt 1987: 139), non-finite verbs whose suffixes are also used as relativizer and case-markings in Australian languages (Austin 1981). In Mparntwe Arrernte (Wilkins 1988: 149), however, regular tense marking is used for simultaneous and sequential: non-past progressive and past completeive respectively, which are then further suffixed with Ergative-Locative-Instrumental *-le* for Same Subject (‘ss’) and Relativizer + Ablative *-rle-nge* for Different Subject (‘ds’).

Whereas in some Papuan languages, e.g. Usan (Reesink 1987: 90–92), Dani (Bromley 1981: 221 f.), Kobon (Davies 1981: 33) non-finite forms signalling either ‘ss’ or ‘ds’ (called “medial verbs” in Papuan linguistics, but “participles” in Bromley), express both sequence and simultaneity, other languages need special affixation to signal simultaneity. Fore (Scott 1978: 123–127) differentiates simultaneity and sequence only for ‘ss’ relations, Amele (Roberts 1988: 48) reduplicates the verb stem for simultaneity both in ‘ss’ and ‘ds’ conjunctions, with ‘ss’ and ‘ds’ signalled by distinctive subject affixes.

In a number of American Indian languages, the contrast between sequence and simultaneity is also marked by suffixation on non-finite verbs. In Pomo languages, such as Maidu (Oswalt 1976) and Kashaya (Oswalt 1977) the distinction is made for both ‘ss’ and ‘ds’ sequences in portmanteau suffixes. The intersection of sequence versus simultaneity with same or different subject marking has also been reported for some Uto-Aztecán languages (Langacker 1977: 189).

Many, if not all, of the languages that express sequential and simultaneous relations together with the switch-reference parameter, also allow more specific temporal expression. As just one example, Nichols (1983: 255) provides data from Dargi (North Caucasus) contrasting a simple anterior converb, used for both ‘ss’ and ‘ds’, with a suffixed one:

- (1) *at”a kižib gal aqhic”ij*  
father sat:down boy got:up  
‘the father sat down and the boy got up’
- (2) *at”a kižib-mu:til gal aqhic”ij*  
father sat:down-when boy got:up  
‘when his father sat down, the boy got up’

Whereas (1) can also have a cause-effect meaning, the converb suffixed with *-mu:til* in (2) is always temporal. Nichols ascribes the cause-effect interpretation of (1) to the mere clause chaining in Dargi, and Chechen, plus the absence of any explicit indication (1983: 255). Similarly, causal, reason, conditional, purpose, and manner are semantic relations, claimed to be available for similar non-finite verb forms in Australian (Austin 1981: 313), Papuan (Davies 1981: 51), Caucasian (Hewitt 1987: 265) and American Indian (Davis 1973: 187) languages.

### 3. Switch-reference

The term **switch-reference** was originally meant to refer to the device that signals either coreferentiality of subjects in consecutive clauses, or its absence (Jacobsen 1983: 153 f.). While the terminology used for this phenomenon may differ in treatments of American Indian and Papuan languages (see Jacobsen 1983: 154), the most general usage is ‘ss’ for same subject and ‘ds’ for different subject, illustrated in representative examples of canonical switch-reference (Haiman & Munro 1983, eds.: ix):

Maricopa:

- (3) *nyaa ’ashvar-k ’iima-k*  
I 1-sing-ss 1-dance-ASPECT  
‘I sang and I danced’
- (4) *Bonnie-sh Ø-ashvar-m ’iima-k*  
Bonnie-SBJ 3-sing-ds 1-dance-ASPECT  
‘Bonnie sang and I danced’

Usan:

- (5) *ye nam su-ab is-omei*  
I tree cut-ss descend-1.SG:PAST  
‘I cut the tree and went down’

- (6) *ye nam su-iné is-orei*  
 I tree cut-DS descend-3.SG:PAST  
 'I cut the tree (it went) down'

Most of the languages with a switch-reference mechanism are verb-final, with switch-reference being marked by verbal affixation (for exceptions cf. Lynch 1983 and Jacobsen 1983).

The origins of switch-reference marking are extremely heterogeneous, as Haiman & Munro (1983: xiii) point out. In American Indian languages the 'ss' and 'ds' markers originate as nominal suffixes, such as deictics and case affixes (cf. Jacobsen 1983). This may be the reason why for these languages the clause with a switch reference marker is claimed to be subordinate to the reference clause. But Winter (1976: 166 f.) suggests a paratactic relationship between the conjoined verbs. For a number of Papuan languages Haiman (1987) has shown that medial desinences are plausibly derived from a coordinating conjunction. Different subject medial verbs in a few languages show (partial) homophony with the subjunctive mood. Various authors have claimed (e.g. Reesink 1983; Roberts 1988) that medial verbs express a coordinating relationship. This does not detract from the possibility that in some languages certain medial verbs are clearly subordinated to the reference clause. This is the case when a suffix is clearly derived from a deictic element, as can be found in Hua (Haiman 1980), Fore (Scott 1978), Chuave (Thurman 1975). The diagnostic for the difference between subordinating and coordinating medials is the scope of various sentential operators expressed in the reference clause. For an extensive discussion, see Haiman (1976). As an example, consider the different scope of the negative in the following (Haiman 1976: 266): in (8) *-na* does not allow negation to be applied to the first clause, which is the normal interpretation in a coordinate string as in (7).

- (7) *ebgi-Ø-na a'-gi-e*  
 kill-ss-and:he not-cook.it-he  
 'he did not kill and cook it'
- (8) *ebgi-ma-na a'-gi-e*  
 kill-REL-and:he not-cook.it-he  
 'he killed it and (but) he didn't cook it'

Switch-reference devices signal the presence or absence of coreferentiality between consecutive subjects. Both in Papuan and American Indian languages proper inclusion of the referent of one subject in the other subject may

trigger 'ss' marking. A change in person (e.g. from 1.PL to 2.SG) is normally marked as 'ds', whereas a change in number (e.g. 1.PL to 1.SG) allows, or even requires, a 'ss' marking in many languages. See for Papuan languages, Reesink (1983) and for American Indian languages Langdon & Munro (1979). In a few languages, the switch-reference device may be used to signal change of time or space. Roberts (1988: 60 f.) gives examples for Papuan Amel:

- (9) *age ceta guldo-co-bil l-i bahim na*  
 3.PL yam carry-DS-3.PL go-ss floor on  
*tac-ein*  
 fill-3.PL:REMPST  
 'they carried the yams on their shoulders and went and filled up the yam store'

Givón (1990: 883 f.) illustrates 'ds' marking for a thematic change even though there is no break in the referential continuity in American Indian Pima-Papago.

Virtually all languages that employ a switch-reference device are claimed to allow a number of semantic relations. The examples adduced for Australian languages (e.g. Austin 1981; Dench 1988) show fewer conditionals and more purpose clauses, while American Indian languages (Jacobsen 1983: 155–161) and Papuan languages show more conditionals, temporals, and cause/reason interpretations. More specific signalling of such relations is done through other means, involving more morphological material.

#### 4. Conditional

Even though **conditionals** may be expressed by non-finite verb forms which also express the more neutral sequential relation, as (10) from Usan illustrates, most languages require a more explicit conjunction.

- (10) *ya dar-a me*  
 water come.down-3.SG:DS not  
*is-inei*  
 descend-1.SG:NEG.FUT  
 '(if) it rains I won't go down'

Thus, Luiseño (Davis 1973: 181–201) has *tó.wili* introducing the protasis with either a finite or non-finite verb, the latter choice determined by non-hypothetical versus hypothetical. The hypotheticality parameter plays a role in many other languages. Hua (Haiman 1978: 581 f.) groups hypothetical conditions together with 'given that' clauses

against contrafactual. The former are marked with a clause-final marker *-mo*, which also indicates potential topics. Givón (1990: 828) suggests that ‘given’ conditionals are probably not a separate semantic category but a pragmatic association with all adverbial subordinate clauses. In widely diverse languages, mainly of the verb-final type, conditional protases (and other adverbial clauses) are marked with a morpheme that also occurs on nominals to indicate definiteness, givenness or topicality. This “conspiracy” led Haiman to claim that “conditionals are topics” (1978). The Nilo-Saharan language Kanuri (Hutchison 1980) has an associative postposition *-(C)à* which fulfils a range of similar functions: a predicate of existence (11), topic marker (12), conditional marker (13), are just a few of the functions that Hutchison illustrates.

- (11) *njî-à*  
water-ASS  
'there is water, it is raining'
- (12) *wú-à nòngñnyí*  
I-ASS I.don't.know  
'as for me, I don't know'
- (13) *sùlúyìn-nà njî súwúdó*  
he.go.out-ASS water he.should.bring  
'if he is going out, he should bring back water'

Other languages that show conditional use of ‘given/topic’ markers are Usan (Reesink 1987), Dani (Bromley 1981), Mapuchero (Smeets 1989).

Not infrequently, verbs of ‘seeing’ or ‘putting down’ function as conditional markers (Usan, Dani). Other sources include: words for (epistemic) modality, copulas, interrogatives, and time words (‘when’), as listed by Closs Traugott (1985). ‘Given/topic’ markers frequently derive from demonstratives, but general locative adpositions are also used in, again, a wide variety of languages: Dani (Bromley 1981), Newari (Genetti 1991: 237), to name but a few.

## 5. Cause/reason

Other adverbial clauses also use subordinators that are at least homophonous, if not identical, with case markers (or adpositions), suggesting that such clauses function as nominals in adverbial adposition phrases. First of all, of course, causal relations are a rather

natural extension of temporal ones via the ‘post hoc ergo propter hoc’ principle (cf. for Papuan Haiman 1980: 409; for Caucasian Hewitt 1987: 264). More explicit causality or reason (reason should strictly speaking be distinguished from cause, but it is often not clear judging from the English glosses ‘because, since’ whether a language makes a morphological distinction) is expressed by case markers (whether suffix or adposition). The ergative-instrument marker is a good candidate, as in Dani (Bromley 1981), Newari, and other Bodic languages (Genetti 1991: 236 f.), Hua (Haiman 1980: 450). Other spatial relators used to express cause/reason are ablative, as in Newari, and some Uto-Aztec languages (Langacker 1977: 195).

As Closs Traugott & König (1991: 194–199) point out, in many languages (including Indo-European) explicit causal conjunctions originate as temporal ones. It seems a fairly safe generalization to say that such interclausal connectors including those for condition and time, occur finally in the first clause in object-verb languages, as suffix, enclitic or postposition, and initially in the second clause in verb-object languages, as preposition, prefix on the verb or proclitic.

## 6. Purpose

**Purpose** does not seem to be conveyed by non-finite verb forms that signal switch-reference, without other morphological means. Austin (1981: 313, 323) gives purposive readings for clauses that are marked for ‘ss’ or ‘ds’ in various Australian languages. Significantly, the purpose clause always follows the main predication, and the ‘ss’ marker is sometimes glossed as Future, so that it is not clear whether purposive is marked as such or just inferable.

The purpose clause typically refers to an event not yet realized at the time of the event expressed by the main predicate. Thus, as Langacker states with respect to Uto-Aztec (1977: 196), “modal and aspectual elements of unrealization figure prominently in these [i.e. purpose] constructions”. Many Papuan languages employ a verb ‘to say’ as a connective signalling purpose. The purpose clause precedes the main predicate and has then features of a direct quotation (e.g. Reesink 1987; De Vries 1990). Consider the Usan example:

- (14) *wuri ai sigeni-om qamb mi eng*  
 they ground dry-IMP say:ss thing the  
*mor big-umirei*  
 home put-3.PL:PAST  
 ‘they put that thing inside in order that  
 the ground would dry up’

These constructions are used for other thought content as well: reason, pretense, negative purpose. Some striking similarities are noted for Andean languages by Adelaar (1990).

A frequent device for expressing purpose is a case marker: genitive, dative, or allative, often as suffix to a non-finite verb form in OV languages, e.g. Newari (Genetti 1991: 238), Rama (Craig 1991: 469), presumably as representatives of the language families they belong to.

## 7. References

- Adelaar, Willem F. H. (1990), “The Role of Quotations in Andean Discourse”. In: Pinkster & Genee (eds.), 1–12
- Austin, Peter (1981), “Switch-Reference in Australia”. *Language* 57, 309–334
- Austin, Peter (1988, ed.), *Complex Sentence Constructions in Australian Languages*. Amsterdam: John Benjamins
- Bromley, H. Myron (1981), *A Grammar of Lower Grand Valley Dani*. Canberra: Research School of Pacific Studies (Pacific Linguistics C.63)
- Closs Traugott, Elizabeth (1985), “Conditional Markers”. In: Haiman, John (ed.), *Iconicity in Syntax*. Amsterdam: John Benjamins, 289–307
- Closs Traugott, Elizabeth & Heine, Bernd (1991, eds.), *Approaches to Grammaticalization*, Vol. 1–2. Amsterdam: John Benjamins
- Closs Traugott, Elizabeth & König, Ekkehard (1991), “The Semantics-Pragmatics of Grammaticalization Revisited”. In: Closs Traugott & Heine (eds.), Vol. I, 189–218
- Craig, Colette (1991), “Ways to Go in Rama: A Case Study in Polygrammaticalization”. In: Closs Traugott & Heine (eds.), Vol. II, 455–492
- Davies, John (1981), *Kobon*. London: Routledge
- Davis, John F. (1973), *A Partial Grammar of Simplex and Complex Sentences in Luiseño*. Ph.D. dissertation, University of California, Los Angeles
- Dench, Alan (1988), “Complex Sentences in Matuthunira”. In: Austin (ed.), 97–139
- Derbyshire, Desmond C. (1985), *Hixkaryana and Linguistic Typology*. Dallas: Summer Institute of Linguistics
- Genetti, Carol (1991), “From Postposition to Subordinator in Newari”. In: Closs Traugott & Heine (eds.), Vol. II, 227–255
- Givón, Talmy (1990), *Syntax: A Functional-Typological Introduction*, Vol. II. Amsterdam: John Benjamins
- Grimes, Joseph E. (1975), *The Thread of Discourse*. The Hague: Mouton
- Haiman, John (1976), “Presuppositions in Hua”. *Chicago Linguistic Society* 12, 258–270
- Haiman, John (1978), “Conditionals are Topics”. *Language* 54, 564–589
- Haiman, John (1980), *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam: John Benjamins
- Haiman, John (1987), “On Some Origins of Medial Verb Morphology in Papuan Languages”. *Studies in Language* 11, 347–364
- Haiman, John & Munro, Pamela (1983, eds.), *Switch-Reference and Universal Grammar*. Amsterdam: John Benjamins
- Haiman, John & Thompson, Sandra A. (1984), “‘Subordination’ in Universal Grammar”. *Berkeley Linguistics Society* 10, 510–523
- Haiman, John & Thompson, Sandra A. (1988, eds.), *Clause Combining in Grammar and Discourse*. Amsterdam: John Benjamins
- Halliday, Michael A. K. & Hasan, Ruqaiya (1976), *Cohesion in English*. London: Longman
- Hewitt, Brian George (1987), *The Typology of Subordination in Georgian and Abkhaz*. Berlin: Mouton de Gruyter
- Houis, Maurice (1968), “Pour une Typologie des séquences de propositions”. *The Journal of West African Languages* 5, 13–24
- Hutchison, John P. (1980), “The Kanuri Associative Postposition: A Case for Subordination”. *Studies in African Linguistics* 11, 321–351
- Jacobsen, William H. (1983), “Typological and Genetic Notes on Switch-Reference Systems in North American Indian Languages”. In: Haiman & Munro (eds.), 151–183
- Langacker, Ronald W. (1977), *Studies in Uto-Aztec Grammar. Vol. I: An Overview of Uto-Aztec Grammar*. Dallas: Summer Institute of Linguistics
- Langdon, Margaret & Munro, Pamela (1979), “Subject and (Switch) Reference in Yuman”. *Folia Linguistica* 13.3/4, 321–344
- Lehmann, Christian (1988), “Towards a Typology of Clause Linkage”. In: Haiman & Thompson (eds.), 181–225

- Longacre, Robert E. (1976), *An Anatomy of Speech Notions*. Lisse: The Peter de Ridder Press
- Longacre, Robert E. (1985), "Sentences as Combinations of Clauses". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description. Vol. II: Complex Constructions*. Cambridge: University Press, 235–286
- Lynch, John (1983), "Switch-Reference in Lenakel". In: Haiman & Munro (eds.), 209–221
- Matthiessen, Christian & Thompson, Sandra A. (1988), "The Structure of Discourse and 'Subordination'". In: Haiman & Thompson (eds.), 275–329
- McGregor, William B. (1990), *A Functional Grammar of Gooniyandi*. Amsterdam: John Benjamins
- Mithun, Marianne (1984), "How to Avoid Subordination". *Berkeley Linguistics Society* 10, 493–509
- Nichols, Johanna (1983), "Switch-Reference in the Northeast Caucasus". In: Haiman & Munro (eds.), 245–265
- Noonan, Michael (1981), "Parataxis in Lango". *Studies in African Linguistics* 12, 45–69
- Oswalt, Robert (1976), "Switch-Reference in Mai-duan: An Areal and Typological Contribution". *International Journal of American Linguistics* 42, 297–304
- Oswalt, Robert (1977), "The Causative as a Reference Switching Mechanism in Western Pomo". *Berkeley Linguistic Series* 3, 46–54
- Pinkster, Harm & Genee, Inge (1990, eds.), *Unity in Diversity. Papers presented to Simon C. Dik on his 50th birthday*. Dordrecht: Foris Publications
- Reesink, Ger P. (1983), "Switch Reference and Topicality Hierarchies". *Studies in Language* 7, 215–246
- Reesink, Ger P. (1987), *Structures and their Functions in Usan, a Papuan Language*. Amsterdam: John Benjamins
- Roberts, John R. (1988), "Amele Switch-Reference and the Theory of Grammar". *Linguistic Inquiry* 19, 45–63
- Schachter, Paul & Otanes, Fe T. (1972), *Tagalog Reference Grammar*. Berkeley etc.: University of California Press
- Scott, Graham (1978), *The Fore Language of Papua New Guinea*. Canberra: Research School of Pacific Studies (Pacific Linguistics B.47)
- Smeets, Ineke (1989), *A Mapuche Grammar*. Ph.D. dissertation, Leiden University
- Thurman, Robert (1975), "Chuave Medial Verbs". *Anthropological Linguistics* 17, 342–352
- Vries, Lourens J. de (1990), "Some Remarks on Direct Quotation in Kombai". In: Pinkster & Genee (eds.), 291–308
- Wilkins, David (1988), "Switch-Reference in Mparntwe Arrernte (Aranda): Form, Function, and Problems of Identity". In: Austin (ed.), 141–176
- Winter, Werner (1976), "Switch-Reference in Yuman Languages". In: Langdon, Margaret & Silver, Shirley (eds.), *Hokan Studies, Papers from the First Conference on Hokan Languages, San Diego, California, April 1970*. The Hague: Mouton, 165–174

*Ger P. Reesink, Leiden (The Netherlands)*

## 113. Negation

1. Introduction
2. Morphological representation
3. Morphemic distinctions
4. Affirmative-negative contrast
5. As a modal concept
6. References

### 1. Introduction

**Negation** is generally regarded as the opposite of **affirmation**; negative statements are also considered to be representing a marked concept, involving the addition of a negative element to a corresponding affirmative state-

ment. This view is clearly based upon the structure of negative sentences occurring in familiar languages. In some of the non-familiar languages, the situation appears to be rather different. There are languages in which the primary distinction is between realis and irrealis moods with negation forming only part of the irrealis mood; in the case of such languages, negation cannot be directly contrasted with affirmation. However, from the point of view of most of the languages, it would be quite convenient to view negative sentences as being derived from affirmative

sentences through the addition of a negative element.

Morphologically, this addition of a negative element to an affirmative sentence involves different types of modifications in the sentence, such as the insertion of an auxiliary, change of the original (lexical) verb into its non-finite form (or its nominalization), neutralization of verbal categories, and so on. These changes appear to support the view that negation is a marked concept. Languages also differ from one another in showing different types of distinctions in their expression of negation; some differentiate between epistemic and deontic statements, some between narrow and wide scope, and some between realis and irrealis negations. Some of these morphological aspects of negation are described below.

## 2. Morphological representation

Languages differ from one another in the way in which negative sentences are formed; as mentioned earlier, we can generally regard these sentences as involving the addition of a negative element to a sentence which otherwise functions as the corresponding affirmative sentence. When the negative element is added to a sentence, its lexical verb may (i) remain unchanged (cf. (1)), or (ii) change into a non-finite form (cf. (2)), or (iii) the negative element itself may function as a finite verb (cf. (3); see Dahl 1979). These three possibilities give rise to three different types of negative sentences as shown below:

Type I: Havyaka dialect of Kannada (Dravidian)

- (1) (a) *avu ka:gada bare-tt-avu*  
they letter write-NPAST-3.PL  
'They (will) write (a) letter'
- (b) *avu ka:gada bare-tt-av-ille*  
they letter write-NPAST-3.PL-NEG  
'The (will) not write (a) letter'

Type II: English

- (2) (a) *John wrote a letter.*
- (b) *John did not write a letter.*

Type III: Fijian (Central Pacific) (Payne 1985: 209)

- (3) (a) *a yaco\_mai ena siga vakaraubuka*  
PAST arrive on day Friday  
*ko Jone*  
ART John  
'John arrived on Friday'

- (b) *a sega ni yaco\_mai ena siga*  
PAST NEG that arrive on day  
*vakaraubuka ko Jone*  
Friday ART John  
'John didn't arrive on Friday'

Notice that in Havyaka the negative element *ille* 'not' is added to the finite verb (which continues to be finite), whereas in English, the use of the negative *not* involves the addition of an auxiliary verb *did* to the sentence; in Fijian, on the other hand, the negative element *sega* 'not' occurs after the past tense marker *a* and has the complementizer *ni* occurring after it; that is, it replaces the verb *yavo\_mai* 'arrive' as the finite element.

Negative sentences of type I occur in many European languages, and are generally considered to represent the canonical form of negation; however, as pointed out by Honda (1996), the other two types are not uncommon; those of type II occur in Cariban and other South American languages, Trans-New Guinean languages, some of the Australian and African (Bantu) languages, and also in Korean and Japanese. Those of type III, on the other hand, occur in some of the Uralic languages, Tungus languages, Nilo-Saharan languages, languages of the Niger-Congo and Kordofanian families, Yuman and Pomo languages, Australian languages, and also in some of the Salishan languages (see Honda 1996 for details).

It is possible to classify negative sentences of type II into two additional types depending upon whether the negative element is associated (a) with the finite auxiliary or (b) with the non-finite lexical verb (Honda 1996). The English negative sentence (2 b), given above, is of type A, since the negative element *not* is associated with the auxiliary *did*; on the other hand, the following sentence (4 b) of Hixkaryana (Derbyshire 1979: 48) exemplifies the second type (type B) of negative sentence:

- (4) (a) *k-amryek-no*  
SBJ(1)-hunt-IMM.PAST  
'I went hunting'
- (b) *amryek-hira w-ah-ko*  
hunt-NEG SBJ(1)-be-IMM.PAST  
'I did not go hunting'

Notice that the negative element is attached to the lexical verb *amryek* 'hunt' which is non-finite; the tense marker and the personal marker are attached to the auxiliary *ah* 'be' and not to the lexical verb.

The above-mentioned cross-linguistic differences in the morphological representation of negation, however, can only provide us with a typology of negative sentences; they are not very helpful in establishing a typology of languages (for negation) because one and the same language may manifest two or more of these distinct types of negative sentences; for example, English exemplifies type I as shown by the translation of (1 b), and also type II A as shown by (2 b); Havyaka shows type I as pointed out in (1 b), and also type II B as in the following:

- (5) (a) *avu indu ka:gada*  
they today letter  
*bare-tt-avu*  
write-NPST-SBJ(3.PL)  
'They will write letter(s) today.'
- (b) *avu indu ka:gada barey-adde*  
they today letter write-NEG  
*iru-tt-avu*  
be-NPST-SBJ(3.PL)  
'They will not write letter(s) today  
(i.e. they will be without writing letter(s)).'

Notice, however, that the distinctions do imply certain basic differences in the way in which the notion of negation is being viewed by different languages. For example, languages in which negation functions as a finite verb (type III) apparently view it as a higher predicate, whereas the ones in which it only occurs as an inflectional marker regard it as part of the system of verbal categories. Some of these distinctions are also meaningful in a different way, as will be shown in the next section.

### 3. Morphemic distinctions

Languages also differ from one another in the kind of morphemic distinctions that they show in their representation of the notion of negation. Some use the same negative element for representing this notion in all kinds of contexts, whereas others use different elements, thereby implying that the notion is not a unified entity. These distinctions can clearly form the basis of a typology of languages that is based upon the representation of negation.

There are languages like Wargamay (Australian) in which the same negative element *ŋa:* is used for negating different types of sentences; it is true that **imperatives** are slightly

different from other sentences in this language in that they take the element *ŋaru* instead of *ŋa:* (see Dixon 1981: 82), but the two elements are clearly relatable; other languages make this distinction more prominent; Mundari (Austroasiatic), for example, uses the element *alo* for negating imperative sentences and *ka* for negating other types of sentences as shown below (Osada 1992: 88):

- (6) (a) *alo-m jomlea*  
NEG.IMP-SBJ.2.SG eat  
'Don't eat it first.'
- (b) *ka-e? jomkedko*  
NEG-SBJ.3.SG eat:PAST  
'He did not eat.'

The occurrence of this distinction between the negation of imperative sentences on the one hand and other types of sentences on the other, appears to be quite frequent; in India, for example, it occurs in Indo-Aryan languages and also in most of the Tibeto-Burman languages; the latter languages make additional distinctions like future/non-future or irrealis/realis. For example, Manipuri uses *nu* in place of the imperative marker *lu* for negating an imperative sentence (and has an additional marker *kum* which is used in the first person) as shown below:

- (7) (a) *cət-lu*  
go-IMP  
'Go!'
- (b) *cət-nu*  
go-PROH  
'Don't go!'
- (8) (a) *cət-si*  
go-JUSS  
'Let us go!'
- (b) *cət-kum-si*  
go-NEG(JUSS)-JUSS  
'Let us not go!'

In the case of indicative sentences, on the other hand, Manipuri makes a distinction between future and non-future negations by using the suffix *loy* in the former case and *de* in the latter case (cf. Bhat & Ningomba 1997). Examples:

- (9) (a) *məhak ləphoy-du ca-roy*  
he banana-that eat-NEG(FUT)  
'he will not eat that banana'
- (b) *məhak ləphoy-du ca-de*  
he banana-that eat-NEG(NFUT)  
'he did not eat that banana'

The morphemic distinction that we generally find in familiar languages is between **sentential negation** and **constituent negation**. In English, for example, the negative morpheme *not* is used for sentential negation whereas the morpheme *no(n)* is used for constituent negation as shown below:

- (10) (a) *He does not smoke.*  
      (b) *He is a non-smoker.*

Sentential negation also gets contrasted with constituent negation in English for expressing the so-called “scope” differences, with the former denoting **wide scope negation** and the latter for denoting **narrow scope negation**. Examples:

- (11) (a) *Someone did not come today* (wide scope).  
      (b) *No one came today* (narrow scope).

Notice, however, that languages may use other devices for expressing this contrast; Havyaka, for example, uses the specific/non-specific distinction occurring in the quantifier for this purpose:

- (12) (a) *indu a:r-o: bayndav-ille*  
          today who-or(SPECIFIC) came-NEG  
          ‘someone did not come today’  
      (b) *indu a:r-u:*  
          today who-and(NON.SPECIFIC)  
          *bayndav-ille*  
          came-NEG  
          ‘no one came today’

The distinction between two types of negative sentences (type I and II B) shown by Havyaka (see 1 b and 5 b) is used by that language for differentiating between sentential negation and verbal negation. Havyaka uses this distinction consistently in all types of sentences.

#### 4. Affirmative-negative contrast

As mentioned earlier, negative sentences are generally contrasted with **affirmative** sentences, and are viewed as involving a modification of affirmative sentences through the addition of a negative element. There are some languages which strongly support this view; for example, Nkore-Kiga (a Bantu language) has a complex system of verbal forms, showing several distinctions of tense and aspect; all these forms can have corresponding negative forms; further, the various affirmative forms can have corresponding participial

forms that are used as subordinates and these can also have corresponding negative forms (see Taylor 1985). Other languages, however, are not so consistent in showing the affirmative-negative correspondence. The general tendency is to **neutralize** some of the distinctions that occur in affirmative sentences when they are changed into negative sentences.

For example, Bemba, another Bantu language, makes a distinction between tomorrow future and after tomorrow future in the affirmative sentences, but in the negative sentences, it neutralizes this distinction (Givón 1978); Russian differentiates between perfective and imperfective aspects in the affirmative, but in the negative the perfective is rarely used; Havyaka differentiates between past and perfect in the affirmative, but in the negative the perfect form is used for denoting an undifferentiated past meaning. This neutralization appears to be correlatable with a typology of languages based upon the “prominence” of verbal categories; that is, languages appear to show a tendency to neutralize, in negative sentences (and also in certain other contexts like nominalizations, and the formation of adjectivals and adverbials), verbal category distinctions which are not very prominent in their grammar and to retain distinctions which are prominent (cf. Bhat 1999).

Another interesting phenomenon that has been noted in connection with this affirmative-negative contrast is the **reversal** of verbal category distinctions called “flip-flop”; in Terêna (Maipuran Arawakan), for example, there are two different verbal forms, called Actual (past, present or definite future) and Potential (indefinite future, habitual and imperative); the following sentences (Ekdahl & Grimes 1964: 268, as quoted by Honda 1996: 159–160) exemplify this contrast:

- (13) (a) *pih-óp-o*  
          go-DIR-ACTUAL  
          ‘he went to where he came from’  
      (b) *pih-ép-a*  
          go-DIR-POT  
          ‘go back to where you came from’

Terêna makes a distinction between two different types of negations, namely *ako* used with Potential verbs and *hyoko* used with Actual verbs; however, the former appears to provide the Actual connotation and the latter the Potential connotation, exemplifying,

thereby, the notion of flip-flop as shown below:

- (14) (c) *ako pih-áp-a*  
 NEG go-DIR-POT  
 ‘he did not go where he came from’
- (b) *hyoko pih-ép-o*  
 NEG go-DIR-ACTUAL  
 ‘do not go back to where you came from’

Honda (1996: 166) provides an interesting explanation for this phenomenon, namely that of viewing ‘x did y’ and ‘x will not do y’ as two different ways of describing one and the same situation in which an event has taken place, and thus will not take place again, and similarly of viewing, ‘x did not do y’ and ‘x will do y’ as two different ways of describing the same situation in which an event has not taken place, and thus has the potential to take place. On the other hand, Kannada (Dravidian) manifests a situation which appears to need a different type of explanation. In this language, past negative meaning is expressed by adding the negative word *illa* ‘not’ to the purposive form of the verb, whereas the non-past negative meaning is expressed by adding the same negative word to the nominalized (neuter singular) form of the verb, as shown below:

- (15) (a) *avanu manege bar-al-illa*  
 he home come-PURP-NEG  
 ‘he has not come home’
- (b) *avanu manege baruyu-d-illa*  
 he home come-NR(3.N.SG)-NEG  
 ‘he will not come home’

The use of the negative word with the purposive form of the verb (which generally has posterior connotation) for denoting past (or prior) meaning in (15a) appears to be rather odd, but a possible explanation for this usage is that the negative word is similar to modal verbs which, in contrast to aspectual verbs, are attached to the purposive (posterior) form of the verb; aspectual verbs, on the other hand, are attached to the prior form of the verb. Examples:

- (16) (a) *avanu manege bar-a-bahudu*  
 he home come-PURP-may  
 ‘he may come home’
- (b) *avanu manege ban-du-bitṭa*  
 he home come-PAST-COMPL  
 ‘he succeeded in coming home’

## 5. As a modal concept

Negation functions as a modal concept even more prominently in certain other languages like Acoma (Kerasan) and Yimas (Papua New Guinean) than in Kannada (cf. 15); the basic verbal distinction, in these languages is between realis and irrealis moods, with negation forming part of the realis mood. In Yimas, for example, the irrealis gets further subdivided into negative, potential and likely, denoted by three different prefixes as shown below (Foley 1991: 251):

- (17) (a) *ta-ka-wa-t*  
 NEG-SBJ.1.SG-go-PFV  
 ‘I did not go’
- (b) *ant-ka-tu-r-um*  
 potential-SBJ.1.SG-kill-PFV-OBJ.3.PL  
 ‘I almost killed them’
- (c) *ka-mpu-a-tput-n*  
 likely-SBJ.3.PL-OBJ.1.SG-hit-PRES  
 ‘they are going to hit me’

In Acoma, on the other hand, negation is used only when the speaker himself is involved in the event that is being negated (as a subject or object); that is, it is used only in contexts in which the speaker can vouch for the non-occurrence of the event; in other contexts, he uses the dubitative (see Miller 1965: 123). Negation is thus viewed, in these languages, as part of the irrealis modal system and is not directly contrasted with the affirmative, as in familiar languages.

## 6. References

- Bhat, D. N. S. (1999), *The Prominence of Tense, Aspect and Mood*. Amsterdam: John Benjamins (Studies in Language Companion Series 49)
- Bhat, D. N. S. & Ningomba, M. S. (1997), *Manipuri Grammar*. Munich: Lincom Europa
- Dahl, Östen (1979), “Typology of Sentence Negation”. *Linguistics* 17, 79–106
- Derbyshire, Desmond C. (1979), *Hixkaryana*. Amsterdam: North-Holland (Lingua Descriptive Studies 1)
- Dixon, Robert M. W. (1981), “Wargamay”. In: Dixon, Robert M. W. & Blake, Barry J. (eds.), *Handbook of Australian Languages, Vol. II*. Canberra: The Australian National University Press, 1–146
- Ekdahl, Muriel & Grimes, Joseph E. (1964), “Terenia Verb Inflection”. *International Journal of American Linguistics* 30, 261–268

- Foley, William A. (1991), *The Yimas Language of New Guinea*. Stanford: Stanford University Press
- Givón, Talmy (1978), “Negation in Language: Pragmatics, Functions, Ontology”. In: Cole, Peter (ed.), *Pragmatics*. New York: Academic Press (Syntax and Semantics 9), 69–112
- Honda, Isao (1996), *Negation: A Cross-Linguistic Study*. Ph. D. dissertation, University of Buffalo
- Miller, Wick R. (1965), *Acoma Grammar and Texts*. Berkeley: University of California Press
- Osada, Toshiki (1992), *A Reference Grammar of Mundari*. Tokyo: Tokyo University of Foreign Studies
- Payne, John R. (1985), “Negation”. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. I*. Cambridge: Cambridge University Press, 197–242
- Taylor, Charles (1985), *Nkore-Kiga*. London: Croom Helm

*D. N. S. Bhat, Mysore (India)*

## 114. Comparison and gradation

1. Preliminary notions
2. Comparative
3. Superlative
4. Equative
5. References

### 1. Preliminary notions

All the languages of the world have at their disposal different means to express **comparison** and **gradation**, but not every language expresses them through morphology. In recent years some relevant works on the typology of comparison were published (Stassen 1985; Xerberman 1999; but see already Jensen 1934), and the semantics of comparative constructions is being investigated by semantically. A reliable survey on comparison and gradation, however, still needs detailed research in several theoretical domains, including morphology (see, for instance, the short entries by Andersen 1992 and Crookston 1994).

In the present article, then, comparison and gradation will be treated to the extent that morphology is involved, and a classification of the means employed for its expression will be sketched.

Viewed conceptually, both comparison and gradation presuppose an entity that some property, state or, more rarely, a more or less dynamic state of affairs applies to. They also presuppose that this state of affairs varies on some scale on which it may, in principle, be measured; i.e. it is a **parameter**. **Gradation** (German *Abstufung*) then is the stepwise modification of the extent to which the parameter applies to the entity, while **compari-**

**son** (German *Vergleich*) assesses this extent with respect to some standard. Taken as a grammatical category, comparison (German *Steigerung*) is the formal modification of some predicative word – most often an adjective – representing a parameter of gradation or comparison, according to the extent to which it applies to its argument, relative to some standard. Similarly, gradation may be manifested in a structural category, in particular of adjectives and verbs.

In a **comparative construction** four elements are identifiable (cf. ex. (1)):

- (a) the element which is compared, the **comparee**, or **topic of comparison**;
- (b) the element that serves as **standard of comparison**;
- (c) the element, called the **pivot**, or **marker of comparison**, that introduces the standard of comparison;
- (d) the predicate that represents the parameter of comparison.

(1) *John is smart-er than Sam*  
 comparee predicate-CMPR pivot standard

Elements (b) and (c) can be absent for different reasons:

- The pivot may be lacking, as some languages possess no element for that function (see 2.1).
- The construction of (1) is in fact very rare among the languages of the world, and the standard is more frequently marked by other morphological means, for instance case markers.
- The standard of comparison can be omitted for semantic and pragmatic reasons.

In fact, although grading involves comparison (Sapir 1951; Lyons 1977: 271), the element or the entity serving as the standard can be either presupposed logically (see, for instance, Panogl 1973 and Bertocchi & Orlandini 1996 for Latin) or by conversational implicature (see Chierchia & McConnell-Ginet 1990: 283–285). Sentences like *I thought he lived in a cleaner house* or *Sheila is less lucky* are rather common in discourse. Comparatives occurring without standard, traditionally labelled **absolute comparatives**, seem much more frequent than those occurring with the standard, but statistics on this point are scarce.

The adjective is the lexical category that typically undergoes comparison. Three degrees are traditionally recognised: positive, comparative, and superlative.

The **positive** degree (from Lat. *positivus* ‘imposing (the standard)’), coincides with the basic form of the adjective itself, e.g. Engl. *a small box, an interesting book*.

The **comparative** degree of some predicate – typically an adjective – marks this predicate as applying to its argument (the comparee) to a higher extent than the standard; e.g. *smaller*. In gradable predicates, esp. in polar (or relative) adjectives, the form that is unmarked for gradation or comparison – for adjectives: the positive degree – semantically involves implicit comparison with a norm. For instance, *an old dog* is a dog that is older than some standard for dogs. It is the function of the comparative degree to allow for the substitution of the implicit default standard of comparison by a particular explicit standard of comparison.

From a cognitive point of view, it is more salient to describe entities that are larger, smaller, than those that are less large, less small and so forth. In grammars the former type is referred to as **comparison of majority**, the latter as **comparison of minority**: *That box is smaller than this vs. Your book is less interesting than hers*. Whereas comparison of majority is expressed by morphological processes in several languages, apparently comparison of minority is expressed only lexically. This implies that there is no affix meaning ‘less’ parallel to that of majority ‘more’ (but see 2.4).

A **comparison of equality** is one that ascribes to the comparee the same value of the parameter of comparison as to the standard. If it is marked formally on the predi-

cate representing the parameter – typically, an adjective – this is called the **equative form** (see 4). In order to describe this construction, five elements have been identified (Haspelmath & Buchholz 1998: 279):

- (2) *My sister is as pretty as you*  
 1            2 3        4 5

where 1 is the **comparee**, 2 is the **parameter marker**, 3 is the **parameter**, 4 is the **standard marker**, 5 is the **standard**.

Very close to the equative is the **similative**, “a construction expressing sameness of manner” (Haspelmath & Buchholz 1998: 278): *Robert is as tall as Maria* is equative, whereas *She sings like a nightingale* is similative. Similative constructions will not be treated here.

By the **superlative** (Lat. *superlativus* ‘raised to the top’), the quality expressed by the adjective is described as being at a very high, possibly the highest degree. However, it must be observed that “the superlative form is often used to denote a high grade, but not necessarily an apical grade, of the graded quality” (Sapir 1951: 146).

In some languages (Classical Arabic, for example), the adjective shows one single form that exerts the function both of the comparative and superlative, without any morphological distinction; this special form is frequently referred to as **elative** (Lat. *elativus* ‘standing out’): *akbar*, from *kabi:r*, for instance, means ‘greater, greatest, very great’ according to the syntactic environment or the context.

The superlative degree in comparison (e.g. *Zembo is the laziest of all the chimps at the zoo*) is to be distinguished from the highest degree in gradation (e.g. *a very old tree*). Languages like English use different forms. The former is expressed by the superlative form of the adjective (the positive form either preceded by *most* or followed by *-est*), generally preceded by the definite article: *the most beautiful, the luckiest*. The latter is marked by an intensifier preceding the positive adjective: *very lucky*. Forms like *\*the very lucky of all men* are ill-formed. Languages like Latin use the superlative in both cases: a superlative like *felicissimus*, for example, occurs both in gradation (*vir felicissimus* ‘very lucky man’) and in comparison (*vir omnium felicissimus* ‘the luckiest man of all’). The two uses of the superlative form are then called **absolute** and **relative superlative**, respectively.

The following points concern formal properties of the structural categories of comparison and gradation. First, the status of com-

parison between derivation and inflection is controversial and may differ according to the language. Second, suppletive expression (see Art. 52) is not rare: Latin *bonus* ‘good’, *meior* ‘better’, *optimus* ‘best’; Hung. *sok* ‘good’, *több* ‘better’. While analytic expression of gradation (e.g. *very tired*) belongs in the realm of syntax, it may also be expressed by compounding (Dutch *dood-moe* (dead-tired) ‘very tired’) and affixing (Latin *per-bonus* ‘very good’, *per-terrere* ‘frighten very much’).

Third, while comparison of adjectives and (derived) adverbs is well-known, it does occur on nouns, too. Cf. Sanskrit *vira-tarā* ‘more hero’, Ancient Greek *kún-teros* ‘(more) shameless’, lit. ‘more dog’; Hungarian *róka* ‘fox’, comparative *rókább* ‘slyer’ (Bhat 1994: 25 f.). In Basque the suffix *-(a)go* can be added to any lexical category: *gizon* ‘man’: *gizanago* ‘more man’; but also *gugana* ‘towards us’: *guganago* ‘more towards us’. In Italian the superlative suffix *-issim-* is added to nominal stems to express the highest degree of a quality: *governissimo* ‘a stable and powerful government’, from *governo* ‘government’; *canzonissima* ‘the best song (among those in competition for an award)’, from *canzone* ‘song’.

Fourth, languages may have semantic, morphological or phonological constraints on the application of comparison to members of a word class or on its expression. The alternation between synthetic and analytic comparison in English is due to a formal constraint. Most of the semantic constraints result from the defining criterion of comparison which essentially involves a gradable parameter. Consequently, adjectives with complementary (contradictory) meaning such as *odd/even* are not used in comparison and gradation, at least not in their literal sense, and therefore often lack the morphological category of comparison. Also, certain adjectives that designate the highest grade tend to avoid comparison and gradation. This is true both for explicitly derived forms such as Dutch *doodmoe*, \**doodmoeërl/\*doodmoest* (Booij 1996: 5) and for lexical terms for the highest degree such as *excellent*, \**very excellent*.

## 2. Comparative

According to the absence or the presence, either optional or obligatory, of a marker of comparison on the lexical category involved, typically adjectives and adverbs (one of the criteria advocated by Xerberman 1999; but see already Jensen 1934), we can subdivide the strategies into four main groups.

### 2.1. No marker

The graded form has no marker. One of the most widespread types of comparative structures is the **juxtaposition** of two positive adjectives, with no marker of comparison and with antonymic value: *X (is) long, Y (is) short = X is longer than Y*. This is very common in isolating languages and can be found also among languages with agglutinative or fusional morphology. As Benveniste had already suggested, this strategy seems to be basic and independent of the morphological type of language in which it occurs:

“La comparaison est implicite dans l'énoncé anti-thétique. Toutes les langues peuvent recourir à ce moyen, même quand elles disposent d'une expression morphologique.” (Benveniste 1948: 126)

This type is illustrated by Samoan in (3) and occurs very frequently among Amazonian and Austronesian languages.

- (3) *Ua loa lenei va'a, ua puupuu lena.*  
be long this boat be short that  
'This boat is longer than that.'

A subtype is one in which a positive adjective is juxtaposed with the negated form of the same adjective, i.e. *X (is) long, Y (is) not long = X is longer than Y*, like in Apalai (Koehn & Koehn 1986: 52):

- (4) *Mopo zumo pyra kyn-exi-ne*  
*Mopo big NEG 3-be-RGMPST*  
*akono zumo.*  
3.brother.in.law big  
'His brother-in-law is bigger than Mopo.'

This strategy has been reconstructed for Proto-Indo-European (Puhvel 1973) and is also attested in Vedic and the Baltic languages (Pinault 1985; 1989), possibly in Latin (Baldi & Cuzzolin 2001: 220). The two subtypes mentioned often cooccur in the same language, like, for instance, in Wari' (Everett & Kern 1997: 193 f.).

### 2.2. Optional lexical morpheme

The form is graded by means of an often optional lexical morpheme, mainly unbound. In this type, too, which is very frequent among the languages of the world, the adjective has no marker of comparison and the standard occurs with a case marker, generally coinciding with some case (genitive, ablative, dative are the most frequent). No language is attested in which the standard of comparison is marked by a special case, although in some languages (for instance Limbu (van Driem

1987: 52) or Tsaxur (Sosenskaja 1999: 561)) the standard is formed by a combination of two, mainly semantic cases that form a “special ending”.

This type is rather widespread particularly among the agglutinative languages: in Turkish either the positive adjective alone or combined with the adverb *daha* is currently employed (Lewis 1988: 54):

- (5) *kurşun-dan (daha) ağır*  
 lead-ABL (more) heavy  
 ‘heavier than lead’

In Udihe, a Manchu-Tunguz language, “the adjective itself does not take any special comparative morpheme” (Nikolaeva & Tolskaya 2001: 180) and the standard is in the ablative case:

- (6) *Ussuri Bikin-digi sunta.*  
 Ussuri Bikin-ABL deep  
 ‘The Ussuri is deeper than the Bikin.’

In Japanese the adjective is normally expressed without any adverb: *Tiroo wa Taroo yori kasikoi desu* ‘Chiroo TOP Taroo from smart is (Chiroo is smarter than Taroo)’ but the emphatic form *no hoo ga* ‘very’ can be added after *yori*, without any basic change in the meaning of the sentence (Hinds 1986: 127 f.).

The adverb meaning ‘more’ and the pivotal element corresponding to *than* may be agglutinated into one phonological word, e.g. in Tümpisa Shoshone (Uto-Aztecán):

- (7) *Üü yuhupi nü wakakwa*  
 you fat me more:than  
 ‘You are fatter than me.’ (Dayley 1989: 288 f.)

### 2.3. Affix

The form is graded by means of an affix. As has been repeatedly noticed, this type, which is quite common in the Indo-European family, is extremely rare among the world’s languages. English employs two morphological processes to form the comparative: in *smaller* there are the lexical morph *small-* and the suffix *-er*, which conveys the meaning ‘more’, whereas in *more interesting* the notion of comparison is expressed lexically by the adverb *more*. In the former case the comparative is expressed synthetically, in the latter analytically. This is in keeping with the general tendency observable among different linguistic families to replace synthetic by analytic expression. In English the use of the suffix *-er*

is limited by phonological constraints (Quirk et al. 1994: 461 f.); it is no longer productive and tends to be replaced by the adverb *more: smaller/more small*.

Other declension categories may interact with the comparative in various ways. The simplest subtype is represented by a language like Old Irish, in which “all the forms of comparison are uninflected, and show no difference of number or gender” (Thurneysen 1980: 232). For instance, the adjective *dian* ‘swift’ has the comparative *déniu* ‘swifter’ for all the numbers and genders. In Latin, oblique case forms of comparative adjectives contain an agglutinative sequence of the comparative suffix plus a case/number suffix which neutralizes gender: *beat-ior-i* ‘happy-CMPR-DAT.SG’, whereas in the so called direct cases (nominative, accusative, and vocative) there is a difference between *beat-ior-em* ‘happy-CMPR-SG.M/F.ACC’ and *beat-ius* ‘happy-SG.N.ACC’: in the latter the marker of comparison also conveys the grammatical information of number, gender and case, with cumulative exponence (cf. Art. 64).

There are two other very rare morphological processes to express gradation. Among the Semitic languages, degree forms exhibit a stem with the structure /aCCaC/ derived from the normal positive form of the adjective, usually /CaCCI:C/; e.g., Arabic *ḥamīd* ‘praiseworthy’, elative *ahmad* ‘more praiseworthy’.

Among the Kartvelian languages the comparative can be formed by **circumfixes**: in Georgian, the comparative is characterised by the circumfix *u-...-es-* with possible vocalic variation in the adjectival stem: *lamezi* ‘beautiful’, comparative *u-lamez-es-i* ‘more beautiful’; *magari* ‘strong’, *u-magr-es-i* ‘stronger’ (Hewitt 1995: 48). This type is no longer productive in Modern Georgian. In Svan, too, the comparative is formed by the circumfix *xo-...-a* and the adjective can also show vocalic variation: *c'rni* ‘red’, *xo-c'ran-a* ‘redder’ (Tuite 1997: 18).

Finally, an interesting case is represented by Modern Irish:

“There is only one degree of morphologically expressed comparison ... It is formed from the base adjective usually by palatalisation of the final consonant and the addition of *-e* ...” (O’Dochartaigh 1992: 74).

This form, however, is preceded by the particle *níos*: *bán* ‘white’, *níos báine (ná)* ‘whiter

(than)'. In Modern Irish, therefore, the comparative form is marked twice, lexically and morphologically.

#### 2.4. Verb

The adjectival category, which clearly exhibits the morphological features typical for nouns in the Indo-European languages, is more verb-like in numerous other languages (a useful analysis in Bhat 1994: 187–209). This phenomenon is particularly frequent among isolating languages, or those showing a low degree of synthesis, and among the native languages of America (but see Bhat 1994: 211–243). We refer to this type as the "verbal type of comparison". Within this type, two subtypes are identifiable. The first involves verbs meaning 'to surpass, to overtake'. Needless to say, the expression of comparison here involves syntax rather than morphology. In the second subtype, the parameter of comparison itself is a verb.

In Saek, a language of the Austro-Thai family spoken between Laos and Thailand, comparison is expressed by the verb *lyyn*<sup>5</sup> 'surpass, overtake' (Morev 1988: 43; numbers in exponence refer to tones):

- (8) *hum<sup>4</sup> myy<sup>3</sup>nii<sup>5</sup> reeng<sup>4</sup> lyyn<sup>5</sup> myy<sup>3</sup>luan<sup>4</sup>*  
wind today strong surpass yesterday  
'today's wind is stronger than yesterday's'

This type is also attested in other South-East Asian languages, genetically related like Zhuang (Moskalev 1971: 198) or unrelated, like Vietnamese (Nguyễn 1997: 122; see also Stassen 1985), but it is not limited to isolating languages: it also occurs in agglutinative or fusional languages with verbs meaning 'to overtake, to exceed'. It is widespread among the Bantu languages and also occurs among the Chadic languages: Hausa (cf. Ščeglov 1970: 239, 273), Miya (cf. Schuh 1998: 314). In Rwanda, for instance,

"to express comparison between two qualities or attributes the verbs *kurusha* and *kuruta* 'exceed, surpass' are used, which agree in person, number and class with the nouns denoting a person or thing that surpasses another person or thing participating in comparison." (Dubnova 1984: 45)

In Amharic (Afro-Asiatic), the verbs *läqqa* 'surpass' or *bällätä* 'exceed' can reinforce the comparative form of the adjective (Hudson 1997: 466):

- (9) *Haylu k-antä yðbält bätam qäccən*  
Haylu from-2.SG.M exceed very thin  
*nä-w.*  
is-he  
'Haylu is much thinner than you.'

This type of comparative, which according to Stassen (1985: 159; on the relationship between comparison and word order see also Andersen 1983 and Romero-Figueroa 1986) tends to be related to SVO languages and supposedly developed in Late Proto-Indo-European (Puhvel 1973), also occurs in languages where the commonest type is the juxtaposed type as in Wari' (Everett & Kern 1997: 194 f.).

In some languages a similar structure is used to express the comparative of minority, employing a lexical item meaning 'behind', cf. (10) from Miskito (Misulmapan) (Suárez 1983: 135).

- (10) *witin jang ninara tukta*  
he I behind child  
'he is younger than I'

A more appropriate translation of (10) would be 'he is less old than I'. A parallel construction is found in Amazonian languages like Sanuma (Borgman 1990: 54).

The comparative construction necessarily has a verbal head in those languages in which property concepts are verbs, such as West Greenlandic (Greenlandic Eskimo):

"Comparative degree is expressed derivationally with a comparative/superlative affix on the (verbal) base expressing the parameter of comparison and, optionally, a case-marked nominal expressing the standard or limit of comparison (this may also be a possessive inflection on the comparative morpheme when it is in nominal form)." (Fortescue 1984: 167 f.):

- (11) *kujataa-nit issin-niru-vuq*  
south-ABL be.cold-more-3.SG.IND  
'it is colder than the south'

"In the older language *niru* could be left out in fact" (Fortescue 1984: 168).

#### 2.5. Diachronic tendencies

Some tendencies are observable in the diachrony of comparative constructions. First of all, some old synthetic superlatives become **opaque** forms and tend to be felt as positive adjectives, for which new comparative and superlative forms are created. In Italian for instance, *intimo* has nowadays almost completely lost its original superlative function of 'innermost' and only means 'intimate', with

a comparative (*più intimo*) and a rather rare *intimissimo*, used in commercials. But even in Latin, especially Late Latin, we find that those superlative forms which were residual and built by means of non-productive rules, tended to be reanalyzed as positive: a form like *extremius* (recorded since 2nd c. A.D.) is a comparative formed from the superlative *extremum* ‘outermost, extreme’.

This phenomenon should not be confused with the tendency to **reinforce** and strengthen comparative forms, a feature which is rather common among the Indo-European languages. For instance, the co-occurrence of two markers of comparison was rather frequent in Late Latin: forms like *magis fortior = fortior/magis fortis* ‘stronger’, *magis beatior = beatior/magis beatus* ‘happier’ are well attested (Hofmann & Szantyr 1965: 166 f.). This phenomenon also occurs in spoken varieties of fusional languages which still retain the synthetic form of the comparative beside the analytic one: see Modern Greek *pio mikróteros* ‘more smaller’ (Holton et al. 1997: 87).

Secondly, a well-known phenomenon in the history of numerous languages is the replacement of synthetic forms by **analytical** ones. This is largely attested in the history of the Romance languages, where the synthetic forms in *-iore(m)* were gradually but nearly completely replaced by the analytical forms deriving from *plus* (Italian *più* and its dialects, French *plus*; Modern Greek *pio*, as in *pio kaló* ‘more beautiful’, ultimately derives from *plus* via Venetian) or from *magis* (Spanish *mas*: *mas hermosa* ‘more beautiful’, Portuguese *mais*: *mais famoso* ‘more famous’, Rumanian *mai*: *mai înalt* ‘higher’), and in several spoken varieties around the Mediterranean.

In Tunisian Arabic the form of the elative (template /aCCaC/: *hasan* ‘good’, *ahsan* ‘better, very good’) is usually replaced by the form *akṭar* (elative form of *kati:r* ‘much’) plus the positive: *hasən akṭor* ‘better’.

### 3. Superlative

Although superlatives are often formed by the same morphological process as comparatives, they also show some typical formations. There are two basic, non-exclusive morphological processes used by the languages of the world. They are, however, mainly employed to form the highest grade

(or absolute superlative), whereas the relative superlative is basically formed by syntactic devices.

#### 3.1. Analytic formation

The form is graded by means of a lexeme. This is probably the commonest means of expressing the highest grade among the world’s languages: almost every language has a word meaning roughly ‘very’ which, preposed or postposed, combines with the adjective (see Klein 1998 on the semantics of adverbs of degree): *very friendly*, Xhosa *inencasa gqitha* ‘delicious very (too delicious)’. In several languages this is the only possible means creating the superlative.

For stylistic, expressive reasons particular adverbs are employed in discourse: e.g. English *terribly*, German *furchtbar* ‘terribly’, Italian *straordinariamente* ‘extraordinarily’ (Austerlitz 1991: 3). In English *terribly good* simply means ‘very good’. Adverbs like those mentioned are often the etymological source of the adverb ‘very’: *very* itself originally meant ‘really’; *sehr*, which is the normal adverb for ‘very’ in present-day German, is the grammaticalised Middle High German adverb *sere*, which in origin meant ‘painful, violent’.

#### 3.2. Synthetic formation

The adjective is graded by means of some morphological process. In this case the inventory of the morphological processes displayed among the languages of the world is also quite rich.

One of the processes by which the superlative is formed is **reduplication** (cf. Art. 57; extremely rarely used in the formation of comparative, as in Upriver Halkomelem; see Galloway 1984: 56). In some cases there is total reduplication of the adjective, as in Samoan: *tele* ‘big’, *teletele* ‘very big’, or Sumerian: *bar* ‘external’, *barbar* ‘very external = foreigner, barbarian’. In colloquial Italian one can reduplicate the adjective with superlative function in predicative position: *È rimasto lì tranquillo tranquillo* ‘He remained there very quiet’, but with an evaluative nuance.

Other cases involve partial reduplication like, for instance, of the first syllable (with some change) in some Mongol languages. In Modern Mongolian the first segment /(C)VC/ of the adjective is reduplicated, but the last phoneme is replaced by /v/, so that /(C)VC/ becomes /(C)Vv/ (**inexact reduplication**: Art. 57): *ulaan* ‘red’, *uv ulaan* ‘very red’,

as in Kalmuk: *xav xar* ‘very black’, whereas in Burjat, related to Mongolian, the phoneme is /b/ instead of /vl/: *jab jagaan* ‘very pink’ (Sanzeev & Todaeva 1993:139). However, formation of the superlative by reduplication appears to be limited to certain adjectival classes.

Among the Indo-European languages the superlative is formed with two types of **affixes**: prefixes and suffixes. The most frequent were various **prefixes** derived from the root \**per-*: *per-iprae-* in Latin (André 1951): *praeflarus* ‘very famous’, *percandidus* ‘very smart’; *pre-*, already documented in Old Russian as *pre-lpri-* (Vjalkina 1995: 321), still occurs with some adjectives in Modern Russian: *mudryj* ‘wise’, *premudryj* ‘very wise’, *krasnyj* ‘red’, *prekrasnyj* ‘very red’; *peri-* in Ancient Greek: *perikallé:s* ‘very beautiful’. In this case, the prefix is not simply added to the positive form of adjective, but rather to an alternate stem.

In Celtic another etymologically unclear prefix *an-* is attested, which already occurs in Gallic (*ande-caros* ‘very red’; Vendryes 1981: 71), and is used in the same function: Irish *an-mhaith* ‘very good’ (from *maith*), Welsh *an-fawr* ‘very big’ (from *mawr*; in both cases the adjectives undergo lenition).

In some modern European languages other types of intensifying prefixed elements occur; although their productivity is rather limited, they frequently occur in colloquial varieties: Dutch *doodmoe* ‘very tired’, lit. ‘dead-tired’ (cf. Italian *stanco morto*), German *saublöde* ‘(sow-silly) very silly’, Italian *straricco* ‘very rich’ (*stra-* from Latin *extra* ‘outside’; cf. already Latin *extraordinarius* ‘extraordinary, beyond the limits of normality’), *arcicontento* ‘very happy’ (*arci-* from Ancient Greek *archi-* ‘chief, prominent’ via Latin).

The most widespread **suffixes** among the Indo-European languages were, however, \*-*isto-* and \*-*tato-* (which bear a morphological similarity with the formatives of ordinals; Benveniste 1948), traces of which still remain in some Indo-European branches. Their morphosyntactic behaviour is almost identical with their comparative counterparts \*-*jos-* and \*-*tero-*.

In some languages a special prefix is added to the comparative form of the adjective or adverb: in Hungarian, for instance, the prefix *leg-* is added to the comparative form: *gyors-abb* ‘more rapid’ becomes *leg-gyors-abb* ‘very rapid, the most rapid’, the adverb *gyors-an*

‘rapidly’ becomes *leg-gyors-abb-an* ‘very rapidly’.

In Modern Irish, the superlative is formed by the synthetic comparative form of the adjective preceded by the particle *is*: *deacair* ‘difficult’, superlative *is deaicre* ‘most difficult’, *deaicre* being the comparative (see 2.3).

#### 4. Equative

Among the graded forms, **equativity** is the one which exhibits the least variety. There are basically two types, analytic and synthetic formation.

##### 4.1. Analytic formation

The parameter marker (see 2) may be expressed analytically by a particle or an adverb (German):

- (12) *Zürich ist so groß wie Wien.*  
‘Zurich is as big as Vienna.’

In some languages the parameter marker can be omitted. This is the case in Italian, for instance, where the presence of the parameter marker is stylistically highly marked:

- (13) *Il mio libro è (così) bello come il tuo.*  
‘My book is as nice as yours.’

In a few cases analytic parameter markers “are semantically more or less transparent and mean something like ‘equally, to the same degree’” (Haspelmath & Buchholz 1998: 284):

- (14) (a) Chinese  
*Tā gēn nǐ yíyàng gāo.*  
she with you one.manner tall  
‘She is as tall as you.’

- (b) Seychelles Creole  
*I ris mem degree ki nu.*  
he rich same extent as we  
‘He is as rich as us.’

##### 4.2. Synthetic formation

Synthetic formations are far less frequent than analytic ones. The parameter marker is expressed synthetically on the parameter by a special morpheme. Among the languages of Europe, such formations occur only in three linguistic groups: Celtic, Finno-Ugric, and Kartvelian, but non-European languages also show them: Tagalog, Indonesian (in these two languages the parameter marker is a prefix; cf. (15) from Indonesian), Greenlandic Eskimo (Haspelmath & Buchholz 1998: 283 f.), Estonian (16):

- (15) *Ayah saya se-tinggi paman saya.*  
father 1.SG EQT-tall uncle 1.SG  
'My father is as tall as my uncle.'
- (16) *Minu õde on minu pikk-une.*  
I:GEN sister COP I:GEN tall-EQT  
'My sister is as tall as me.'

The standard is not usually in the case in which it occurs after the comparative, and in some languages there is a special case marker (Ancash Quechua):

- (17) *Pani-i-mi qam-naw*  
sister-1.SG-DIRECT.EVIDENCE you-EQT  
*shumaq.*  
pretty  
'My sister is as pretty as you.'

In Old Irish the equative suffix was *-ithir/-idir* (of unclear origin; Thurneysen 1980: 237f.); the choice between the two was motivated by phonology (cf. (18 a–b)): "the former as a rule after monosyllables, the latter after polysyllables" (Thurneysen 1980: 233):

- (18) (a) *léir* 'eager'  
*léirthir* 'as eager'  
(b) *erlam* 'ready'  
*erlamidir* 'as ready'

In Modern Irish the equative form is no more productive and has been replaced by the construction [*comh* + the positive form of the adjective + *le*]: *comh cliste le* 'as clever as'. In Modern Welsh, instead, the parameter marker precedes the synthetic equative form in *-ed* (Middle Welsh *-hed*) and not the simple positive form: *oer* 'cold', *cyn oered a* 'as cold as' (not \**cyn oer a*), even though this happens with some adjectives that do not have the *-ed* form.

## 5. References

- Andersen, Paul K. (1983), *Word Order Typology and Comparative Constructions*. Amsterdam: Benjamins
- Andersen, Paul K. (1992), "Gradation". In: Bright, William (ed.), *International Encyclopedia of Linguistics, Vol. II*. New York, Oxford: Oxford University Press, 79
- André, Jacques (1951), "Les adjectifs et adverbes à valeur intensive en *per-* et *prae-*". *Revue des Études Latines* 29, 121–147
- Austerlitz, Robert (1991), "The European Absolute Superlative, an Orphan of Grammar, of the Lexicon, and of History". In: Ivir, Vladimir & Kalogera, Damir (eds.), *Languages in Contact and Contrast*. Berlin, New York: Mouton de Gruyter, 1–13
- Baldi, Philip & Cuzzolin, Pierluigi (2001), "Towards a New Historical Syntax of Latin". In: Moussy, Claude (ed.), *De lingua Latina novae quaestiones. Actes du Xème Colloque Internationale de Linguistique Latine, Paris-Sèvres, 19–23 avril 1999*. Louvain: Peeters, 201–225
- Benveniste, Émile (1948), *Noms d'agent et noms d'action en indo-européen*. Paris: Maisonneuve
- Bertocchi, Alessandra & Orlandini, Anna (1996), "Quelques aspects de la comparaison en latin". *Indogermanische Forschungen* 101, 195–232
- Bhat, D. N. S. (1994), *The Adjectival Category*. Amsterdam, Philadelphia: Benjamins
- Booij, Geert E. (1996), "Inherent versus Contextual Inflection and the Split Morphology Hypothesis". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology 1995*. Dordrecht: Kluwer, 1–16
- Borgman, Donald M. (1990), "Sanuma". In: Derbyshire, Desmond C. & Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages, Vol. II*. Berlin, New York: Mouton de Gruyter, 17–248
- Chierchia, Gennaro & McConnell-Ginet, Sally (1990), *Meaning and Grammar*. Cambridge/MA: Massachusetts Institute of Technology
- Crookston, Ian (1994), "Comparative Constructions". In: Asher, Ronald E. & Simpson, J. M. Y. (eds.), *The Encyclopedia of Language and Linguistics, Vol. II*. Oxford etc.: Pergamon Press, 624–629
- Dayley, Jon P. (1989), *Tümpisa (Panamint) Shoshone Grammar*. Berkeley etc.: University of California Press
- Driem, George van (1987), *A Grammar of Limbu*. Berlin etc.: Mouton de Gruyter
- Dubnova, Yelena Zinov'evna (1984), *The Rwanda Language*. Moscow: Nauka
- Everett, Daniel L. & Kern, Barbara (1997), *Wari*'. London, New York: Routledge
- Fortescue, Michael (1984), *West Greenlandic*. London etc.: Croom Helm
- Galloway, Brent D. (1984), *A Grammar of Upriver Halkomelem*. Berkeley etc.: University of California Press
- Haspelmath, Martin & Buchholz, Oda (1998), "Equative and Similative Constructions in the Languages of Europe". In: van der Auwera, Johann (ed.), *Adverbial Constructions in the Languages of Europe*. Berlin, New York: Mouton de Gruyter, 277–334

- Hewitt, George B. (1995), *Georgian*. Amsterdam, Philadelphia: Benjamins
- Hinds, John (1986), *Japanese*. London, New York: Routledge
- Hofmann, Johann Baptist & Szantyr, Anton (1965), *Lateinische Syntax und Stilistik*. München: Beck
- Holton, David & Mackridge, Peter & Philippaki-Warburton, Irene (1997), *Greek*. London: Routledge
- Hudson, Grover (1997), "Amharic and Argobba". In: Hetzron, Robert (ed.), *The Semitic Languages*. New York: Routledge, 457–485
- Jensen, Hans (1934), "Der steigernde Vergleich und sein sprachlicher Ausdruck". *Indogermanische Forschungen* 52, 108–139
- Klein, Henny (1998), *Adverbs of Degree in Dutch and Related Languages*. Amsterdam, Philadelphia: Benjamins
- Koehn, Edward & Koehn, Sally (1986), "Apalai". In: Derbyshire, Desmond C. & Pullum, Geoffrey K. (eds.), *Handbook of Amazonian Languages, Vol. I*. Berlin, New York: Mouton de Gruyter, 33–127
- Lewis, Geoffrey L. (1988), *Turkish Grammar* [reprinted edition with corrections]. Cambridge: Cambridge University Press
- Lyons, John (1977), *Semantics*. Cambridge: Cambridge University Press
- Morev, Lev Nikolaevič (1988), *Jazyk sèk*. Moskva: Nauka
- Moskalev, Aleksej Alekseevič (1971), *Grammatika jazyka čízuan*. Moskva: Nauka
- Nguyêñ, Đinh-Hoà (1997), *Vietnamese*. Amsterdam, Philadelphia: Benjamins
- Nikolaeva, Irina & Tolskaya, Maria (2001), *A Grammar of Udihe*. Berlin, New York: Mouton de Gruyter
- O'Dochartaigh, Cathair (1992), "The Irish Language". In: MacAulay, Donald (ed.), *The Celtic Languages*. Cambridge: Cambridge University Press, 11–99
- Panagl, Oswald (1973), "Präsuppositionen und die Syntax der lateinischen Komparation". *Salzburger Beiträge zur Linguistik* 1, 361–375
- Pinault, Georges (1985), "Négation et comparaison védique". *Bulletin de la Société de Linguistique* 80.1, 103–144
- Pinault, Georges (1989), "Les voies de la comparaison védique". In: Valentín, Paul (ed.), *La comparaison*: *Linguistica Palatina, Colloquia III*. Paris: Presses de l'Université de Paris-Sorbonne, 53–74
- Puhvel, Jaan (1973), "Nature and Means of Comparison in Proto-Indo-European Grammar". *The Journal of Indo-European Studies* 1, 145–154
- Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffry & Svartvik, Jan (1994), *A Comprehensive Grammar of the English Language*. London: Longman
- Romero-Figueroa, Andrés (1986), "Warao Comparatives". *Studies in Language* 10.1, 97–108
- Sanzeev, G. D. & Todaeva, B. X. (1993), "Mongolskie jazyki". In: *Jazyki Azii i Afriki, V: Altajskie jazyki*. Moskva: Nauka, 98–186
- Sapir, Edward (1951), "Grading: A Study in Semantics". In: Mandelbaum, David G. (ed.) *Selected Writings of Edward Sapir in Language, Culture and Personality*. Berkeley, Los Angeles: University of California Press, 122–149 [originally appeared (1944) in *Philosophy of Science* 11, 93–116]
- Ščeglov, Jurij K. (1970), *Očerk grammatiki jazyka xausa*. Moskva: Nauka
- Schuh, Russell G. (1998), *A Grammar of Miya*. Berkeley etc.: University of California Press
- Sosenkaja, Tat'jana B. (1999), "Sravnitel'nye konstrukcii". In: Kibrik, Alexander Evgen'evič (ed.), *Èlementy caxurskogo jazyka v tipologičeskem osvěščenii*. Moskva: Nasledie, 560–581
- Stassen, Leon (1985), *Comparison and Universal Grammar*. Oxford: Blackwell
- Suárez, Jorge A. (1983), *The Mesoamerican Indian Languages*. Cambridge: Cambridge University Press
- Thurneysen, Rudolf (1980), *A Grammar of Old Irish*, trans. and revised by D. Binchy and Osborn Bergin. Dublin: Dublin Institute for Advanced Studies
- Tuite, Kenin (1997), *Svan*. München, Newcastle: Lincom
- Vendryes, Joseph (1981), *Lexique étymologique de l'irländais ancien*. A. Dublin: Institute for the Advanced Studies
- Vjalkina, L. V. (1995), "Prilagatel'noe". In: *Drevnerusskaja grammatika XII–XIII vv.* Moskva: Nauka, 294–326
- Xerberman, K. R. (1999), "Komparativnye konstrukcii v sravnennii: K voprosu ob otноšenii grammatiki k ètimologii i jazykovoj tipologii". *Voprosy jazykoznanija* 2, 92–107

*Pierluigi Cuzzolin, Bergamo (Italy)*  
*Christian Lehmann, Erfurt (Germany)*

## XV. Morphologische Typologie und Universalien

### Morphological typology and universals

#### 115. Approaches to morphological typology

1. Introduction
2. Language and thought
3. The notion ‘morphological type’
4. Morphemes and categories
5. Head-marked vs. dependent-marked structure
6. Typology and evolution
7. References

##### 1. Introduction

The term ‘morphological typology’ invokes in the first instance the famous tripartite division into isolating, agglutinating and flectional (or fusional) languages associated above all with the work of Wilhelm von Humboldt. This classical typology of the beginning of the 19th century had absorbed the earlier models of Adam Smith and the *Encyclopédistes* based respectively on a synthetic-analytic division of languages or one into langues analogues and langues transpositives. It was in its turn to be reinterpreted in structural terms by Edward Sapir and is indirectly reflected in the discussion on methods of morphological analysis which pre-occupied American taxonomic linguists around the middle of the 20th century. Morphology figures prominently in a recent typology based on whether the morphological exponents of syntactic relations are located on the head or on the dependent of a construction.

##### 2. Language and thought

The classic morphological typology developed by the Schlegel brothers and Wilhelm von Humboldt must be understood within the context of a linguistic philosophy whose central concern was the relationship between language, thought and reality (Manchester 1985). In order to appreciate the relevant issues, we should remind ourselves that in the Age of Reason the diversity of the languages

of the world was perceived as a superficial phenomenon behind which lay a universe of innate concepts or ‘eternal ideas’. Although these were reflected only very imperfectly in any one language, they could be reconstructed through the systematic comparison of the grammatical and semantic structures of different languages. Language types would result from relating the observable structures to a universal set of logical concepts by a method which was in essence deductive and constructive (Coseriu 1972: 214). Wilhelm von Humboldt (1836) supported this rationalist position to the extent that he acknowledged the dependence of thought on concepts but departed from it in claiming that concepts are not universals but are specific to individual languages. He argued that this must be the case, since languages are the product of their histories and can be shown to differ from one another in both form and content. For him, then, every language encapsulates a unique segmentation of the external world and of the universe of human experience. Cognition is achieved through the individual speech act in which the speaker uses his language creatively in a context-related utterance.

For Humboldt, therefore, the relationship between form and meaning is language-specific and linguistic comparison does not reveal any universal logic. What is universal, however, is the dependence of concept formation on ‘articulated sound’. Concepts, he claimed, are of two kinds, basic (or lexical) and relational. Every language gives overt expression to basic concepts in the form of lexical elements, but individual languages differ in the way they handle relational concepts and it is in this that lies the essential foundation of Humboldtian typology (Humboldt 1836; Schleicher 1869: 1–30). Relational concepts comprise functions typically associated (in those languages which have them) with two kinds of affix: derivational affixes, which

are employed in building a lexicon from a limited stock of lexical bases (or roots), and grammatical affixes, which indicate the relations between lexical items within the linear structure of the sentence.

This classic typology distinguishes three, marginally four, language types to one or other of which, ideally, any language should be assignable. In point of fact, however, no language is totally consistent typologically so that in practice the types are to be taken as referring rather to prototypes which are never fully realised (see 3). Humboldt would seem to have vacillated between these two positions. When, therefore, a linguist speaks of a language as being isolating or agglutinating, the reference is in fact rather to the prototype and ignores the language's less 'prototypical' features. **Isolating languages** were seen as either giving no overt expression to **relational meaning** at all or as employing for this purpose the same units as they use for encoding lexical concepts. Thus, for instance, in Chinese the verb 'give' is also used as a dative marker. **Agglutinative (or agglutinative) languages** were seen as attaching the individual elements expressing relational concepts to the lexical base serially (Turkish *evler-in* 'house-PL-GEN (of the houses)'). In this type, relational concepts are thus given overt expression and the semantic structure is directly reflected on a one-to-one basis in the formal structure of the word. The result is a word which is relatively unconstrained as regards the nature and the number of its morphological segments and which still to some extent reveals the lexical origins of its affixes. The extreme situation is reached in the so-called polysynthetic languages, which even permit the incorporation of such constituents as lexical objects into the verb form. The question of whether or not these languages should be treated as a subtype of the agglutinative ones or should be counted as a separate type was never resolved. **Flectional (or inflecting, or fusional) languages** were seen as encoding relational concepts by means of two mechanisms: either by internal modification of the lexical base (English *took*, past tense of *take*) or by adding 'meaningless' affixes to the base (English *goats*, plural of *goat*). In this latter the *-s*, notwithstanding its grammatical meaning, was treated as 'meaningless' since it lacks any association with a lexical base. The former mechanism was considered to be a truly symbolic representation of the structure of the underlying thought be-

cause relational meaning is achieved through the modification of lexical elements, which are basic. Where the latter mechanism was concerned it was considered important that the added elements, often fused with one another and with the lexical root, should have no resemblance to or association with any lexical root.

As regards the relationship between language and thought, Humboldt saw the above three types as increasingly successful attempts to represent and facilitate thought. The inflecting type was considered to be 'the most perfect' because, in the internal inflection of the lexical base, relational meaning is represented in a manner quite distinct from lexical meaning and at the same time fully integrated with it. The agglutinative type failed to reach such perfect integration because here the lexical origins of the relational elements were still in many instances transparent. It was, however, considered superior to the isolating type, which was seen as entirely lacking any overt expression of relational meaning (the use of word order as a means of expressing this was obviously outside the domain of the typology).

A generation earlier, Adam Smith had advocated a division into compounded and uncompounded languages, a distinction which was continued in Friedrich Schlegel's division into synthetic and analytic. Interestingly in the present context, Smith praised the synthetic (and, in particular, inflectional) languages for their economy but gave preference to the analytic type because, since in this each relational concept is represented by one invariable element, it was more conducive to clear thinking. Although Friedrich Schlegel must have been familiar with the work of Adam Smith (Coseriu 1968), there is no evidence of his views having entered the discussion about degrees of linguistic perfection. Neither was direct reference ever made to the *Encyclopédistes'* distinction between **langues analogues**, which have the subject-verb-object order of logical predicates, and **langues transpositives**, which have some other order plus inflectional morphology to mark subject and object (Ramat 1995).

Humboldt's evaluation may rather have been influenced by the highly sophisticated and very modern analyses of word structure by the Sanskrit grammarians, which had just become accessible. For, while the Greek and Latin grammarians had contented themselves with listing morphologically related forms in

nominal and verbal paradigms, the Sanskrit grammarians segmented off the individual affixes and devised a systematic means of dealing with stem and affix alternants through sandhi rules and a system of vowel gradations. Their work, in fact, made explicit the structure not only of Sanskrit words but also, indirectly, of the inflecting type in general and without it the classic typology could not have been developed.

Despite their intimate knowledge of languages and their considerable insight into language structure, neither the Schlegel brothers nor Humboldt concerned themselves with systematic linguistic analysis. It was only a generation later that August Schleicher was to introduce the term ‘morphology’ into linguistics and to devise for the first time a systematic notational system capable of representing word structures cross-linguistically (Schleicher 1959; cf. Bynon 1986). Schleicher’s analysis involved not only the segmentation of word-forms into lexical root plus affixes but also the systematic representation of what were later to be termed morpheme alternants. In this he was undoubtedly influenced by the methods of the Sanskrit grammarians. The concept of the morpheme as the smallest meaningful unit of grammatical structure was, however, only to come later with structural linguistics, and one must also await the arrival of modern linguistics for an adequate exposition of morphological typology.

It was Sapir (1921: 59–156) who made explicit in modern linguistic terms both the formal and the semantic components implicit in the Humboldtian typology. He operated with three parameters, each representing one aspect of variation within and across languages. Of the two parameters which relate to the morphological form of the word, the first is concerned with its degree of internal complexity based on how much syntactic information is subsumed within it. The range of possibilities proposed, which he terms index of **synthesis**, extends from analytic through synthetic to polysynthetic. **Analytic structure** implies a morphologically simple word so that the burden of formal signalling falls on the syntax. **Synthetic structure** implies words which, although of moderate length, are morphologically complex and whose formal structure includes the markers of syntactic relations (subject, object, etc.). **Polysynthetic structure**, finally, implies maximally complex words in which entire syntactic constituents

are treated as part of a single word (as, for instance, when a lexical direct object is incorporated into the verb). Later, Greenberg (1960) was to quantify word complexity for a number of languages, although he based his analysis on only a very limited set of text samples (see Art. 117). For English the degree of synthesis (that is the average number of morphemes per word) is given as 1.68, for Sanskrit 2.59, and for the polysynthetic Eskimo language, 3.72.

Sapir’s second parameter relating to morphological form is concerned with the technique employed by the language for combining lexical and relational meaning in the syntactic string. Here the possibilities are isolation, agglutination, fusion and symbolic linking (or internal flection). Isolating technique implies the overt signalling of relational concepts by means of lexical bases and its covert signalling by means of word order or distributional constraints. In Vietnamese, for instance (cf. Art. 142), which is often cited as the prototypical isolating language, ‘we’ is made up of ‘I’ plus an element which signals plural, subject and object are distinguished by position, and word class is indicated by distribution. Thus, the syntactic status of *tôi* in the following sentence is determined by its place relative to the other elements (Comrie 1989:43):

- (1) *khi tôi ở’ên nhà ban tôi ...*  
     when I come house friend I  
     ‘when I came to my friend’s house ...’

Agglutinating technique means that the individual relational concepts combine serially with the lexical base, as we have already seen in Turkish *ev-ler-in* ‘house-PL-GEN’. Fusional technique implies such close integration of lexical and relational elements that morphological segmentation becomes impossible (English *brought* from *bring*), while symbolic technique refers to the signalling of relational concepts within the lexical base itself (English *took* from *take*). These two latter techniques correspond to the traditional flectional type. Greenberg’s (1960) figures for the degree of fusion, measured in terms of morphophonemic alternations per morpheme juncture, are 0.34 for English, 0.09 for Sanskrit, and 0.03 for Eskimo. Thus, although Sanskrit words are on average more complex than English words, English turns out according to his sample to be more fusional.

Being concerned with the mechanics of concept formation, Sapir’s remaining param-

eter is perhaps the most basic of the three in the present context of language and thought. Sapir asks two questions: firstly, does a language keep its relational concepts free from any admixture with lexical bases or not (giving *pure-relational* versus *mixed-relational*), and secondly does it keep its lexical concepts pure or does it build these up from the aggregation of inseparable elements (giving *simple* versus *complex*) (Sapir 1921: 138). To illustrate the first of these alternatives, there are languages in which plural number is not as a rule overtly encoded but must be understood from the context. These languages do, nevertheless, have the possibility of employing a lexical item (such as a word meaning ‘many’ or ‘several’) to indicate plurality where this is felt necessary. A language employing this concrete means of encoding plurality will thus be situated at the mixed-relational end of the scale whereas a language like English, which combines the relatively concrete concept of number with the abstract relation of subject (encoded in subject-verb agreement, as in *the boy runs* versus *the boys run*, and even more so in *the sheep is grazing* versus *the sheep are grazing*), will rank near the pure-relational end. Its place on the continuum is thus determined by the degree to which a concept is grammaticalised. To illustrate this second parameter, in English the word *head* not only denotes the body part but also the person in charge of a business or a school (*‘the head forbids ...’*), and the verb *to head* is derived from the noun without the addition of any overt marker; ‘simple’ forms such as these, which are prevalent in isolating languages, may be contrasted with formations like *breadth* or *depth* or *length* which, being made up of an aggregate of inseparable elements, are labelled ‘complex’. Between the two universally attested extremes, lexical bases at the concrete end and the syntactic relations subject-of and object-of at the abstract end, are situated closer to the former such notions as agentive noun (*teacher*) and diminutive (German *Pferd-chen* ‘little horse’) and to the latter dative, plural and past tense.

Applied to particular languages, Sapir’s typology classes Turkish as “synthetic”, “agglutinative”, and “complex pure-relational”, Ancient Greek and Latin also as “synthetic” but differing from Turkish in being preponderantly “fusional” (neither the first person singular nor the accusative case, for instance, have a single invariant representation) and sometimes “symbolic”, and in that their con-

cept formation is “complex mixed-relational”. English differs from Greek and Latin in being, like Vietnamese, “analytic” and “isolating” although in this latter concept formation is “simple pure-relational”.

Sapir’s typology has made explicit, as Humboldt had intended, the full range of possibilities available to language for the formation of propositions. Comrie (1989: 42–52), in line with the American structuralists, has adopted and systematised Sapir’s two parameters which relate to formal structure but has abandoned the index which parametrises concept formation and is therefore most immediately concerned with language and cognition. The Cologne UNITYP model (Seiler 1990; 1995) perpetuates these cognitive issues most systematically and with maximal explicitness by combining the study of linguistic organisation at different levels of structure with the ways in which different languages ‘apprehend’ (that is conceptualise) the universe of thought and reality.

### 3. The notion ‘morphological type’

Typological classification comes up against the problem of whether or not a language can in fact be meaningfully assigned to a particular type or whether **language types** are simply ideals which are never fully realised. If this latter is the case, then individual languages can only be described in relative terms by reference to the ideal types or prototypes.

Humboldt appears to have oscillated between these two positions whereas, within the Prague School, Skalická defined a language type as a theoretical construct comprising a range of properties which tend to go together (are ‘mutually favourable’) cross-linguistically (Sgall 1995). He in fact extended the scope of his enquiry beyond both the classic typology and the morphological domain to include aspects of word phonology and of word order. Where morphology is concerned, he characterised any given language in terms of the relative contribution made to its structure by the different morphological techniques (so that for him a language can, for instance, be at the same time both inflecting and agglutinating). The fact that only one technique predominates in the structure of most languages was to be interpreted in terms of linguistic economy, since the use of either grammar words or affixes or alternation or word order as the means of encoding

relational meaning is less costly than the simultaneous deployment of several types. That is to say, his position is ultimately a functional one, a perspective which is characteristic of Prague School theory. Sapir had already recognised the difficulty of assigning the entire morphology of a language to a single type and had allowed for the possibility of its belonging at one and the same time to more than one type where a particular parameter is concerned. Thus, in terms of technique, he characterizes the Polynesian languages as agglutinative-isolating and Cambodian as fusional-isolating.

Neither with regard to any particular parameter nor in terms of mutually favourable properties would it then appear possible in practice to assign a language in its entirety to a type. More recently the emphasis has therefore shifted away from the typological characterisation of specific languages to a global perspective in which the entire range of linguistic variation is examined in terms of linguistic universals. Typical examples are Greenberg's word order universals and those partial typologies which focus on specific areas of structure that represent the conceptualisation of particular cognitive domains. Thus causative constructions, voice systems, possessive constructions, grammatical-relation changing mechanisms are all at the forefront of current interest (see Shibatani & Bynon 1995). As a result, the philosophical questions which underly morphological typology can no longer be contained within the domain of morphology alone.

#### 4. Morphemes and categories

In the heyday of taxonomic linguistics, any kind of cross-linguistic comparison was held to be suspect because the objective procedures of analysis to which linguists were then committed did not yield categories that went beyond the domain of a single language or language state. Linguists were, on the other hand, perfectly aware of the fact that some of the difficulties they were facing in analysing particular structures were closely tied up with questions of typology. Bazell spoke of the then popular view which held that "inflectional languages are characterised by indeterminacy of segments, isolating languages by indeterminacy of classes [that is categories, TB], while agglutinative languages are free of either indeterminacy" (Bazell

1966: 30). Thus, in an inflecting language such as Latin, even seemingly unproblematic word-forms like *puellarum* 'girl:GEN.PL' lack a determinate segmentation, and in an isolating language like Vietnamese the units that form the input to the syntactic rules are not always those minimal units which are easily identifiable through segmentation. Only agglutinative languages appear to be fully determinate with regard to both segmentation and classification.

Broadly speaking, in the taxonomic approach, the higher units of structure (sentence, phrase, word) are analysed as being composed of units at the level immediately below, sentences consisting of phrases, phrases of words, and words of morphemes. The composition of the morpheme itself is more contentious, since morphemes and phonemes belong to different hierarchies – the former to the 'grammatical' hierarchy already mentioned, the latter to a separate phonological hierarchy comprising phonological word, syllable, phoneme. This argues against an analysis which sees **morphemes** as being composed of **phonemes**. The two hierarchies are, rather, linked by a relation of 'representation', morphemes ideally being represented by sequences of phonemes. In practice, however, there are complications. The morpheme *wife* for instance does not have a unique representation because in the plural the fricative is voiced (singular [waif], plural [waivz]). It is necessary in cases such as these, which are typical of fusional languages, to introduce some intermediate unit to take care of the alternation. Two competing solutions were proposed, either the **morph** or the **morphophoneme** (Hockett 1954; 1961). The former results from the morphological segmentation of the word, consists of phonemes and, together with its co-allomorphs, represents a given morpheme. The latter is a phoneme-sized abstract unit out of which morphemes are composed and which is represented by phonemes. The morpheme *wife* is thus either represented by the two allomorphs [waif] and [waiv-], each of which consists of phonemes, or it is composed of the morphophonemes /w-ai-F/ in which F represents the alternating phonemes /f/ and /v/ whose distributions here are said to be 'lexically predictable' (see Fig. 115.1, based on Hockett).

In the case of a prototypically agglutinative language an intermediate unit would not seem necessary but in practice even such well

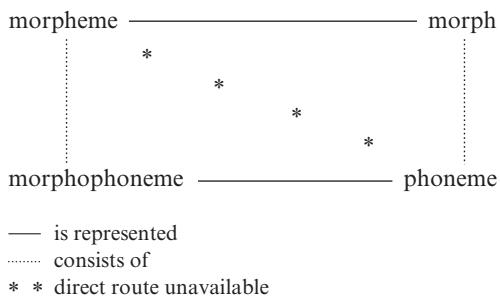


Fig. 115.1: Morpheme, morph, morphophoneme, and phoneme

behaved languages as Turkish have morpheme alternants, mainly because of vowel harmony. However, whichever route is chosen, morphological segmentation is determinate and results in units which represent morphosyntactic categories on a one-to-one basis. In the case of fusional languages on the other hand morphological segmentation is problematic and controversial and the morphological segments stand for cumulations of morphosyntactic categories. The *-o* of *amo* 'I love' for instance jointly represents first person, singular number, indicative mood, present tense and active voice (compare *amas* 'you love', *amamus* 'we love', *amem* 'I would love', *amabam* 'I loved', and *amor* 'I am loved'). Thus, even if linguists can agree on an optimal segmentation, the resulting segments do not directly match the categories that are relevant to syntactic and semantic analysis. A morpheme-based morphological analysis, in other words, will work well for the agglutinating type but appears redundant when applied to the isolating type and inadequate when applied to the flectional type. This latter type, in particular, would appear to fare better with a word-and-paradigm model which relates whole words to functional categories (Robins 1960; Matthews 1970). Clearly, it is only in agglutinative languages that there is a systematic correlation between morphological segments and morphosyntactic categories, a situation that calls into question the universal application of a morpheme-based analysis.

However, instead of developing the typological aspect of these problems, linguists were at the time pre-occupied rather with demonstrating that a single descriptive model, whether morph or morphophoneme based, could be made to work for all the

word-structures of a single language. Subsequently the generalisation of one or other variant of the morpheme-based approach became a matter of theoretical preference, the Generative school opting for the morphophoneme. To each morpheme they assigned a single dictionary spelling ('underlying' or morphophonemic representation) from which the various surface representations could be derived by rule. Cross-linguistic considerations, including the old typology, were not central issues at the time, at least not in America.

In order to illustrate the absence of congruence between morphemes and categories, let us take a simple case like that of plural number in English. The syntactic rules treat *cats*, *dogs*, *horses*, *oxen*, *men*, *mice*, *feet*, *sheep* and *children* as 'the same', irrespective of the morphological shape taken by the plural marker. It was argued by taxonomists that a plural morpheme can in fact be obtained by segmentation and classification and justified by the complementary distribution of its 'phonologically' and 'lexically' conditioned variants. But such an analysis lumps together the productive and phonologically predictable variants *-s*, *-z* and *-iz*, which form a natural unit, with the highly idiosyncratic pluralisation mechanisms found in *mice*, *men*, *sheep*, *fungi*, *cacti* and many others which need separate listing. Some of us will remember the five (or was it six?) competing morpheme-based analyses of *took* (the past tense of *take*), and with regard to the assignment of segments to classes, Matthews's *cantarebbero* (Italian, 'that I may sing', from *cantare*) is a rather extreme case of multiple exponency. There can be no doubt that morpheme-based analyses become increasingly cumbersome the more fusional and symbolic the structure in question.

Problems such as these led Bazell (1966) to suggest that a typology might be based on those areas of structure which most lead to disagreement between linguists. In a language like Latin, controversy relates mainly to the segmentation of complex words because the natural unit of segmentation is the word. In a language such as Turkish on the other hand, controversy centres on word boundaries ('is a given element a suffix or a particle or postposition?') because the natural unit of segmentation is the morph(eme). Finally, in Vietnamese, while segmentation readily yields morph(eme)-size units, difficulties arise when these are related to syntactic classes. In short, the morpheme as an ana-

lytic unit is better suited to those languages, or structures within languages, in which morphological segments and morphosyntactic categories maximally coincide.

### 5. Head-marked vs. dependent-marked structure

The approach to be considered in this section also concerns the interface between morphology and syntax, but here the focus is on where precisely the marker of a syntactic relation is situated. Nichols (1986) distinguishes two types, ‘head-marking’ and ‘dependent-marking’. Head-marking means that the morphological indicator of a syntactic relation is affixed to the head of a construction, dependent-marking that it is affixed to the dependent. *John's house*, for instance, is dependent-marked, the marker of the possessive relation being attached to the dependent (*John*), whereas in Hungarian a structure like *az ember haz-a* ‘the man house-3.SG (the man's house)’ is head-marked. As the terminology suggests, the approach is in terms of dependency rather than constituency and the assignment of head and dependent is as follows (Nichols 1986: 57):

head	dependent
possessed noun	possessor
noun	modifying adjective
pre-/postposition	object of pre-/postposition
predicate	arguments and adjuncts
auxiliary verb	lexical (‘main’) verb
main clause	subordinate clause

Tab. 115.1: Head and dependent marking

In most Indo-European languages the structure is basically dependent-marking, but the subject relation is as a rule head-marked. Thus, in the clause, it is the arguments of the verb that bear the marker (for example the case affix) of their relation to the verb as head, but the verb itself also shows agreement with the subject. In the noun phrase, the adjective agrees with its head noun, in adpositional phrases case-marking on the noun (and possibly on other constituents) signals the relation of the noun phrase to the pre- or postposition, which is the head.

Thus, in a pure dependent-marking language, the clause is endocentric:

### (2) NP-NOM NP-ACC NP-DAT ... V

For example Chechen (a Caucasian language, Nichols 1986: 108):

- (3) *cu stag-a zudčun-na*  
DEM.OBL person-ERG woman-DAT  
*kni:ga-Ø j-elira*  
book-NOM CL<sub>j</sub>-give.PAST  
‘the man gave the woman a book’

In a head-marking language, on the other hand, the clause is exocentric, the verb-form on its own having the status of a complete clause:

- (4) (NP<sub>1</sub> NP<sub>2</sub> NP<sub>3</sub> ...) V-PRO<sub>1</sub>-PRO<sub>2</sub>-PRO<sub>3</sub>  
(in which -PRO- indicates a pronominal clitic and the numbers indicate the matching lexical noun phrases).

For example Abkhaz (Caucasian, Nichols 1986: 108):

- (5) *a-xàc'a a-pħ'òs a-ś°q°'ð*  
the-man the-woman the-book  
*Ø-lò-y-te-yt'*  
it-to.her-he-give.PAST-FIN  
‘the man gave the woman a book’

In head-marked clause structure the syntactic relations thus hold between the pronominal elements and the verb, the matching lexical noun phrases forming endocentric relationships with the pronominal elements. In a dependent-marking language on the other hand the syntactic relations are between the lexical noun phrases and the verb.

This typology makes explicit a basic difference in the morphological signalling of syntactic relations. The overall characterisation of any particular language is obtained by calculating the proportion of its head-marked to its dependent-marked structures. Nichols argues that the marking patterns of entire language families are remarkably stable over time, although changes in either direction do occur and it is here that this typology has considerable explanatory potential (see 6).

### 6. Typology and evolution

August Schleicher interpreted morphological typology in evolutionary terms as the step-wise progression of language from isolation via agglutination to flexion. He likened this hierarchy to the *systema naturae* of contemporary biology, in which all organic structures were seen as having developed from inorganic ones and complex organisms from

simpler ones. The isolating, agglutinating and flectional language types thus represented for him the same natural succession as did rocks, plants and animals, each higher order subsuming the one below it. In both schemes the crucial factor was a direct correlation between ‘adjacency in the system’ and ‘sequence in evolution’ and it was this that guaranteed in his thinking their non-arbitrariness (see Bynon 1986 for further discussion and references).

When languages are looked at *in toto*, this position is clearly without empirical support. This is for two principal reasons. Firstly, since no language is in fact typologically pure, any such statement about a language must be suspect. Secondly, it has been remarked that such languages as present-day French and English are in the process of developing increasing numbers of isolating structures (Sapir’s ‘drift’ towards the invariable word, the isolating technique of marking subject and object, etc.), which suggests a cycle rather than a hierarchy. Schleicher’s claimed evolutionary progression through types, with flection as the summit, is compatible with a philosophy of history which postulates the division of this into two distinct periods, one of creative evolution in prehistoric times followed by a one of decay in historic times. The creative peak was reached in the classic period of the Indo-European and Semitic languages (Sanskrit, Ancient Greek, Hebrew), whereas the development of the Romance, Germanic, and Slavonic languages was seen as representing a period of decay in which the feeling for the inherited morphology has broken down (Schleicher 1859; 1869: 1–30).

The possibility of morphological decay being the beginning of a new cycle, with new isolating structures replacing eroded inflections, was never considered at the time. It was of course well known that the modern European languages have, to varying degrees, replaced their inherited synthetic structures by analytic ones, but this was perceived as a development within, and not beyond, the flectional type. Today most historical linguists would probably accept some version of a typological cycle, from isolation via agglutination to flection and back to isolation, but limited of course to individual structures and not applied to entire languages.

As early as 1822, when discussing the development of grammatical forms, Humboldt ([1822] 1905: 303) made reference to the so-

called “agglutination theory”, which he attributed to Horne Tooke. This claimed that certain words, in particular those of fairly general meaning, would gradually lose their independent use, material connotations and free form status and become first grammar words and then affixes. While it is perfectly reasonable to look for the source of new morphology in the syntax of a previous language state, the assumption of a direct transition from synthetic sequence to single word by merely suppressing the word boundaries is too simplistic. Recent research has demonstrated that there are complex intervening processes, described under the term ‘grammaticalisation’ (or grammaticalisation), and that these are in fact extremely interesting because they have repercussions which extend well beyond the word which has been grammaticalised.

The process whereby a lexical word becomes grammaticalised involves in the first instance some ‘bleaching’ of its semantic content and some reduction in its phonological form. The English future marker *will*/*'ll* (in *I will*/*'ll go*) for instance derives from a full verb in Old English which meant ‘to want’. Similarly, the French future endings 1st sg. *-ai*, 2nd sg. *-as*, 3rd sg. *-a*, etc., go back to the present tense forms of Latin *habere* ‘to have’ found in such constructions as *cantare habeo* ‘sing:INF have:1.SG’ with modal and future meaning (compare in English *I have to go to London tomorrow*). The grammaticalisation of both these words entailed the syntactic and semantic reanalysis of the entire constructions in which it occurred. Thus, in Old English, the ancestor of English *will* behaved much like its cognate *wollen* in present-day German. It could take a lexical object (as in Modern German *ich will ein Haus* ‘I want a house’) and its subject and that of the embedded verb could differ (compare Modern German *ich will, dass er kommt* ‘I want him to come’), etc. Today English *will*/*'ll* is an auxiliary and can only occur with a main verb infinitive. It is conceivable that, at some point in the future, the weak variant *'ll* will become altogether dissociated from the strong form *will* and be reanalysed as a future particle, enclitic to the preceding subject. Such a reanalysis has occurred in French, all connection between the *-ai* of *je chanterai* ‘I will sing’ and the *ai* of *j'ai* ‘I have’ having become completely lost, and here the process has resulted in morphologisation proper, with the auxiliary becoming a person-

number affix on the verb as the head of its construction. In English the link between *will* and its ‘weak form’ *'ll* has not (yet?) broken and both retain their auxiliary status.

Nichols argues that the direct conversion of a syntactic string into morphology, by the suppression of word boundaries, can only occur when the grammaticalised word is the head of its construction. However, although this is the case with an auxiliary in relation to the main verb there may be factors which hinder agglutination (see below). In circumstances when the grammaticalised element is not the head it tends to change its position in the sentence from where it should by syntactically to the prosodically weakest position. This weak position is normally immediately after the first word or the first constituent of the sentence. The grammaticalised element is then said to be in ‘second position’, the favoured place for sentence clitics. (The finite verb in German main clauses is in second position: *Hans kommt morgen*, or *morgen kommt Hans*, both meaning ‘tomorrow John will come’). Nichols predicts that highly grammaticalised elements will, however, ultimately attach themselves to the head of their construction. Thus in the clause it is the verb which becomes the ultimate landing site for clitic pronouns and particles, whereas in the noun phrase it is the noun which attracts to itself demonstratives and articles, which will finally end up as affixed gender markers (Greenberg 1978).

For instance, while it has for long been accepted that the personal affixes of the verb are etymologically personal pronouns that have lost their referential meaning and become agglutinated to it, it has in fact taken more than a century to establish a proper framework that makes explicit the successive stages of the grammaticalisation process. Lehmann (1982) has, on the basis of cross-linguistic and diachronic evidence, established a succession of stages that lead from anaphoric agreement to grammatical agreement. At the one extreme, in anaphoric agreement, a lexical noun phrase is resumed within a discourse by means of a referential pronoun (as in *The man ...; he ...*), which has the syntactic status of a noun phrase. Progressive weakening of the pronominal element affects both its word status and its referential meaning. At the other extreme, in grammatical agreement, we find a morphological marker on the verb with the purpose

of signalling a syntactic relation within the clause. In English, for instance, there is obligatory agreement between the subject noun phrase and the verb (as in *He comes*). Givón has argued that the source construction ultimately developing into agreement morphology must be a stylistically marked structure containing both an anaphoric pronoun and its lexical antecedent (or ‘postcedent’; see Bynon 1993 for details and further references). In the following highly colloquial French sentence for instance, *l(e)* and *ce filme* are coreferential:

- (6) *Tu l' as vu ce filme?*  
 you it have seen this film  
 ‘Have you seen this film?’

It may be supposed that over time such a marked sequence with a right-dislocated topic will become reanalysed as a simple clause, if indeed this has not already happened. The syntactic agreement of the verb with one (or more) of its arguments would, in other words, develop out of an anaphoric agreement holding between a dislocated lexical noun phrase in the role of topic or focus and its resumptive pronoun inside the clause.

Such may indeed be the ultimate outcome, but it is worth noting that at present the construction is head-marked, in both the varieties of French illustrated above and in certain varieties of Spanish which require an obligatory pronominal clitic with animate noun phrases, as in:

- (7) *lo vi a Juan*  
 him saw.I to John  
 ‘I saw John’

At this stage, when a matching lexical noun phrase is not present, the clitic has referential meaning (*tu l'as vu* ‘you have seen it’, *lo vi* ‘I saw him’). The new agreement morphology has thus developed, in these instances at least, by way of head-marking.

In the nominal domain, Lewy (1964) has observed a progression in some European languages from stem-inflection via word-inflection to ‘isolating inflection’. The first phase is found in Latin *capit-is* ‘head-GEN.SG’ of *caput* ‘head’, in which *capit-* is not a word and cannot occur on its own. His second phase is found in German *des Kopf-es* ‘DEF: GEN head-GEN’, with both the noun and the article inflecting for case, gender and number. The third phase is found in French

*de la tête* ‘of DEF:SG.F head’ all three invariable), in which the preposition signals the syntactic relation and the article the grammatical gender of the noun.

The question when grammaticalisation will lead to new morphology and when it will lead to isolating structure is complex, and the determining factors are not fully understood. Nichols (1986) suggests very plausibly that agglutinative morphology is most likely to develop in consistent verb-initial (head-first) and verb-final (head-last) languages because in these, modifiers of the head are located on ‘the other side of’ the site of agglutination and thus do not hinder this process. For, it will be remembered, in a ‘well-behaved’ verb-final language the nominal object precedes the verb and nominal modifiers precede the noun. Agglutination of some cliticised element can thus occur on the side opposite to a modifier, that is post-nominally (case-suffix) and post-verbally (tense-aspect-modality marker). The likeliness of subsequent fusion occurring would appear to depend on phonological factors (prosodic and phonotactic).

Isolating structures on the other hand seem more likely to develop in verb-medial and otherwise inconsistent languages where intervening modifiers tend to impede the ultimate ‘migration’ (Nichols 1986) of highly grammaticalised elements to the head of their construction.

## 7. References

- Bazell, Charles E. (1966), “Linguistic Typology”. In: Strevens, Peter D. (ed.), *Five Inaugural Lectures*. London: Oxford University Press
- Bynon, Theodora (1986), “August Schleicher: Indo-Europeanist and General Linguist”. In: Bynon, Theodora & Palmer, Frank R. (eds.), *Studies in the History of Western Linguistics*. Cambridge: Cambridge University Press, 129–49
- Bynon, Theodora (1993), “Pronominal Attrition, Clitic Doubling and Typological Change”. In: Abraham, Werner (ed.), *Grammaticalization and Reanalysis*, (*Folia Linguistica Historica* 13), 27–64
- Comrie, Bernard (201989), *Language Universals and Linguistic Typology*. Oxford: Blackwell
- Coseriu, Eugenio (1968), “Adam Smith und die Anfänge der Sprachtypologie”. In: Brekle, Herbert E. & Lipka, Leonhard (eds.), *Wortbildung, Syntax und Morphologie: Festschrift zum 60. Geburtstag von Hans Marchand*. The Hague: Mouton, 46–54
- XV. Morphologische Typologie und Universalien
- Coseriu, Eugenio (1972), *Die Geschichte der Sprachphilosophie von der Antike bis zur Gegenwart, Vol. II: Von Leibniz bis Rousseau*. Tübingen: Narr
- Greenberg, Joseph H. (1960), “A Quantitative Approach to the Morphological Typology of Language”. *International Journal of American Linguistics* 26, 178–194
- Greenberg, Joseph H. (1978), “How does a Language Acquire Gender Markers?”. In: Greenberg, Joseph H. et al. (eds.), *Universals of Human Language, Vol. III*. Stanford: Stanford University Press, 47–82
- Hockett, Charles F. (1954), “Two Models of Grammatical Description”. *Word* 10, 210–31 [reprinted in: Joos, Martin (1957, ed.), *Readings in Linguistics*. University of Chicago Press, 386–399]
- Hockett, Charles F. (1961), “Linguistic Elements and their Relations”. *Language* 37, 29–73
- Humboldt, Wilhelm von (1905 [1822]), “Ueber das Entstehen der grammatischen Formen und ihren Einfluss auf die Ideenentwicklung”. In: Leitzmann, Emil (ed.), *Gesammelte Schriften*. Berlin: Behr, 285–313
- Humboldt, Wilhelm von (1836), *Ueber die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Dümmler [english version: (1990), *On language*, translated by Peter Heath. Cambridge University Press]
- Lehmann, Christian (1982), “Universal and Typological Aspects of Agreement”. In: Seiler, Hansjakob & Stachowiak, Franz J. (eds.), *Apprehension, Vol. II*. Tübingen: Narr, 201–267
- Lewy, Ernst (1964), *Der Bau der europäischen Sprachen*. Tübingen: Niemeyer
- Manchester, Martin (1985), *The Philosophical Foundations of Humboldt's Linguistic Doctrines*. Amsterdam: Benjamins
- Matthews, Peter H. (1970), “Recent Developments in Morphology”. In: Lyons, John (ed.), *New horizons in linguistics*. Penguin, 96–114
- Nichols, Johanna (1986), “Head-marking and Dependent-marking Grammar”. *Language* 62, 56–119
- Ramat, Paolo (1995), “Typological Comparison: Towards a Historical Perspective”. In Shibatani & Bynon (eds.), 1–30
- Robins, Robert H. (1960), “In Defence of WP”. *Transactions of the Philological Society* 1959, 116–144 [reprinted in: Robins, Robert H. *Diversions of Bloomsbury*. Amsterdam: North-Holland, 47–77]

- Sapir, Edward (1921), *Language*. New York: Harcourt, Brace & World
- Schleicher, August (1859), "Zur Morphologie der Sprache". *Mémoires de l'Académie des Sciences de St.-Petersburg, series 7, vol. I*, 1–38
- Schleicher, August (²1969), *Die deutsche Sprache*. Stuttgart: Cotta
- Seiler, Hansjakob (1990), "Language Typology in the UNITYP Model". In: Bahner, Werner (ed.),

*Proceedings of the XIVth International Congress of Linguists*. Berlin: Akademie-Verlag, 155–167

Sgall, Petr (1995), "Prague School Typology". In: Shibatani & Bynon (eds.), 49–84

Shibatani, Masayoshi & Bynon, Theodora (1995, eds.), *Approaches to language typology*. Oxford: Oxford University Press

Theodora Bynon, London (United Kingdom)

## 116. Types of morphological structure

1. General remarks
2. The parts of speech
3. Inflection and derivation
4. References

### 1. General remarks

Everything sayable in a spoken human language can be thought of as either the words or music of a solo performer. We are here addressing these two questions: What are possible ways, in a given language, of forming **words**, and what can they mean? What are possible musical tunes in a given language, and what can they mean?

Formally, the contrast between "words" and "music" is the familiar contrast between segmental and suprasegmental messages, but these labels imply as well a conceptual contrast between "what you say" and "how you say it", between linguistic and paralinguistic expression, between thoughts and feelings, perhaps even the distinction between the signs of a language and prelinguistic symptoms. And this implied contrast is misleading: not only is it true that in all languages with phonemic stress or phonemic tone, the formal dimension we are calling "music" clearly encodes "what you say" rather than "how you say it". In fact, in all languages with non-random intonation patterns, "music" is cognitive rather than affective. Yet there is an intuitive sense in which many people may feel that there is something right about the metaphor of words and music, and its familiar non-technical implications. Nor is this intuition entirely unjustified, as the following considerations may attest.

The most fundamental property of human languages is that they are systems of com-

munication which employ signs rather than symptoms. Both signs and symptoms have meanings, but the distinguishing property of signs is that we can use them not only to express, but to conceal our thoughts, and to tell lies. Involuntary symptoms cannot help but tell the truth: in fact, they are the truth. In this (profound) sense, language is a matter of convention, politeness, and etiquette. Sincerity is possible, but telling it like it is is only one choice among several; and, as Talleyrand suggested, not the favourite choice.

To the question whether "words" are symptomatic, we can confidently answer "except possibly in glossolalic utterances, never". That is, there are no segmentally represented words in any recognized, reproducible, coded language, from the most passionately uttered expletive to the most sincere protestations of undying love which the speaker does not consciously choose, both over other words or over silence. All "words" are signs.

If we ask the same question about "music", our answer must be much more tentative. Granted: A good actor can mimic and thus exercise control over every shade and tone of voice, and all of us can do this to varying degrees as individuals, and are required to do it to a large extent as competent speakers of different human languages. (For example, the familiar rising intonation of questions is not only a symptom of curiosity, but a part of the grammar of many languages.) Nevertheless, if we want to learn how a speaker really feels, we try to attend to factors like hoarseness, quavering, and intensity, musical aspects of the message like pitch and amplitude, rather than the actual words that are spoken. Rightly or wrongly,

we believe we can read the speaker's moods and personality from the music of his utterance, even when he is speaking a language we do not understand. We do this because we share the sense that the music of an utterance has more potential for sincerity (or sliding out from under both the speaker's conscious control, and the grammatical control of language-specific conventions) than the words do. Summing up: it is as if we believed that all words are signs and that there are no symptomatic words, but there are tune-symptoms (as well as conventional, culturally constructed tune-signs).

Another fundamental principle of morphological structure, that of **double articulation**, asserts that the smallest meaningful units (morphemes) of language, are made up of smaller meaningless units, or phonemes. In real music, as well, a melody whose "meaning" may be the feelings it evokes, or its aesthetic appeal in general, is made up of individual notes which convey no mood and have no intrinsic beauty. Does this structural principle apply, however, to the "music" of language? If so, what are the musical morphemes, and what are the phonemes they consist of? Double articulation allows us to distinguish suprasegmental phonemes (tone in Chinese or stress in English) from morphemes (the melody which signals a specific meaning like interrogation, or the added amplitude which signals emphasis). So it would seem that double articulation is a fact about both "word" and "music" in spoken languages. But there are still three important differences between the two.

First, musical morphemes, from one-note melodies like high tone for future tense in Hua or singular number in Shilluk (Sapir 1921: 80), to melodic clichés like the stereotyped (L)HM of "Oh Susan" in English (Ladd 1978; Liberman 1979) are formally rudimentary: they are extremely short, there are very few of them, and the range of things they can express is pitifully small, perhaps not even substantially different from the range of emotions that can be communicated by facial expressions. There is no language, for example, in which it is possible to express ideas like "my parents were poor but honest" through musical means alone.

There is a further limitation on musical morphemes: like facial expressions, but unlike musical phonemes, and unlike segmental morphemes, they tend to be motivated. This

also suggests that music in language is less a sign, more a symptom.

Finally, there is abundant evidence that on the segmental plane, words are worn down, first to grammatical affixes, and then to phonemes, by a general process of erosion (Givón 1979; Haiman 1985; Heine et al. 1991; Hopper & Traugott 1993; Hopper 1994). It seems that there is hardly any evidence of a parallel **grammaticalization** process on the suprasegmental plane: phonemic tone, where its origins have been reconstructed with any confidence, seems to arise by phonologization of originally epiphenomenal variations induced by contrasting segmental contexts which were subsequently neutralized: high tone after voiceless consonants, low tone after voiced in Chinese, and so on. **Tone** does not ever seem to have originated as a morphemic melody which then became ritualized. **Phonemic stress** is typically derivative of segmental structures such as syllable weight. Hyman (1977: 44) has suggested that lexical stress patterns may be frozen or ritualized sentence intonation patterns:

"Intonation becomes grammaticalized as word-stress when the suprasegmental features of pitch, duration, and intensity that would have characterized a word in isolation are encoded within the word, and thus come to function in words not in isolation." (Hyman 1977: 44)

This ingenious hypothesis does assert a parallelism between grammaticalization in segmental and suprasegmental domains, and would suggest that morphemic sentence-intonation patterns in general might result in phonemic word-stress patterns. But if "utterance prosodies will be generalized to clause, phrase, and word boundaries" as Hyman suggests (1977: 55), then one might propose (for example) that word-initial ictus would occur in languages whose word-order was such that focussed items came sentence-initially, and word-final ictus may occur in those languages where the focussed element came last: but the evidence for such a plausible correlation is too sketchy to carry conviction, in fact, non-existent.

Morphological structure then is basically the structure of **words**, and the ways in which words of different types are identified: formally, we distinguish between affixation and compounding, and, within affixation, among prefixation, suffixation, infixation, and circumfixing (see Art. 54). Functionally, we distinguish among different kinds of compound-

ing (endocentric vs. exocentric; among endocentric compounds, between coordinate (*dvandva* or *amredita*) and attributive (*tatpu-rusha*) compounds and between inflection and derivation, still best summarized in Bloomfield (1933) (cf. Art. 86). The parts of speech may be (partially) identified through their phonological structure, as well as through the nature of the affixes with which they occur (cf. Art. 69–72).

## 2. The parts of speech

Wittgenstein once compared words of a language with the pedals on a car: formally indistinguishable, but doing totally different kinds of work. The analogy is well taken. Assuming that the **parts of speech** correspond (albeit roughly) to the brake, clutch, and accelerator, it is striking how formally undifferentiated noun and verb stems are in spoken languages. Even where they are formally distinguished, the phonological flagging is only partial. For example, finite verbs in Classical Greek invariably have retracted stress, while only some nouns do; nouns and adjectives in English have retracted stress while otherwise homophonous verbs do not (*reject*, *degenerate*, etc.), but the pattern does not extend to nouns and verbs in general; in Hua, all proper nouns end in the glottal stop, but only some common nouns do, and all verb stems end in one of the vowels /i,o,u/ but only some common nouns do (Haiman 1980). Very generally, pronouns (like function words in general) tend to be short and relatively unstressed, while common nouns can be either long or short, a function of Zipf's Law (Zipf 1935). In none of the cases is a part of speech clearly and unambiguously marked by its phonological structure alone. Even the distinction between words and phrases, signalled in English by the difference between the compound stress rule and the nuclear stress rule (*blackbird* and *nobody* vs. *black bird* and *no body*, cf. Bloomfield 1933: 228 f.) is contingent rather than absolute: the compound word can only be stressed on the first syllable, but the phrase may be stressed on either word.

Nevertheless, some of these essentially fortuitous phonological distinctions may become ritualized (that is, emancipated from their original motivation) in morphology and syntax.

The relative lightness of pronouns interacts with Behaghel's law of increasing bulk (Behaghel 1932) in the Romance languages to

produce the familiar contrast between [Verb + NP object] and [pronoun object + Verb]. The same syntactic contrast – emancipated from its original phonological motivation – arises between incorporated and non-incorporated object noun phrases in languages like Cree (Wolfart 1973) and Guarani (Velazquez-Castillo 1993) (noun phrase incorporation, like phonological reduction, is a function of frequency in both of these languages, cf. Sapir 1911: 264).

The change in word order also reflects a change in the status of the compound expression: [Verb + NP] is a phrase, but [NP + Verb] is a word, as in the English pairs *keep house* (nuclear stress) vs. *housekeep* (compound stress). Whether or not word order is changed, phrases degenerate into compounds through **institutionalization**, which is generally once again a consequence of frequency (cf. Art. 80).

In other languages such as Hungarian and Irish, the relative lightness of pronouns leads to their acquiring affix status as opposed to stem status, leading to the contrast between [noun phrase + postposition affix] and [postposition stem + pronoun affix], as in Hungarian *élet-ben* 'life-in' vs. *ben-em* 'in-my (in me)'.

For the most part, however, the parts of speech are most reliably flagged by purely morphological rather than phonological criteria: nouns by affixes of case, verbs by affixes of tense, and so forth. The stems themselves may be quite indeterminate, their eventual syntactic function being decided and signalled by the affixes with which they are allowed to occur in speech. In Hua, for example, the potential topic suffix *-mo* occurs on phrases which are functioning as noun phrases in the utterances where they appear: this means that it is allowed on verbs which are functioning as protasis clauses (but not on final verbs which are functioning as complete utterances or on relative clauses which are functioning as adjectives), and it is allowed on nouns in the nominative, ergative, locative, or benefactive cases (but not on nouns in the genitive case, which is invariably attributive, or on nouns in the vocative "case", which invariably are complete utterances).

## 3. Inflection and derivation

"Inflection is not a way of making new words but a way of doing something to words we already have. It is manipulative, not creative." (Bolinger 1975: 117)

There are no languages known in which the distinctions between derivational and inflectional morphemes are signalled, even partially, by their phonological properties. Morphologically, there are partial flaggings: it has been repeatedly noted that derivational affixes tend to occur closer to the stem than do inflectional ones (which of course, begs the question); and in English, inflectional affixes where they are agglutinative, are exclusively suffixes, while derivational affixes are both prefixing and suffixing (begging the same question). As is widely known, the distinction between the two kinds of morphemes is relatively hard to draw in any way. In the recent theoretical treatment of the subject (Bybee 1985: chapter 4), it has in fact been proposed that the distinction should not be clearcut at all, but that there exists a lexical/derivational/inflectional continuum, with the identification of an affix being dependent on a consideration of both its meaning and form. In the spirit of Bybee's suggestion, a fairly sharp distinction can be drawn between derivational and inflectional morphology, based on the fact that – unlike inflectional categories – derivational “categories” can not be given semantic definitions. The contrast may be illustrated by a simple experiment. Given an inflectional category like **plural** in English, we can say with confidence that the category is expressed by affixation in some cases (*dog-s*), by synthesis in others (*geese*), or by suppletion in others (*people*); we can even say the category is expressed in problematic cases by periphrasis (*more than one mongoose*). The reason for our confidence is that there exists a semantic definition of the category which is independent of its morphological expression. This we cannot do with “derivational categories”, for which only a purely formal definition is possible. By this criterion, the comparative in English is inflectional, aspect in Russian is derivational, and so forth. This is to restate Bolinger's observation cited in the epigraph to this section: derivational morphology is creative in the sense that it creates new meanings in every case, and thus no semantic common definition can be provided for a derivational affix.

#### 4. References

- Behaghel, Otto (1932), *Deutsche Syntax, Vol. IV.* Heidelberg: Carl Winter  
 Bloomfield, Leonard (1933), *Language*. New York: Holt, Rhinehart & Winston

#### XV. Morphologische Typologie und Universalien

- Bolinger, Dwight (1975), *Aspects of Language*. New York: Harcourt, Brace, Jovanovich [1968]  
 Bybee, Joan L. (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins  
 Givón, Talmy (1979), *On Understanding Grammar*. New York: Academic Press  
 Haiman, John (1980), *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam: Benjamins  
 Haiman, John (1985), *Natural Syntax*. Cambridge: University Press  
 Heine, Bernd & Claudi, Ulrike & Hünnemeyer, Friederike (1991), *Grammaticalization: A Conceptual Framework*. Chicago: The University of Chicago Press  
 Hopper, Paul J. (1994), “Phonogenesis”. In: Pagliucca, William (ed.), *Perspectives on Grammaticalization*. Amsterdam: Benjamins, 29–45  
 Hopper, Paul & Traugott, Elizabeth Closs (1993), *Grammaticalization*. Cambridge: Cambridge University Press  
 Hyman, Larry M. (1977), “On the Nature of Linguistic Stress”. In: Hyman, Larry M. (ed.), *Studies in Stress and Accent*. Los Angeles: Department of Linguistics, University of Southern California (Southern California Occasional Papers in Linguistics 4), 37–82  
 Ladd, Robert (1978), “Stylized Intonation”. *Language* 54, 517–539  
 Liberman, Mark (1979), *The Intonational System of English*. New York: Garland  
 Sapir, Edward (1911), “The Problem of Noun Incorporation in American Languages”. *American Anthropologist* 13, 250–282  
 Sapir, Edward (1921), *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & World  
 Velazquez-Castillo, Maura (1993), *The Grammar of Inalienability: Possession and Noun Incorporation in Paraguayan Guarani*. Ph.D. dissertation, University of California, San Diego  
 Wolfart, H. Christoph (1973), *Plains Cree: A Grammatical Study*. Philadelphia: American Philosophical Society  
 Zipf, George Kingsley (1935), *The Psychobiology of Language*. Boston: Houghton Mifflin

*John Haiman, St. Paul (U.S.A.)*

## 117. Quantitative Typologie

1. Die quantitativen morphologischen Indizes von J. H. Greenberg
2. Zur Weiterentwicklung des Greenbergschen Ansatzes
3. Taxonomische morphologische Sprachklassifikation und Suche nach Zusammenhängen
4. Quantifizierung der Eigenschaften grammatischer Subsysteme
5. Zitierte Literatur
  
1. Die quantitativen morphologischen Indizes von J. H. Greenberg

Die quantitative morphologische Typologie, wie sie sich in den vergangenen 30–40 Jahren entwickelt hat, kann als Nachfolgerin und ‘Erbin’ der morphologischen Sprachtypologie des 19. Jahrhunderts betrachtet werden. Das zeigt sich besonders deutlich bei J. H. Greenberg, dessen Arbeit “A Quantitative Approach to the Morphological Typology of Language” (Greenberg 1960) den eigentlichen Beginn der Entwicklung der quantitativen morphologischen Typologie bezeichnet. Ineichen bezeichnet diesen Aufsatz als “wegweisend, weil er seinerzeit die Dimensionen der quantifizierenden Typologie in den Grundzügen absteckte” (Ineichen 1979: 84). Es sollte freilich beachtet werden, daß auch andere Autoren Vorschläge zur quantifizierenden Erfassung morphologischer Eigenschaften erarbeitet haben, so etwa Thorndike (1943), Hamp (1958), Kelemen (1970) und Robins (1965). Ein ganzes System von quantifizierbaren Eigenschaften des Morphems ist von Skalička (1935), Slávičková (1968) und Sil'nikič (vgl. Silnitsky 1998) entwickelt worden.

Greenberg betrachtet die morphologische Wortstruktur als “something of fundamental importance to the over-all characterization of a language” (Greenberg 1960: 180) und stellt sich die Aufgabe, “to re-examine the nineteenth-century approach to the problem” (Greenberg 1960: 180; vgl. auch 1960: 179). Das Verbindungsglied zwischen Greenbergs Ansatz und der älteren morphologischen Typologie ist die Sprachtypologie Edward Sapir's, an die Greenberg unmittelbar anknüpft (s. auch Art. 115):

“My [morphological] typology is a pragmatic approach using Sapir's attribute space with some modification.” (Greenberg 1957: 72; vgl. auch 1960: 187 f.)

Ungeachtet aller wichtigen Modifikationen an den Verfahren der ‘klassischen’ morphologischen Typologie, die Sapir vornimmt, läßt auch er sich leiten von einer “logic of metonymy in which a few parts (i. e. a low number of morphosyntactic features) represent the whole” (Voegelin et al. 1960: 198).

Greenberg (1960) schlägt zehn Indizes vor, mit deren Hilfe verschiedene morphologische Eigenschaften natürlicher Sprachen in Form quantitativer Begriffe erfaßt werden können. Damit formalisiert er aber nicht nur den Ansatz von Sapir, sondern erweitert ihn um die Dimension einer pragmatischen Frequenzuntersuchung, da sich die Indizes auf die Vorkommenshäufigkeit von morphologischen Einheiten in Texten bestimmter Länge beziehen.

Die zehn Indizes von Greenberg entsprechen insgesamt fünf morphologischen Parametern, von denen drei von Sapir übernommen sind. Zwei der Sapischen Parameter sind formaler Natur, und so bezieht sich Greenberg ohne weiteres auf sie, nämlich die Parameter “Synthesis” und “Technique”, die allerdings bei Sapir nur eine nachgeordnete Rolle spielen.

Auf Sapis Parameter “Synthesis” bezieht sich der bei Greenberg an erster Stelle stehende “index of synthesis” (M/W), der die “gross complexity of the word” mißt, d. h. “the degree of complexity exhibited on the basis of the number of subordinate meaningful elements it contains” (Greenberg 1960: 182). Dieser Index bezieht die Anzahl der Morpheme eines Textes auf die Anzahl der Wörter. Sein kleinstmöglicher Wert – für den Fall, daß W = M – ist 1; je größer er ist, desto ‘synthetischer’ ist die betreffende Sprache.

Auf Sapis Parameter “Technique” bezieht sich Greenbergs “index of agglutination” (A/J), der die Anzahl der agglutinierenden morphematischen Junkturen zur Anzahl der morphematischen Junkturen überhaupt ins Verhältnis setzt. Hier ist 1 der höchste Indexwert. Je weiter sich der für eine Sprache ermittelte Indexwert von 1 entfernt, um so weniger **agglutinierend**, d. h., um so **fusionierender** ist diese Sprache. Der Unterschied zwischen agglutinierenden und fusionierenden Sprachen ist der “between languages in which the subordinate elements are added to the root elements mechanically, i. e., without ei-

ther of the elements being modified (...), and those involving a process of fusion by which the constituent elements become difficult to recognize and separate" (Greenberg 1960: 183).

Der Hauptgesichtspunkt von Sapiers Schema ist der des Ausdrucks bestimmter Begriffe: "radical concepts" (I), "derivational concepts" (II), "concrete relational concepts" (III), "pure relational concepts" (IV) (vgl. hierzu Sapir 1949, Ch. V, v. a. S. 101) (s. Art. 115). Greenberg kritisiert Sapir: "..., Sapir seems to be talking about concepts, but in reality his test is a formal, not a semantic one" (Greenberg 1960: 184). Daher modifiziert Greenberg diesen Sapierschen Hauptparameter, indem er sich zu einem formalen Vorgehen entschließt (vgl. Greenberg 1960: 184): "..., the present treatment is based on the possibility of the exhaustive division of morphemes into three classes, root, derivational, and inflectional" (Greenberg 1960: 186).

Die Quantifizierung von Sapiers formalisiertem Hauptparameter führt Greenberg zur Aufstellung von drei Indizes: der "compounding index" (R/W) bezieht die Anzahl der **Wurzeln** in einem Text auf die Anzahl der Wörter; der "derivational index" (D/W) setzt die Anzahl der **Ableitungsmorpheme** eines Textes zur Anzahl der Wörter in Beziehung, und der "gross inflectional index" (I/W) bezieht die Anzahl der **Flexionsmorpheme** in einem Text auf die Anzahl der Wörter.

Den drei Sapierschen Parametern fügt Greenberg zwei weitere hinzu. Der erste davon "is the order of subordinate elements in relation to the root", "a topic discussed by Sapir as important for the morphological structure of a language but not included in his final formulation" (Greenberg 1960: 186). Die beiden diesem Parameter entsprechenden Indizes – der "prefixal index" P/W und der "suffixal index" S/W – beziehen die Anzahl der **Präfixe** bzw. der **Suffixe** in einem Text auf die Anzahl der Wörter.

"The final parameter has to do with the devices employed for relating words to each other. It therefore brings in syntactical as well as morphological considerations. There are three devices that language may use, inflectional morphemes without concord, significant order, or concord (agreement)." (Greenberg 1960: 187)

So ergeben sich drei weitere Indizes: der "isolational index" O/N bezieht die Anzahl der lediglich durch die Stellung signalisierten syntaktischen Konnexions ("order") auf die

Anzahl der syntaktischen Konnexions ("nexus"). Der "pure inflectional index" Pi/N bezieht die Anzahl der durch Flexion (aber nicht durch Kongruenz) signalisierten syntaktischen Konnexions auf die Anzahl aller Konnexions. Der "concordial index" Co/N schließlich setzt die Anzahl der durch Kongruenz ("concord") signalisierten Konnexions in einem Text zur Anzahl aller Konnexions in ein Verhältnis.

Wie aus den vorstehenden Angaben ersichtlich ist, haben wir es bei der Anwendung der Greenbergschen Indizes mit dem Fall einer indirekten oder abgeleiteten Messung zu tun, die darauf beruht, daß die Resultate von jeweils zwei direkten Messungen zueinander in Beziehung gesetzt werden. Die direkten Messungen bestehen in der Abbildung einer Eigenschaft auf eine Maßskala, im vorliegenden Fall, wo die Anzahl von Morphemen, Wörtern, Junkturen usw. ermittelt wird, auf die Menge der nichtnegativen Zahlen.

In ihrer Gesamtheit bilden die zehn Indizes einen "multidimensional attribute space" (Greenberg 1957: 76), in dem jede untersuchte Sprache lokalisiert werden kann. Die Auswahl der gerade hier gewählten Attribute steht unter der Annahme, "that they represent closely related aspects of the same general class of phenomena which may reasonably be expected to reveal significant mutual connections" (Greenberg 1957: 76). Greenberg wendet sein Verfahren an, indem er die Werte seiner zehn Indizes für acht Sprachen – Sanskrit, Angelsächsisch (Altenglisch), Persisch, Englisch, Jakutisch, Swahili, Anna-mitisch, Eskimo – anhand von Textstichproben berechnet, die jeweils 100 Wörter lang sind. Es handelt sich also um eine **pragmatische Frequenzuntersuchung**. Bei einer solchen Untersuchung wird die Häufigkeit einer Einheit aus Texten ermittelt, während sich die systematische Häufigkeit auf ein Inventar von Einheiten bezieht.

## 2. Zur Weiterentwicklung des Greenbergschen Ansatzes

Greenbergs Aufsatz von 1960 bezeichnet den eigentlichen Beginn der Entwicklung der quantitativen morphologischen Typologie (vgl. etwa die Anwendung der Greenbergschen Indizes auf die Klassifikation der romanischen Sprachen bei Contreras 1963). Diese Entwicklung bezog ihren Impetus weitgehend aus der Auseinandersetzung mit

Greenbergs Ansatz. Sie ist in verschiedenen Richtungen verlaufen, von denen einige im folgenden behandelt werden sollen.

Einige Linguisten haben Greenbergs **Indizes** um weitere Indizes ergänzt. So schlägt Cowgill (¹1976: 123) einen “index of infixes per word” (Inf/W) vor. Auch andere Typologen – etwa Ju. Ja. Glazov (1965) und M. Mejlah (1973) – verwenden in ihren Untersuchungen den “Infigierungsindex”. V. B. Kasevič & S. E. Jachontov (1982, Hrsg.: 37) betrachten die Nichtunterscheidung von autosemantischen und synsemantischen Wörtern (s. Art. 27) als wichtigen Mangel von Greenbergs Typologie und schlagen deshalb einen “Index der Analytizität” (Aux/W) vor, der die Anzahl der synsemantischen Wörter in einem Text auf die Anzahl aller Wörter bezieht (vgl. auch Sil'nickij et al. 1986: 106). Sil'nickij (vgl. Silnitsky 1993: 159) kommt allerdings zu dem Ergebnis, daß dieser zusätzliche Index keinerlei signifikante klassifikatorische Funktion erfülle und deshalb nicht weiter zu berücksichtigen sei. Von Pierce (1962) stammt ein Index C/W, der die Anzahl der Konsonanten in einem Text auf die Anzahl der Wörter bezieht, also die durchschnittliche Anzahl von Konsonanten pro Wort mißt (s. Art. 44).

Kasevič & Jachontov (1982, Hrsg.) modifizieren gleichfalls einige Bestimmungen der durch die Indizes erfaßten sprachlichen Einheiten, beispielsweise indem sie den Begriff der Agglutination weiter fassen als Greenberg und die Worteinheit anders bestimmen (vgl. Greenberg 1982: 3). Dies führt natürlich zu einer Veränderung der Indexwerte. M. Mejlah (1973: 165) berücksichtigt in seinem System nur die ersten vier von Greenbergs fünf Parametern. Den fünften vernachlässigt er deswegen, weil die zu ihm gehörenden Indizes “im Bereich der reinen Morphologie” fakultativ, in dem der Syntax hingegen unzureichend seien.

A. L. Krober (1960) machte die von Greenberg für acht Sprachen berechneten Indexwerte zur Grundlage einer “rank order” zwischen diesen Sprachen, ohne Greenbergs Vorgehen selbst in irgendeiner Hinsicht zu modifizieren. Jede Sprache bekommt bezüglich jeden Indexes einen Rang zwischen 1 und 8 zugeordnet, der sich jeweils aus dem Vergleich der Werte für einen Index ergibt. Beispielsweise erhält das Eskimo im Hinblick auf den “index of synthesis” den Rang 1, weil es für diesen Index unter allen betrachteten Sprachen den höchsten Wert aufweist. Die

Gruppierungen, die sich aus der sich so ergebenden Rangtabelle ablesen lassen, versucht Krober (1960) zu den Klassen der morphologischen Sprachtypologie des 19. Jahrhunderts in Beziehung zu setzen und – für die indogermanischen Sprachen – entwicklungsgeschichtlich zu interpretieren.

Eine weitere, vieldiskutierte Frage ist die nach der **Stichprobengröße**. Greenberg selbst hatte sich mit Texten im Umfang von jeweils 100 Wörtern begnügt. Es fragt sich, ob aufgrund derartig kleiner Stichproben signifikante Daten gewonnen werden können, d. h. Aussagen, die jeweils für die betreffende Sprache gelten. Auf dieses Problem macht bereits Cowgill aufmerksam (vgl. auch van Son & Peters 1989: 20):

“Another weakness of the material used here is the small size of the samples, 100 words of text from each language. ... One of the urgent desiderata of current typology is to determine the minimum size of sample needed to be reasonable sure of having an accurate picture of a language.” (Cowgill ¹1976: 116f.)

Mit diesem Problem hat sich J. E. Pierce beschäftigt. Pierce will folgende Frage lösen: “how large a sample is necessary for computation of indices and how should this sample be secured, i. e., a series of small samples drawn from different texts or one large one, and if the former, how small” (Pierce 1966: 44). Zu diesem Zweck werden drei von Greenbergs Indices – M/W, R/W, S/W – für drei verschiedene Stichproben aus türkischen Texten berechnet: (a) für 14 Stichproben von je 100 Wörtern, wobei diese Stichproben vier Stilarten und sechs Autoren repräsentieren; (b) in zehn Stichproben von je 200 Wörtern aus denselben Quellen wie unter (a), jedoch so, daß eine Überschneidung vermieden wurde; (c) in zehn Stichproben von je 100 Wörtern aus einem zusammenhängenden, 1000 Wörter umfassenden Text.

Pierce (1966) testet die Differenzen zwischen den Indexmittelwerten seiner drei Stichprobenmengen auf ihre Signifikanz. Die Differenz zwischen den Mittelwerten der ersten und der dritten Stichprobenmenge erweist sich als hochsignifikant, d. h., mehrere 100-Wörter-Stichproben aus mehreren verschiedenen Texten sind für das Türkische als Ganzes repräsentativer als ebensolange Stichproben aus einem zusammenhängenden Text. Die Differenz zwischen den Mittelwerten der 100- und der 200-Wörter-Stichproben ist nichtsignifikant. “This would seem to indi-

cate that Greenberg's 100-word-sample size, arbitrary though it may have been, was large enough, i. e. nothing of significance was gained by increasing the sample size to 200-words, ..." (Pierce 1966: 47; u. a. auf Pierce beruft sich Sil'nickij zur Begründung seines Entschlusses, die Werte von Greenbergs Indizes für Texte von ca. 100 Wörtern Umfang zu berechnen; vgl. Silnitsky 1993: 140). Davon unberührt bleibt natürlich der Umstand, daß statistische Signifikanz nur dann erreicht werden kann, wenn man nicht nur, wie Greenberg das getan hat, eine einzige Stichprobe untersucht.

Letzteres Problem ist von A. V. Stepanov aufgegriffen worden, der in einer 1995 erschienenen Arbeit konstatiert, die Anzahl der bisher untersuchten Textstichproben sei nicht ausreichend gewesen, "to make conclusions about real statistical distribution of index values for any index" (Stepanov 1995: 143). Um diesem Mangel abzuhelpfen, schlägt Stepanov vor, der Untersuchung der Greenbergschen Indizes für jede hierfür ausgewählte Sprache ein großes Textkorpus zugrundezulegen, das alle möglichen Textformen und -arten berücksichtigen solle: "... the corpus will include hundreds of texts from all existing genres, sources, historical periods etc. as one large sample" (Stepanov 1995: 144).

In bezug auf ein solches Korpus soll nun nicht einfach für einen gegebenen Index dessen Durchschnittswert berechnet werden, vielmehr soll die Verteilungsfunktion eines jeden Indexes durch die Parameter Mittelwert, Standardabweichung und relativer Fehler beschrieben werden. Dieses Vorgehen ermögliche es, zu analysieren, in welcher Weise zufällige und systematische Faktoren das von dem jeweiligen Index repräsentierte morphologische Merkmal in einer bestimmten Sprache beeinflußten. Jedesmal nämlich, wenn die empirische Verteilung nicht einer Normalverteilung entspreche, trage für diese "Störung" ein systematischer Faktor die Verantwortung, der anschließend ermittelt werden müsse. Die exemplarischen Analysen von Texten aus einigen semitischen Sprachen vermögen allerdings nicht zu überzeugen, da Stepanov seine empirisch gewonnenen Verteilungen nicht mit Hilfe der einschlägigen statistischen Verfahren testet, sondern nur impressionistisch auswertet. Seine an und für sich anregenden Vorschläge bedürfen daher weiterer Präzisierung.

Eine wichtige, leider aber noch immer nicht hinreichend beachtete Modifikation des Greenbergschen Ansatzes stammt von V.

Krupa (1965). Dieser kritisiert, daß Greenberg nicht die mathematischen Eigenschaften seiner Indizes untersucht habe. Diese seien inkommensurabel, da ihre Werte in drei unterschiedliche Intervalle fielen:

- (a) Die Werte der Indizes A/J, O/N, Pi/N und Co/N fallen in das geschlossene Intervall  $<0; 1>$ , weil  $A \leq J$ ,  $O \leq N$ ,  $Pi \leq N$  und  $Co \leq N$ .
- (b) Die Werte der Indizes M/W und R/W fallen in das halbgeschlossene Intervall  $<1; 0>$  – wobei 0 natürlich nur eine theoretische Größe ist –, weil  $M \geq W$  und  $R \geq W$ .
- (c) Die Werte der Indizes I/W, D/W, P/W, Inf/W und S/W bewegen sich – theoretisch – zwischen 0 und 1, d. h. innerhalb des halbgeschlossenen Intervalls  $<0; 1>$ , weil  $I \geq W$ ,  $D \geq W$ ,  $P \geq W$ ,  $Inf \geq W$ ,  $S \geq W$  oder  $I < W$ ,  $D < W$ ,  $P < W$ ,  $Inf < W$  und  $S < W$ .

So ist also beispielsweise der höchste Wert von A/J gleichzeitig der niedrigste Wert von M/W; d. h., die Werte beider Indizes können nicht unmittelbar miteinander verglichen werden. Um diese Komplizierung zu vermeiden, schlägt Krupa (1965) vor, die Indizes so zu konstruieren, daß sämtliche Werte in das Einheitsintervall  $<0; 1>$  fallen. Diese Bedingung ist für die unter (a) genannten Indizes bereits erfüllt. Die übrigen Indizes sind wie folgt zu ersetzen: M/W → W/M; R/W → W/R, D/W → D/M, I/W → I/M; P/W → P/M; Inf/W → Inf/M; S/W → S/M.

Der neue Index W/M mißt nicht, wie M/W, direkt den Grad des **Synthetismus**, sondern den des **Analytismus**. Dieser ist am höchsten, wenn  $M = W$ . Der Synthetismus berechnet sich komplementär als  $1-W/M$ . W/R ist ein direktes Maß der Einwurzeligkeit von Wörtern. Der Komplementärindex  $1-W/R$  ist ein Maß für den Grad der Komposition. Die Ersetzung von D/W durch D/M bedeutet, daß die Anzahl der Derivationsmorpheme nicht mehr auf die Anzahl der Wörter, sondern die der Morpheme im Text bezogen wird; entsprechend bei den übrigen Indizes.

Obwohl die Werte aller Indizes jetzt in das Einheitsintervall fallen, gliedern sich die Indizes in drei Gruppen, und zwar unter dem Gesichtspunkt der Fähigkeit bzw. der Unfähigkeit, die Extremwerte der Skala tatsächlich zu erreichen. Beispielsweise kann der Index A/J sowohl den Wert 0 wie den Wert 1 annehmen, da es in einem Text gar keine agglutinierenden morphematischen Junkturen zu

Index	Bereich	Intervall	Autor
1 W/M	Analytismus	(0; 1>	G./mod. Krupa
2 A/J	Agglutination	<0; 1>	Greenberg
3 W/R	Komposition	(0; 1>	G./mod. Krupa
4 D/M	Derivation	<0; 1)	G./mod. Krupa
5 I/M	Flexion	<0; 1)	G./mod. Krupa
6 P/M	Präfixation	<0; 1)	G./mod. Krupa
7 S/M	Suffixation	<0; 1)	G./mod. Krupa
8 O/N	Morphosyntax	<0; 1>	Greenberg
9 Pi/N	Morphosyntax	<0; 1>	Greenberg
10 Co/N	Morphosyntax	<0; 1>	Greenberg

W = Anzahl der Wörter  
 M = Anzahl der Morpheme  
 A = Anzahl der agglutinierenden Konstruktionen  
 J = Anzahl der morphematischen Junkturen  
 R = Anzahl der Wurzelmorpheme  
 D = Anzahl der Derivationsmorpheme  
 I = Anzahl der Flexionsmorpheme  
 P = Anzahl der Präfixe  
 S = Anzahl der Suffixe  
 O = Anzahl der Orders  
 N = Anzahl der Nexus  
 Pi = Anzahl der durch Flexion signalisierten syntaktischen Konnexions  
 Co = Anzahl der Kongruenzen

Tab. 117.1: Morphologische Indizes der Sprachtypologie

geben braucht bzw. die Anzahl der agglutinierenden morphematischen Junkturen gleich der Anzahl der morphematischen Junkturen überhaupt sein kann. Der Index D/M hingegen kann nur den Extremwert 0 erreichen – etwa in analytischen Sprachen ohne jegliche Derivationsmorpheme –, nicht aber den Extremwert 1, da es keine Texte geben kann, die nur aus Derivationsmorphemen beständen (weitere Einzelheiten vgl. bei Krupa 1965: 34). Eine Übersicht über die Indizes zeigt Tab. 117.1. Die für 20 Sprachen berechneten Werte der Greenberg-Krupaschen Indizes sind in Tab. 117.2 zusammengestellt (nach Altmann & Lehfeldt 1973: 40).

Mit den von Krupa modifizierten Indizes arbeiten etwa Altmann & Lehfeldt (1973: 39 ff.), van Son & Peters (1989) u. a. Aber noch in dem von V. B. Kasevič und S. E. Jachontov (1982, Hrsg.) herausgegebenen Band, der ganz der Anwendung von Greenbergs Ansatz gewidmet ist, werden Greenbergs Indizes in ihrer ursprünglichen Form verwendet und die Werte der Indizes ohne Zuhilfenahme mathematischer Verfahren interpretiert. Sil'nickij et al. (1986) verwenden den „index of synthesis“ und den „compounding index“ in der von Krupa vorgeschlagenen Form und ersetzen im Nenner der Indizes O/N, Pi/N und Co/N N durch W (so auch

	W/M	A/J	W/R	D/M	I/M	P/M	S/M	O/N	Pi/N	Co/N
Sanskrit	0.39	0.09	0.88	0.24	0.32	0.06	0.46	0.16	0.46	0.38
Bengali	0.53	0.46	0.92	0.15	0.28	0.01	0.42	0.57	0.29	0.14
Altpersisch	0.41	0.20	0.98	0.17	0.41	0.08	0.50	0.23	0.39	0.38
Neopersisch	0.66	0.34	0.97	0.07	0.26	0.01	0.32	0.52	0.29	0.19
Altgriechisch (Homer)	0.48	0.10	0.99	0.10	0.41	0.03	0.48	0.48	0.27	0.26
Neugriechisch	0.55	0.40	0.98	0.07	0.37	0.02	0.42	0.53	0.21	0.26
Altenglisch	0.47	0.11	1.00	0.09	0.42	0.03	0.48	0.15	0.47	0.38
Neuenglisch	0.60	0.30	1.00	0.09	0.32	0.02	0.38	0.75	0.14	0.11
Jakutisch	0.46	0.51	0.98	0.16	0.38	0.00	0.53	0.29	0.59	0.12
Suaheli	0.39	0.67	1.00	0.03	0.31	0.45	0.16	0.40	0.19	0.41
Vietnamesisch	0.94	–	0.93	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Eskimo	0.27	0.03	1.00	0.34	0.47	0.00	0.73	0.02	0.46	0.38
Türkisch (Schrift)	0.43	0.60	1.00	0.11	0.43	0.00	0.54	0.43	0.67	0.20
Türkisch (Mündl.)	0.57	0.67	0.96	0.06	0.38	0.00	0.44	0.69	0.16	0.03
Gotisch	0.43	0.19	0.97	0.13	0.42	0.04	0.52	0.37	0.34	0.29
Altkirchenslaw.	0.44	0.20	1.00	0.15	0.41	0.05	0.51	0.41	0.33	0.26
Hethitisch	0.51	0.42	1.00	0.12	0.36	0.01	0.48	0.35	0.32	0.33
Rigveda	0.39	0.08	0.91	0.19	0.38	0.07	0.48	0.26	0.48	0.27
Asoka	0.40	0.26	0.82	0.17	0.34	0.03	0.49	0.40	0.18	0.42
Griechisch (N. T.)	0.41	0.12	0.97	0.11	0.47	0.07	0.51	0.34	0.32	0.34

Tab. 117.2: Typologische Indizes für 20 Sprachen

bereits Kasevič & Jachontov 1982, Hrsg.; vgl. auch Silnitsky 1993; in dieser Arbeit werden Krupas Modifikationen der Greenbergschen Indizes ignoriert, obwohl der Autor auf S. 139 davon spricht, er habe "certain modifications" von Krupa berücksichtigt). Letztere Modifikation berührt freilich nicht nur die Form der drei Indizes, da die Anzahl der syntaktischen Konnektionen nicht gleich der Anzahl der in einem Text vorhandenen Wörter zu sein braucht.

Die von Krupa vorgeschlagene Modifikation der Greenbergschen Indizes bedeutet für die quantitative morphologische Typologie einen wichtigen Fortschritt, wird durch sie doch die Vergleichbarkeit der Werte dieser Indizes gewährleistet. Daneben sind aber bei der Indexbildung noch weitere Gesichtspunkte zu beachten, von denen wir hier nur zwei erwähnen wollen: (a) Der Gesichtspunkt der Interpretierbarkeit verweist darauf, daß man in der Lage sein muß, die Werte, die ein Index annehmen kann, qualitativ-linguistisch zu interpretieren. Die Werte müssen also "verständlich" sein. Diesem Aspekt tragen bereits die unmodifizierten Indizes von Greenberg Rechnung, wie wir oben gesehen haben. Je größer beispielsweise der Wert ist, den der "index of synthesis" annimmt, desto 'synthetischer' ist die betreffende Sprache, usw. (b) Um die Werte, die ein bestimmter Index in zwei Stichproben aus einer Sprache oder aus zwei Sprachen annimmt, "statistisch auf Unterschiedlichkeit testen zu können, d. h. einen Signifikanztest durchführen zu können, muß man auch berechnen können, mit welcher Wahrscheinlichkeit der Index einen bestimmten Wert annimmt, d. h., man muß die Wahrscheinlichkeitsverteilung des Index kennen" (Altmann & Grotjahn 1988: 1031; vgl. auch Altmann & Lehfeldt 1980: 29). Von der Lösung dieses Problems kann heute um so weniger die Rede sein, als das Problem bisher noch nicht einmal als solches erkannt worden zu sein scheint. Solange es nicht gelöst ist, wird man sich damit begnügen müssen, hinreichend große Stichproben zu analysieren, weil man im Falle großer Stichproben beim Testen auf die Normalverteilung übergehen kann. Bei kleinen Stichproben kann man für einige Indizes eventuell die Binomialverteilung benutzen. Von dem Problem der Bestimmung der Stichprobengröße war bereits oben die Rede (zu weiteren Problemen der Indexbildung vgl. auch Altmann 1978).

Eine für die Interpretation der Werte von Greenbergs Indizes äußerst wichtige Frage ist die nach deren gegenseitigem Verhältnis: Sind die von den Indizes erfaßten Eigenschaften voneinander logisch unabhängig, oder gibt es zumindest zwischen einigen von ihnen bestimmte logische Implikationen? Greenberg selbst weist auf dieses Problem hin, wenn er sagt: "In general, the lower the first, or synthetic, index, the fewer the morph junctures which occur and the less the importance of this second index [d. h. des "index of agglutination" A/J] in characterizing the language" (Greenberg 1960: 185 f.). Vgl. auch Cowgill (1976: 130): "If a language has a high morpheme-per-word index, it will automatically have high indices in at least two, and probably all four, of the indices 4 through 7" (vgl. auch Kasevič & Jachontov 1982, Hrsg.: 307).

Krupa, der selber auf die Notwendigkeit hinweist, die "mutual relationships linking the individual indices" (1965: 35) zu studieren, vernachläßigt diesen Gesichtspunkt, wenn er das Einheitsintervall in vier Bereiche einteilt und den Wert, den ein Index annimmt, ohne Bezugnahme auf die Werte anderer Indizes beurteilt; vgl.:

"Thus, for example, a value of 0.88 for W/M is highly analytical, since  $0.76 < 0.88 < 1$ , while a value of 0.27 for the same index is moderately synthetic, since  $0.26 < 0.27 < 0.50$ .

Instances where the resulting value of an index is equal to 1 or 0 may be called purely A or purely  $\neg A$ ." (Krupa 1965: 35)

Was wir bei Krupa finden, ist also eine absolute, nicht-relationale Klassifikation, die nicht die Relationen zwischen den verschiedenen Indizes berücksichtigt.

Den Schritt hin zu einer relationalen Klassifikation unternehmen V. Krupa und G. Altmann in einer 1966 erschienenen Arbeit. 'Relationale' Klassifikation bedeutet, daß eine Sprache nicht aufgrund des absoluten Wertes eines Indexes hinsichtlich der durch diesen Index quantifizierten Eigenschaft eingestuft wird, sondern unter Berücksichtigung (mindestens) des Wertes eines anderen Indexes, von dem gezeigt worden ist, daß er mit dem fraglichen Index logisch verknüpft ist. In dieser Weise korrelieren beispielsweise die Indizes W/M und D/M, und zwar negativ: Wenn W/M den Wert 1 annimmt, dann muß D/M den Wert 0 aufweisen. Betrachten wir die Werte von D/M isoliert, dann sind, wenn wir dem Vorschlag von Krupa (1965) folgen, fast

alle von Greenberg untersuchten Sprachen "highly non-derivative", weil diese Werte fast immer in das Intervall  $<0.00; 0.25>$  fallen. Berücksichtigt man jedoch die Korrelation von D/M zu W/M – je analytischer eine Sprache, um so geringer ihr 'Derivationsgrad' –, sind ganz ähnliche absolute Werte von D/M u. U. verschieden zu bewerten. Im Swahili ist D/M = 0.03, im Vietnamesischen D/M = 0.00. Der Wert 0.03 bedeutet für das Swahili, daß es eine "extremely non-derivative language" ist, weil es nämlich eine "highly synthetic language" mit W/M = 0.39 ist. Ein solcher Wert würde eigentlich einen höheren 'Derivationsgrad' erwarten lassen. Im Vietnamesischen ist W/M = 0.94, d. h., diese Sprache ist "highly analytical", so daß "the zero degree of derivation is a normal (average) event" (Krupa & Altmann 1966: 36).

### 3. Taxonomische morphologische Sprachklassifikation und Suche nach Zusammenhängen

In einem weiteren Schritt wurden von G. Altmann und W. Lehfeldt (1973) die Indizes von Greenberg-Krupa als Grundlage für eine – exemplarisch zu verstehende – morphologische **Sprachklassifikation** benutzt sowie, im Zusammenhang damit, für ein Verfahren zur systematischen Untersuchung der Beziehungen zwischen den von Greenberg-Krupa gemessenen morphologischen Merkmalen.

Bei der Konstruktion der morphologischen Sprachklassifikation orientieren sich Altmann und Lehfeldt an den Verfahren der v. a. in der Biologie entwickelten Numerischen Taxonomie. Danach werden die Indizes von Greenberg-Krupa als Elemente eines Profilvektors betrachtet, die die Koordinaten einer Sprache in einem zehndimensionalen Raum darstellen. So können wir in Tab. 117.2 jede Zeile als einen solchen Vektor interpretieren. Dieser Ausgangspunkt ermöglicht eine Klassifikation der so charakterisierten Sprachen, wobei man zuerst die Ähnlichkeit bzw. die ihr komplementäre 'Entfernung' der Sprachen hinsichtlich der Indexwerte berechnet und dann auf der Grundlage der Ähnlichkeits- bzw. Entfernungskoeffizienten die Klassen aufstellt (vgl. hierzu auch Haarmann 1976: 46 ff.).

Zur Messung der Entfernung je zweier Sprachen bedienen sich Altmann & Lehfeldt (1973: 39 ff.) der euklidischen Distanz, die die

Differenz zwischen den Werten aller Indizes berücksichtigt. Dieses Maß darf allerdings, strenggenommen, zur Entfernungsmessung nur dann verwendet werden, wenn die berücksichtigten Koordinaten logisch voneinander unabhängig sind. Diese Voraussetzung trifft im Falle der Greenberg-Krupa-Indizes nicht allgemein zu, wie oben gezeigt wurde (vgl. dazu u. a. auch Mejlah 1973: 165 f.). Daher müssen alle Ergebnisse, die aus der Interpretation der auf der Grundlage der euklidischen Distanz gemessenen Entfernungen gewonnen werden, als nur vorläufig betrachtet werden.

Das von Altmann & Lehfeldt (1973) verwendete Distanzmaß lautet:

$$(1) d(A, B) = \left[ \sum_{i=1}^n (a_i - b_i)^2 \right]^{\frac{1}{2}}$$

Hierbei bezeichnen A und B die jeweils betrachteten Sprachen, a und b die Werte des Indexes i in diesen beiden Sprachen. Beispielsweise errechnet sich die Distanz zwischen dem Swahili und dem Eskimo als:

$$(2) d(\text{Swahili}, \text{Eskimo}) = [(0.39 - 0.27)^2 + (0.67 - 0.03)^2 + \dots (0.41 - 0.38)^2]^{\frac{1}{2}} = 1.14$$

Berechnet man die Distanzen für alle Sprachpaare aus Tab. 117.1, so erhält man eine Distanzmatrix, die die Grundlage für den eigentlichen Klassifikationsvorgang abgibt (vgl. Altmann & Lehfeldt 1973: 42 f.; eines solchen Verfahrens bedienen sich auch Sil'nickij et al. 1986 für die Klassifizierung von 31 Sprachen, wobei sie von den Korrelationskoeffizienten zwischen diesen Sprachen ausgehen. Vgl. dort S. 124 f. einen Vergleich mit der Klassifikation von Altmann & Lehfeldt und mit der traditionellen morphologischen Sprachklassifikation).

Für die Klassifikation ist wesentlich, daß die Anzahl der Klassen nicht von vornherein festgelegt wird. Altmann & Lehfeldt (1973) bedienen sich eines sog. Cluster-Algorithmus. Dieser fügt, beginnend mit der kleinsten Entfernung, alle Sprachen in einem Iterationsverfahren zu immer größeren Gruppen zusammen. Das Ergebnis läßt sich in Form eines Baumdiagrammes darstellen (vgl. Abb. 117.1, nach Altmann & Lehfeldt 1973: 41).

Erst nach Abschluß des Klassifikationsvorganges wird die Anzahl der Klassen bestimmt. Benötigt man viele Klassen mit jeweils wenigen, dafür aber relativ ähnlichen Sprachen, so wird man den Schnitt etwa bei einer Distanz von 0.5 anlegen (6 Klassen). In-

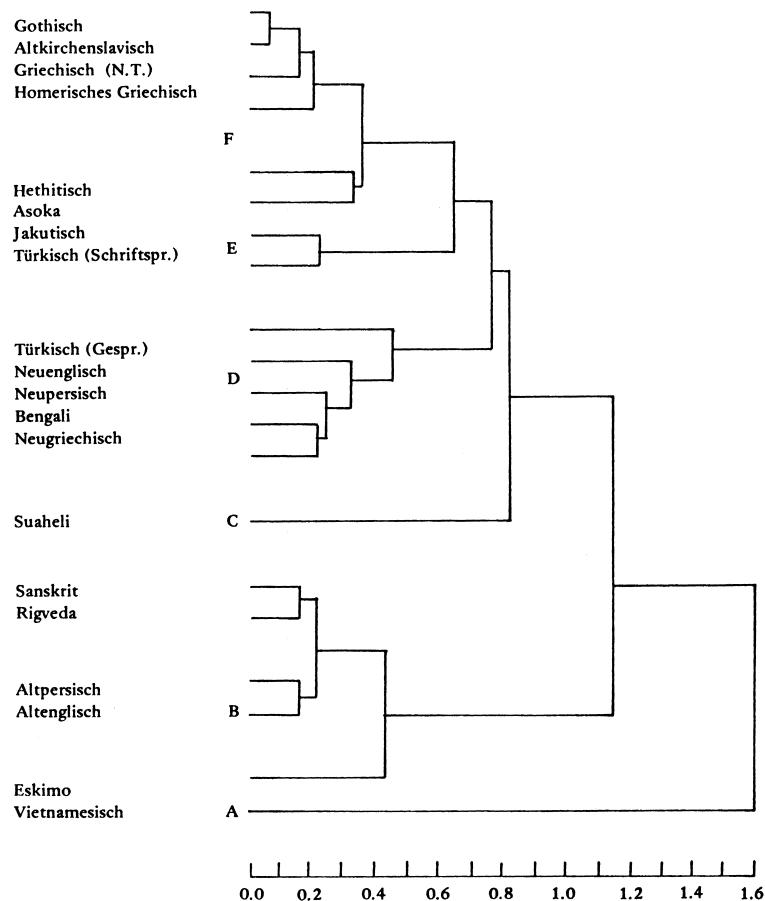


Abb. 117.1: Hierarchische Klassifikation von 20 Sprachen

teressiert dagegen eine globale Einteilung, so wäre ein Schnitt bei 1.0 denkbar, der nur noch drei Klassen ergibt (eine Klasse umfaßt jeweils alle Sprachen, die an einem der ‘Zweige’ hängen).

Eine weitere morphologische Sprachklassifikation, die auf morphologischen Indizes beruht, ist von G. B. Sil'nickij (vgl. Silnitsky 1993) vorgelegt worden. Obwohl diesem Autor Krupas Modifikation von Greenbergs Ansatz bekannt ist, operiert er, wie bereits erwähnt, mit den ursprünglichen Indizes von Greenberg, wobei allerdings im Nenner der Indizes O/N, Pi/N und Co/N N durch W ersetzt wird (dazu s. o.). Hinzukommt der oben bereits genannte “Index der Analytizität”. Sil'nickij klassifiziert 31 Sprachen, indem er die Werte der Korrelationskoeffizienten auswertet, die für jedes Sprachenpaar auf der Grundlage der genannten insgesamt elf Indizes berechnet worden sind. Die Klassifikation, zu der der Autor auf diese Weise ge-

langt, weist “a relatively high degree of correspondence with the traditional genealogical and typological language classes” (Silnitsky 1993: 159) auf (vgl. S. 156 einen Vergleich mit der Klassifikation von Altmann & Lehfeldt 1973).

Mit einer bloßen Sprachklassifizierung, auf wie vielen Merkmalen sie auch immer beruhen mag, kann sich die Sprachtypologie nicht zufriedengeben. Ihr geht es darüber hinaus (a) um die Aufdeckung der empirischen Zusammenhänge zwischen den bei der Klassifizierung berücksichtigten Merkmalen sowie (b) um die Frage, welche Indizes ausschlaggebend für die Klassenbildung gewesen sind. Auf diese Weise wird die interne Struktur der Klassen erhellt.

Was das erste Problem betrifft, so berechnen Altmann & Lehfeldt (1973: 44–46) die **Korrelationskoeffizienten** zwischen den gemessenen Merkmalen und stellen die höchsten positiven wie negativen Korrelationen ei-

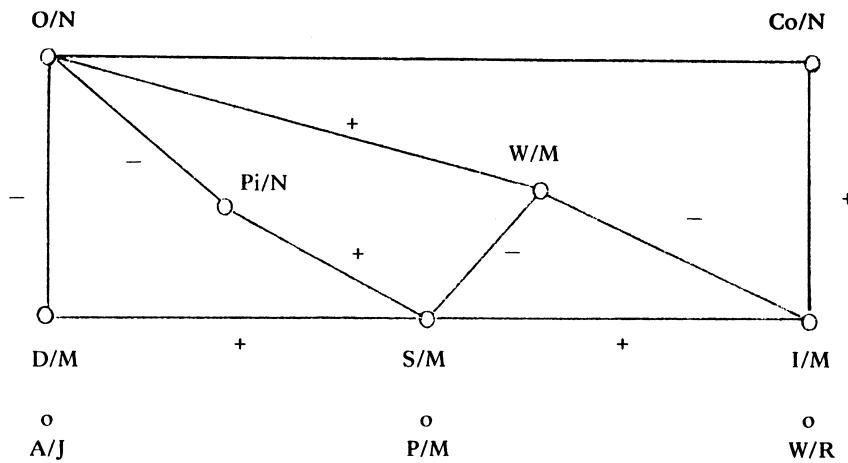


Abb. 117.2: Graph der Merkmalszusammenhänge

nes jeden Indexes in Form eines Graphen dar (vgl. Abb. 117.2).

Hier sieht man, daß drei Merkmale vollkommen unabhängig voneinander sind: A/J, P/M und W/R. Bei den übrigen sind zwei Gruppen zu erkennen: O/N – W/M sowie D/M – S/M – I/M – Pi/N – Co/N. Die Merkmale dieser Gruppen sind untereinander positiv, die Gruppen negativ korreliert. Die positive Korrelation etwa zwischen S/M und Pi/N zeigt, daß Sprachen, deren Wörter viele Suffixe enthalten, dazu tendieren, syntaktische Konnektionen durch Flexion zu signalisieren (vgl. Sil'nickij et al. 1986: 109; diese Autoren haben für 31 Sprachen verschiedener Familien- und Klassenzugehörigkeit die Korrelationskoeffizienten zwischen den – z.T. modifizierten (vgl. 2) – Greenbergschen Indizes ausgewertet und sind dabei zu ähnlichen Ergebnissen wie Altmann & Lehfeldt 1973 gelangt; vgl. auch Silnitsky 1993).

Die Kenntnis der Regression eines Merkmals auf ein anderes erlaubt es, das Intervall abzuschätzen, in dem der Wert des zweiten Merkmals liegen wird, wenn der des ersten gegeben ist. Die konkreten Werte für die Indizes aus Tab. 117.1 haben Krupa & Altmann (1966) berechnet.

Zum zweiten Problem: Um herauszufinden, welche Merkmale in den einzelnen Klassen ähnlich ausgeprägt sind, konstruiert man einen sog. **klassifikatorischen Schlüssel** (vgl. Altmann & Lehfeldt 1973: 46–48), indem man ‘von oben her’ in das Baumdiagramm hineingeht und bei jeder Verzweigung in der Tabelle der Ausgangswerte (Tab. 117.1) nach einem oder mehreren Merkmalen sucht, das/

die die jeweilige Zerlegung ergibt/ergeben. In unserem Falle etwa unterscheidet das Kriterium  $W/M > 0.90$  die erste Klasse, die nur das Vietnamesische umfaßt, eindeutig von allen übrigen Klassen. Gesucht werden, mit anderen Worten, hinreichende, nicht notwendige Unterscheidungsmerkmale. Eine vollständige Analyse ergibt:

### (3) Klasse

- A  $W/M > 0.90$
- B  $W/M \leq 0.90, O/N < 0.27$
- C  $W/M \leq 0.90, O/N \geq 0.27, P/M > 0.10$
- D  $W/M \leq 0.90, O/N \geq 0.27, P/M \leq 0.10, S/M < 0.45$
- E  $W/M \leq 0.90, O/N \geq 0.27, P/M \leq 0.10, S/M \geq 0.45, Pi/N > 0.50$
- F  $W/M \leq 0.90, O/N \geq 0.27, P/M \leq 0.10, S/M \geq 0.45, Pi/N \leq 0.50$

Schließlich kann man die numerischen Aussagen noch in ‘Normalsprache’ rückübersetzen, um zu anschaulichen, einprägsamen Charakterisierungen der Klassen zu gelangen. Sie können etwa folgendermaßen lauten:

- A hochgradig analytisch, maximal isolierend
- B gering isolierend, relativ freie Wortfolge
- C stark präfigierend, stark agglutinierend
- D schwach präfigierend, schwach suffigierend
- E stärker suffigierend, viele flexivische Kongruenzen, stark agglutinierend
- F wenige rein flexivische Kongruenzen

Es ist bezeichnend, daß sich die quantitative morphologische Typologie bisher weitge-

hend darauf beschränkt hat, einen einzigen Ansatz, den von Greenberg, weiterzuentwickeln. Bei allen Fortschritten, die hierbei erzielt worden sind, ist zweierlei nicht zu übersehen: (a) Zahlreiche Eigenschaften der Greenberg-Krupaschen Indizes, ihre Reliabilität und Validität, sind bisher gar nicht oder nur ungenügend untersucht worden. Auch sind uns die statistischen Eigenschaften der Indizes, etwa ihre Stichprobenverteilung, kaum bekannt. (b) Bei den Greenberg-Krupaschen Indizes geht es um die Quantifizierung und Messung einer kleinen Menge relativ einfach zu bestimmender morphologischer Merkmale. Es sollte danach gestrebt werden, komplexe morphologische Eigenschaften zu quantifizieren und zu messen. Einen solchen Ansatz wollen wir im folgenden abschließend vorstellen.

#### 4. Quantifizierung der Eigenschaften grammatischer Subsysteme

Alle Indizes, die auf Greenberg zurückgehen, beruhen auf pragmatischen Stichproben, d. h. auf Zählungen an Texten, und erfassen Relationen, in denen die Elemente dieser Stichproben zueinander stehen. In mehrfacher Hinsicht anders verfährt der Ansatz, der von Lehfeldt (1985; vgl. auch Raster 1980) entwickelt, ausgearbeitet und von ihm und seinen Mitarbeitern im Rahmen eines Projektes zur "Vergleichenden Morphologie der slawischen Sprachen" erprobt wurde. Dieser sog. "analytisch-synthetisch-funktionale" Ansatz beschreibt zunächst einen Ausschnitt der Formenbildung einer Sprache nach einer bestimmten Konzeption, um anschließend zu einer ganzheitlichen Charakterisierung des betreffenden Teilsystems zu kommen.

Der deskriptive Ansatz zerlegt die Formenbildung zunächst in drei Komponenten – **Endungsreihen** ("Flexionsparadigmen"), **morphonologisches Alternationsverhalten** ("morphonologische Paradigmen") und – soweit in der fraglichen Sprache vorhanden – **Akzentuierungsmuster** ("Akzentparadigmen") – und ermittelt die Zahl der sprachlichen Mittel auf jeder dieser drei Ebenen. So findet man im Präsens des Ukrainischen beispielsweise drei Reihen verbaler Endungen, fünf Möglichkeiten der Stammalternationen (keine Veränderung, Verkürzung um den vokalischen Auslaut, Alternation des Auslautkonsonanten usw.) und drei Akzentuierungsmuster (Stammakzentuierung, Endungskzentuierung, Wechselakzentuierung). Dies ist der analytische Teil der Beschreibung. Jedes konkrete Verb realisiert (mindestens) eine spezifische Kombination dieser zur Verfügung stehenden Teile, wobei nach Zusammenhängen zwischen den verschiedenen Klassen von Ausdrucksmitteln sowie verschiedenen Stammklassen und den jeweils realisierten Kombinationen der grammatischen Mittel gesucht wird. Die Beschreibung dieser Zusammenhänge ist der synthetische Teil des Ansatzes. Im abschließenden funktionalen Teil wird der Informationsgehalt untersucht, der jedem konkreten Formenbildungsmittel zukommt, also seine funktionale Belastung, und zwar bezogen auf das jeweilige Teilsystem der Formenbildung.

Die Quantifizierung der Eigenschaften des auf diese Weise beschriebenen Systems erfaßt mehrere Aspekte. Unmittelbar gegeben ist die Zahl der Formenbildungsmittel für die genannten drei Ebenen, im ukrainischen Präsens etwa 3 (Endungsreihen), 5 (Alternationsmuster) und 3 (Akzentuierungsmuster). Theoretisch kann es in dieser Sprache demnach  $3 \times 5 \times 3 = 45$  Möglichkeiten der Präsensbildung geben; tatsächlich kommen jedoch nur 22 verschiedene Kombinationen vor, wie die systemische Untersuchung aller Verben gezeigt hat. Das Minimum verschiedener Kombinationen liegt natürlich bei 5. Diese drei Werte, d. h. die minimale, die tatsächliche und die maximale Zahl verschiedener Kombinationen, setzt das sog. "Maß der Verbundenheit" zueinander in Beziehung, um dieses Subsystem der Grammatik ganzheitlich zu charakterisieren. Es hat folgende Form:

$$(4) M_V = \frac{K_{\max} - K_{\text{real}}}{K_{\max} - K_{\min}}$$

Die Werte dieses Maßes liegen im Einheitsintervall  $<0; 1>$  und sind somit für vergleichende Zwecke nutzbar. Das *tertium comparationis* liegt dabei auf der Inhaltsseite ("Präsensinhaltsparadigma"). Niedrige Werte des Indexes bedeuten eine Beschränkung auf wenige Formenbildungsmuster und signalisieren zugleich einen hohen Grad an implikatorischen Beziehungen zwischen den Elementen, hohe Werte zeigen eine freiere Kombinierbarkeit aller Mittel untereinander, so daß sich aus der Kenntnis eines Mittels (beispielsweise der Endung) nur schlecht Voraussagen über die anderen Mittel (etwa die Betonung) machen lassen.

Russisch 0,44	Ukrainisch 0,575	Weißrussisch 0,60	Slowakisch 0,67	Slowenisch 0,69	Serbokroatisch 0,72
Oborskisch 0,78	Tschechisch 0,93	Polnisch 0,92	Bulgarisch 1,00	Makedonisch 1,00	

Tab. 117.3: Verbundenheit der Präsensformenbildung in slawischen Sprachen

Für das Ukrainische ergibt sich konkret der folgende Wert:  $(45-22)/(45-5) = 0,575$ . Damit liegt diese Sprache zusammen mit den beiden anderen ostslawischen Sprachen im unteren Bereich, wie die Vergleichszahlen zeigen (s. o. Tab. 117.3)

Diese numerischen Daten lassen sich nun wieder sprachwissenschaftlich interpretieren. Die Tatsache, daß die drei ostslawischen Sprachen am unteren Ende der Werteskala eine komplette Gruppe bilden, läßt sich auf den Einfluß der Betonungsmuster zurückführen, wie sie in allen diesen Sprachen vorhanden sind. Das heißt, anders formuliert, daß in den ostslawischen Sprachen die Betonungsmuster durch ihr „Hinzutreten zu den Endungsreihen und den morphonologischen Alternationsmustern die Verbundenheit des Systems der Präsensformenbildung schwächen. Im Serbokroatischen und im Slowenischen hingegen, die gleichfalls verschiedene Betonungsmuster kennen, wird durch diese die Verbundenheit nicht nur nicht verringert, sondern sogar verstärkt (vgl. hierzu Lehfeldt 1998).

Einen weiteren Aspekt erfaßt das „Maß der Prädiktivität“, das wir hier nicht im einzelnen behandeln wollen, da es bisher nur für wenige Sprachen berechnet worden ist. Es setzt die jeweils zur Auswahl stehenden Ausdrucksmittel in Beziehung zu den Wortstämmen und fragt danach, wie gut sich die Verwendung dieser Mittel aufgrund der Stammklassenzugehörigkeit des Lexems vorhersagen läßt. Selbst wenn sich beispielsweise alle einzelnen Komponenten miteinander kombinieren, das Maß der Verbundenheit also seinen Maximalwert annimmt, kann es ja durchaus sein, daß eine jede konkrete Stammklasse immer nur eine ganz bestimmte Kombination zuläßt, so daß in dieser Hinsicht dennoch vollkommene Eindeutigkeit gegeben ist. Dieser Index gibt damit implizit auch Auskunft über die Frage, wie leicht oder schwer das Formenbildungssystem dieser Sprache zu erlernen bzw. zu beherrschen ist: Bei einem hohen Grad an „lexikalischer Spe-

zialisierung“ beispielsweise muß zu jedem Lemma dessen Formenbildung separat gelernt werden.

Es versteht sich, daß auch dieser Ansatz nur ein erster Schritt zu einer Quantifizierung der morphologischen Struktur einer Sprache sein kann, zumal auch hier noch nicht alle theoretischen Probleme geklärt sind (so wird der Index nachweislich von Entscheidungen auf der Phonemebene unmittelbar beeinflußt). Immerhin ist damit jedoch ein Weg vorgezeichnet, auf dem eine quantitative morphologische Sprachtypologie weiter ausgebaut werden kann.

## 5. Zitierte Literatur

- Altmann, Gabriel (1978), „Zur Verwendung der Quotiente in der Textanalyse“. *Glottometrika* 1, 91–106
- Altmann, Gabriel & Grotjahn, Rüdiger (1988), „Linguistische Meßverfahren“. In: Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. (Hrsg.), *Soziolinguistik: Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft, Bd. II*. Berlin, New York: De Gruyter, 1026–1039
- Altmann, Gabriel & Lehfeldt, Werner (1973), *Allgemeine Sprachtypologie*. München: Fink (Uni-Taschenbücher 250)
- Altmann, Gabriel & Lehfeldt, Werner (1980), *Einführung in die quantitative Phonologie*. Bochum: Brockmeyer
- Contreras, Heles (1963), „Una clasificación morfo-sintáctica de las lenguas románicas“. *Romance Philology* 16, 261–268
- Cowgill, Warren (1976), „A Search for Universals in Indo-European Diachronic Morphology“. In: Greenberg, Joseph H. (Hrsg.), *Universals of Language*. Cambridge/MA: MIT Press [1963], 114–141
- Glazov, Jurij Ja. (1965), „K probleme tipologičeskogo schodstva dravidijskich i tjurskich jazykov [On the problem of typological similarity between Dravidian and Turkish languages]“. *Lingvistiches-*

- kaja tipologija i vostočnye jazyki.* Moskva, 205–212
- Greenberg, Joseph H. (1957), “The Nature and Use of Linguistic Typologies”. *International Journal of American Linguistics* 23, 68–77
- Greenberg, Joseph H. (1960), “A Quantitative Approach to the Morphological Typology of Language”. *International Journal of American Linguistics* 26 [<sup>1</sup>1954], 178–194
- Haarmann, Harald (1976), *Grundzüge der Sprachtypologie*. Stuttgart usw.: Kohlhammer (Urban-Taschenbücher 242)
- Hamp, Eric P. (1958), “The Calculation of Parameters of Morphological Complexity”. *Proceedings of the Eighth International Congress of Linguists*, Oslo, 134–142
- Ineichen, Gustav (1979), *Allgemeine Sprachtypologie: Ansätze und Methoden*. Darmstadt: Wissenschaftliche Buchgesellschaft
- Kasevič, Vadim B. & Jachontov, Sergej E. (1982, Hrsg.), *Kvantitativnaja tipologija jazykov Azii i Afriki* [A quantitative typology of Asian and African languages]. Leningrad
- Kelemen, József (1970), “Sprachtypologie und Sprachstatistik”. In: Dezső, L. & Hajdú, P. (Hrsg.), *Theoretical Problems of Typology and the Northern Eurasian Languages*. Amsterdam, 53–63
- Kroeber, Alfred L. (1960), “On Typological Indices I: Ranking of Languages”. *International Journal of American Linguistics* 26, 171–177
- Krupa, Viktor (1965), “On Quantification of Typology”. *Linguistics* 12, 31–36
- Krupa, Viktor & Altmann, Gabriel (1966), “Relations between Typological Indices”. *Linguistics* 24, 29–37
- Lehfeldt, Werner (1985), *Sprjaženie ukrainskogo glagola* [Ukrainian verbal inflection]. München: Sagner
- Lehfeldt, Werner (1998), “Das System der Präsensformenbildung im Polnischen im Vergleich mit den anderen slavischen Sprachen”. In: Rothe, Hans & Thiergen, Peter (Hrsg.), *Polen unter Nachbarn, Polonistische und komparatistische Beiträge zu Literatur und Sprache. XII. Internationaler Slavistenkongreß in Krakau 1998*. Köln usw.: Böhlau, 29–53
- Mejlach, Michail (1973), “Indeksy morfoložičeskoj tipologii [Indices for a morphological typology]”.
- Problemy grammatičeskogo modelirovaniya.* Moskva, 155–170
- Pierce, Joe E. (1962), “Possible Electronic Computation of Typological Indices for Linguistic Structures”. *International Journal of American Linguistics* 28, 215–226
- Pierce, Joe E. (1966), “Sampling and Typological Indices in Language”. *Linguistics* 24, 43–50
- Raster, Peter (1980), “Paradigmatische Strukturen von Verbalsystemen indogermanischer Sprachen”. *Kwartalnik Neofilologiczny* 27, 259–286
- Robins, Robert H. (1965), “Some Typological Observations on Sundanese Morphology”. *Lingua* 15, 435–450
- Sapir, Edward (1949), *Language: An Introduction to the Study of Speech*. New York [<sup>1</sup>1921]
- Silnitsky, George (1993), “Typological Indices and Language Classes: A Quantitative Study”. *Glotto-metrika* 14, 139–160
- Silnitsky, George (1998), “Correlational Analysis of the Indo-European Morphological System”. *Journal of Quantitative Linguistics* 5, 81–95
- Sil'nickij, Georgij B. & Jachontov, K. S. & Jachontov, Sergej E. (1986), “Primenenie korrelacion-nogo i faktornogo analiza k tipologii jazykov [Correlation and factor analysis in language typology]”. *Aktual'nye voprosy derivatologii i derivatografii*, Vladivostok, 105–125
- Skalička, Vladimír (1935), *Zur ungarischen Grammatik*. Prag: Facultas philosophica Universitatis Carolinae
- Slavičková, Eleonora (1968), “Toward a Typological Evaluation of Related Languages”. *Travaux linguistiques de Prague* 3, 281–289
- Son, Jos van & Peters, Toine (1989), “Quantificeerde taalclassificatie”. *TABU, Bulletin voor Taalwetenschap* 19.1, 16–30
- Stepanov, Arthur V. (1995), “Automatic Typological Analysis of Semitic Morphology”. *Journal of Quantitative Linguistics* 2, 141–150
- Thorndike, Edward L. (1943), “Derivation Ratios”. *Language* 19, 27–37
- Voegelin, Charles F. & Ramanujan, R. A. & Voegelin, Florence M. (1960), “Typology of Density Ranges I: Introduction”. *International Journal of American Linguistics* 26, 198–205
- Sebastian Kempgen, Bamberg (Deutschland)  
Werner Lehfeldt, Göttingen (Deutschland)

## 118. Cross-linguistic generalizations and their explanation

1. Classical morphological typology
2. Cross-linguistic generalizations in morphology
3. Explanations of cross-linguistic generalizations in morphology
4. References

### 1. Classical morphological typology

One of the earliest and well known cross-linguistic generalization is the morphological typology brought up by the 19<sup>th</sup> century German philologists August and Friedrich von Schlegel, (cf. F. von Schlegel 1808; A. von Schlegel 1818), and Wilhelm von Humboldt (Humboldt 1825; 1836). They proposed a typology of languages based on their dominant morphological characteristics (cf. Art. 10). Languages were called  **fusional** if they employ bound morphs that indicate two or three grammatical functions in one form and if this is a significant property of the morphology of that language.  **Agglutinative** languages differ from fusional languages in that they come close to a one-to-one correspondence between grammatical function and bound form, one grammatical form indicating one grammatical meaning.  **Isolating** languages ideally lack bound forms all together, i.e. they express grammatical meanings by means of concatenated independent words. A fourth type, the  **polysynthetic** type, was added later by W. von Humboldt in order to cover also some American Indian languages in this typology. Polysynthetic languages are extreme in their morphological behavior in that they bind all kinds of sometimes quite specific grammatical, adverbial, and even nominal meanings morphologically at the verb.

This fourfold morphological typology (cf. Greenberg 1974: 36–39; Comrie 1989: 42–46) was established around three or four ideal types, and it was not assumed by the proponents of this concept that there was a single language that fits perfectly one of these types. On the contrary, Humboldt was quite aware of the fact that languages usually show more than one of these ideal characteristics in their grammar and that they are therefore only exponents to some degree of a certain type. Recent typological studies on the distribution of these types among the languages of the world revealed that pure isolating and pure agglutinative languages are rare. Most of the languages employ a mix of fusional and aggluti-

native strategies. There are, in addition, doubts whether pure isolating languages exist at all. Vietnamese which is often cited as an isolating language has extensive compounding and reduplication (cf. Greenberg 1966 a: 93; Bybee 1985: 45 f.). The concept of ideal types of languages as a whole or of subsystems of a language has been given up in modern linguistic typology, because these ideal types become arbitrary if there are no real languages corresponding to them. In addition, these ideal types cannot be explained or motivated by intra and extra linguistic factors.

The classical morphological typology was holistic in the sense that languages were classified as wholes, a result that was desired at that time, because it allowed to formulate evaluative statements about entire languages. The typology was used to support the intellectual prejudices of that time with regard to languages (and cultures). The morphological classification served as an argument for the idea that the classic languages such as Greek, Latin, and Sanskrit are superior over others, because they belong to the fusional type. The reasoning was circular: because these languages were assumed to be superior from the beginning and because they belong to the fusional type it was concluded that languages of this type represent the most advanced languages of the world. Accordingly, the isolating and the agglutinative type were seen to represent less perfect stages in the historical development of a language (cf. Greenberg 1974: 36–39). In addition, the evaluative ranking of language types from isolating, agglutinative to fusional was taken as a historical path for the potential development of languages. This idea was brought up particularly by August Schleicher (cf. Greenberg 1974: 37). Languages begin at the isolating stage and may proceed to the more advanced stages, and if they are lucky they reach the fusional stage. The idea that languages can be evaluated as superior to others on the basis of their linguistic structure in general and on the basis of their morphological structure in particular has been abandoned at all by modern 20<sup>th</sup> century linguistics because of lack of empirical evidence. However, the idea that there is a historical connection between the isolating, the agglutinating and fusional morphological type can be found in a different form in all modern theories of gram-

maticalization. Based on abundant evidence from historical linguistics it could be shown that grammatical/inflectional categories derive historically from independent lexical units accompanied by various phonetic and semantic changes (cf. also Art. 146). There is a non-reversible, unidirectional path from free words to grammatical affixes for many grammatical categories (cf. Lehmann 1985; 1995; Croft 1990: 230–244; Givón 1979: 209; and many others). So there is a historical connection between isolating, agglutinating, and fusional, not on the level of entire languages and grammatical systems, but on the level of individual constructions, words, and affixes representing a similar semantic/grammatical concept. Grammatical affixes, either inflectional or derivational, are not created out of nothing by speakers of a language, but represent stages in a historical/evolutionary process of linguistic change. This insight is an important constraint on all attempts to explain cross-linguistic generalizations in morphology (cf. 3).

The classical morphological typology refers only to the formal side of the expression of grammatical meanings. The semantic side of the form-function relation of morphological forms was not taken into account. No attempt was made to answer the question which types of meanings are preferably expressed morphologically (derivational and inflectional), and which types of meanings are expressed preferably lexically in the languages under consideration. A first attempt in this direction is Sapir (1921). Extra-linguistic explanations of cross-linguistic generalizations in morphology can be searched for and found only if the meaning/function of the morphemes is taken into consideration. Therefore, no extra-linguistic explanation could be given for the existence of these types of languages nor their distribution among the languages of the world.

Furthermore, the classical morphological typology did not correlate the parameter “morphological expression type” (the main ideal types isolating, agglutinative, fusional, and polysynthetic can in fact be interpreted as points on two different scales, the number of morphemes per morph and the number of morphemes per word (cf. Greenberg 1960)) with other typological parameters such as word order, alignment type, head and dependent marking, etc. There is no quantitative typological study investigating this question and it seems that contemporary typologists do not expect any significant correlation be-

tween morphological type and other typological parameters (cf. e.g. Bauer 1988: 170). Some indication that this assumption is correct can be found in Nichols' investigation of the correlation of the parameter “morphological complexity” with other typological parameters in a sample of more than 170 languages (cf. Nichols 1992). Nichols' parameter “morphological complexity” reflects the morphological distinction between agglutinative/fusional versus isolating only indirectly by counting the amount of morphological marking of head-dependent relations. Isolating languages have rather low values on this scale, fusional and agglutinative languages have higher scores. Inflectional categories such as tense, aspect, and mood, which do not indicate the syntactic relations between head and dependent are ignored in her investigation (cf. Nichols 1992: 64). The results seem to support the assumption: languages generally avoid the extreme types of complexity, i.e. very high and very low values, but have a high score in the middle values. There is no correlation between morphological complexity and word order (cf. Nichols 1992: 112), there is a slight correlation of head marking with low complexity and dependent marking with high complexity (cf. Nichols 1992: 98 f.), and a slight correlation of high complexity with ergative alignment (cf. Nichols 1992: 109 f.). Otherwise, the parameter morphological complexity has practically no predictive value with regard to other typological parameters.

Cross-linguistic generalizations in contemporary morphology are largely generalizations over form-function relations in morphological units. The discussion of some aspects of the classical 19<sup>th</sup> century morphological typology has shown that purely formal generalizations are dead end roads, they hardly allow external linguistic explanations and they seem to have no significance with regard to other typological parameters. In the following section (cf. 2), I would like to present an overview of the types of cross-linguistic generalizations in morphology. Section 3 presents different types of language external explanations of these generalizations.

## 2. Cross-linguistic generalizations in morphology

### 2.1. Inflectional versus derivational morphology

The notional distinction between **inflectional** and **derivational** morphology is well established in linguistics, but hard to measure and

to operationalize (cf. Art. 38 for the main criteria distinguishing inflectional and derivational morphology). Nevertheless, cross-linguistic comparison came up with some generalizations dealing with this distinction. It was already Eugene Nida (1946) who observed that languages usually have more derivational than inflectional morphemes. This fact was formulated in Greenberg's Universal #29 more precisely: "If a language has inflection, it always has derivation." (Greenberg 1966a: 93). This implicational universal states that there are languages which have derivational morphology, but no inflectional morphology (and not vice versa), otherwise, languages have both types of morphology to varying degrees. There is another implicational universal dealing with the derivation/inflection distinction, Universal #28: "If both the derivation and inflection follow the root, or they both precede the root, the derivation is always between the root and the inflection" (Greenberg 1966a: 93; see also Bybee 1985: 33). This universal describes the observation that derivational morphology is usually placed closer to the word stem than the inflectional morphology (see also 2.2).

## 2.2. Cross-linguistic generalizations concerning the order of morphemes

The **order** of morphemes in the structure of words is another field for cross-linguistic generalizations. It has frequently been observed that there is an asymmetry between prefixes and suffixes. Suffixing (no matter whether inflectional or derivational), however, is the preferred way to indicate grammatical/derivational categories cross-linguistically. Exclusively suffixing languages are not uncommon, but exclusively prefixing languages are rare (cf. Greenberg 1966a: 92).

More interesting than the distribution of pure structural positions of affixes with regard to the root/stem is the question whether there is a preferred order of morphemes with respect to the inflectional/derivational categories they indicate. Bybee (1985) found clear empirical evidence in her sample of languages that a) derivational morphology is closer to the word stem than inflectional morphology (this is a confirmation of Greenberg's Universal #28), and that inflectional categories of the noun and the verb have the following order, cf. the templates in (1 a-b).

- (1) (a) [Verb root/stem – aspect – tense – mood – person]
- (b) [Noun root/stem – number – case]

The templates in (1 a–b) mean that inflectional categories of aspect such as perfective/imperfective are closer to the verb stem than tense and mood categories. Person marking of the subject/object and personal agreement with the subject is always found in the outermost layer (cf. Bybee 1985: 33–35). As for the noun, it can be observed that number marking always precedes case (this is a confirmation of Greenberg's Universal #39, cf. Greenberg 1966a: 95). It is also the inflectional categories closest to the stem, i.e. aspectual morphemes in verbs and number morphemes in nouns which cause morphophonemic changes in the stem or undergo morphophonemic changes themselves more frequently than the inflectional categories farther away from the stem/root (cf. Bybee 1985: 36 f.).

## 2.3. Implicational universals

**Implicational universals** are tools to describe cross-linguistic generalizations with regard to morphological marking asymmetries, i.e. markedness relations within paradigms and across paradigms (cf. also Art. 29). A paradigm internal implicational universal is e.g. Greenberg's Universal #34 with respect to nominal morphology: "No language has a trial number unless it has a dual. No language has a dual unless it has a plural" (Greenberg 1966a: 94). Universal #34 states that there are no languages with a marked trial number but no dual and plural, and that there are no languages with a marked dual without having a marked plural. The same implicational hierarchy holds also for personal pronouns (cf. Helmbrecht 2002: Ch. 7). Other paradigm internal universals are #35 which states the unmarkedness of the singular, or #44 which claims the unmarkedness of gender categories for third person pronouns.

There are also cross-linguistic generalizations which make claims about the co-occurrence of morphological categories in a language, or even more specific in certain word classes. An example of this type of cross-linguistic generalization is Greenberg's Universal #30 with respect to verbal inflection: "If the verb has categories of person-number or if it has categories of gender, it always has tense-mode categories" (Greenberg 1966a: 93). Universal #30 states that there are languages with verbs inflected for tense-mode categories and not for person, but that there are no languages with person marking on the verb without having morphological tense-

mode categories. Most of the languages have both types of categories marked on the verb (cf. also Bybee's statistics 1985: 31). Other universals of this type are #37 which states that gender marking on nouns implies the marking of number and not vice versa.

### 3. Explanations of cross-linguistic generalizations in morphology

Affixes indicating derivational and inflectional categories derive almost always historically from independent lexical material which undergoes certain changes with regard to its phonological shape (syntagmatic coalescence, paradigmatic attrition, and fusion of morphemes), with regard to its distribution (fixation of word order, condensation, paradigmaticization, and obligatorification) and with regard to its semantic content (semantic generalization (expansion), and desemantication) (cf. Lehmann 1995: 122–179). Hence grammatical affixes, inflectional and derivational, represent only a certain stage in this evolutionary process which is characterized by the loss of phonological autonomy and morphosyntactic freedom in favor of a fixed morphological position and semantic generality (see also Lüdtke 1980, ed.; 1985). Since affixes are the outcome of a process of **grammaticalization** (affixes may also be borrowed, but I will disregard this possibility here for the sake of simplicity and brevity) they are never intentional ad hoc coinages by the speakers created to express some grammatical meaning such as tense and aspect, which may be lacking in the grammar of their language. Such a freedom on the side of the speaker may exist on other levels of language such as the lexical and syntactic level, where concepts may easily receive a new wording via metaphorical or metonymic extensions in new syntactic constructions. This freedom does not hold with respect to grammatical affixes. The speaker intention can therefore be excluded as direct causal mechanism for the emergence of grammatical affixes and the cross-linguistic patterns associated with it (cf. also Keller 1994; Croft 2000: 66–71). Explanations for the morphological patterns that are the basis of the cross-linguistic generalization discussed above (cf. 2) were proposed in the literature: they are derived from the physiological and cognitive conditions of human beings in the world, and from the conditions of language

use. An important methodological requirement for an explanation of a cross-linguistic generalization in morphology is that the proposed causal mechanisms can be shown to be effective in every single speech act and utterance contributing to the conventionalization of an innovation/alternative expression (cf. Bybee 1988: 357). In the next few sections, some of the proposals will be presented.

#### 3.1. Innateness

It is a central idea of Chomskyan generative grammar that the core categories and principles of universal grammar are **innate**. Every human being is genetically equipped with universal grammar and this is a precondition for the capability of children to acquire their specific language (cf. Chomsky 1965). It is argued that the principles of universal grammar cannot be learned out of the defective language input in normal child-adult communication, particularly because of the lack of negative evidence in adult linguistic responses. Therefore, it is concluded that there is an *a priori* necessity for a genetic representation of universal grammar out of which all specific grammars can be constructed (cf. Chomsky 1965; 1968; Hoekstra & Kooij 1988). The categories and principles of universal grammar are necessarily abstract. It should be possible to write the grammatical rules of each specific language with the categories and principles of universal grammar. Chomsky distinguished formal and substantive universals. The former describe the general design of the grammar, its rules and its components (modules), the latter describe the content of the rules such as the syntactic categories and bar levels of X-bar syntax (cf. Jackendoff 1977). Later, universal grammar was enriched by parameter sets which allow for variation in universal grammar (cf. Chomsky 1981).

Universal grammar is basically a syntactic theory. It contains very abstract claims about syntactic categories and relations in terms of phrase structure rules. Since there are practically no substantive, i.e. absolute universals with regard to morphology this cannot be part of the universal grammar. In the view of universal grammar, morphology is language specific and serves syntax. Since morphology is not part of universal grammar, it is not assumed to be innate. Hence innateness is not utilized to explain morphological generalizations of the types discussed in 2.

### 3.2. Natural language processing explanations

Research in the field of **language processing** deals with the principles of comprehension and production of language. Hypotheses and explanations in this field are usually based on data drawn from psycholinguistic experiments on real time language use. Linguistic communication among human beings is heavily constrained by their physiological and cognitive limits such as limits with regard to articulation speed, short-term memory, and others. These limitations in the human capacity of language use can be associated with the ease and difficulty of language processing, i.e. there are linguistic structures which may be easier to process than others. Differences in processing ease were taken then as a source for the explanation of differences in the typological distribution of word order types (cf. Hawkins 1983).

An explanation in terms of processing ease is offered for the above mentioned (cf. 2.2) asymmetry with regard to prefixing and suffixing (cf. Hawkins & Cutler 1988). Starting point for Hawkins & Cutler is the hypothesis that the harmony principle across word order universals would predict that head initial languages (VO order, prepositions etc.) should prefer prefixing (with the affix as head of the word), and that head final languages should prefer suffixing. The latter prediction is indeed true, there is a significant preference for suffixing in head final languages and there are even many languages exclusively suffixing. As for head initial languages, the prediction is not true. There is a preponderance of suffixing and even the number of languages exclusively prefixing is much lower than the number of languages exclusively suffixing. The explanation for this obvious asymmetry lies, according to Hawkins & Cutler, in the processing of words. Psycholinguistic evidence shows that with regard to the recognition of words the beginning is more salient than the other parts of it and that stems are processed before their inflectional morphemes thus separating lexical recognition from processing of the relational syntactic/semantic information. The information coded in stems is more important than the information coded in affixes. It makes therefore sense for affixes to be postposed to the stem in the less salient position rather than to be preposed to it. There is an important objection to this type of explanation in morphology brought forward by Hall (1988) and Bybee

(1988). The processing explanation may indeed describe how inflected words are processes in the process of comprehension, but it cannot serve as explanation of the rise of suffixes. Affixes derive historically from (perhaps inflected) lexical material. The explanation in terms of processing priorities cannot be applied to this situation and does certainly not explain why the stem in question is reduced to an affix. Stems which developed into affixes are fixed in their syntactic position long before they become real affixes. The rigidification and the position of the source stem cannot be explained within this processing approach.

### 3.3. Cognitive/semantic explanations

The two generalizations concerning the relative order of derivational and inflectional morphology and the relative order of different inflectional categories with regard to the stem presented in 2.2 are explained by Bybee (1985) by two interacting semantic principles, **relevance** and **lexical generality**. These two factors also determine the likelihood that a semantic notion will be encoded as a derivational/inflectional category. The semantic notion must be highly relevant to the meaning of the stem to which it attaches, and it must be a very generally applicable semantic notion, otherwise it will not apply to enough lexical stems to be inflectional (cf. Bybee 1985: 19).

Relevance is seen as a cognitive principle, because it depends on cognitive and cultural salience. It is defined in the following way: "A meaning element is relevant to another meaning element if the semantic content of the first directly affects or modifies the semantic content of the second" (Bybee 1985: 13). With other words, things which belong together conceptually stand close together structurally. Bybee can show that the order of inflectional categories with respect to the verb stem can be explained with this cognitive principle. The inflectional category aspect affects the meaning of the verb stem more than tense, mood, and person in this order. Aspect modifies the verbal meaning with regard to the internal temporal structure/perspective of the event designated by the verb. Tense has less effect on the meaning of the verb, because it places the whole event in time with respect to the time of the utterance, and mood refers to the way speakers present the truth of the proposition in discourse and real world contexts without

changing the designated proposition itself. Person marking on the verb (subject agreement) has practically no impact on the meaning of the verb stem as such, since the event of the verb remains the same no matter which person is the participant. The relevance principle does not only explain the generalizations with regard to the order of morphemes in the internal structure of words, but also the implicational Universal #30 (cf. 2.3) which claims that languages may have morphological tense/mood categories on the verb without person agreement (with the subject), or both of these, but may not have person marking on the verb without also having tense/mode categories. Since person marking is semantically/cognitively least relevant for the meaning of the verb, it is the last category to be morphologically fused with the verb.

The same relevance principle can also explain the order of noun morphology. Number is an inherent category of the noun which has a great impact on the referent of the noun phrase. The category number is often not obligatory and rather derivational in the languages of the world. Case is an always inflectional category of the noun indicating the relation of the noun phrase to other constituents of the clause (for the distinction between inherent and relational inflectional categories cf. also Anderson 1985).

The same reasoning holds for the relative order of inflectional and derivational categories. Bybee argues that derivational morphology modifies the meaning of the verb stem more drastically than inflectional categories and this is the reason why derivational morphology is closer to the verb stem than inflectional morphology. This is also the reason why voice and valence changing morphology, which have a great impact on the event structure designated by the verb are predominantly coded as derivational morphology, and are closer to the verb stem than inflectional categories (cf. Bybee 1985: 81–109).

The semantic/cognitive principle of relevance invoked by Bybee in order to explain cross-linguistic ordering regularities in morphology has the advantage that it can also be employed to explain the emergence of these structural patterns. Bybee argues convincingly that the same principle also governs the order of the lexical items which are the historical source for derivational and inflectional affixes (cf. Bybee 1985: 38–43). And further, the process of the grammaticalization from a lexical item to a grammatical af-

fix can only commence because the meaning of the lexeme is highly relevant to the meaning of the lexeme it accompanies. This relevance leads to an increase in frequency (of co-occurrence, see 3.4) and the processes of semantic generalization and phonological attrition and fusion. Bybee stresses the fact that not all lexical items which stand closely together will finally fuse with each other. This process is only possible if the meaning of one of the items is highly relevant to the meaning of the other.

The principle of relevance applied by Bybee to explain morphological patterns is reminiscent of Haiman's principle of structural isomorphism which states that the structure of the linguistic utterance reflects the structure of the concept expressed by it (cf. Haiman 1985: 102–108). This is an iconic principle because it states a similarity between the linguistic sign and the concept it designates (cf. also Art. 30 on iconicity). Within the domain of morphology, iconic relationships between the linguistic sign and the meaning may be observed, but iconicity can no longer be made responsible for the creation of these signs. The principle of relevance produces diagrammatic, i.e. iconic relationships between grammatical meanings and lexical meanings and their formal expression in syntax. The iconic relationships then become fossilized in morphology and finally erode because of the principle of economy (see 3.4).

#### 3.4. Discourse-pragmatics: frequency and economy

Lexical items that are in the process of becoming a grammatical affix undergo a generalization of their meaning. This is the result of the usage of these forms in more and more pragmatic and linguistic contexts. This is also the precondition for the further increase of textual **frequency** of these forms including an increase in type as well as token frequency. The increase of textual frequency is the source for markedness relations between grammatical forms. Frequent forms tend to be phonologically short and morphologically simple, forms which are used rarely are longer and morphologically complex (cf. Greenberg 1966 b: 65–69; Croft 1990: 156–160). The causal mechanism behind this observation is **economy**. Speakers tend to shorten and to simplify frequent forms, this principle is also called Zipf's Law (cf. Zipf 1935; Mańczak 1980). It has the effect that the most frequent grammatical values have

zero or minimal expression. Paradigm-intrernal markedness patterns such as the ones often observed in number marking of nouns (and pronouns) are explained by economy. The singular is the least marked, mostly zero marked, value compared to the plural and dual (if present at all) which are more marked usually by separate forms. This pattern corresponds to the overall textual frequency of the different values of the grammatical category number. Singular is by far the most frequent number value in nouns (and pronouns) before plural and dual. The dual is textually a rare category value. With regard to number marking in nouns and pronouns, the economy principle explains language specific markedness relations as well as the typological distribution of number categories (cf. Universal #34 and #35 in 2.3). The observation that certain category values such as the singular are almost always zero marked cross-linguistically was taken as evidence that these category values are the most natural ones. **Naturalness** (cf. Art. 31), however, is not a language external explanation but a statement of the fact that certain category values are less marked than others in the languages of the world.

Frequency and economy are not the last steps in the search for language external explanations of (cross-linguistic) markedness relations. A rise in textual frequency may reflect the pragmatic salience of the meaning expressed by the lexical item or grammatical affix. The person hierarchy as part of the empathy hierarchy is a good example for discourse salience. Reference to the two basic speech act participants speaker and hearer in normal conversation is much more frequent than references to other discourse participants (cf. Helmbrecht 2002: Ch. 7), because they are the two individuals who intend to talk about something relevant to both of them in one way or other. The pragmatic salience of reference to the speech act participants is therefore the reason for the high frequency of the appropriate referential expressions. And further, the most economic way to perform the reference to speech act participants is to have and use simple and short forms such as the ones given in a paradigm of personal pronouns. The salience of this type of reference, i.e. its textual frequency and the economic motivation explain markedness relations within pronominal paradigms as well as the almost universal distri-

bution of the speaker/hearer (singular) person categories (cf. Helmbrecht 2002).

Economy on the other hand reflects the limits in the cognitive and physiological capacities of humans with regard to linguistic communication: a) complex information has to be encoded in a linear way, b) the speed of articulation is much slower than the capability of cognitive inferencing, and c) the short term memory is restricted with regard to the quantity and duration of information it may store. It is therefore an important goal of speakers to express as fast and brief as possible what they want to say, otherwise they loose the attention of the hearer. Economy is therefore effective on all levels of grammar and discourse.

#### 4. References

- Anderson, Stephen R. (1985), "Inflectional morphology". In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III*. Cambridge: Cambridge University Press, 3–56
- Bauer, Laurie (1988), *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press
- Bybee, Joan (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins
- Bybee, Joan (1988), "The Diachronic Dimension in Explanation". In: Hawkins (ed.), 351–381
- Chomsky, Noam (1965), *Aspects of the Theory of Syntax*. Cambridge/MA: MIT Press
- Chomsky, Noam (1968), *Language and Mind*. New York: Harcourt, Brace & World
- Chomsky, Noam (1981), *Lectures on Government and Binding*. Dordrecht: Foris Publications
- Comrie, Bernard (1989), *Language Universals and Linguistic Typology*. Oxford: Basil Blackwell
- Croft, William (1990), *Typology and Universals*. Cambridge: Cambridge University Press
- Croft, William (2000), *Explaining Language Change: An Evolutionary Approach*. Harlow: Pearson Education Limited
- Givón, Talmy (1979), *On Understanding Grammar*. New York etc.: Academic Press
- Greenberg, Joseph H. (1960), "A Quantitative Approach to the Morphological Typology of Language" [1954]. *International Journal of American Linguistics* 26, 178–194
- Greenberg, Joseph H. (1966a), "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". In: Greenberg, Jo-

- seph H. (ed.), *Universals of Grammar*. Cambridge/MA: MIT Press, 73–113
- Greenberg, Joseph H. (1966b), *Language Universals, with Special Reference to Feature Hierarchies*. The Hague: Mouton (Janua Linguarum, Series Minor 59)
- Greenberg, Joseph H. (1974), *Language Typology: A Historical and Analytical Overview*. The Hague, Paris: Mouton
- Haiman, John (1985), *Natural Syntax: Iconicity and Erosion*. Cambridge: Cambridge University Press
- Hall, Christopher J. (1988), “Integrating Diachronic and Processing Principles in Explaining the Suffix Preference”. In: Hawkins (ed.), 321–344
- Hawkins, John A. (1983), *Word Order Universals*. New York: Academic Press
- Hawkins, John A. (1988, ed.), *Explaining Language Universals*. Oxford: Basil Blackwell
- Hawkins, John A. & Cutler, Anne (1988), “Psycholinguistic Factors in Morphological Asymmetry”. In: Hawkins (ed.), 280–305
- Helmbrecht, Johannes (2002), *Personal Pronouns: Form, Function, and Grammaticalization*. Habilitationsschrift, University of Erfurt
- Hoekstra, Teun & Kooij, Jan G. (1988), “The Innativeness Hypothesis”. In: Hawkins (ed.), 31–52
- Humboldt, Wilhelm von (1825), *Über das Entstehen der grammatischen Formen und ihren Einfluß auf die Ideenentwicklung*. Berlin: Königliche Preussische Akademie der Wissenschaften (Abhandlung der historisch-philologischen Klasse)
- Humboldt, Wilhelm von (1836), *Über die Kawisprache auf der Insel Java. Vol. I: Über die Verschiedenheit des menschlichen Sprachbaus*. Berlin: Königliche Preussische Akademie der Wissenschaften
- Jackendoff, R[ay] (1977), “X Syntax: A Study of Phrase Structure”. *Linguistic Inquiry* (Monograph 2)
- Keller, Rudi (1994), *Sprachwandel: Von der unsichtbaren Hand in der Sprache*. Tübingen, Basel: Francke
- Lehmann, Christian (1985), “Grammaticalization: Synchronic Variation and Diachronic Change”. *Lingua e Stile* 20, 303–18
- Lehmann, Christian (1995), *Thoughts on Grammaticalization*. München, Newcastle: LINCOM
- Lüdtke, Helmut (1980, ed.), *Kommunikations-theoretische Grundlagen des Sprachwandels*. Berlin, New York: W. de Gruyter
- Lüdtke, Helmut (1985), “Zum Stand der Forschung in der Sprachwandeltheorie”. In: Kürschner, Wilfried & Vogt, Rüdiger (eds.), *Sprachtheorie, Pragmatik, Interdisziplinäres. Akten des 19. Linguistischen Kolloquiums, Vechta 1984, Vol. II*. Tübingen: Niemeyer, 3–9
- Mańczak, Witold (1980), “Frequenz und Sprachwandel”. In: Lüdtke (ed.), 37–79
- Nichols, Johanna (1992), *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press
- Nida, Eugene (1946), *Morphology*. Ann Arbor: University of Michigan Press
- Sapir, Edward (1921), *Language*. New York: Harcourt, Brace & World
- Schlegel, August Wilhelm von (1818), *Observations sur la langue et la littérature provençales*. Paris
- Schlegel, Friedrich von (1808), *Über die Sprache und Weisheit der India: Ein Beitrag zur Begründung der Alterthumskunde; nebst metrischen Uebersetzungen indischer Gedichte*. Heidelberg: Mohr und Zimmer
- Zipf, George (1935), *The Psychobiology of Language: An Introduction to Dynamic Philology*. Cambridge/MA: MIT Press

*Johannes Helmbrecht, Erfurt (Germany)*

## XVI. Systeme morphologischer Struktur: Sprachskizzen

### Systems of morphological structure: Illustrative sketches

#### 119. English (Indo-European: Germanic)

1. Introduction
2. Morphological operations
3. Inflection
4. Derivation
5. Compounding
6. Cliticization
7. References

##### 1. Introduction

English is spoken throughout the world, being the first language of roughly 350 million people. The main dialect divisions are between Britain, the USA, Australia/New Zealand and South Africa. Within these regions there is considerable further differentiation, especially in Britain. It is a member of the West Germanic group of languages, having separated from its nearest neighbours (Frisian, Dutch and German) some 1,500 years ago with the Anglo-Saxon colonization of Britain.

There has been considerable scholarship on present-day morphology and on its historical development. The most detailed description of word formation and compounding is Marchand (1969) (who can be consulted for earlier references). Also useful are Adams (2001), Bauer (1983), Carstairs-McCarthy (2002), Plag (2003) and Szymanek (1989). Handbooks such as Quirk et al. (1985) and references cited there provide information about inflection. Almost every theoretical discussion of morphology within the generative framework contains some discussion of English. Particularly important works in this tradition are Aronoff (1976), Di Sciullo & Williams (1987), Hoeksema (1986), Lees (1963), Levi (1978), Lieber (1980), Malicka-Kleparska (1987), Scalise (1984), Selkirk (1982), Siegel (1979), Szymanek (1985).

##### 2. Morphological operations

The **morphological operations** found in English are affixation (prefixing and suffixing), ablaut, compounding and cliticization. There is very limited use of stress alternation to signal derivation (see 4.3.1). In addition, there is the zero operation (null affixation or conversion; cf. Art. 45, 90). Affixation, ablaut and zero formation are used alike for inflection and derivation. There is a limited degree of reduplication, mainly in what might be called “expressive morphology” (cf. Pullum & Zwicky 1987), e.g. *silly-billy*. Cf. also examples from baby-talk, e.g. *choo-choo* ‘train’.

Affixation, compounding and cliticization are recursive. Inflectional affixes do not allow stacking, however. It is claimed (e.g. Fabb 1988) that some derivational suffixes are restricted to absolute final position.

###### 2.1. Affixation

All inflectional **affixes** are suffixes. Both prefixes and suffixes are used for derivation. Expletives and swearwords can be inserted inside morphemes, giving the impression of infixation, as in *abso-bloomin'-lutely*, or *fan-bloody-tastic*. These are claimed to obey phonological restrictions defined in terms of foot structure (McCarthy 1982). However, it has also been argued that this is not a genuine morphological process, but an example of ‘expressive morphology’, hence, more akin to word play (Pullum & Zwicky 1987).

Affixes are divided into two broad classes on phonological and morphological grounds. The Class I ('+-boundary', Root) affixes frequently trigger alternations in the lexical phonology, particularly stress shift. They are often called “stress-sensitive” (Siegel 1979). The Class II ('#-boundary', Word) affixes only trigger processes in the phrase phonol-

ogy, and specifically never affect stress (“stress neutral”). In the theory of Lexical Phonology (Kiparsky 1982) it is claimed that all Class I affixation precedes Class II affixation. This is disputed by many (Aronoff & Sridhar 1983; Fabb 1988; cf. Spencer 1991: 179–182).

Typical examples of the Class I/II distinction are the negative prefixes *in-* and *un-*. The former obligatorily undergoes assimilation before consonant-initial stems: *impossible*, *illegal*, *irregular*, while *un-* only undergoes optional, phrase-level assimilation: *unpardonable*, *unlikely*, *unruly*. A great many of the Root (Class I) affixes attach to bound stems. By its nature, it is difficult or impossible to identify the syntactic category of a bound stem. Many such affixes also attach to free morphemes or morphologically complex words, e.g. *-able/ible* in *readable* as opposed to *possible*. It is sometimes found that a prefix will ‘potentiate’ a suffix. Thus, a certain number of nominalizations in *-ment* are formed from denominal verbs derived by prefixation of *en-* (*enslavement*, *enthronement*, *entombment*).

With the exception of [i<sup>ŋ</sup>] all the regular inflectional affixes have coronal consonants. The formatives [z, d] have three main allophones. After any sibilant (/s, z, ʃ, ʒ, tʃ/) [z] takes the form [əz]; after a non-sibilant voiceless sound it is [s]; elsewhere it is [z]. After /t, d/ the [d] formative takes the shape [əd]; after any voiceless sound except /t/ it is [t]; elsewhere it is [d].

It is not uncommon for affixation to be conditioned by the phonological shape of the stem (see Art. 47). For instance, the nominalizing suffix *-al* attaches only to stems with final main stress (e.g. *arrival*, *referral* but not *\*differal*). Of some theoretical interest is a case such as the inchoative/causative suffix *-en* ([n]) as in *weaken*. This is constrained to attach to monosyllabic adjective stems ending in at most one obstruent optionally preceded by a sonorant. Thus, we can form *lengthen*, *sweeten*, *loosen*, *shorten*. Interestingly, stems which violate this restriction ending in *-st* or *-ft* can still accept the affix provided the /t/ subsequently deletes: *moisten*, *soften* (cf. Halle 1973; Siegel 1979). This restriction does not apply to other homophones of *-en*. For instance, we can have denominal adjectives such as *flaxen* (from *flax*) and past participles such as *fallen*, *blown*, *torn*.

Many affixes derive from full words and in a number of cases it is not clear whether a given formative is a genuine affix or an element in a compound. Marchand (?1969: 356–358) distinguishes a class of “semi-suffixes”, such as *-wise* (as in the rather colloquial *morphology-wise this is a puzzling construction*). An oft-cited example of the shift from compound to affix status in progress is that of *-man* as in *postman*, *chairman* and so on. Evidence of the reduced, affixal status comes from the pronunciation with a reduced vowel or even a syllabic nasal. (Contrast, say, *superman* in this regard, with a full, though unstressed, vowel). Strauss (1982) has argued that formatives such as *anti-* and *hypo-* are really elements of neo-classical compounds, rather than prefixes.

## 2.2. Other types of operation

Compounding and cliticization are dealt with in 5 and 6. Other types of word formation are **blending** (e.g. *smog* from *smoke* + *fog*; cf. Art. 91) and **clipping** (e.g. *mike* from *mic[rophone]*; cf. Art. 92). **Backformation** (frequently through the mechanism of folk etymology) provides a wealth of new words, e.g. the verb *aggress* (cf. *aggressor*, *aggression*, *aggressive*).

**‘Stub’ compounds**, formed by coalescence of parts of words, are common in proper names, especially of institutions, though the process does not reach the proportions found in Russian. Except for proper names they tend to have a marked colloquial nuance. Examples are *Caltech* (Californian Institute of Technology), *Linac* (Linear accelerator), *sit-com* (situation comedy). It is often difficult to distinguish such formations from blends, clippings, and other forms of (orthographically influenced) word creation.

**Acronyms** (cf. Art. 92) are extremely common and have considerable sociological importance in political and advertising spheres. More often than not the source of the acronym is forgotten, so that it was common, for instance, to hear journalists speak about *SALT* talks, even though *SALT* itself stands for Strategic Arms Limitation Talks. Occasionally words are simply invented (apparently the case with the trademark *Kodak*).

## 3. Inflection

Modern English has lost much of the inflectional system inherited from Old English, principally due to phonological changes af-

fecting suffixes. The following inflectional categories are found:

- nouns: ‘number’;
- adjectives: ‘comparison’;
- verbs: ‘tense’, ‘agreement’, ‘gerund’.

### 3.1. Nouns

**Nouns** used in a non-count sense take a plural inflection. The regular form is a suffix whose underlying phonological form is generally given as [z] (though alternative analyses are possible). This is a borrowing from Norman French. Irregular plurals may be formed by ablaut (e.g. *tooth* ~ *teeth*, *man* ~ *men*), or by zero affixation (e.g. *sheep* ~ *sheep*). A handful of words retain the Germanic *-en* desinence: *oxen*, *children*, *brethren*. A good many foreign borrowings, especially from classical languages, retain the original plural form: *alga* ~ *algae*, *phenomenon* ~ *phenomena*, *cherub* ~ *cherubim*, *mafioso* ~ *mafiosi*. However, in many cases there is a regular alternative: *concerto* ~ *concertos*, but usually *concerti grossi* (from *concerto grosso*). The irregularity of plurals of the type *criterion* ~ *criteria* is indicated by the frequent (mis)use of the plural as a singular form. This is especially noteworthy with the form *datum* ~ *data* in which the form *data* (*this data shows that...*) is used as a singular even in linguistics articles.

### 3.2. Adjectives

The **comparative** is the suffix *-er* and the **superlative** is *-est*. Their attachment is conditioned phonologically: the result may not extend beyond a single foot in underlying phonological form. Hence, we have *green-er*, *happi-er*, *nobl-er* but not *\*porous-er*. Otherwise, comparison is analytic, with *more*, *most*, which is also a permissible alternative for most adjectives taking *-er/est*. Some adjectives take (partially) suppletive comparative/superlative forms: *far* ~ *further* ~ *furthest*; *bad* ~ *worse* ~ *worst*.

The formation of manner adverbs by *-ly* is generally regarded as derivation (since it involves a category change). However, if we regard such adverbs as distributional variants of adjectives (dominated by a verb phrase in syntax; cf. Emonds 1976), then we could say there was no class of adverb, and *-ly* could be regarded as an inflection.

### 3.3. Verbs

There is a word class division into **lexical verbs** and **auxiliary verbs**. Some belong to both classes (*have*, *be*, *do*, *need*, *dare*).

Regular lexical verbs have only four distinct forms:

- base form;
- 3rd person singular subject agreement form (present indicative only): [z];
- past tense/past participle/passive participle: [d];
- present participle/gerund: [iŋ].

The base form is used for the imperative and the little-used subjunctive. Combined with the particle *to* it forms the infinitive.

The 3rd person agreement form is regular for nearly all verbs (including most of the ‘strong’ ones). Some dialects have lost this affix, thus regularizing the verb paradigm. In other dialects the 3rd person form has spread to other person/number forms, also regularizing the paradigm.

The past tense and past/present participle forms are distinct only in irregular verbs. The term “past participle” refers to the form used with the auxiliary verb *have* to form the perfect tenses, the “passive participle” forms the passive voice with the auxiliary *be* (or sometimes *get*). These forms are never distinct, even in irregular verbs. It frequently functions as an adjective, including with intransitive verbs provided they are unaccusative, e.g. *a freshly killed chicken*, *a well travelled man*.

The [iŋ] form is derived from the base form in all cases. It functions as a present participle (with continuous tense forms, using the auxiliary *be*) and as a gerund(ive). It frequently functions as an adjective (*the falling rain*).

There are some two hundred or so ‘strong’ verbs, whose past tense/participle forms are irregular in various ways. In some cases there is a distinct suffix, [t], as in *keep* ~ *kept* ~ *kept*. In others there is differentiation between past tense and past/present participle forms, e.g. by ablaut in the past tense and suffixation of [n] in the participle: *speak* ~ *spoke* ~ *spoken*; *fall* ~ *fell* ~ *fallen*; *write* ~ *wrote* ~ *written*. In others, both inflected forms are ablauts: *sing* ~ *sang* ~ *sung*. In some cases there is no change: *cut* ~ *cut* ~ *cut*. Dialectally, it is common for strong verbs to drift into the regular class. The standard language has doublets such as *light* ~ *lit* ~ *lit*, alongside the regular *light* ~ *lighted*.

There is a handful of verbs with additional irregularities, including *be* (which has the largest number of distinct forms, eight), *do*, *have*.

The modal auxiliary verbs show the most deviation: they have only a base form, though some analyse forms such as *could*, *would*, *might* as past tense forms of *can*, *will*, *may*.

#### 4. Derivation

The three major lexical categories, noun, verb, adjective, undergo predominantly **affixal derivation** to derive new words. The morphological processes vary in productivity from the highly circumscribed, historically fossilized, to the more or less fully productive.

Morphological conversion (or zero affixation) frequently relates noun-verb pairs. It is often difficult to decide which of the pair is primary. Virtually any underived noun can be used as a verb provided it can be pragmatically interpreted (cf. Downing 1977). Verb-to-noun conversion is also frequent.

Derivational processes can be roughly divided in those which have an essentially grammatical function, as opposed to those which add semantic elements. Among the former will be simple nominalizations of verbs and adjectives, causative or inchoative verbs derived from nouns and adjectives and so on. Among the latter will be diminutives and augmentatives, the feminine suffix *-ess*, and so on. Frequently, an affix with a predominantly grammatical function (simply category changing) will form polysemous words with other, more specialized meanings. Nominalizations are notorious for this, though it happens with other categories, too. For instance, *artistic* can mean simply ‘pertaining to the artist’ as in *she has an artistic temperament*, though this usage is fairly rare. (We could not say, for instance, \**the average artistic income varies from month to month*, meaning ‘the income of the average artist’). However, the most common meaning of the word is something like ‘having the character of an artist or work of art’ (*she’s very artistic, he laid out the furniture very artistically*).

The great majority of prefixes simply add semantic content to the base. As a result very few prefixes are genuinely derivational in the sense of being capable of changing category. For this reason, suffixes will be dealt with in 4.1–4.3, separately from prefixes, 4.4.

##### 4.1. Words derived by suffixation from nouns

###### 4.1.1. Derived nouns

There are few genuinely productive **suffixes**. Diminutives (cf. Art. 99) are a poorly repre-

sented class in English, and they tend to have additional connotations. The suffix *-ette* (acquired through French) is found though mainly to form proprietary names or other specialized uses (e.g. *kitchenette*). The Germanic *-ling* (e.g. *duckling, gosling*) is unproductive and rare, as is *-let* (*piglet, starlet*). The suffix *-y/ie* is particularly common in babytalk (*doggy, horsey*) and in forming hypocoristics (*Billy, Lizzie*, from *William, Elizabeth*).

Unproductive suffixes with largely grammatical function include *-dom*, *-hood*, *-ship* (e.g. *kingdom, serfdom; boyhood, brotherhood; fellowship, friendship*). A productive suffix is *-(ia)na*, meaning ‘things connected with the person named in the stem’, as in *Wagneriana, Victoriana*. A frequently found suffix is *-ism*, particularly attached to proper names to refer to a doctrine associated with the person named in the stem: *Marxism-Leninism, Taoism, Darwinism, Satanism*; cf. also *gangsterism*. This may have just the general function of creating an abstract noun, irrespective of the semantics of the base: *totemism, egoism*. Usually, however, *-ism* will attach to adjectives, e.g. *Lutheranism, feminism*. Doublets sometimes occur, so that while *Spencerianism* refers to the tenets of the nineteenth century thinker Herbert Spencer, the term *a spencerism* could be coined (and sometimes is coined) ad hoc, to refer to specific quirks of behaviour associated with a particular person called “Spencer”.

Corresponding to *-ism* is the agentive suffix *-ist*, as in *Marxist-Leninist*. This sometimes seems to condition allomorphy as in *flautist* (from *flute*). Related in meaning to doctrinal *-ist* is *-ite*, frequently added to personal names (*Benthamite, Thatcherite*). Agentives can also be formed from nouns by *-ian*: *grammarian, musician, librarian*. A number of adjectival suffixes can be used to form nouns referring to people (e.g. *American, Davidson-ian, Iraq-i*).

Other such suffixes include *-acy* (*piracy*), *-age* (*parentage*), *-ance/lence* (occasionally: *tenancy*), *-arian* (occasionally: *Parliamentarian*), *-ate* (as in *consulate*), *-cy* (*baronetcy*), *-eer* (*engineer*), agentive *-er* (as in *potter, jeweller, petitioner*), *-er* specifying provenance (*Londoner*), *-(e)ry* (*jewellery, Jewry*), *-ese* (as in *journalese*), *-ess* to form feminines (*lioness*), *-ful* (as in *handful*), *-in(e)* (to form names of chemicals, e.g. *chlorine, aspirin*), *-ing* (as in *bedding, towelling, matting*), *-y* (as in *ministry, mastery*).

In Spencer (1988) it is noted that a lexicalized phrase such as transformational grammar or theoretical linguistics productively licenses a ‘personal noun’ formed from the agentive of the head of the phrase: *transformational grammarian, theoretical linguist*. This is a form of paradigmatic word formation, giving rise to so-called “bracketing paradoxes” (cf. Art. 24, 82).

#### 4.1.2. Derived adjectives

Adjectives can be derived with relatively few semantic connotations by the suffix *-al* (*accidental*) and *-ic(al)*. The latter very often attaches to bound roots with Latin(ate) origins, e.g. *economic(al)*, though it may also attach to words, as in *patriotic, athlete*. Note that *-ic* attracts stress to the preceding syllable.

There are several affixes for forming adjectives from names of countries, including *-(i)an* (*African, Moldavian*; cf. also *Elizabethan, Skinnerian*), *-ese* (*Chinese*), *-i* (used mainly with Middle East countries, *Israeli, Iraqi*), *-ish* (*Danish*). Note the contrast between these and *-ic*, meaning ‘having the properties of’ (*Asiatic, Hellenic*) or referring to a group (*Germanic, Cushitic, Slavonic*). The suffixes *-ern/erly* apply to points of the compass (*western/westerly*).

Other suffixes include *-able* (as in *comfortable, fashionable*), *-ary* (*cautionary, revolutionary*), *-ate* (*affectionate, passionate*), *-en* (*wooden*), *-esque* (*picturesque, Kafkaesque*), *-ful* (*careful, playful*: note this does not attach to verbs, as claimed by Rooper 1987: 227f.; cf. Art. 83), *-ine* (*alpine, elephantine*), *-less* (*boneless, fearless*), *-oid* (*spheroid, humanoid*), *-ous* (*glorious, poisonous*, and with truncation of a nominal suffix, *ambitious, superstitious, contentious*), *-some* (*bothersome, troublesome*), *-sy* (*bitsy, folksy* colloquially only), *-ward(s)* (*homeward, sideways*), *-y* (*catty, dirty*). Marchand (1969) includes the following ‘semi-suffixes’: *-like* (*child-like*), *-worthy* (*blameworthy, praiseworthy*), *-wise* (*clockwise*; these formations are usually used adverbially: *crosswise, lengthwise*).

The *-ed* suffix forms what appear to be idiosyncratic past participles of nonexistent verbs, as in *birdbrained, knock-kneed, short-sleeved, four-sided*. Moreover, these appear to be so-called “bracketing paradoxes”. As far as I can tell, there is only an etymological connection with the past participle form, however. This appears to be an instance of paradigmatic word formation (see Spencer 1991: 417; cf. Art. 24, 82).

#### 4.1.3. Derived verbs

There are no important productive suffixes for creating denominal verbs. Some unproductive examples are causatives such as *-ify, -ize* (as in *beautify, glorify* and *motorize, harmonize*), though these are principally found with adjectives. Scientific terminology abounds in terms such as *chlorinate, hydrate* (cf. also *orchestrate*). With a prefix we also have examples such as *encapsulate, defenestrate* (see Plag 1999).

#### 4.2. Words derived by suffixation from adjectives

##### 4.2.1. Derived adjectives

Suffixation seems to be limited to *-ish* (*greenish*), with occasional cases such as *good-ly* and *weary-some*. If numerals are regarded as adjectives, then the formation of ordinals from cardinals by *-th* would count here (*fifth, sixth*).

##### 4.2.2. Derived nouns

A variety of suffixes are used to create abstract nouns meaning something like ‘quality of Adjective’. The most productive (and well-known) is *-ness*. This can often be found even when there is a conventional alternative. Thus, alongside *grammatical-ity* we might hear *grammatical-ness*. The suffix *-ity* is restricted to the latinate or French derived vocabulary. However, it is extremely productive with certain types of affixed adjective, notably those in *-able, -ic, -al, -id* (cf. Aronoff 1976: 53; Marchand 1969: 314). The suffix *-ism* was mentioned in 4.1.1 as attaching to adjectives as well as nouns to indicate a doctrine. Likewise, *-ist* attaches to a small number of adjectives (*specialist, socialist*), especially in its meaning of ‘professional specialist’, e.g. *Russianist, orientalist*.

Bauer (1983: 222) describes a kind of consonant mutation in the formation of nouns such as *dependence* from *dependent*. Marchand (1969: 248), however, analyses these as derived from the corresponding verbs by affixation of *-ance/lence*. The evidence points in both directions. On the one hand, *elegance, exuberance, obsolescence* can only be derived from adjectives. On the other hand, *appearance, clearance, maintenance* have no adjective in *-ance/lence* from which they could be derived.

Other suffixes include *-er* (as in *fiver, tenner*, meaning ‘five-/ten-pound note’), *-hood* (*falsehood, likelihood*), *-th* (*length* (with ab-

laut), *warmth*), -y (principally from adjectives in *-ant/ent*, with accompanying spirantization of /t/ to /sl/, spelled c; note that Marchand 1969: 249, lists this under *-ancy/ency*, while Bauer 1983: 222 lists it under *-cy*). Examples are *discrepancy*, *infancy*, *stridency*, *valency*. Bauer (1983: 222) includes the suffix *-dom*, but I can only find two examples: *freedom*, *wisdom*.

#### 4.2.3. Derived verbs

The two most important suffixes are *-ify* and *-ize*, both with causative meanings. However, although there are examples such as *solidify*, *falsify* in which we seem to have straightforward derivation from an adjective, the vast majority of verbs in *-ify* are formed from bound roots of indeterminable syntactic category, e.g. *economize*, *mechanize*, *optimize*. Similarly, *-ize* attaches both to nouns and adjectives, though the basis of choice is not always clear. A good many adjectives in *-al* accept *-ize* (*centralize*, *equalize*, *vocalize*). The inchoative/causative *-en* has been mentioned in 2.1.

### 4.3. Words derived by suffixation from verbs

#### 4.3.1. Derived nouns

Verbs form mainly process or result **nominalizations**, or agents/instruments and patients. In addition to conversion, the principal suffixes giving nominalizations are *-age* (*breakage*), *-al* (*arrival*), *-ancelence* (*dependence*) (but cf. 4.2.2), *-ation/ition/ution/tion/lion* (*declaration*, *addition*, *solution*, *absorption*, *reception*) (see Aronoff 1976: 100–104), *-ment* (*amusement*). A small number of disyllabic verbs with final stress form nouns with initial stress (*transpórt* – *tránsport*). All verbs form a nominalization with *-ing*, and the infinitive is often used with nominal force. Agent/instrument nominals are formed by *-er/or* (*baker*, *stabilizer*, *actor*, *elevator*). Patients are formed by *-ee* (*payee*), *-ling* (*hireling*). (Note that “patient” here means any obligatory internal argument.)

The morphosyntax of ‘derived nominalizations’ has been the subject of intense debate in syntax and morphology. A description with theoretical discussion can be found in Malicka-Kleparska (1987) (see also Art. 34, 83, 89). English has no systematic way of distinguishing between nominalizations representing events or processes and those representing results or states. Thus, it is not un-

common to encounter polysemy of the type shown by the word *construction*, which can mean ‘act of constructing’, (*the construction of the dam took two years*), ‘abstract result’ (*the construction of the dam cost millions*) ‘concrete result’ (*the construction is eighty meters high*), ‘manner’ (*the construction of the dam is very unusual*) (cf. Bauer 1983: 185–189).

#### 4.3.2. Derived adjectives

Suffixes deriving adjectives are principally latinate, frequently triggering morphophonemic alternations typical of latinate affixation. Examples are *-ive* (*permissive*, *creative*), *-atory* (*circulatory*, *explanatory*), and *-ant/ent* (*repentant*, *persistent*). Some include *-ful* here though this generally attaches to nouns (*resentful* and *rueful* are the only clear cases of deverbal formation). The most important suffix of this derivational type, and the only fully productive one, is *-able/ible*. This attaches to transitive verbs to give the meaning ‘such that can be Verbed’ and is especially found with negated stems (*washable*, *unreconstructable*, *invincible*). Thus, it is predicated of the object of the verb stem and is sometimes called the “objective potential form”. There is a homophonous latinate suffix which may alter the stress pattern and induce idiosyncratic meaning changes (e.g. *lamentable* from *lamént*; n.b. minimal pairs such as *uncompáráble* vs. *incómparable*). The nominalization is regularly formed with *-ity* (*washability*, *invincibility*).

#### 4.3.3. Derived verbs

In some languages verbs can be derived from verb stems to form diminutives, causatives or aktionsart alternants. None of these is found in English.

### 4.4. Prefixation

**Prefixation** hardly ever gives rise to a category change. The main exceptions are unproductive *be-*, *en-*, creating verbs (as in *be-head*, *bedevil*, *enslave*, *enlarge*, *entangle*), and productive *de-*, as in *de-frost*. A good many prefixes attach promiscuously to two or three major lexical categories. Which category is selected seems to depend primarily on the meaning of the prefix. Thus, *counter-*, with a loose meaning of ‘opposite’ attaches to nouns, verbs, or adjectives (*counter-espiionage*, *counter-attack*, *counter-productive*). On the other hand, *un-* with the meaning ‘opposite of’ attaches only to adjectives (*unhappy*)

while in its reversative meaning it attaches only to verbs (*untie*). It never attaches to nouns.

Prefixes differ as to whether they induce morphophonemic alternations (such as stress shift onto the prefix) or not. As with the suffixes, a good many formatives have a dual character, such that in one of their senses or uses they do not affect the phonology, while in the other sense or use they do. Examples of this are found with the prefixes *anti-* and *hyper-*. Thus, we find *anti-coágulant*, *anti-héros*, *anti-wár*, with stress on the base, but *antíthesis*, *antípodés*, etc. with retraction of stress to the second syllable of the prefix. Similarly, we find *hyper-áctive*, *hyper-véntilate*, *hyper-thýroidism*, but *hypérnymy*, *hypérbole*. When the prefix itself is stressed it almost always forms part of a neo-classical borrowing or coining (and hence tends to adopt the stress pattern of the Greek original). When the prefix is not so stressed it is being used as a genuine (often productive) English morpheme.

In addition, a number of prefixes double as genuine morphemes with determinate semantics and cranberry morphemes. These are often distinguished phonologically. A well-known case is that of *re-*. In its productive use it attaches to verb stems with the meaning ‘do again’. However, it is also a very common element in latinate compound verbs (see 4.5). In its latter use it is unstressed and causes voicing of /s/ to /z/; cf. *assign*, *consign* but *resign* (= ‘capitulate’). However, in its reiterative sense it bears secondary stress and does not trigger voicing: *rè-sign* (‘sign again’). The prefix *de-* behaves similarly.

The phonological difficulties with prefixes are paralleled by the difficulty of deciding whether we are dealing with genuine prefixation or a type of compounding. The problem is that many prefixes exhibit a looser cohesion with their base than would be expected from genuine affixes. The most salient such property is that Class II (Word) prefixes can be conjoined, which is impossible with suffixes, or with the strongly cohesive Class I (Root) prefixes. Thus, we can say *pro- and anti-war demonstrations*, *hypo- or hyper-active children*. This is even possible (in scientific discourse at least), with prefixes attaching to bound bases (that is, bases which are not themselves words): *sub- and supra-liminal (levels of excitation)*.

In favour of a prefixation analysis is the observation that when attached to verbs

these formatives narrow the semantic and syntactic properties of the verb stem (Carlson & Roeper 1980). Thus, we can say *to calculate their time of arrival*, *to calculate when they will arrive*, *to miscalculate their time of arrival* but not *\*to miscalculate when they will arrive*. Similarly, prefixation by *out-* creates a transitive verb, including when the original base is an intransitive verb, as in *outbid*, *outlast*, *outshine*, or even a noun, as in *outdistance*, *outwit*, and in the formula illustrated by *to out-Reagan Reagan*.

Prefixation accompanied by suffixation often gives rise to what are sometimes called “bracketing paradoxes”, in which the morphological constituent structure seems to be at variance with that implied by the phonology or the semantics (cf. Art. 82; Spencer 1991: ch.10). A classic case (cf. Pesetsky 1985) is that of the comparative of adjectives negated by *un-*. The restriction on the *-er* comparative morpheme preventing it from attaching to polysyllabic bases (see 3.2 above) is systematically violated by words such as *unfairer*, *unhappier*. Phonologically the suffix attaches, therefore, to the unprefixed root: [[*un*[*fair-er*]], [[*un*[*happy-er*]]. However, these words mean ‘more unfair/unhappy’, in which the scope of the suffix includes the negative prefix, implying a semantic structure [[*unfair**er*], [[*un-happy**er*]]. Likewise, if we consider a word such as *international*, we see that this must derive morphologically from affixation of *inter-* to the adjective *national* (there is no intermediate form *\*internation* which could undergo suffixation by *-al*). But the meaning of ‘betweenness’ provided by the prefix *inter-* modifies the base nation, and it is this semantic composite which is rendered adjectival by suffixation of *-al*. So semantically the word is [[*inter-nation**al*]], while morphologically it is [[*inter*[*nation-al*]]]. Cases of this sort are legion.

#### 4.5. Bound roots and affixes

The influx of words of Romance origin has meant that a good deal of Romance morphological structure has been borrowed, especially in the form of allomorphic variation. However, since borrowed morphological complex words have often suffered drastic semantic shifts, the link between form and meaning is often difficult or impossible to discern. Consequently, it is easy to find cases of latinate cranberry morphemes (cf. Aroff 1976).

A well-known example of this phenomenon is the set of latinate verbs formed from a prefix, such as *ad-*, *ab-*, *con-*, *de-*, *in-*, *re-*, *sub-*, and bound root, such as *-ceive*, *-fer*, *-late*, *-mit*. Prefixes show allomorphic variation in the final consonant, e.g. *con-fer*, *col-late*, *com-mit*; *ad-vise*, *ac-cept*, *as-sume*, *af-firm*. Roots show wider variation, apparent when the verbal root undergoes derivation, e.g. *con-ceive* ~ *con-cept-ion*, *de-ceive* ~ *de-cept-ion*, *per-ceive* ~ *per-cept-ion*, *re-ceive* ~ *re-cept-ion*; *ad-mit* ~ *ad-miss-ion*, *com-mit* ~ *com-miss-ion*, *per-mit* ~ *per-miss-ion*, *trans-mit* ~ *trans-miss-ion*. Even a fairly cursory inspection of such examples shows that there is no common semantic feature either to the prefixes or to the roots in such cases. Thus, morphemes which are identifiable on morphological grounds, cannot be given a semantic characterization (Aronoff 1976).

The phenomenon is not restricted to pre-fixed verbs, however. There is a fair number of suffixed forms in which it is difficult to assign a meaning to the constituent morphemes, though there is some evidence from morphophonemic alternations that the words are morphologically complex. Some examples, all forming adjectives, include *-ant* (*malignant* ~ *malignancy*, *pregnant* ~ *pregnancy*, *vacant* ~ *vacancy*), *-ile* (*hostile* ~ *hostility*, *servile* ~ *servility*, *virile* ~ *virility*), *-alible* (*amiable* ~ *amiability*, *capable* ~ *capability*, *possible* ~ *possibility*).

The last example illustrates another common phenomenon, that of dual membership. This *-alible* suffix is generally regarded as a Class I or Root suffix, because it can cause stress shift as in *compáre* ~ *cómparable* (cf. also *incómparable*, not \**uncómparable*), or trigger allomorphy (e.g. *defend* ~ *indefensible* (*actions*)). However, there is another variety of this suffix which attaches to words (not bound roots), and which is stress neutral (Class II, Word affix), as in *compáre* ~ *com-páráble* (cf. *uncompáráble*, not \**incompáráble*), or *defend* ~ *undefendable* (*position*).

The most obvious phonological properties of these suffixes or suffix like elements relate to stress. For instance, there are a good many adjectives ending in a syllable *-ic*, which is never stressed when word final, and which attracts the stress to the previous syllable, whether it is heavy or light (cf. *misánthropy* ~ *misanthrópic*). Likewise, a good many adjectival suffixes seem to play no role in stress assignment (i.e. they are extrametrical, cf. Hayes 1981). Thus, in a whole host of words

ending in *-al*, *-ar*, *-ant*, *-ous*, stress is assigned to the penultimate if heavy and the antepenultimate otherwise (e.g. *pésonal*, *anecdótal*, *dialéctal*, cf. Chomsky & Halle 1968: 81f.). This is also the case where the base does not seem to be a word in its own right (or even a morpheme): *cóporal*, *sacerdótal*, *fraternal*.

## 5. Compounding

**Compounding** is very productive and a major source of new coinings. Three interesting types are noteworthy: Germanic compounding, which can be subdivided into root (primary) compounding and synthetic (verbal) compounding and neo-classical compounding. We also find exocentric types, including appositional (copulative) compounds and bahuvihi compounds.

### 5.1. Germanic compounding

The Germanic compounding type is an endocentric construction with precisely one head and one modifier. Only words or lexicalized phrases are compounded, never roots. Inflected forms are only regularly compounded in the form of noun plurals. The head defines the syntactic and semantic properties of the whole compound. The process is recursive.

Usually, but not invariably, stress falls on the left element of the compound (so-called “compound stress”). It is very difficult, however, to see any real pattern to compound stressing (for an interesting account see Zwicky 1986). It is therefore difficult to distinguish between genuine compounds and syntactically formed complex nominals. For instance, while *fire brigade* is presumably a compound, *London fire brigade* would generally be analysed as a noun modifier (*London*) syntactically modifying a head noun (which happens itself to be a compound). However, it is difficult to find solid criteria on which to base such judgements.

Closely related to noun-noun compounds are constructions in which a denominal adjective modifies a noun. For instance, corresponding to *atom bomb* we have *atomic bomb*, with essentially the same meaning and use. Levi (1978) regards these constructions as sufficiently similar to warrant parallel analysis and uses the term “complex nominal” to cover both. However, it is not possible to construct analogs to synthetic compounds in this way. Thus, while *child abuse* and *infantile abuse* have essentially the same

extension, the former means ‘the abuse of children’, while the latter means ‘abuse which takes place during childhood’. Moreover, it is impossible to replace *atom smasher*, *atom smashing* with *atomic smasher*, *atomic smashing*. More research is needed on the relationship between compounding, nominal modifiers and denominal relational adjectives.

Compounds are spelt either as one word, two separate words or hyphenated, on an entirely ad hoc basis which seems to reflect, if anything, degree of lexicalization. A good many words which are synchronically monomorphemic were originally compounds (e.g. *husband*, *nostril*, *window*).

### 5.2. Root compounding

Nouns (N) and adjectives (A) freely compound with each other. Prepositions (P), which semantically behave like adverbials in compounds, freely appear as modifiers, though not as heads. Verb (V) headed compounds are rare or nonexistent, with the exception of P V compounds. Some examples are given below:

- N + N: *horserace* (type of race), *racehorse* (type of horse)
- A + N: *bighead*, *nervous system*, *hummingbird*
- P + N: *undercoat*, *through-road*
- V + N: *go-cart*, *drawbridge*
- N + A: *trigger-happy*, *birdbrained*, *earth shattering*, *house proud*
- A + A: *ice cold*, *bright red*
- P + A: *off-white*, *inborn*
- P + V: *overlook*, *undermine*, *upgrade*, *output*

Many of these groups have their own subtypes. Marchand (1969: 60–82) gives a detailed description. In some cases of compounding with P it is difficult to distinguish the construction from prefixation. For instance, should *over-ripe* be regarded as a P + A compound or as a prefixed form of *ripe*? A number of established (lexicalized) compounds may be difficult to categorize because of conversion. For instance, if we assume that *pass* is basically a verb, then we could analyse the verb to *by-pass* as a result of direct compounding, i.e. [*by*[*vpass*]], or as a combination of P + N derived from V, with the whole compound then undergoing conversion back to V, i.e. [*v*[*Nby*[*N*[*vpass*]]]].

A number of verbal derivatives of the type *withstand*, *undergo*, *overhang*, go back to original P + V compounds. When the head

V is a strong verb it inflects as though it were a simplex verb, e.g. *withstood*, *underwent*, *overhung*. This distinguishes such cases from superficially similar constructions in which a verb first undergoes conversion to a noun, then undergoes compounding, and then is converted back into a verb. As in nearly all such cases of double conversion, the resulting verb is inflected as a weak verb. A much-discussed example is the contrast between *withstood* and *grandstanded*. The latter has the form [v[v[N[Aggrand][N[vstand]]]]]ed] (cf. Kiparsky 1982).

A curious type of compound is illustrated by words such as *bird-brained*, *three-wheeled*, *short-sleeved*, *five-sided*. These appear to be headed by a past participle. However, there is no base verb from which such a past participle can be formed. They are in fact formed by affixing *-ed* directly to (lexicalized) phrases denoting body parts, clothing parts, external geometrical features and other inalienably possessed parts (cf. Spencer 1991: 417). The *-ed* formative is therefore similar to the Hungarian phrasal affix *-úl/ú* which only attaches to noun headed phrases to form an adjectival phrase, as in *kék szémű* ‘blue-eyed’ (*kék* ‘blue’, *szém* ‘eye’) or *magyari nyelvű* ‘pertaining to the Hungarian language’ (*magyari* ‘Hungarian’, *nyelvű* ‘language’).

### 5.3. Synthetic compounding

It is possible to form a N + N or N + A compound in which the head is a deverbal nominal or a participle, and the modifier bears some sort of syntactic relation to the head. The clearest cases are those in which the modifier is the direct object of the verb, as in *truck driver*. This can be related to (X) *drives trucks*. The resulting combination bears a certain resemblance to noun incorporation in a good many groups of languages such as Iroquoian, Polynesian, Chukotko-Kamchatkan and so forth (cf. Art. 88). The construction poses the theoretical problem of ensuring that syntactic constraints, such as satisfaction of argument structure, are met in the compound (cf. Art. 34, 83, 87). A descriptive is deciding what counts as a case of synthetic compounding.

The commonest types of uncontroversial synthetic compound are agentives of the *truck driver* type, and gerund/participle forms such as *truck driving*. The latter retain certain aspectual properties of verbs (cf. Roeper 1988). Many authors would include passive participle headed constructions, such as

*moth-eaten, hand-made, pan-fried*). In the latter the modifier has the function of an adverbial (e.g. agentive *eaten by moths*, instrumental *made by hand*, locative *fried in a pan*). Some authors (e.g. Selkirk 1982) include compounds with nominalizations of verbs other than -ing gerunds, such as *slum-clearance*, *self-deception*, *troop deployment*, as well as compounds of deverbal adjectives such as *water repellent*, *machine readable*, *disease inhibitory*.

Synthetic compounds can easily be confused with root compounds formed from a deverbal noun whose base can be used intransitively. For instance, in addition to *truck driver* we could coin *motorway driver* meaning ‘one who drives (regularly) on motorways’. (This construction has primary stress on *motorway*, so it is clearly a compound). However, this is not a synthetic compound, rather, it is a root compound, whose head is a derivative of *drive* used intransitively. With the handful of verbs which must be used transitively, it is all but impossible to form such root compounds. For instance, while we can say *omelette maker* we could not say *pan maker* meaning ‘one who makes (e.g. omelettes) in a pan’. This is because *make* is very difficult to use intransitively.

#### 5.4. Neo-classical compounding

In scientific language it is extremely common to find compounds formed from Latin or especially Greek roots, neither of which exist as separate words. In most cases, these are separated by the meaningless intermorph *-o-*. In the forms which are commonly used this process appears not to be recursive, at least not in the way that Germanic compounding is.

Many of the roots can appear as first or second member, though some appear to be positionally fixed. Thus, we can have *telegram*, *gramophone*, *phonogram*, *telephone*, but not *\*gramotel* (since the *tel* element only appears in first position). There is a fine dividing line between compounding of this kind and prefixation of neo-classical roots. It is difficult, for instance, to determine whether *electro-* should be treated as a kind of prefix or as an obligatorily initial compound element. These questions are discussed in more detailed by Bauer (1983: 213–216).

#### 5.5. Exocentric compounds

English has a number of types of unheaded compounds. Well-known examples are those such as *pickpocket*, which bear a close resem-

blance to Romance compounds such as French *porteparole* ‘spokesman’, Italian *portavoce* ‘megaphone’, or Spanish *portavoz* ‘spokesman’. Marchand (2<sup>nd</sup> 1969: 380) points out that those referring to people tend to have a pejorative sense. Animal names such as *wagtail* or other constructions, such as *breakwater* lack such overtones. Marchand (2<sup>nd</sup> 1969: 386–389) also discusses bahuvrihi compounds of the type *hunchback*, *paleface*, which he distinguishes from other exocentric types such as *pickpocket*. There seems to be no particular morphological (as opposed to etymological) reason for drawing such a distinction, however. Both types are fairly marginal.

A common source of exocentric compound is the noun derived by conversion of a verb + particle combination, illustrated by *runaway*, *drop-out* (referring to people) and *blackout*, *spin-off*. This type is highly productive.

Appositional compounds are found of the type *manservant*, *learner-driver*, *Austria-Hungary*. These appear to be double headed, in that a *learner-driver* is both a *learner* and a *driver*. Another source of appositional compounds which has begun to make itself felt is V + V compounding to form modifiers, of the kind *cook-chill (meals)*, *drink-drive (campaign)*, *fly-drive (service)*.

## 6. Cliticization

English exhibits what Zwicky (1977) has called “simple” clitics and “special” clitics, though most if not all clitics have ‘special’ forms, and cannot be automatically assumed to be mere phonological reductions. The function words which typically appear as **clitics** are: articles, personal pronouns, prepositions, auxiliary verbs (including the infinitival particle *to*), conjunctions.

The following words appear as simple clitics with reduced vowels in British English (with some dialectal variation): *a/an*, *the*; *I*, *me*, *my*, *you*, *your*, *he*, *him*, *his*, *she*, *her*, *we*, *us*, *they*, *them*, *their*; *to*, *of*, *at*, *for*; *and*, *or*, *nor*, *that*, *as*. The pronouns *it*, *its* lack full (non-clitic) forms. In American dialects *who* also reduces to /hə/ (cf. Kaisse 1985: 62–70).

The articles have the form /eɪ/, /ə/, /æn/, /ən/ and /ði:/, /ðɪ/, /ðə/. In the standard dialect the forms /æn/, /ən/ and /ði:/, /ðɪ/ are found before words beginning with vowels (a case of phonologically conditioned allomorphy at the

phrase phonology level). The forms /eɪ/ and /ði:/ are otherwise only found when emphasized. The forms /ə(n)/ and /ðɪ, ðə/ are proclitics.

In my dialect the pronouns reduce to the following forms, which can serve either as proclitics or enclitics: *I* – /a/, *me, my* /mɪ/, *you, your* – /jə/, *he, him* – /ɪm/, *his* – /ɪz/, *she, her* – /ʃɪ/, *we, us* – /wɪ/, *they, them* – /ðɪ/, *their* – /ðə/. Reduction of *my* is not characteristic of so-called RP, though in regional accents and in conservative ‘upper-class’ accents of Britain it is often reduced as shown. In addition, *them* (though not *they, their*) may lose its first consonant. Hence, the phrase *then they put them on their knees* could be pronounced as /ðenðɪputəmonðəni:z/ but not as /ðenɪputəmonəni:z/.

The commoner prepositions undergo vowel reduction when unstressed and can be proclitics or enclitics: /tə, əv, ət, fə/. They contrast with prepositions such as *in, on, off*. Hence, we can pronounce *the smell of the sea* as /ðəsmeləvðəsi:/ but we cannot pronounce *a wind off the sea* as /əwindəfðəsi:/.

The conjunctions *and, or*, and less frequently *nor* are reduced to /ən(d), ə, nə/. The reduced form of *and* is often heard as a syllabic nasal. They typically appear as enclitics, though they can be proclitic if they begin a phonological phrase or tone group. The reduced forms of *that, as* are /ðət, əz/. The conjunction *that* contrasts with the homophonous demonstrative in that the vowel of the latter can never be completely reduced to schwa.

The reduced auxiliary verbs been discussed in various places in the generative literature (cf. Kaisse 1985). They appear in the following forms: *have – l(ə)vɪ, has – l(ə)zɪ, had – l(ə)dɪ, am – l(ə)mɪ, are – lə/, is – l(i)zɪ, do – ldu, də/, does – lðəzɪ, will, shall – l(ə)lɪ, would, should – l(ə)dɪ.*

The forms with optional schwa differ significantly from the forms without it. In some cases the clitic triggers allomorphy, e.g. *you – you're /ju: – jɔ:(r)/, we – we're /wi: – wɜ:/* (British English). More importantly, the schwa-less forms have a distribution which is morphosyntactically very restricted. The schwa-initial allomorphs can be thought of as simple clitic forms of special clitics, while those without schwa are special bound allomorphs of special clitics, behaving more like affixes than independent words. The schwa

initial allomorphs can encliticize to more or less any kind of word:

- (1) *The people we met at the zoo[əv] phoned.*

However, the schwa-less allomorphs are restricted to attaching to pronouns which are the unique subjects of their clauses (cf. Kaisse 1985: 44 f.):

- (2) (a) *You[v] seen him.*

(b) *The people who met you[əv]/\*you[v] left.*

It has been suggested (Spencer 1991: 383) that the cliticized forms of subject pronouns are better regarded as special inflected forms of the pronoun (cf. the discussion of *n't* below).

Infinitival *to* encliticizes to auxiliaries and certain matrix verbs, a phenomenon often referred to in syntactic circles as ‘*wanna* contraction’. This results in allomorphic alternations over and above normal phonological reductions, especially in American dialects, e.g. *going to* becomes /gʌnə/. There has been considerable debate on the syntactic determinants of this kind of cliticization, the general view being that it is impossible if an infinitival subject has been extracted by wh-movement, but permissible if that subject is PRO or NP-trace (in Government-Binding terms). Hence, we find contraction in (3) but not in (4) (but cf. e.g. Postal & Pullum 1986: 107, on ‘liberal dialects’ which permit (4 b)):

- (3) *Who do you wanna see?*

- (4) (a) *Who do you want to see Tom?*

(b) \**Who do you wanna see Tom?*

The consensus appears to be that a verb can license procliticization of *to* if it governs *to* syntactically (cf. Bouchard 1986; Postal & Pullum 1986; Lightfoot 1986; also Aoun et al. 1987: 573–575).

Most auxiliary verbs serve as host to *n't*, the contracted form of *not*. Often this conditions allomorphy: *will – won't, shall – shan't*. Zwicky & Pullum (1983) have argued that these are inflected forms and not the result of genuine (special) cliticization.

A formative which behaves like a clitic but which lacks a corresponding full form is the possessive *'s*. This attaches to the final word of a noun phrase, showing the same allomorphy as the plural suffix: *the man we met in Wales[əz]* address. Zwicky (1987) has observed that truncation occurs when this for-

mative is attached to a word already ending in any suffix whose underlying form is /z/. e.g. *Katz[əz] reactions* but *the two cat[s]/\*[səz] reactions*.

## 7. References

- Adams, Valerie (2001), *Complex Words in English*. London: Longman
- Aoun, Joseph & Hornstein, Norbert & Lightfoot, David & Weinberg, Amy (1987), "Two Types of Locality". *Linguistic Inquiry* 18, 537–577
- Aronoff, Mark (1976), *Word Formation in Generative Grammar*. Cambridge/MA: MIT Press
- Aronoff, Mark & Sridhar, S. (1983), "Morphological Levels in English and Kannada; or, Atarizing Reagan". In: *Papers from the Nineteenth Meeting of the Chicago Linguistics Society, Parasession on the Interplay of Phonology, Morphology, and Syntax*, 3–16 [also in Gussmann (1987, ed.), 9–22]
- Bauer, Laurie (1983), *English Word-Formation*. Cambridge: Cambridge University Press
- Bouchard, Denis (1986), "Empty Categories and the Contraction Debate". *Linguistic Inquiry* 17, 95–104
- Carlson, Greg & Roeper, Thomas (1980), "Morphology and Subcategorization: Case and the Unmarked Complex Verb". In: Hoekstra, Teun & van der Hulst, Harry & Moortgat, Michael (eds.), *Lexical Grammar*. Dordrecht: Foris, 123–164
- Carstairs-McCarthy, Andrew (2002), *An Introduction to English Morphology*. Edinburgh: Edinburgh University Press.
- Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York: Harper and Row
- Di Sciullo, Anna-Maria & Williams, Edwin (1987), *On the Definition of Word*. Cambridge/MA: MIT Press
- Downing, Pamela (1977), "On the Creation and Use of English Compound Nouns". *Language* 53, 810–842
- Emonds, Joseph (1976), *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*. New York: Academic Press
- Fabb, Nigel (1988), "English Suffixation is Constrained Only by Selectional Restrictions". *Natural Language and Linguistic Theory* 6, 527–539
- Gussmann, Edmund (1987, ed.), *Rules and the Lexicon: Studies in Word-Formation*. Lublin: Redakcja Wydawnictw Katolickiego Uniwersytetu Lubelskiego
- Halle, Morris (1973), "Prolegomena to a Theory of Word-Formation". *Linguistic Inquiry* 4, 3–16
- Hayes, Bruce (1981), *A Metrical Theory of Stress Rules*. Bloomington: Indiana University Linguistics Club [Revision of unpublished MIT PhD dissertation. Also published by Garland Publishers, New York, 1985]
- Hoeksema, Jack (1986), *Categorial Morphology*. New York: Garland Publishers
- Kaisse, Ellen, M. (1985), *Connected Speech*. Orlando: Academic Press
- Kiparsky, Paul (1982), "From Cyclic Phonology to Lexical Phonology". In: Hulst, Harry van der & Smith, Norval (eds.), *The Structure of Phonological Representations, Part 1*. Dordrecht: Foris
- Lees, Robert (1963), *The Grammar of English Nominalizations*. The Hague: Mouton [First appeared as (1960), *International Journal of American Linguistics* 26.3, Part II]
- Levi, Judith (1978), *The Syntax and Semantics of Complex Nominals*. New York: Academic Press
- Lieber, Rochelle (1980), *The Organization of the Lexicon*. Unpublished MIT PhD Dissertation [Distributed by Indiana University Linguistics Club, 1981]
- Lightfoot, David, W. (1986), "A Brief Response". *Linguistic Inquiry* 17, 111–113
- Malicka-Kleparska, Anna (1987), "Potential Forms and Lexicons". In: Gussmann (ed.), 103–119
- Marchand, Hans (1969), *The Categories and Types of Present-Day Word-Formation*. Munich: C. H. Beck Verlagsbuchhandlung
- McCarthy, John (1982), "Prosodic Structure and Expletive Infixation". *Language* 58, 574–590
- Pesetsky, David (1985), "Morphology and Logical Form". *Linguistic Inquiry* 16, 193–246
- Plag, Ingo (1999), *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter
- Plag, Ingo (2003), *Word Formation in English*. Cambridge: Cambridge University Press
- Postal, Paul & Pullum, Geoffrey (1986), "Misgovernment". *Linguistic Inquiry* 17, 104–112
- Pullum, Geoffrey & Zwicky, Arnold (1987), "Plain Morphology and Expressive Morphology". *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 13, 330–340
- Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan (1985), *A Grammar of Contemporary English*. London: Longman

- Roeper, Thomas (1987), "Implicit Arguments and the Head-Complement Relation". *Linguistic Inquiry* 18, 267–310
- Roeper, Thomas (1988), "Compound Syntax and Head Movement". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology* 1, Dordrecht: Foris, 187–228
- Scalise, Sergio (1984), *Morphology in Generative Grammar*. Dordrecht: Foris
- Selkirk, Elisabeth (1982), *The Syntax of Words*. Cambridge/MA: MIT Press
- Siegel, Dorothy (1979), *Topics in English Morphology*. New York: Garland [Revision of MIT PhD dissertation, 1974]
- Spencer, Andrew (1988), "Bracketing Paradoxes and the English Lexicon". *Language* 64, 663–682
- Spencer, Andrew (1991): *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell
- Strauss, Stephen (1982), *Lexicalist Phonology of English and German*. Dordrecht: Foris
- Szymanek, Bogdan (1985), *English and Polish Adjectives: A Study in Lexicalist Word Formation*. Lublin: Redakcja Wydawnictw Katolickiego Uniwersytetu Lubelskiego
- Szymanek, Bogdan (1989), *Introduction to Morphological Analysis*. Warsaw: Pan'stowe Wydawnictwo Naukowe
- Zwickly, Arnold (1977), *On Clitics*. Bloomington: Indiana University Linguistics Club
- Zwickly, Arnold (1986), "Forestress and After-stress". *Ohio State University Working Papers in Linguistics* 32, 46–62
- Zwickly, Arnold (1987), "Suppressing the Zs". *Journal of Linguistics* 23, 133–148
- Zwickly, Arnold & Pullum, Geoffrey (1983), "Cliticization vs. Inflection: English n't". *Language* 59, 502–513

*Andrew Spencer, Essex (Great Britain)*

## 120. Deutsch (Indogermanisch: Germanisch)

1. Vorbemerkungen
2. Die Flexion der Nomina
3. Die Flexion der Verben
4. Wortbildung
5. Zitierte Literatur

### 1. Vorbemerkungen

#### 1.1. Allgemeines

Die westgermanische Sprache Deutsch wird von ca. 103 Millionen Sprechern als Muttersprache gesprochen (Zahlen nach Hawkins 1987: 110). Ihr zusammenhängendes Kerngebiet umfaßt Deutschland, Österreich, Liechtenstein und den deutschsprachigen Teil der Schweiz (insgesamt ca. 90 Millionen Sprecher) mit muttersprachlichem Schulunterricht der Standardsprache sowie angrenzende Gebiete in vielen Nachbarländern mit unterschiedlichem offiziellem Status des Deutschen. Hinzu kommen verschiedene Sprachinseln inner- und außerhalb Europas. Die **dialektale Gliederung** ist sehr ausgeprägt (vgl. Wiesinger 1983). Allgemeines Verständigungsmittel ist die normierte (aber meist regional gefärbte) **Standardsprache** (vgl. Besch 1983). Diese ist Gegenstand der folgenden

Darstellung, und zwar in ihrer schriftnächsten, gesprochenen Norm (mit Realisierung aller e geschriebenen /ə/).

Als einer der ersten "Meilensteine" in der morphologischen Beschreibung des Neuhochdeutschen kann Adelung (1782) gelten, u.a. wegen seiner "modernen" Einteilung der Flexionsklassen des Substantivs (Adelung 1782: 397–468). Diese wurde jedoch bald verdrängt durch die eigentlich historisch motivierte, aber lange Zeit auch für die traditionelle synchrone Beschreibung verbindliche Unterscheidung von Jacob Grimm zwischen "starker" und "schwacher" Flexion bei Nomina und Verben (Grimm [1819] 21870: 509, 755–758). Als die erste Anwendung des **Strukturalismus** auf die deutsche Morphologie kann die Klassifikation der Substantive von Fourquet gelten (Fourquet 1952: 33–40); erst ein Jahrzehnt später kamen die grundlegenden Aufsätze von Bech (1963) und Werner (1965; 1969) sowie die Wortbildungslehre von Fleischer hinzu (Fleischer [1969], jetzt Fleischer & Barz 21995). Die erste umfassende und zugleich einflußreichste generative Darstellung ist Wurzel (1970). Die derzeit größte Bestandsaufnahme zur schriftsprachlichen Wortbildung ist das Projekt Deutsche Wortbildung (1973–1992). Daneben stehen Gersbach & Graf (1984–1985) und das große Morpheminvantor von Augst (1975 a).

## 1.2. Zur typologischen Einordnung

Trotz der Größe und Inhomogenität seiner Sprachgemeinschaft und einer bewegten äußeren Sprachgeschichte ist das Deutsche in morphologischer Hinsicht erstaunlich konservativ. So hat es mehr von den **synthetischen Zügen** des Urgermanischen bewahrt und sogar noch ausgebaut als die anderen germanischen Sprachen (außer dem Isländischen und Färöischen), insbesondere in der Nominalflexion (s. z. B. Dal 1942; Hawkins 1986). Solche synthetischen Züge sind eine relativ große Zahl an Grammemen (z. B. bei den Nomina 4 Kasus, 2 Numeri, 3 Genera) sowie so hochgradig flektierende Verfahren zu ihrem Ausdruck wie Ablaut und Umlaut (Binnenflexion, oft noch zusätzlich zu Endungen) und eine äußerst reiche Allomorphik mit z. T. sehr komplizierten Determinierungsverhältnissen. Daraus ergeben sich zahlreiche Probleme der Segmentierung und Klassifizierung, die für den beschreibenden Linguisten (und erst recht für die Sprachbenutzer selbst) nur noch durch den Ansatz von synchron nicht mehr ableitbaren Suppletivformen zu lösen sind, die einzeln individuell gelernt werden müssen. Insbesondere in der Wortbildung unterstreicht dies in formaler Hinsicht die große Bedeutung des Wortes als eigenständige linguistische Einheit im Deutschen.

Andererseits gibt es auch Ansätze zur Einführung des **analytischen Ausdrucksverfahrens** durch Flexionsabbau auf verschiedenen Wegen und durch die Entstehung von Hilfsverb- und Präpositionalkonstruktionen und sogenannter "Artikelflexion" (Askedal 1996; Wurzel 1996). In keinem Fall sind jedoch dadurch echte isolierende Konstruktionen mit 1:1-Zuordnung zwischen Grammemen und Wörtern entstanden. Vielmehr wirken häufig syntaktisch weit voneinander entfernte Wortformen beim Ausdruck der Grammeme zusammen (vgl. Fourquet 1970 a: 25; Werner 1979: 977). Dabei markieren diese Wortformen häufig Anfang und Ende von (verschiedenen definierten) Satzkonstituenten. Es ist zu erwägen, ob dieses an den verschiedensten Stellen des Sprachsystems auftretende "klammernde Verfahren" (Ronneberger-Sibold 1997) nicht einen funktionalen Zusammenhang zwischen verschiedenen morphosyntaktischen Eigenschaften des Deutschen stiftet, die vor dem Hintergrund eines "klassischen" Kontinuums zwischen Flexion und Isolation merkwürdig inkonsistent wirken (vgl. 2.2 zur Flexion der Nominalphrase sowie 3.2 und 4.3 zur "Satzklammer").

## 2. Die Flexion der Nomina

### 2.1. Die Flexion der Substantive

Die Substantive werden nach Kasus (Nominitiv, Genitiv, Dativ, Akkusativ) und Numerus (Singular und Plural) flektiert, indem die grundsätzlich formal (und wohl auch inhaltlich) unmarkierte **Grundform** des Nominativs Singular unmittelbar verändert wird: Der Plural wird durch Endungen und/oder Umlaut ausgedrückt (*Tag-e*, *Gäst-e*, *Väter*), der Kasus nur durch Endungen (*Tag-(e)s*, *Tag-e-n*). Der Singular wird nie formal markiert. Ob man deshalb ein **Null-Morphem** {SG} ansetzen will, hängt von der Einstellung zur Zero-Problematik ab (kritisch z. B. Kloke 1982: 151 f., auch zu anderen Null-Lösungen beim Substantiv; vgl. auch Art. 45). Auch das eng mit dieser Frage zusammenhängende Problem der grundsätzlichen **Trennbarkeit von Kasus- und Numerusflexion** (wie in *Tag-e-n* angedeutet) ist umstritten (vgl. z. B. Rettig 1972: 18 f.; Mugdan 1977: 72 f.).

Das Prinzip der Grundformflexion ist durchbrochen in Fremdwörtern mit Stammflexion, die häufig den Plural mit dem einheimischen Suffix *-en* bilden (*Pizz-a* – *Pizz-en*). Es ist umstritten, ob *Pizz-en* oder die Konkurrenzform *Pizza-s* mit dem "peripheren" Pluralsuffix *-s*, aber Grundformflexion stärker ins deutsche System integriert ist (für *-s* etwa Wurzel 1984: 94; für *-en* etwa Köpcke 1993: 152 ff., Wegener 2002 und Kloke 1982: 171, wo diese Analyse sogar auf bestimmte heimische Substantive auf *-e* ausgedehnt wird: *Straß-e* – *Straß-en* wie *Firm-a* – *Firm-en*). Weiteres zur Integration der **Fremdwörter** ins deutsche Flexionssystem s. in Rettig (1972: 66–104), Augst (1975 b: 52–62), Köpcke (1993: 143 ff.), Wegener (1995).

Die **Allomorphik** ist besonders ausgeprägt beim Morphem {PL} mit der Minimalliste *-el* / *-Ø* (*Tag-e/Spaten-Ø*), Umlaut + *-el-Ø* (*Gäst-e/Väter-Ø*), *-enl-n* (*Strahl-en*, *Funke-n*, *Nadel-n*), (Umlaut)...*-er* (*Männ-er*, *Leib-er*), *-s* (*Auto-s*) allein für das zentrale System (Bech 1963; Werner 1969: 93). Die Alternanz zwischen *-e* und *-Ø* bzw. *-en* und *-n* ergibt sich bei Substantiven automatisch aus einer morphphonologischen Regel, die die Folge [ə(Sonant)ə] vermeidet, indem sie das [ʃ] der Endung tilgt. Diese Strukturbeschränkung, die auch an vielen anderen Stellen der deutschen Flexion und Wortbildung auftritt, hat zur Folge, daß die prototypische Pluralform eines deutschen Substantivs ein Trochäus mit [ə] als Kern der zweiten Silbe ist, ein wichtiges

Strukturmerkmal in verschiedenen output-orientierten Beschreibungen des Flexionssystems, die von der Gestalt des flektierten Wortes ausgehen (Köpcke 1993; Neef 1996; Eisenberg 1998). Auch die Beschränkung auf *-s* als einzig mögliches Pluralallomorph nach unbetontem Vollvokal (*Auto-s*) bewirkt, daß ebenfalls bei diesen Wörtern, die nicht dem Erwortschatz angehören, der Trochäus der Singularform im Plural erhalten bleibt. Selbst bei dem einzigen Kasussuffix des Plurals *-n/-Ø* {DAT} sind die Allomorphe so verteilt, daß diese lautliche Beschränkung eingehalten wird: *Tag-e-n*, *Funke-n-Ø*, *Wagen-Ø-Ø*, *Auto-s-Ø*.

Die Kasussuffixe im Singular sind {GEN} *-(e)s* (*Tag-es*, *Vater-s*), *-(e)n* (*Mensch-en*, *Hase-n*), *-ns* (*Funke-ns*), *-Ø* (*Frau*) sowie {DAT} und {AKK} *-(e)n* (*Mensch-en*, *Hase-n*). (-e im Dativ Singular (*dem Tag-e*) ist in der Gegenwartssprache als archaisch markiert.)

Die Kombinierbarkeit der verschiedenen Kasus- und Numerusallomorphe innerhalb von Paradigmen unterliegt starken Beschränkungen. Z. B. impliziert *-(e)n* im Genitiv Singular dasselbe Suffix auch im Dativ und Akkusativ Singular und im ganzen Plural. Auf diese Weise konstituieren sich zahlreiche **Flexionstypen** wie z. B. der durch die Paradigmen von *Mensch/Affe* repräsentierte.

Welchem Flexionstyp ein gegebenes Substantiv angehört, hängt in unterschiedlich hohem Maße ab von seinem Genus, dem Ausgang seiner Grundform und bestimmten semantischen Merkmalen, unter denen [ $\pm$ Lebewesen] das wichtigste ist. So flektieren z. B. Maskulina auf *-e*, die ein Lebewesen bezeichnen, immer wie *Affe*; Feminina flektieren nie im Singular, im Plural dagegen immer, und zwar auf *-n*, falls die Grundform auf *-e* ausgeht (*Rose*), mit Umlaut + *-el-Ø* sonst (*Gänse*, *Mütter*) usw. Zuordnungen dieser Art gelten in manchen Fällen mit Sicherheit, in anderen nur mit einer gewissen Wahrscheinlichkeit (entsprechende statistische Angaben s. in Duden '1998: 222 ff.; zur psychologischen Realität bei den Sprachbenutzern Köpcke 2000 und – als Testfall für die Entscheidung zwischen konnektionistischen und regelgeleiteten Wissensmodellen – Bartke 1998).

Mit Hilfe der besprochenen Regularitäten lassen sich die zahlreichen Flexionstypen zu **Flexionsklassen** gruppieren. Die traditionelle, auf Jacob Grimm zurückgehende und vorwiegend diachron begründete Einteilung umfaßt die Klassen **„schwach“** mit {GEN.SG}

*-(e)n* oder *-Ø* (bei Feminina) und {PL} *-(e)n* (*Affe*, *Mensch*, *Rose*, *Frau*), **„stark“** mit {GEN.SG} *-(e)s* oder *-Ø* (bei Feminina) und {PL} nicht auf *-(e)n* (*Tag*, *Gast*, *Mann*, *Auto*, *Gans*) und **„gemischt“** mit {GEN.SG} *-(e)s* und {PL} *-(e)n* (*Strahl*, *Pantoffel*) (Duden '1966: 172–175). (Die „schwachen“ Feminina (*Rose*, *Frau*) können auch zur „gemischten“ Klasse gerechnet werden.)

Zusätzlich zu oder anstelle von dieser sehr groben Einteilung sind zahlreiche weitere Klassifikationen vorgeschlagen worden. Ein wichtiges Ziel ist dabei, daß die Angabe der Klasse zusammen mit dem Genus, der Grundform und der Bedeutung im Lexikon erlaubt, jedes Substantiv mit Sicherheit oder zumindest hoher Wahrscheinlichkeit richtig zu flektieren. Eine gleichzeitig exhaustive und redundanzfreie Klassifikation dieser Art bietet Wurzel (letzte Fassung Wurzel 1998). Für didaktische Zwecke empfiehlt sich allerdings wohl eine weniger komprimierte Darstellung wie z. B. diejenige von Wegener (1995). (Überblicke über die zahlreichen Klassifikationsvorschläge s. in Rettig 1972: 41–65, Bettelhäuser 1976: 53–107 und Mugdan 1977: 108–116. Später erschienen sind Jørgensen 1980, Leys 1986, Carstairs 1986, D. Bittner 1990 und 2003.)

Das **Genus** (Maskulinum, Femininum, Neutr. trum) ist als inhärentes Merkmal fest mit dem Substantiv verbunden und muß i. a. mit diesem gelernt werden. Am Substantiv selbst wird es nie ausgedrückt, sondern nur an seinen Attributen und Stellvertretern (vgl. Werner 1975, auch zu den verschiedenen Funktionen des Genus). Es ist jedoch in manchen Fällen möglich, aus bestimmten morphologischen, phonologischen und semantischen Merkmalen eines Substantivs auf sein Genus zu schließen. So sind z. B. alle Substantive mit dem Suffix *-heit* feminin, desgleichen z. B. alle Baumbezeichnungen, deren Grundform auf *-e* ausgeht (*Linde*) (s. Duden '1998: 199–207). In den meisten Fällen aber ist eine solche Zuordnung des Genus entweder gar nicht oder nur mit einer gewissen Wahrscheinlichkeit möglich (vgl. Köpcke 1982; Köpcke & Zubin 1996). Es ist umstritten, inwieweit solche Wahrscheinlichkeiten die Sprachbenutzer bei der Genuszuweisung an Neologismen und nonsense-Wörter leiten (z. B. Arndt 1970; kritisch Gregor 1983: 59–62).

## 2.2. Die Flexion der Nominalphrase

Trotz seiner großen Komplexität ist das Flexionssystem des Substantivs bemerkenswert ineffizient bei der Unterscheidung der Formen innerhalb des Paradigmas. Von den acht

Formen sind maximal vier verschieden, z. B. beim Typ *Gast*, und von diesen vieren bezeichnen nur zwei eindeutig eine bestimmte Kasus-Numerus-Kombination (*Gastes*, *Gästen*). Andere Flexionstypen wie z. B. *Frau* oder *Mensch* enthalten überhaupt nur zwei unterschiedliche Formen (s. die Aufstellung in Mugdan 1977: 70). Diese **Ambiguität** betrifft den Kasus stärker als den Numerus: Bei fast allen Flexionstypen ist der Plural in allen Kasus eindeutig bezeichnet (vgl. Art. 156 zur Numerusprofilierung), aber bei keinem einzigen Typ sind auch nur in einem der beiden Numeri alle Kasusformen eindeutig. Hier ist die disambiguierende Wirkung der mit dem Substantiv kongruierenden Konstituenten vonnöten. In erster Linie sind dies seine vorangestellten Begleiter und Attribute. Bis auf wenige Fälle von Flexionslosigkeit (oder Null-Allomorphik) (*ein, kein, mein ...*) werden diese Wörter in **Kongruenz** mit dem Substantiv durch Endungen nach Kasus, Numerus und (nur im Singular) nach Genus flektiert. Dafür stehen im wesentlichen zwei Flexionsarten zur Verfügung, die traditionell **„pronominal“** oder **„stark“** und **„schwach“** genannt werden. In (*ein*) *gut-er, trocken-er Wein* vs. *dies-er/d-er gut-e, trocken-e Wein* ist *-er* das pronominale, *-e* das schwache Allomorph des Morphems {NOM.SG.M}. (Eine morphologische Segmentierung der Endungen wie beim Substantiv ist hier nicht möglich; Vater 1979.) Die Endungen des bestimmten Artikels unterscheiden sich zwar phonologisch von denen der mehrsilbigen starken Formen (sehr deutlich z. B. *die* versus *dies-e*) aber kaum hinsichtlich ihrer Distinkтивität. Die **Auswahl** aus den Allomorphen ist lexikalisch (durch die Wortart) und syntaktisch determiniert: Bei dem Substantiv vorangestellten, attributiv gebrauchten Adjektiven/Partizipien steht die schwache Flexion nach einem flektierten Determinans, die pronominale in allen anderen Fällen. Prädikativ und adverbial gebrauchte sowie dem Substantiv nachgestellte attributive Adjektive/Partizipien sind unflektiert (zu einigen Gebrauchsschwankungen s. <sup>6</sup>Duden 1998: 284–297 und sehr ausführlich Ljungerud 1955: 161–306). Daraus ergibt sich, daß die meisten erweiterten deutschen Nominalphrasen durch ein pronominal flektiertes Element eingeleitet werden. Dessen Endung drückt zusammen mit derjenigen des Substantivs soweit wie möglich eindeutig Kasus, Numerus und Genus der ganzen Nominalphrase aus (Gallmann 1990). Die schwachen Endungen sind zur Disambiguierung

nur notwendig bei Null-Plural am Substantiv: *der gut-e Lehrer – der gut-en Lehrer*. Trotz dieses Zusammenwirkens zwischen pronominaler Flexion und Substantivflexion bleiben bei allen substantivischen Flexionstypen bestimmte Synkretismen bestehen, z. B. zwischen Nominativ und Akkusativ außer im Maskulinum Singular (ausführlich dazu Pike 1965: 212–216 im tagmemischen Rahmen). Solche formalen Übereinstimmungen lassen sich als Abbild bestimmter funktioneller Gemeinsamkeiten z. B. zwischen Nominativ und Akkusativ interpretieren (B. Wiese 1996).

Der engen Beziehung, die pronominale und substantivische Flexion durch diese gegenseitige Disambiguierung miteinander eingehen, versuchen verschiedene Beschreibungen gerecht zu werden, die von dem traditionellen Konzept der Kongruenz innerhalb der Nominalphrase abweichen, so etwa Admoni (1982: 78) (**„kooperierende Monoflexion“**), Werner (1979: 976–980) (**diskontinuierliche Morphe**, fußend auf Fourquet 1970 a: 25) und Wurzel (1984: 90–93) (**Artikelflexion des Substantivs** zum Ausdruck des Kasus neben Endungsflexion und Umlaut zum Ausdruck des Numerus). Letzteres ist problematisch wegen der starken Kasusambiguität vieler pronominaler Endungen wie z. B. *-er* {NOM.SG.M}, {GEN.SG.F}, {DAT.SG.F}, {GEN.PL}. Hier wirkt erst das Substantiv mit seinem Genus und Numerus disambiguierend.

Unabhängig von diesen verschiedenen Interpretationen ist ein wichtiger Gesichtspunkt die syntagmatische Stellung von pronominaler Flexion und Substantivflexion innerhalb der Nominalphrase: Diese wird von ihnen wie durch zwei Grenzsignale eingeklammert bis auf die nachgestellten Attribute wie den Relativsatz usw., die jedoch ihre eigenen Grenzsignale haben (Ronneberger-Sibold 1997: „**nominale Kongruenzklammer**“).

Die Paradigmen des **Relativpronomens** und des **Personalpronomens** der 3. Person, die als echte Pro-Nomina typischerweise allein eine Nominalphrase füllen, ähneln zwar denen der Determinanten, sind aber stärker differenziert, insbesondere an den Stellen, an denen in der komplexen Nominalphrase die Substantivflexion disambiguierend wirkt (Frey 1975). Die Personalpronomina der 1. und 2. Person, eher **„Rollenwörter“** als echte Pronomina, flektieren nicht nach Genus, sondern nur (suppletiv) nach Kasus und nach Numerus, wobei allerdings zu bemerken ist, daß *wir, ihr* keine echten Plurale von *ich, du* sind (Howe 1996).

**Flexionsabbau** findet im gegenwärtigen Deutsch v. a. im Bereich des Genitivs statt.

Es ist (mit Moser 1973) zu unterscheiden zwischen dem Schwund der Genitivendung am Substantiv in bestimmten Fällen, meist nach eindeutig für ‘GEN’ markiertem Begleiter (*des Mai* statt *des Mais* oder gar *des Maien*) einerseits (Belegmaterial in Appel 1941, Rowley 1988; jüngste Behandlung in Wurzel 1991) und dem Ersatz des Kasus durch eine äquivalente Konstruktion andererseits. I. a. ist dies eine Präpositionalphrase: *sich an jemanden erinnern* statt *sich jemandes erinnern*, *das Buch von dem Lehrer* statt *das Buch des Lehrers* (Statistik in Pfeffer & Lorentz 1979). Obligatorisch ist die letztgenannte Konstruktion bereits bei unerweiterten Nominalphrasen mit einem Appellativum als Kern: *der Genuß von Wein* statt *\*der Genuß Weins*, vermutlich wegen der uneindeutigen Genitivformen in vielen Flexionstypen, v. a. bei den Feminina (*\*der Genuß Milch*) (Teuber 2000). Unter typologischem Gesichtspunkt wird dieser Kasusersatz i. a. mit der Entwicklung vom synthetischen zum analytischen Sprachbau in Verbindung gebracht (z. B. Ronneberger-Sibold 1980: 84–88). Es ist jedoch bemerkenswert, daß sich gegenwärtig (v. a. in der Sprache der Medien) auch eine **synthetische Ersatzkonstruktion** stark ausbreitet, nämlich das Kompositum (*die Brandt-Reise*, vgl. 4.4).

Interessant sind in diesem Zusammenhang auch die sog. **Verschmelzungsformen** des bestimmten Artikels mit gewissen Präpositionen (*im*, *zum*, *zur* usw.) (vgl. Dedenbach 1987; Nübling 1992). Synchron strukturalistisch kann man diese Formen als Portmanteaumorphe beschreiben; in diachroner Perspektive stellen sie einen ersten Schritt zur Entwicklung neuer **synthetischer Wortformen** dar. Die Erscheinung gehört in den weiten Bereich der **En-** bzw. **Proklise**, die in der gesprochenen Umgangssprache und erst recht den Dialekten sehr viel weiter verbreitet ist als in der hier beschriebenen Schriftsprache.

### 3. Die Flexion der Verben

Im Paradigma der Verben unterscheidet man zwischen **einfachen (synthetischen)** und **zusammengesetzten (analytischen)** Formen und bei beiden wiederum zwischen finiten (nach Person/Numerus flektierten) und infiniten Formen.

#### 3.1. Die synthetischen Verbformen

Die finiten synthetischen Verbformen flektieren nach Person/Numerus (1., 2., 3. Person Singular und Plural), Tempus (Präsens, Präteritum) und Modus (Indikativ, Konjunktiv I (Präsens), Konjunktiv II (Präteritum), Imperativ) (zu den tempusneutralen Bezeichnungen “Konjunktiv I und II” s. 3.3). Die infiniten Formen sind der Infinitiv, das Partizip I (Präsens) und II (Perfekt).

Die Flexion erfolgt durch Affixe und/oder Vokalwechsel: (*ich*) *sag-e* – (*ich*) *sag-t-e*; (*wir*) *nehm-en* – (*wir*) *nahm-en* – (*wir*) *nähm-en* – *ge-nomm-en* – *nimm*). Eine formal unmarkierte Grundform wie bei Substantiven und Adjektiven existiert standardsprachlich (noch) nicht; ihre Entwicklung ist aber in der gesprochenen und geschriebenen Umgangssprache weit fortgeschritten beim Imperativ: *geh* (Duden 6<sup>1998</sup>: 170). Eine andere, nicht standardsprachliche endungslose Form ist der sog. Inflektiv wie z. B. *würg* in der Comic-Literatur (Teuber 1999). Bemerkenswerterweise ist der semantisch unmarkierte Infinitiv (die Nennform) nicht von einer solchen Entwicklung betroffen.

Die in 2.1 angesprochenen Probleme der Segmentierbarkeit der Formen, der Null-Lösungen, speziell der Alternanz von /ə/ mit /Ø/ (vgl. Reed 1973) sowie des Formensynkretismus sind in der Flexion der Verben nicht grundsätzlich verschieden von der der Substantive. (Eine ausführliche, rein synchrone Beschreibung der Fakten speziell für den Ausländerunterricht bietet Darski 1999.)

Ein spezifisches Problem der Verbalflexion ist dagegen der **Ablaut** und die daraus erwachsenden Schwierigkeiten bei der **Klassifizierung**. Die Verben lassen sich nach der Bildung des Präteritums und des Partizips II in zwei große Gruppen einteilen: die regelmäßigen sog. schwachen und alle anderen. Die **schwachen Verben** bilden ihr Präteritum und Partizip Perfekt durch Anfügung des Flexivs *-t* an den unveränderten Präsensstamm; an dieses treten im Präteritum die Personalendungen (*sag-t-e*, *sag-t-est* usw.); im Partizip II kommt (außer bei den nicht anfangsbetonten Verben und deren Präfigierungen) noch das Präfix *ge-* hinzu (*ge-sag-t*, aber *entsag-t*, (*ein-)studier-t*). Der Konjunktiv Präteritum der schwachen Verben ist systematisch homonym mit dem Indikativ (bzw. existiert nicht als synthetische Form). Die Formen sind also morphotaktisch transparent und völlig regelmäßig vom Präsensstamm aus durch Addition von Endungen zu bilden. Von der reinen **Agglutination** trennt sie im wesentlichen nur das morphologisch determinierte Allomorph *-e* der 3. Person Singular: *sag-t-e* im Präteritum gegenüber *sag-t* im Präsens. Daran än-

dert auch die alternative Segmentierung *lieb-te-Ø* (etwa bei Halle 1970: 323) nichts. Vermutlich ist es diesen Eigenschaften zu verdanken, daß dieser Flexionstyp im modernen Deutschen den einzigen produktiven darstellt: Nach ihm flektieren die meisten existierenden Verben und alle Neologismen (zur Belegstärke der verschiedenen Flexionstypen s. Mater 1968).

Im Vergleich dazu ist die Gruppe der nicht-schwachen Verben klein. Traditionell wird sie aufgegliedert in die Gruppe der ca. 20 „unregelmäßigen“ und die eigentlichen starken. Das wichtigste Merkmal der **starken Verben** ist, daß sie das Präteritum statt durch Anfügen von *-t(-)* durch **Ablaut** markieren. Hinzu kommen einige Endungsallomorphe. Das Partizip Perfekt kann ebenfalls Ablaut gegenüber dem Präsens und/oder dem Präteritum zeigen, außerdem hat es die Endung *-en* statt *-(e)t*: *ge-nomm-en* vs. *ge-sag-t*.

Ablaut ist im Vergleich zum Umlaut eine bedeutend kompliziertere Form der Binnendifflexion: Beim Umlaut alternieren ja nur zwei Vokale, die sich gegenseitig implizieren: Mit *a* kann nur *ä* und mit *ä* nur *a* wechseln. Beim Ablaut in der Verbalflexion alternieren dagegen bis zu 4 Vokale miteinander (unter Einschluß des diachron auf einen alten Umlaut zurückgehenden sogenannten ***e-i-Wechsels*** in der 2./3. Person Singular Präsens Indikativ und der 2. Person Singular Imperativ: *nehmen – nimmst – nimmt – nimm*). Die Kombinierbarkeit dieser Vokale unterliegt zwar gewissen Beschränkungen, volle Vorhersagbarkeit ist jedoch nicht annähernd erreicht. Zusätzlich zum Ablaut tritt bei den meisten starken Verben mit Präsensstammvokal /a/, /a:/, /a<sup>u</sup>/ **Umlaut** ein (in der Distribution bis auf den Imperativ parallel zum *e-i-Wechsel*: *laufen – läuft* wie *nehmen – nimmt*). Ferner zeigt Umlaut der Konjunktiv Präteritum aller starken Verben mit umlautfähigem Stammvokal. Gewöhnlich lautet hier der Vokal des Indikativ Präteritum um (*nahm – nähme*), manchmal jedoch auch der des Partizips Perfekt (*schalt – schölte* nach *gescholten*) oder ein dritter (der des ehemaligen Plural Präteritum Indikativ, vgl. Art. 156: *warf – würfe*).

Zur Bestimmung eines **Ablauttyps** sind also fünf Kennformen notwendig, die alle mit verschiedenen Vokalen besetzt sein können, z. B. in *nehmen – nimmt – nahm – nähme – genommen*. Mindestens das Präsens und der Indikativ Präteritum müssen verschiedene Vokale zeigen: *reiten – reitet – ritt – ritte – geritten*. Hinzu kommen bei einigen Verben

**konsonantische Veränderungen:** (*wir*) *leiden* – (*wir*) *litten*. Auf diese Weise ergeben sich sehr viele, großenteils nur schwach belegte Ablauttypen.

Die Zusammenfassung dieser Typen zu **Ablautklassen** ist schwierig, weil die Auswahl aus den Typen nur schwach determiniert ist, v. a. durch den/die auf den Ablautvokal folgenden Konsonanten (vgl. Halle 1970; Ulvestadt 1956; Fourquet 1970 b; Wurzel 1970: 63–80; Fabricius-Hansen 1977; Veith 1984). Im Prinzip stehen sich – ähnlich wie bei den Substantiven (vgl. 2.1) – die beiden Möglichkeiten gegenüber, lieber wenige Klassen zu bilden mit vielen Mitgliedern und komplizierten Regeln für die Bestimmung der korrekten Formen innerhalb der Klassen und trotzdem noch relativ zahlreichen Ausnahmen, (extrem z. B. Ségral & Scheer 1998) oder viele Klassen mit jeweils wenigen Mitgliedern und wenigen, einfachen Regeln. Der – didaktisch durchaus begründbare – Extremfall ist die durchgehende Beschreibung durch Suppletivformen (vgl. Augst 1975 b: 231–281; Darski 1999: 30–38).

Zur Gruppe der „**unregelmäßigen**“, d. h. synchron suppletiv zu wertenden Verben zählen traditionell die sog. „Rückumlaut-Verben“ (*nennen – nannte*), ferner *wissen* und alle Modalverben (bis auf *wollen* alle ehemalige Präteritopräsentia, dazu Literaturüberblick in Soeteman 1967) und schließlich die synchron keiner Gruppe mehr zuzuordnenden Verben *gehen, stehen, tun, haben, werden* sowie das am stärksten suppletive *sein*.

Statt der recht groben Einteilung in schwache, starke und „unregelmäßige“ kann man die Verben auch auf einer Skala der Regelmäßigkeit mit sehr viel feiner abgestuften Übergängen anordnen (A. Bittner 1996 unter Ausnutzung bestimmter implikativer Beziehungen zwischen verschiedenen Flexionsformen; Köpcke 1999 mit einem auf Schemata basierenden Ansatz). Dabei fällt eine starke positive Korrelation zwischen **Textfrequenz** und **Unregelmäßigkeit** auf (Werner 1987: 599; Harnisch 1988; zu Textfrequenzen der einzelnen Formen s. auch Ortmann 1975–1978). In der Tat treten auch im gegenwärtigen Deutsch noch starke Verben bei sinkender Frequenz leicht zum regelmäßigen schwachen Typ über (*melkte* statt *molk*). Dies kann zu „gemischten“ Paradigmen führen: Das Partizip lautet immer noch *gemolken* (zu entgegengesetzten Tendenzen im Niederländischen – mit etwas anders strukturiertem Ablautsystem – s. Hempen 1988; zu extremen Formen der Irregularisierung im Hochfrequenzbereich s. Nübling 2000; allgemein zu Flexionsschwankungen s. Theobald 1992).

### 3.2. Die analytischen Verbformen

Die sog. **analytischen Verbformen** bestehen aus einer finiten Form der sog. Hilfsverben *haben, sein* und *werden* und mindestens einer

infiniten Form (dem – nicht nach Kasus/Numerus/Genus flektierten – Partizip Perfekt oder dem Infinitiv) des sog. Hauptverbs. Traditionell gelten als analytische Verbformen: das Perfekt (*hat genommen, ist gekommen*), das Plusquamperfekt (*hatte genommen, war gekommen*), das Futur I (*wird kommen, nehmen*) und II (*wird gekommen sein, genommen haben*) und das Passiv (eventuell aufgegliedert in Vorgangs- und Zustandspassiv: *wird genommen – ist genommen*). Mehrere analytische Formen können sich überlagern, sodaß Folgen von zwei und sogar (allerdings wenig gebräuchlich) von drei infiniten Verbformen innerhalb einer analytischen Form entstehen können: *wird genommen worden sein* (Futur II Passiv). Das Partizip Perfekt von *werden* lautet in solchen Fällen *worden*. (S. Duden 1998: 115–117 für eine vollständige Aufstellung aller nur denkbaren Formen einschließlich der analytischen Infinitive bis hin zu (?) *geliebt haben werden*.)

Durch Modusflexion der finiten Verben können von allen diesen Formen die entsprechenden Konjunktive gebildet werden, z. B. *hat genommen, ist gekommen* (Indikativ Perfekt) – *habe genommen, sei gekommen* (Konjunktiv I Perfekt) – *hätte genommen, wäre gekommen* (Konjunktiv II Perfekt, der sog. Konjunktiv Plusquamperfekt, zur Bedeutung s. u. 3.3 (ausführlich dazu Glinz 1970: 102–141).

Im Vergleich zu den synthetischen sind alle analytischen Formen relativ junge **Grammatikalisierungsprodukte** aus Verbalperipherasen (Diewald 1997). Die Grammatikalisierung hat sich dabei kaum formal ausgewirkt (bis auf das erwähnte Partizip *worden* sind alle beteiligten Verbformen synchron dieselben wie im freien syntaktischen Gebrauch), sondern nur semantisch im Sinne einer unterschiedlich weit vorangeschrittenen Idiomatisierung (zur grammatischen Idiomatik Moskal'skaja 1975: 60).

Es wird daher eine rege Diskussion über die **synchrone Kompositionnalität** der analytischen Formen geführt. Nach der kompositionellen Interpretation in ihrer reinsten Form sind die analytischen Verbformen normale Syntagmen, deren Bedeutung sich nach den allgemeinen syntaktischen Regeln aus den Einzelbedeutungen der verknüpften Lexeme und ihren grammatischen Bestimmungen ergibt. Z. B. wären *die Rose ist verblüht, die Rose ist geknickt* und *die Rose ist rot* nach dieser Auffassung semantisch und syntaktisch völlig parallel zu analysieren als Aussagen über den gegenwärtigen Zustand der Rose (Rupp 1965: 191–194 [1967: 158–160]: “sein-Perspektive”). Zu dieser Auffassung

tendieren – mit verschiedenen Modifikationen und in verschiedenen Grammatikmodellen – Bartsch (1980), Ballweg (1988), Fabricius-Hansen (1986), Klein (1999). Nach der traditionellen, nicht-kompositionellen Interpretation sind die Hilfsverben, anders als die entsprechenden Vollverben, demantisierte Funktionswörter zum Ausdruck bestimmter grammatischer Kategorien des Hauptverbs. Dessen infinite Formen haben ihrerseits allenfalls distinktive Funktion (z. B. zur Unterscheidung zwischen Passiv und Futur), aber nicht die grammatische Bedeutung, die sie in anderen Kontexten tragen. Diese Position vertritt in größter Strenge Fourquet (1969: 65), ferner Reis (1976), Askedal (1982), Grønvik (1986), Radtke (1998). Offensichtlich ist die kompositionelle Interpretation sehr viel leichter auf die mit *sein* gebildeten analytischen Formen anwendbar (wie etwa das zitierte *ist verblüht*) als z. B. auf die mit *haben* gebildeten, insbesondere bei intransitivem Hauptverb (*hat geschlafen*).

In der Wortstellung unterscheiden sich die analytischen Verbformen nicht von vergleichbaren freien Syntagmen wie z. B. den Modalverbgefügen. Beide bilden eine “Satzklammer” (außer im eingeleiteten Nebensatz), z. B.: *Er hat den Brief geschrieben wie er will den Brief schreiben* (s. auch 2.2 und 4.3 zum “klammernden Verfahren”).

Vorschläge zur Erweiterung des traditionellen Inventars an analytischen Verbformen betreffen bestimmte teilweise **grammatikalisierte peripherastische Konstruktionen** wie z. B. das sog. **bekommen-Passiv** (*er bekommt die Krone aufgesetzt*; vgl. Brincker 1971: 117–126; Leirbukt 1997), verschiedene **“Verlaufsformen”** (*er ist am beim Schreiben, etwas ist im Kommen*; Duden 1966: 72), die sog. **tun-Peripherase** (*angeln tut er*; Erben 1969) und die sogenannten **doppelten Perfekt- und Plusquamperfektformen** (*hat/hatte genommen gehabt, ist/war gekommen gewesen*; Thieroff 1992).

### 3.3. Die grammatischen Funktionen (“Bedeutungen”) der Verbformen

Die Entstehung der analytischen Verbformen lässt sich ihrerseits als Teil einer umfassenden Umgestaltung des verbalen Kategoriengefüges betrachten, die im Gegenwartsdeutschen noch keineswegs abgeschlossen ist. Daher sind die grammatischen Funktionen der einzelnen Formen innerhalb des Gesamtsystems schwieriger exakt zu bestimmen als in weitgehend stabilen verbalen Kategoriengefügen wie etwa dem des klassischen Lateinischen. Die beiden wichtigsten Entwicklungsstränge könnte man stark vergröbernd einerseits im Ausbau der Kategorie Tempus im Indikativ sehen, andererseits in ihrem Abbau im Konjunktiv. Dieser hat nämlich seine Funktion

als Modus der Unterordnung in Abhängigkeit vom Tempus des übergeordneten Satzes weitgehend eingebüßt, sodaß seine Tempusformen Konjunktiv I und II heute mehr dem Ausdruck von verschiedenen Arten der "Irrealität" im weitesten Sinne als dem Ausdruck von Tempus dienen. Daraus ergeben sich für das System des Gegenwartsdeutschen die folgenden Probleme (Forschungsüberblicke und Stellungnahmen s. in Radtke 1998; Thieroff 1992; Abraham & Janssen 1989, Hrsg.; Der Begriff Tempus 1969).

Ist das **Futur** eine Tempusform, oder ist **werden** in Fügungen wie *er wird kommen* **Modalverb** parallel zu *er kann, soll, will ... kommen*, und das nicht nur in seiner "inferentiellen" oder "epistemischen" Verwendung zum Ausdruck einer vorsichtigen Vermutung, sondern generell (Saltveit 1960; Vater 1975; Ballweg 1987; Redder 1999)?

Ist der **Indikativ Präsens** eine echte **Tempusform** zum Ausdruck der Gegenwärtigkeit oder Nicht-Vergangenheit, oder ist er eine **tempusneutrale Form**, die erst durch ihren jeweiligen Kontext Zeitbezug erhält? Ist folglich das sogenannte historische Präsens ein stilistisch markierter Sonderfall oder eine ganz normale Verwendung des Präsens?

Sind das **Perfekt** und das **Plusquamperfekt** "vollendete" **Aspektformen** des Präsens bzw. Präteritums oder eigene Tempusformen? Wenn das zweite der Fall ist, worin besteht dann der Tempusunterschied zwischen Perfekt und Präteritum? Oder liegt der wesentliche Unterschied gar in der Erzählhaltung (Weinrich 1971)? Als weitere Erschwernis kommt hinzu, daß umgangssprachlich (besonders in Süddeutschland) das Perfekt das Präteritum bereits weitgehend als Tempusform ersetzt hat, ein typologisch weit verbreiterter Vorgang, der mit dem grundsätzlichen Primat des Aspekts vor dem Tempus zu tun haben könnte (Leiss 1992).

Beim **Konjunktiv** liegt, wie schon angedeutet, das Hauptproblem darin, daß der Konjunktiv II zwar formal vom Präteritalstamm gebildet wird (*er kam – er käm-e*), sich inhaltlich jedoch in seiner Hauptfunktion als Ausdruck des Irrealis/Potentialis auf gegenwärtige oder zukünftige Sachverhalte bezieht (*wenn er jetzt/nächste Woche käme, wäre das zu früh*) (dazu z. B. Graf 1977: 69–109). Hauptfunktion des Konjunktiv I ist die (für wahr gehaltene) indirekte Rede (*er sagt(e), er komme schon/gleich*). Um diesen Widerspruch zwischen Form und Inhalt aufzulösen, ist die Kategorie der **Distanz** vorgeschla-

gen worden. Danach drücken präteritales *kam* und irreales *käme* beide größere temporelle bzw. modale Distanz des Sprechers vom Sachverhalt aus als präsentliches *kommt* und als wahr wiedergegebenes *komme* (Thieroff 1992: 274–299; etwas anders Feuillet 1993: 185 ff., 198 ff.).

Beim sog. **Würde-Konjunktiv** (*würde kommen*) schließlich existiert die unmittelbar parallele Form des Indikativs *\*würde kommen* nicht (mehr). Formal handelt es sich um den Konjunktiv II des Futur I (*wird kommen*), inhaltlich hat die Form aber höchstens in ihrer Nebenfunktion als Futur des Präteritums Zukunftsbezug (Fabricius-Hansen 2000). Hauptsächlich wird sie als Ersatzform für die synthetischen Formen des Konjunktiv II (*käme*) verwendet (Werner 1965: 123–125 [1970: 377–381]; Bausch 1979: 157–216). In der gesprochenen Umgangssprache wird der Konjunktiv I in der indirekten Rede häufig entweder durch den Konjunktiv II bzw. die Umschreibung mit *würde* ersetzt (*er sagt(e)/hat gesagt, er käme gleich/würde gleich kommen*), obwohl diese eigentlich der Wiedergabe einer nicht für wahr gehaltenen Rede dienen sollten, noch häufiger aber durch den Indikativ (*er sagt(e)/hat gesagt, er kommt gleich*) (Duden 6<sup>1998</sup>: 166). Insgesamt scheint die Entwicklung auf einen Zustand mit nur zwei Zeitstufen des Konjunktivs (*er würde kommen – er wäre gekommen/würde gekommen sein*) hinauszulaufen, die ausschließlich für den Irrealis bzw. Potentialis verwendet werden.

Zur Angemessenheit der traditionellen Einordnung des **Imperativs** als Modus (nicht eher als Satzart?) und als finit (nur die 2. Person Singular hat eine eigene, vom Indikativ distinkte Form) s. Donhauser (1986).

#### 4. Wortbildung

Einer der auffälligsten synthetischen Züge des Deutschen ist seine reiche Wortbildung, reich sowohl in Bezug auf seine Wortbildungsmöglichkeiten als auch in Bezug auf deren produktive Verwendung. Die Wortbildungsmöglichkeiten umfassen nach traditioneller Auffassung (mindestens) die beiden **Wortbildungarten** Derivation und Komposition, ferner für die Derivation eine sehr große Zahl an **Wortbildungsmitteln** (Suffixe, Präfixe usw.) und viele **Wortbildungsmodelle** (deverbale Substantive, desubstantivische Verben, Komposita aus Substantiv + Substantiv

usw.). Durch Wortbildungsart, -mittel und -modell ist jeweils ein **Wortbildungstyp** bestimmt. Die starke Produktivität vieler Typen zeigt sich nicht nur an der großen Zahl von entsprechenden Wortbildungen im Lexikon, sondern vor allem auch an den zahlreichen sog. Augenblicksbildungen, die nie im Lexikon erscheinen. Ein weiteres Anzeichen für die Vitalität der Wortbildung im Deutschen ist das ständige Entstehen neuer Wortbildungstypen, teilweise bei gleichzeitigem Verlust der Produktivität von älteren (vgl. 4.3).

#### 4.1. Ableitung, Konversion, Kürzung

Die Wortbildungsmittel der **Derivation** sind in erster Linie **Suffixe** (*-lich*, *-bar* usw.), in zweiter Linie **Präfixe** (*un-*, *ver-* usw.) (s. Art. 54). Beide können auch zusammen auftreten (*un-ver-änder-lich*). Davon zu unterscheiden sind die (weniger zahlreichen) **Zirkumfixe (diskontinuierlichen Morphe)** (*Ge-heul-e*). **Umlaut** (*fallen* – *fällen*) und **Ablaut** (*bind-en* – *Band* – *Bund*) sind als selbständige Wortbildungsmittel kaum noch produktiv (eventuell *Einstieg* zu *einstieg-en*; Polenz<sup>2</sup>1980: 173). In Verbindung mit bestimmten Suffixen wie *-lich* (*schwärz-lich*), *-in* (*Ärzt-in*) (Doleschal 1992) und den Diminutivsuffixen ist der Umlaut jedoch eingeschränkt produktiv: *Väterchen* (aber *Onkel-chen*); *Fräu-lein* (aber *Frauchen*). Sehr stark ausgeprägt ist diese Art von Stammallomorphik bei den Fremdwörtern: *Kollis-ion* zu *kollid-ier-en* usw. (vgl. Hoppe et al. 1987).

Die Ableitungssuffixe stehen i. a. vor den Flexiven (*Veränder-ung-en*) (bis auf den Typ *Kind-er-lein*). Sie bestimmen Wortart, Flexionstyp und – bei Substantiven – häufig auch das Genus (vgl. 2.1).

Die **implizite Ableitung**, auch **Konversion** oder **Nullableitung** genannt (Kastovsky 1969; kritisch Naumann<sup>3</sup>2000: 15f.) wird auch als eigene Wortbildungsart behandelt (z. B. Duden<sup>6</sup>1998: 420). Es handelt sich um einen Wechsel der Wortart (eventuell verbunden mit semantischer Modifikation) ohne offene Wortbildungsmittel. Das konvertierte Wort nimmt dabei die Flexion der neuen Wortart an: (*das/lein*) *Hoch* – (*des/leines*) *Hoch-s* usw. Daher ist es unter morphologischem Gesichtspunkt problematisch, substantivisch gebrauchte Adjektive/Partizipien wie (*der*) *Abgeordnet-e* – (*ein*) *Abgeordnet-er* – (*des/leines*) *Abgeordnet-en* mit erhaltener Adjektivflexion als Konversion zu betrachten.

Ein besonderes Problem stellt die Konversion von Verben zu Substantiven dar, da hier das Flexiv des

Infinitivs *-(e)n* nicht immer unverbleibt wie in *versuch-en* – *Versuch* (traditionell als Form von **Rückbildung** behandelt, z. B. Henzen<sup>3</sup>1965: 240), sondern auch mitkonvertiert werden kann: (*das*) *Versuchen* – (*des*) *Versuchen-s* usw. In diesem Fall wird *-en* häufig als Derivativ in Konkurrenz mit *-ung* (*Versuch-ung*) usw. gewertet (Sandberg 1976; zur Problematik Polenz<sup>2</sup>1980: 170). Umgekehrt nimmt bei Konversion zum Verb das neue Verb selbstverständlich die Infinitivendung an: *Fisch* – *fisch-en*. Manchmal ist in solchen Fällen die Ableitungsrichtung synchron nicht eindeutig: Ist *Liebe* von *lieben* abgeleitet oder umgekehrt (Bergenholtz & Mugdan 1979; Becker 1990: 49–52)? Zum ganzen Themenkomplex des Wortartenwechsels s. Eschenlohr (1999) und Vogel (1996).

Nicht mit der Rückbildung zu verwechseln ist die **Wortkürzung**, bei der im Gegensatz zur Derivation und Komposition keinerlei semantische oder grammatische Veränderung eintritt, sondern lediglich ein längerer Ausdruck durch einen kürzeren ersetzt wird, etwa *Limonade* durch *Limo*, *Oberkellner* durch *Ober*, *Lastkraftwagen* durch *LKW*, *unbekanntes Flugobjekt* durch *Ufo* usw. (Kobler-Trill 1994; Steinhauer 2000). Die so gekürzten Wörter sind keine durchsichtigen Wortbildungen, sondern (außer beim Typ *Ober* mit Kürzung an der Morphemgrenze) **neue Wurzeln** mit bestimmten, stark präferierten phonologischen Eigenschaften (Ronneberger-Sibold 1996; zum besonderen Typ der sogenannten *i*-Bildungen (*Am-i* < *Amerikaner* + *-i*) auch Neef 1996: 278–284 und R. Wiese 1996: 62–65).

Obwohl die Wortkürzung häufig als eigene Wortbildungsart betrachtet wird (z. B. Duden<sup>6</sup>1998: 420), dürfte es sinnvoller sein, sie zusammen mit anderen, noch weniger regulären Verfahren wie verschiedenen Arten der **Wortkreuzung** (*Sial* < *Silizium* × *Aluminium*) und der **Verfremdung** (*Vileda* < *Wie Leder*) als eine Art von **Wortschöpfung** zu betrachten. Während bei der regulären Wortbildung das Produkt automatisch aus dem sprachlichen Ausgangsmaterial und der Anwendung eines Wortbildungstyps resultiert, wird bei der Schöpfung ein nicht im Einzelnen vorgezeichneter Weg vom vorhandenen Ausgangsmaterial zu einem erwünschten Ziel (meistens einer Lautgestalt mit bestimmten lautlichen Eigenschaften) gesucht (Ronneberger-Sibold 2000).

#### 4.2. Komposition

Durch **Komposition** werden hauptsächlich **Determinativkomposita** gebildet, bei denen das determinierte **Grundwort** oder Kernwort rechts vom determinierenden **Bestimmungswort** steht: In *Gartenhaus* ist *Haus* durch *Garten* näher bestimmt, bei *Hausgarten* ist es um-

gekehrt. Das Grundwort, meist ein Substantiv oder Adjektiv (zu den Verben vgl. 4.3), bestimmt – wie die Ableitungssuffixe – die Wortart sowie gegebenenfalls das Genus des ganzen Kompositums und trägt allein die Flexion; die Flexion des Bestimmungswortes entfällt (*Hochhaus* gegenüber *hohes Haus*). Dies unterscheidet – zusammen mit den sog. Fugenelementen (s. u.) und einer bestimmten Akzentkontur (ein Hauptakzent, i. a. auf dem Bestimmungswort) – die Komposita von freien syntaktischen Fügungen, wo eine solche Verwechslung von der Wortfolge her überhaupt möglich wäre. Komposita können ihrerseits Basis für Ableitungen sein (*jungfräulich*) oder/und zu größeren Komposita zusammengefügt werden, sodaß sich sehr lange Wörter mit komplizierter innerer Struktur ergeben können wie etwa (((*Wohnungs*)-bau)-(*förderungs*))-gesetz) (Duden '1998: 433; einen Vergleich mit den geringeren Möglichkeiten des Englischen s. in Marchand 1960).

Besteht der Bestimmungsteil eines Kompositums zwar aus freien Wörtern, ist selbst aber kein freies Kompositum, so spricht man von **Zusammenbildung**: ((*Ein-familien*)-haus); \**Einfamilie*. Dasselbe ist auch möglich bei Ableitungen: ((*rot-wang*)-ig); \**Rotwange*, \**wangig*. Besonders produktiv ist dieses Verfahren bei der Substantivierung von Verben samt ihren Ergänzungen und Bestimmungen: (*das*) *Von-der-Hand-in-den-Mund-Leben*, (*die*) *Inbetriebnahme*.

Als **exozentrische** oder **Possessivkomposita** (auch Bahuvrihi-Bildungen) werden häufig Komposita wie *Rotkäppchen* bezeichnet, weil die Bedeutung des Gesamtwortes nicht in der des Grundwortes enthalten ist: Rotkäppchen ist kein rotes Käppchen, sondern eine Person, die ein solches Käppchen besitzt. Allerdings läßt sich sinnvollerweise argumentieren, daß dies lediglich eine metonymische Verwendung eines normalen, endozentrischen Determinativkompositums darstellt (s. die Diskussion in Naumann 3'2000: 44 f.). Mit größerem Recht könnte man den Typ *Vormittag* 'die Zeit vor dem Mittag' als exozentrisches Kompositum bezeichnen, etwa im Gegensatz zu *Vor-dach*, das tatsächlich ein Dach bezeichnet (Duden '1998: 493).

Ein Beispiel für die wenigen echten **Kopulativkomposita** ist etwa der Typ *blau-'grün* 'blau und grün gemustert', durch den Endakzent unterschieden von dem determinativen Typ '*blaugrün* 'grün mit blauem Einschlag'.

Eine formale Besonderheit vieler Komposita (und einiger Ableitungen z. B. auf *-haft*) mit einem Substantiv als Bestimmungswort sind die sogenannten **Fugenelemente** -(e)s- (*Wohnung-s-bau*), -e- (*Hund-e-hütte*), -er- (*Hühn-er-bein*), -(e)n- (*Linde-n-blatt*), -(e)ns- (*Herz-ens-angelegenheit*) und Subtraktion eines auslautenden -e (*Kirsch-Ø-kuchen*). Obwohl diese Elemente diachron häufig (aber keineswegs immer) aus Genitivformen hervorgegangen sind wie z. B. in *Tag-es-licht* (Fleischer & Barz 2'1995: 136), lassen sie sich synchron nicht mehr grundsätzlich als Flexionsformen interpretieren wegen semantischer und distributioneller Unverträglichkeiten: Ein Gänsebraten besteht nur aus einer Gans, -s nach -ung wie in *Wohnungs-* ist distributionell ausgeschlossen. Jedoch ist vielfach auch flexivische Interpretation möglich (man spricht von flexionseigenen Fugenelementen) und wohl auch intendiert, z. B. bei Neubildungen mit sogenanntem Binnenplural wie *Parteienaffäre* vs. *Parteiaffäre* (im ersten Fall sind mehrere Parteien involviert, im zweiten nur eine). Wo dies nicht möglich/intendiert ist, kann man die Fugenelemente entweder als funktionsleerte Allomorphe des Erstgliedes auffassen, weil sie bis zu einem gewissen Grade von ihm determiniert werden und bei seinem Tonvokal Umlaut auslösen können, oder aber als selbstständige Interfixe (Dressler et al. 2000). Die zweite Interpretation entspricht eher ihrer Funktion, die Kompositionsfuge zu markieren und dadurch gleichzeitig die Einheit des Kompositums im Gegensatz zur freien syntaktischen Fügung anzudeuteten (zu Determinierung und Funktion ausführlich Fuhrhop 1998: 187–220; Ramers 1997).

#### 4.3. Halbaffixe, trennbare Verben

Bei manchen Wortbildungen wie z. B. *hochmodern* oder *gedankenlos* ist die Einordnung als Kompositum oder Ableitung schwierig: Einerseits sind *hoch* und *los* selbstständige Wörter, andererseits ist in den fraglichen Bildungen ihre Bedeutung "abgeschwächt", und sie sind produktiv fast wie Derivationsaffixe: *hochkomplex*, *hochzufrieden* ..., *brotlos*, *freudlos* ... Man spricht deshalb von **Halbpräfixen** bzw. **-suffixen** oder von **Prä-** bzw. **Suffixoiden**. Es handelt sich hier deutlich um eine Übergangsform bei dem diachronen Prozeß der Entstehung von Affixen aus Kompositonsgliedern (vgl. Urbaniak 1983).

Von den Prä- oder Suffixoiden unterscheiden sich die sogenannten **Konfixe** der gelehrt-

ten Wortbildung wie z. B. *Mikro-* oder *-meter* (als Bezeichnung für ein Meßgerät): Sie haben zwar eine sehr präzise lexikalische Bedeutung, können aber trotzdem nicht wie Lexeme frei vorkommen (Duden '1998: 415 f.; weiteres zur Lehn-Wortbildung Hoppe 1999).

Ein weiterer wichtiger diachroner Prozeß ist die inhaltliche und formale **Unverbierung** syntaktisch häufig benachbarter Wortformen. Durch ihn entstehen die meisten **Verbal-komposita** wie *achtgeben* (nach neuer Rechtschreibung *Acht geben*), *festhalten*, *kennenlernen* (nach neuer Rechtschreibung *kennen lernen* s. u.). Ihre große Besonderheit ist, daß selbst bei "alten", stark idiomatisierten Komposita wie *achtgeben* die formale Unverbierung sich auf diejenigen syntaktischen Stellungstypen beschränkt, in denen die beiden Konstituenten auch nach den allgemeinen syntaktischen Regeln meistens nebeneinander stehen (*weil sie das Gesagte festhält*); in den anderen Stellungstypen erscheinen die Kompositionsglieder getrennt: *Sie hält das Gesagte fest*). Das Grundverb nimmt sozusagen sein Bestimmungswort nicht in alle Positionen mit. Vielmehr bildet es mit ihm eine "Satzklammer" (vgl. 3.2). Bei Kontaktstellung behalten sogar die Infinitivpartikel *zu* sowie das *ge-* des Partizips Perfekt ihre ursprüngliche Stellung vor dem Grundverb bei: Sie stehen also zwischen den Kompositionsgliedern (*festgehalten*, *festzuhalten*). In allen diesen Fällen liegt der Hauptakzent (wie auch bei den meisten nominalen Komposita) auf dem Bestimmungswort. Ist dieses ein Adjektiv, so unterscheidet der Akzentsitz das Kompositum formal von der (oft noch parallel vorhandenen) syntaktischen Fügung: *'festzuhalten : fest zu 'halten*. Bei einem Substantiv oder Verb als Bestimmungswort versagt dieses Kriterium, weil diese Wörter auch in der syntaktischen Fügung den Hauptakzent der Gruppe tragen. Hier hängt die Einschätzung als Kompositum oder syntaktische Fügung allein vom Grad der Idiomatisierung ab. Daraus gilt es hier (ähnlich wie bei den Halbaffixen) die für solche graduellen diachronen Prozesse typischen Übergangsformen, die schwer eindeutig einzuordnen sind. Dies führt zu den notorischen Orthographieproblemen. Z. B. schrieb man vor der Rechtschreibreform *radfahren*, aber *Auto fahren*; heute wird *Rad fahren* getrennt geschrieben, dafür aber auch stärker idiomatisiertes *Acht geben* und *kennen lernen*.

Teilweise anders als diese echten verbalen **Distanzkomposita**, aber nicht einheitlich, ver-

halten sich die – synchron nicht immer eindeutig zu identifizierenden – sog. **verbalen Pseudokomposita** wie *frühstückt* (*gefrühstückt*), *schutzimpfen* (*schutzgeimpft*, *?er schutzimpft*, *\*er impft schutz*), die in Wirklichkeit Ableitungen aus nominalen Komposita (*Frühstück*, *Schutzimpfung*) sind (vgl. Asdahl Holmberg 1976; zur wiederholt erwähnten Nähe der Verbalkomposita zur Inkorporation Wurzel 1996: 504–506).

Die Distanzkomposita werden manchmal unter die **Zusammenrückungen** gezählt. Meistens reserviert man diesen Begriff jedoch für solche durch Unverbierung entstandenen Bildungen wie *daran*, *dabei*, *anstelle*, *Vergißmeinnicht*, bei denen die letzte Konstituente nicht die Wortart der ganzen Bildung bestimmt.

Von den echten Distanzkomposita unterscheidet man die **Partikelverben** wie *abrehen*, *andrehen*, *zurückdrehen*, deren Erstglied ein Adverb oder eine Präposition ist. Syntaktisch und akzentuell verhalten sie sich ganz wie die Distanzkomposita. Deshalb werden sie mit diesen auch als **trennbare Verben** zusammengefaßt. Das abtrennbares Erstglied bezeichnet man in beiden Fällen als **Verbzusatz** oder **Präverb**. Der Unterschied liegt in der Art ihres Entstehens: Während die Entstehung von Distanzkomposita im wesentlichen nach wie vor ein gradueller, diachroner, synchron unproduktiver Prozeß ist, auch wenn er in der Gegenwart abläuft, werden die Partikelverben synchron massenweise nach Bedarf und ohne irgendwelche Zwischenstufen gebildet: Aus dem diachronen Prozeß ist eine synchrone produktive Wortbildungsregel geworden. Hand in Hand damit geht in den meisten Fällen die Entwicklung vom Kompositionsglied zum Halbpräfix durch Idiomatisierung, vor allem bei den Präpositionen. Eine ihrer zahlreichen Funktionen ist der Ausdruck von **Aktionsarten** (*einschlafen* – *aus-schlafen*), ohne daß es jedoch zu einem ausgebauten System wie etwa im Russischen käme (vgl. Steinitz 1981). Manche dieser Halbpräfixe können "fest" oder "unfest" mit ihrem Grundverb verbunden sein, meistens mit deutlicher semantischer Differenzierung: *um'fahren* (*er um'fährt*) – *'umfahren* (*er fährt ... 'um*). Der Hauptakzent liegt dabei bei der unfesten Verbindung auf dem Halbpräfix (wie bei allen Verbzusätzen), bei der festen auf dem Verb (wie bei den echten Präfixverben wie z. B. *be'fahren*).

#### 4.4. Bedeutung und Funktion von Wortbildungen

Die **Bedeutung** von Wortbildungen wird traditionell durch Paraphrasen wiedergegeben, z. B. *Gartenhaus* ‘Haus, das sich in einem Garten befindet’, *eßbarer Apfel* ‘Apfel, der gegessen werden kann’ (zur tatsächlichen Nutzung solcher Paraphrasen in Texten s. Dederding 1982). Die Paraphrasen erlauben einerseits eine Klassifizierung von Wortbildungen nach ihrer Wortbildungsbedeutung zu den von Weisgerber (1964) sogenannten **Wortständen** – z. B. die Zusammenstellung aller Eignungsadjektive (*eßbar*, *löslich* usw.), aller instrumentativen Adjektive (*telefonisch*, *gewaltsam* usw.) –, andererseits die Ermittlung der syntaktischen und textuellen Funktion von Wortbildungen. Ersetzt man z. B. die *Schale mit Blumen bepflanzen* durch die mögliche Paraphrase *Blumen in die Schale pflanzen*, so erkennt man als eine wesentliche Funktion der verbalen Präfixbildung mit *be-* die **Transposition** des Verbs in eine andere Valenzklasse (vgl. Günther 1974; Eroms 1980). Dieses wiederum eröffnet andere Möglichkeiten der Textgestaltung (z. B. die Thematisierung von *Schale* als Subjekt eines Passivsatzes). Häufig ist mit der Transposition (die notwendigerweise immer bei wortklassenverändernder Wortbildung eintritt) eine **Verdichtung** des Ausdrucks verbunden, z. B. beim Ersatz von *Leistungen, die in der Schule erbracht werden* durch *schulische Leistungen* (vgl. Eichinger 1982).

Die Verdichtung ist sicherlich Hauptmotiv bei den vor allem in der Mediensprache sehr häufigen Augenblicksbildungen von Determinativkomposita zum Ersatz für Konstruktionen mit nachgestelltem Genitiv- oder Präpositionalattribut: *die Brandt-Reise* statt *die Reise Brandts* oder *von Brandt* (möglich auch noch *Brandts Reise*). (Allgemein zur Funktion der verschiedenen Wortbildungstypen im Text siehe Eichinger 2000.)

Sehr umstritten ist der Gebrauch von Paraphrasen als Tiefenstrukturen für Wortbildungen in verschiedenen generativen Wortbildungstheorien. Eines der ersten Beispiele ist die Ableitung von *Holzschuppen* aus *Schuppen, der aus Holz hergestellt ist* bzw. *Schuppen, in dem sich Holz befindet* bei Motsch (1970). Mit der Entwicklung der generativen Grammatik ist auch dieses Beschreibungsverfahren verändert/verfeinert worden (vgl. die Darstellungen/Anwendungen in Kürschner 1974; Fanselow 1981; Toman<sup>2</sup> 1987; Olsen 1986). Allen diesen Ansätzen ist gemeinsam, daß sie Wortbildungsgesetze nicht als grundsätzlich verschieden von syn-

taktischen Regeln betrachten. Die Gegenposition vertritt z. B. Erben (42000: 22–23) und das von ihm betreute Projekt ‘Deutsche Wortbildung’ (dazu auch Ortner/Ortner 1984: 124–130; sehr differenziert auch Polenz<sup>2</sup> 1980, Herbermann 1981 und – mit Rückgriff auf den Analogie-Begriff – Becker 1990).

Zu diesem allgemeinen Streit zwischen ‘Syntaktikern’ und ‘Lexikalisten’ in der Wortbildung läßt sich aus morphologischer Sicht wohl sagen, daß das Deutsche auf einer Skala zwischen den beiden Extremen weiter auf der lexikalistischen Seite anzusiedeln wäre als z. B. das Englische, eben wegen der ausgeprägten formalen Merkmale – Flexion, Akzent, Stammallomorphik, Fugenelemente –, die das Wort im Gegensatz zum freien Syntagma als eine eigenständige, potentiell auf-listbare linguistische Einheit identifizieren.

#### 5. Zitierte Literatur

- Abraham, Werner & Janssen, Theo (1989, Hrsg.), *Tempus – Aspekt – Modus: Die lexikalischen und grammatischen Formen in den germanischen Sprachen*. Tübingen: Niemeyer (Linguistische Arbeiten 237)
- Adelung, Johann Christoph (1782), *Umständliches Lehrgebäude der Deutschen Sprache, zur Erläuterung der Deutschen Sprache für Schulen, Bd. I*. Leipzig: Breitkopf [Nachdruck 1971 Hildesheim, New York: Olms]
- Admoni, Wladimir G. (41982), *Der deutsche Sprachbau*. München: Beck [<sup>1</sup>1960 Leningrad]
- Appel, Elsbeth (1941), *Vom Fehlen des Genitiv-s*. München: Beck (Arbeiten zur Entwicklungspsychologie 21)
- Arndt, Walter W. (1970), “Nonrandom Assignment of Loanwords: German Noun Gender”. *Word* 26, 244–253
- Asdahl Holmberg, Märta (1976), *Studien zu den verbalen Pseudokomposita im Deutschen*. Göteborg: Universität Göteborg (Göteborger Germanistische Forschungen 14)
- Askedal, John Ole (1982), “Zur semantischen Analyse deutscher Auxiliarkonstruktionen unter dem Aspekt der Unterscheidung von summariver und nicht-summariver Bedeutung”. *Deutsch als Fremdsprache* 19, 159–167
- Askedal, John Ole (1996), “Überlegungen zum Deutschen als sprachtypologischem ‘Mischtyp’”. In: Lang & Zifonun (Hrsg.), 369–383
- Augst, Gerhard (1975 a), *Lexikon zur Wortbildung*, Bd. 1–3. Tübingen: Narr (Forschungsberichte des Instituts für deutsche Sprache 24, 1–3)

- Augst, Gerhard (1975 b), *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*. Tübingen: Narr (Forschungsberichte des Instituts für deutsche Sprache 25)
- Ballweg, Joachim (1987), "More Future!". In: Dyhr, Mogens & Olsen, Jørgen (Hrsg.), *Festschrift für Karl Hyldgaard-Jensen zum 70. Geburtstag am 3. Februar 1987*. Kopenhagen: Reitzel (Kopenhagener Beiträge zur Germanistischen Linguistik, Sonderband 3), 11–24
- Ballweg, Joachim (1988), *Die Semantik der deutschen Tempusformen*. Düsseldorf: Schwann (Sprache der Gegenwart 70)
- Bartke, Susanne (1998), *Experimentelle Studien zur Flexion und Wortbildung*. Tübingen: Niemeyer (Linguistische Arbeiten 376)
- Bartsch, Werner (1980), *Tempus, Modus, Aspekt: Die systembildenden Ausdrücke beim deutschen Verbalkomplex*. Frankfurt a. M.: Diesterweg (Schule und Forschung: Neusprachliche Abteilung 4210)
- Bausch, Karl-Heinz (1979), *Modalität und Konjunktivebrauch in der gesprochenen deutschen Standardsprache: Sprachsystem, Sprachwandel und Sprachvariation im heutigen Deutsch, Teil I: Forschungslage, theoretische und empirische Grundlagen, morphologische Analyse*. München: Hueber (Heutiges Deutsch I, 9, 1)
- Bech, Gunnar (1963), "Zur Morphologie der deutschen Substantive". *Lingua* 12, 177–189
- Becker, Thomas (1990), *Analogie und morphologische Theorie*. München: Fink (Studien zur Theoretischen Linguistik 11)
- Bergenholtz, Henning & Mugdan, Joachim (1979), "Ist Liebe primär? Über Ableitung und Wortarten". In: Braun, Peter (Hrsg.), *Deutsche Gegenwartssprache: Entwicklungen, Entwürfe, Diskussionen*. München: Fink (Kritische Informationen 79), 339–354
- Besch, Werner (1983), "Dialekt, Schreibdialekt, Schriftsprache, Standardsprache". In: Besch et al. (Hrsg.), 961–991
- Besch, Werner & Knoop, Ulrich & Putschke, Wolfgang & Wiegand, Herbert Ernst (1983, Hrsg.), *Dialektologie: Ein Handbuch zur deutschen und allgemeinen Dialektforschung, 2. Halbband*. Berlin, New York: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft 1.2)
- Bettelhäuser, Hans-Jörg (1976), *Studien zur Substantivflexion der deutschen Gegenwartssprache*. Heidelberg: Winter (Monographien zur Sprachwissenschaft 2)
- Bittner, Andreas (1996), *Starke "schwache" Verben – schwache "starke" Verben: Deutsche Verbflexion und Natürlichkeit*. Tübingen: Stauffenburg (Studien zur deutschen Grammatik 51)
- Bittner, Dagmar (1990), "Strukturprinzipien der neuhochdeutschen Substantivflexion: Vorschlag eines Modells". In: Boretzky, Norbert & Enninger, Werner & Stolz, Thomas (Hrsg.), *Spielarten der Natürlichkeit – Spielarten der Ökonomie. Beiträge zum 5. Essener Kolloquium über "Grammatikalisierung: Natürlichkeit und Systemökonomie" vom 6. 10.–8. 10. 1988 an der Universität Essen, 2. Band, 2. Halbband*. Bochum: Brockmeyer (Bochum-Essener Beiträge zur Sprachwandelforschung VIII, 2), 207–225
- Bittner, Dagmar (2003), *Von starken Feminina und schwachen Maskulina. Die neuhochdeutsche Substantivflexion – Eine Systemanalyse im Rahmen der natürlichen Morphologie*. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung (ZAS Papers in Linguistics 31)
- Brinker, Klaus (1971), *Das Passiv im heutigen Deutsch: Form und Funktion*. München: Hueber; Düsseldorf: Schwann (Heutiges Deutsch I.2)
- Butt, Matthias & Fuhrhop, Nanna (1999, Hrsg.), *Variation und Stabilität in der Wortstruktur: Untersuchungen zu Entwicklung, Erwerb und Varietäten des Deutschen und anderer Sprachen*. Hildesheim usw.: Olms (Germanistische Linguistik 141–142, 1988)
- Carstairs, Andrew (1986), "Macroclasses and Paradigm Economy in German Nouns". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39, 3–11
- Dal, Ingerid (1942), "Systemerhaltende Tendenzen in der deutschen Kasusmorphologie". *Norsk tidskrift for sprogvitenskap* 12, 199–212 [Nachdruck in: Dal, Ingerid (1971), *Untersuchungen zur germanischen und deutschen Sprachgeschichte*. Oslo usw.: Universitetsforlaget, 158–170]
- Darski, Józef (1999), *Bildung der Verbformen im Standarddeutschen*. Tübingen: Stauffenburg (Stauffenburg-Linguistik)
- Dedenbach, Beate (1987), *Reduktions- und Verschmelzungsformen im Deutschen: schwache Formen bei Artikeln und Pronomina*. Frankfurt a. M.: Lang (Europäische Hochschulschriften 1.1016)
- Dederding, Hans-Martin (1982), *Wortbildung, Syntax, Text: Nominalkomposita und entsprechende syntaktische Strukturen in deutschen Patent- und Auslegeschriften*. Erlangen: Palm & Enke (Erlanger Studien 34)
- Der Begriff Tempus (1969) = Der Begriff Tempus – eine Ansichtssache? Düsseldorf: Schwann (Wir-kendes Wort: Beihefte 20)

- Deutsche Wortbildung (1973–1992) = *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache: Eine Bestandsaufnahme des Instituts für deutsche Sprache, Forschungsstelle Innsbruck.* [S. Kühnhold & Wellmann (1973), Wellmann (1975), Kühnhold & Putzer & Wellmann (1978), Kühnhold & Prell (1984), Ortner & Müller-Bollhagen et al. (1991), Pümpel-Mader & Gassner-Koch & Wellmann (1992)]
- Diewald, Gabriele (1997), *Grammatikalisierung: Eine Einführung in Sein und Werden grammatischer Formen.* Tübingen: Niemeyer (Germanistische Arbeitshefte 36)
- Doleschal, Ursula (1992), *Movierung im Deutschen: Eine Darstellung der Bildung und Verwendung weiblicher Personenbezeichnungen.* Unterschleißheim: Lincom Europa (Edition Linguistik 1)
- Donhauser, Karin (1986), *Der Imperativ im Deutschen: Studien zur Syntax und Semantik des deutschen Modussystems.* Hamburg: Buske (Bayreuther Beiträge zur Sprachwissenschaft 6)
- Dressler, Wolfgang U. & Libben, Gary & Stark, Jacqueline & Pons, Christiane & Jerama, Gonia (2000), “The Processing of Interfixed German Compounds”. *Yearbook of Morphology* 1999, 185–220
- Duden <sup>2</sup>1966 = *Duden: Grammatik der deutschen Gegenwartssprache*, bearb. von Paul Grebe. Mannheim usw.: Bibliographisches Institut (Der Große Duden, Band 4)
- Duden <sup>6</sup>1998 = *Duden: Grammatik der deutschen Gegenwartssprache*, bearb. von Peter Eisenberg et al. Mannheim usw.: Bibliographisches Institut (Der Große Duden, Band 4)
- Eichinger, Ludwig M. (1982), *Syntaktische Transposition und semantische Derivation: Die Adjektive auf '-isch' im heutigen Deutsch.* Tübingen: Niemeyer (Linguistische Arbeiten 113)
- Eichinger, Ludwig M. (2000), *Deutsche Wortbildung: Eine Einführung.* Tübingen: Narr (Narr Studienbücher)
- Eisenberg, Peter (1998), *Grundriss der deutschen Grammatik, Bd. I: Das Wort.* Stuttgart: Metzler [<sup>1</sup>1986]
- Engel, Ulrich & Grebe, Paul & Rupp, Heinz (1969, Hrsg.), *Festschrift für Hugo Moser zum 60. Geburtstag am 19. Juni 1969.* Düsseldorf: Schwann
- Erben, Johannes (1969), “Tun als Hilfsverb im heutigen Deutsch”. In: Engel et al. (Hrsg.), 46–52
- Erben, Johannes (<sup>4</sup>2000), *Einführung in die deutsche Wortbildungslehre.* Berlin: Schmidt (Grundlagen der Germanistik 17) [<sup>1</sup>1975]
- Eroms, Hans-Werner (1980), ‘*Be-'Verb und Präpositionalphrase: Ein Beitrag zur Grammatik deutscher Verbalpräfixe.* Heidelberg: Winter (Monographien zur Sprachwissenschaft 9)
- Eschenlohr, Stefanie (1999), *Vom Nomen zum Verb: Konversion, Präfigierung und Rückbildung im Deutschen.* Hildesheim etc: Olms (Germanistische Linguistik. Monographien 3)
- Fabri, Ray & Ortmann, Albert & Parodi, Teresa (1998, Hrsg.), *Models of Inflection.* Tübingen: Niemeyer (Linguistische Arbeiten 388)
- Fabricius-Hansen, Cathrine (1977), “Zur Klassifizierung der starken Verben im Neuhochdeutschen”. *Deutsche Sprache* 5, 193–205
- Fabricius-Hansen, Cathrine (1986), *Tempus fugit: Über die Interpretation temporaler Strukturen im Deutschen.* Düsseldorf: Schwann (Sprache der Gegenwart 64)
- Fabricius-Hansen, Cathrine (2000), “Die Geheimnisse der deutschen würde-Konstruktion”. In: Thieroff et al. (Hrsg.), 139–153
- Fanselow, Gisbert (1981), *Zur Syntax und Semantik der Nominalkomposition: Ein Versuch praktischer Anwendung der Montague-Grammatik auf die Wortbildung im Deutschen.* Tübingen: Niemeyer (Linguistische Arbeiten 107)
- Feuillet, Jack (1993), *Grammaire structurale de l'allemand.* Bern usw.: Peter Lang (Collection contacts, Série III: Etudes et documents 20)
- Fleischer, Wolfgang (1969), *Wortbildung der deutschen Gegenwartssprache.* Leipzig: Enzyklopädie
- Fleischer Wolfgang & Barz, Irmhild (<sup>2</sup>1995), *Wortbildung der deutschen Gegenwartssprache.* Tübingen: Niemeyer
- Fourquet, Jean (1952), *Grammaire de l'allemand.* Paris: Hachette
- Fourquet, Jean (1969), “Das Werden des neuhochdeutschen Verbsystems”. In: Engel et al. (Hrsg.), 53–65
- Fourquet, Jean (1970 a), *Prolegomena zu einer deutschen Grammatik.* Düsseldorf: Schwann (Sprache der Gegenwart 7)
- Fourquet, Jean (1970 b), “Les verbes forts de l'allemand”. *Les langues modernes* 64, 470–477
- Frey, Eberhard (1975), “The Economy of the German Gender-Number-Case Inflection”. *International Review of Applied Linguistics in Language Teaching* 13, 185–197
- Fuhrhop, Nanna (1998), *Grenzfälle morphologischer Einheiten.* Tübingen: Stauffenburg (Studien zur deutschen Grammatik 57)

- Gallmann, Peter (1990), *Kategoriall komplexe Wortformen: Das Zusammenwirken von Morphologie und Syntax bei der Flexion von Nomen und Adjektiv*. Tübingen: Niemeyer (Germanistische Linguistik 108)
- Gersbach, Bernhard & Graf, Rainer (1984–1985), *Wortbildung in gesprochener Sprache: die Substantiv-, Verb- und Adjektiv-Zusammensetzungen und -Ableitungen im "Häufigkeitswörterbuch gesprochener Sprache", Bd. 1–2*. Tübingen: Niemeyer (Idiomatica 12, 13)
- Glinz, Hans (1970), *Deutsche Grammatik I: Satz – Verb – Modus – Tempus*. Frankfurt a.M.: Athenäum (Studienbücher zur Linguistik und Literaturwissenschaft 2)
- Graf, Rainer (1977), *Der Konjunktiv in gesprochener Sprache: Form, Vorkommen und Funktion untersucht an Tonbandaufnahmen aus Baden-Württemberg, Bayrisch-Schwaben, Vorarlberg und Liechtenstein*. Tübingen: Niemeyer (Idiomatica 5)
- Gregor, Bernd (1983), *Genuszuordnung: Das Genus englischer Lehnwörter im Deutschen*. Tübingen: Niemeyer (Linguistische Arbeiten 129)
- Grimm, Jacob (21870), *Deutsche Grammatik, Erster Theil* (besorgt durch Wilhelm Scherer). Berlin: Dümmler [1819 Göttingen: Dieterich; Nachdruck 1967 Hildesheim: Olms]
- Grønvik, Ottar (1986), *Über den Ursprung und die Entwicklung der aktiven Perfekt- und Plusquamperfektkonstruktionen des Hochdeutschen und ihre Eigenart innerhalb des germanischen Sprachraums*. Oslo: Solum; Paderborn: Schöningh
- Günther, Hartmut (1974), *Das System der Verben mit BE- in der deutschen Sprache der Gegenwart: Ein Beitrag zur Struktur des Lexikons der deutschen Grammatik*. Tübingen: Niemeyer (Linguistische Arbeiten 23)
- Halle, Morris (1970), “Die Konjugation im Deutschen”. In: Steger (Hrsg.), 319–331 [orig.: Halle, Morris (1953), “The German Conjugation”. *Word* 9, 45–53]
- Harnisch, Rüdiger (1988), “Natürliche Morphologie und morphologische Ökonomie”. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 41, 426–437
- Hawkins, John A. (1986), *A Comparative Typology of English and German: Unifying the Contrasts*. Austin/TX: University of Texas Press
- Hawkins, John A. (1987), “German”. In: Comrie, Bernard (Hrsg.), *The World's Major Languages*. London, Sydney: Croom Helm, 110–138
- Hempen, Ute (1988), *Die starken Verben im Deutschen und Niederländischen: Diachrone Morpholo-*gie. Tübingen: Niemeyer (Linguistische Arbeiten 214)
- Henzen, Walter (31965), *Deutsche Wortbildung*. Tübingen: Niemeyer (Sammlung kurzer Grammatiken germanischer Dialekte B Ergänzungsserie 5)
- Herbermann, Clemens Peter (1981), *Wort, Basis, Lexem und die Grenze zwischen Lexikon und Grammatik: Eine Untersuchung am Beispiel der Bildung komplexer Substantive*. München: Fink
- Hoppe, Gabriele (1999), *Das Präfix -ex: Beiträge zur Lehn-Wortbildung*. Tübingen: Narr (Studien zur deutschen Sprache 15)
- Hoppe, Gabriele & Kirkness, Alan & Link, Elisabeth & Nortmeyer, Isolde & Rettig, Wolfgang & Schmidt, Günther D. (1987), *Deutsche Lehnwortbildung: Beiträge zur Erforschung der Wortbildung mit entlehnten WB-Einheiten im Deutschen*. Tübingen: Narr (Forschungsberichte des Instituts für deutsche Sprache 64)
- Howe, Stephen (1996), *The Personal Pronouns in the Germanic Languages: A Study of Personal Pronoun Morphology and Change in the Germanic Languages from the First Records to the Present Day*. Berlin, New York: de Gruyter (Studia Linguistica Germanica 4)
- Jørgensen, Mogens Wied (1980), “Zur Anzahl der Klassen in der deutschen Substantivdeklination”. In: Dyhr, Mogens & Hyldgaard-Jensen, Karl & Olsen, Jørgen (Hrsg.), *Festschrift für Gunnar Bech zum 60. Geburtstag am 23. März 1980*. Kopenhagen: Universität Kopenhagen, Institut für germanische Philologie (Kopenhagener Beiträge zur Germanistischen Linguistik, Sonderband 1), 228–261
- Kastovsky, Dieter (1969), “Wortbildung und Nullmorphem”. *Linguistische Berichte* 2, 1–13 [Nachdruck in: Lipka & Günther (1981, Hrsg.), 145–157]
- Klein, Wolfgang (1999), “Wie sich das deutsche Perfekt zusammensetzt”. In: Klein, Wolfgang & Musan, Renate (Hrsg.), *Das deutsche Perfekt. Zeitschrift für Literaturwissenschaft und Linguistik* 40, Heft 113, 52–85
- Kloeke, Wus van Lessen (1982), *Deutsche Phonologie und Morphologie: Merkmale und Markiertheit*. Tübingen: Niemeyer (Linguistische Arbeiten 117)
- Kobler-Trill, Dorothea (1994), *Das Kurzwort im Deutschen: Eine Untersuchung zu Definition, Typologie und Entwicklung*. Tübingen: Niemeyer (Germanistische Linguistik 149)
- Köpcke, Klaus-Michael (1982), *Untersuchungen zum Genussystem der deutschen Gegenwartssprache*. Tübingen: Niemeyer (Linguistische Arbeiten 122)

- Köpcke, Klaus-Michael (1993), *Schemata bei der Pluralbildung im Deutschen: Versuch einer kognitiven Morphologie*. Tübingen: Narr
- Köpcke, Klaus-Michael (1999), "Prototypisch starke und schwache Verben der deutschen Gegenwartssprache". In: Butt & Fuhrhop (Hrsg.), 45–60
- Köpcke, Klaus Michael (2000), "Starkes, Schwaches und Gemischtes in der Substantivflexion des Deutschen: Was weiß der Sprachbenutzer über seine Deklinationsparadigmen?". In: Thieroff et al. (Hrsg.), 171–183
- Köpcke, Klaus Michael & Zubin, David (1996), "Prinzipien für die Genuszuweisung im Deutschen". In: Lang & Zifonun (Hrsg.), 473–491
- Kühnhold, Ingeburg & Prell, Heinz-Peter (1984), *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache: Morphem- und Sachregister zu Band I-III*. Düsseldorf: Schwann (Sprache der Gegenwart 62)
- Kühnhold, Ingeburg & Putzer, Oskar & Wellmann, Hans (1978), *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache, 3. Hauptteil: Das Adjektiv*. Düsseldorf: Schwann (Sprache der Gegenwart 43)
- Kühnhold, Ingeburg & Wellmann, Hans (1973), *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache, 1. Hauptteil: Das Verb*. Düsseldorf: Schwann (Sprache der Gegenwart 29)
- Kürschner, Wilfried (1974), *Zur syntaktischen Beschreibung deutscher Nominalkomposita: Auf der Grundlage generativer Transformationsgrammatiken*. Tübingen: Niemeyer (Linguistische Arbeiten 18)
- Lang, Ewald & Zifonun, Gisela (1996, Hrsg.), *Deutsch – typologisch*. Berlin, New York: de Gruyter (Institut für deutsche Sprache: Jahrbuch 1995)
- Leibukst, Oddleif (1997), *Untersuchungen zum "Bekommen"-Passiv im heutigen Deutsch*. Tübingen: Niemeyer (Germanistische Linguistik 177)
- Leiss, Elisabeth (1992), *Die Verbalkategorien des Deutschen: Ein Beitrag zur Theorie der sprachlichen Kategorisierung*. Berlin, New York: de Gruyter (Studia linguistica germanica 31)
- Leys, Odo (1986), "Zur Geschichte und zur Formulierung der Pluralregeln der deutschen Substantive". In: Cox, Heinrich L. & Vanacker, V. F. & Verhofstadt, E. (Hrsg.), *Wortes anst – verbi gratia. Donum natalicium Gilbert A. R. De Smet*. Leuven: Acco, 303–308
- Linguistische Studien III (1973) = *Linguistische Studien III. Festgabe für Paul Grebe zum 65. Geburtstag, Teil 1*. Düsseldorf: Schwann (Sprache der Gegenwart 23)
- Lipka, Leonhard & Günther, Hartmut (1981, Hrsg.), *Wortbildung*. Darmstadt: Wissenschaftliche Buchgesellschaft (Wege der Forschung 564)
- Ljungerud, Ivar (1955), *Zur Nominalflexion in der deutschen Literatursprache nach 1900*. Lund: Gleerup; Kopenhagen: Munksgaard (Lunder germanistische Forschungen 31)
- Marchand, Hans (1960), "Die Länge englischer Komposita und die entsprechenden Verhältnisse im Deutschen". *Anglia* 78, 411–416 [Nachdruck in: Kastovsky, Dieter (1974, Hrsg.), *Studies in Syntax and Word-formation. Selected Articles by Hans Marchand on the Occasion of His 65th Birthday*. München: Fink (Internationale Bibliothek für allgemeine Linguistik 18), 199–205]
- Mater, Erich (1968), *Deutsche Verben, Bd. V: Flexionsklassen*. Leipzig: Bibliographisches Institut
- Moser, Hugo (1973), "Beobachtungen zum heutigen Kasussystem und Kasusgebrauch". In: Scholler, Harald & Reidy, John (Hrsg.), *Lexicography and Dialect Geography. Festgabe für Hans Kurath*. Wiesbaden: Steiner (Zeitschrift für Dialektologie und Linguistik, N. F. 9), 165–186
- Moskal'skaja, Ol'ga I. (1975), *Grammatik der deutschen Gegenwartssprache*. Moskau: "Hochschule"
- Motsch, Wolfgang (1970), "Analyse von Komposita mit zwei nominalen Elementen". In: Bierwisch, Manfred & Heidolph, Karl Erich (Hrsg.), *Progress in Linguistics*. Den Haag: Mouton, 208–223 [Nachdruck in: Lipka & Günther (1981, Hrsg.), 212–232]
- Mugdan, Joachim (1977), *Flexionsmorphologie und Psycholinguistik*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 82)
- Naumann, Bernd (2000), *Einführung in die Wortbildungslehre des Deutschen*. Tübingen: Niemeyer (Germanistische Arbeitshefte 4) [1972]
- Neef, Martin (1996), *Wortdesign: Eine deklarative Analyse der deutschen Verbflexion*. Tübingen: Stauffenburg (Studien zur deutschen Grammatik 52)
- Nübling, Damaris (1992), *Klitika im Deutschen: Schriftsprache, Umgangssprache, alemannische Dialekte*. Tübingen: Narr (ScriptOralia 42)
- Nübling, Damaris (2000), *Prinzipien der Irregularisierung: Eine kontrastive Analyse von zehn Verben in zehn germanischen Sprachen*. Tübingen: Niemeyer (Linguistische Arbeiten 415)
- Olsen, Susan (1986), *Wortbildung im Deutschen: Eine Einführung in die Theorie der Wortstruktur*. Stuttgart: Kröner (Kröners Studienbibliothek 660)

- Ortmann, Wolf Dieter (1975–1978), *Hochfrequente deutsche Wortformen, Bd. 1–4*. München: Goethe-Institut, Arbeitsstelle für wissenschaftliche Didaktik, Projekt Phonothek
- Ortner, Lorelies & Müller-Bolhagen, Elgin & Ortner, Hanspeter & Wellmann, Hans & Pümpel-Mader, Maria & Gärtner, Hildegarde (1991), *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache, 4. Hauptteil: Substantivkomposita*. Berlin: de Gruyter (Sprache der Gegenwart 79)
- Ortner, Hanspeter & Ortner, Lorelies (1984), *Zur Theorie und Praxis der Kompositaforschung: Mit einer ausführlichen Bibliographie*. Tübingen: Narr (Forschungsberichte des Instituts für deutsche Sprache 55)
- Pfeffer, J. Alan & Lorentz, James P. (1979), “Der analytische Genitiv in Wort und Schrift”. *Muttersprache* 89, 53–70 [Nachdruck in: Pfeffer, J. Alan (1982, Hrsg.), *Probleme der deskriptiven deutschen Grammatik*. Heidelberg: Groos (Studies in Descriptive Linguistics 7), 7–33]
- Pike, Kenneth L. (1965), “Non-linear Order and Anti-redundancy in German Morphological Matrices”. *Zeitschrift für Mundartforschung* 32, 193–221
- Polenz, Peter von (1980), “Wortbildung”. In: Althaus, Hans Peter & Henne, Helmut & Wiegand, Herbert Ernst (Hrsg.), *Lexikon der Germanistischen Linguistik, Bd. I*. Tübingen: Niemeyer, 169–180 [¹1973]
- Pümpel-Mader, Maria & Gassner-Koch, Elisabeth & Wellmann, Hans (1992), *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache, 5. Hauptteil: Adjektivkomposita, Partizipialbildungen*. Berlin: de Gruyter
- Radtke, Petra (1998), *Die Kategorien des deutschen Verbs: Zur Semantik grammatischer Kategorien*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 438)
- Ramers, Karl Heinz (1997), “Die Kunst der Fuge: Zum morphologischen Status von Verbindungselementen in Nominalkomposita”. In: Dürscheid, Christa & Ramers, Karl Heinz (Hrsg.), *Sprache im Fokus. Festschrift für Heinz Vater zum 65. Geburtstag*. Tübingen: Niemeyer, 33–45
- Redder, Angelika (1999), “‘werden’ – funktional-grammatische Bestimmung”. In: Redder, Angelika & Rehbein, Jochen (Hrsg.), *Grammatik und mentale Prozesse*. Opladen: Westdeutscher Verlag, 295–336
- Reed, Carroll E. (1973), “Verb Conjugation in Modern Standard German”. In: Linguistische Studien III, 163–170
- Reis, Marga (1976), “Zum grammatischen Status der Hilfsverben”. *Beiträge zur Geschichte der deutschen Sprache und Literatur* (Tübingen) 98, 64–82
- Rettig, Wolfgang (1972), *Sprachsystem und Sprachnorm in der deutschen Substantivflexion*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 32)
- Ronneberger-Sibold, Elke (1980), *Sprachverwendung – Sprachsystem: Ökonomie und Wandel*. Tübingen: Niemeyer (Linguistische Arbeiten 87)
- Ronneberger-Sibold, Elke (1996), “Preferred Sound Shapes of New Roots: On Some Phonotactic and Prosodic Properties of Shortenings in German and French”. In: Hurch, Bernhard & Rhodes, Richard (Hrsg.), *Natural Phonology: The State of the Art*. Berlin, New York: Mouton de Gruyter (Trends in Linguistics, Studies and Monographs 92), 261–292
- Ronneberger-Sibold, Elke (1997), “Typology and the Diachronic Evolution of German Morphosyntax”. In: Fisiak, Jacek (Hrsg.), *Linguistic Reconstruction and Typology*. Berlin, New York: Mouton de Gruyter (Trends in Linguistics, Studies and Monographs 96), 313–335
- Ronneberger-Sibold, Elke (2000), “Creative Competence at Work: the Creation of Partial Motivation in German Trade Names”. In: Doleschal, Ursula & Thornton, Anna M. (Hrsg.), *Extragrammatical and Marginal Morphology*. München: Lincom Europa, 87–105
- Rowley, Anthony (1988), “Zum Genitiv des ganz besonderen Typs”. *Muttersprache* 98, 58–68
- Rupp, Heinz (1965), “Zum deutschen Verbalsystem”. *Neuphilologische Mitteilungen* 66, 179–199 [Nachdruck in: *Jahrbuch des Instituts für deutsche Sprache 1965/66* (1967), Düsseldorf: Schwann (Sprache der Gegenwart 1), 148–164]
- Saltveit, Laurits (1960), “Besitzt die deutsche Sprache ein Futur?”. *Der Deutschunterricht* 12, 46–65
- Sandberg, Bengt (1976), *Die neutrale -(e)n-Ableitung der deutschen Gegenwartssprache: Zu dem Aspekt der Lexikalisierung bei den Verbalsubstantiven*. Göteborg: Universität Göteborg (Göteborgger germanistische Forschungen 15)
- Ségéral, Philippe & Scheer, Tobias (1998), “A Generalized Theory of Ablaut: The Case of German Strong Verbs”. In: Fabri et al. (Hrsg.), 28–59
- Soeteman, Carl (1967), “Präteritopräsentia revisa”. In: *Jahrbuch des Instituts für deutsche Sprache 1965/66*. Düsseldorf: Schwann (Sprache der Gegenwart 1), 137–147
- Steger, Hugo (1970, Hrsg.), *Vorschläge für eine strukturelle Grammatik des Deutschen*. Darmstadt: Wissenschaftliche Buchgesellschaft (Wege der Forschung 146)

- Steinhauer, Anja (2000), *Sprachökonomie durch Kurzwörter: Bildung und Verwendung in der Fachkommunikation*. Tübingen: Narr (Forum für Fachsprachen-Forschung 56)
- Steinitz, Renate (1981), *Der Status der Kategorie Aktionsart in der Grammatik oder Gibt es Aktionsarten im Deutschen?* Berlin/DDR: Zentralinstitut für Sprachwissenschaft (Linguistische Studien A 76)
- Teuber, Oliver (1999), "fasel, beschreib, erwähn – der Inflektiv als Wortform des Deutschen?". In: Butt & Fuhrhop (Hrsg.), 7–26
- Teuber, Oliver (2000), "Gibt es zwei Genitive im Deutschen?". In: Thieroff et al. (Hrsg.), 171–183
- Theobald, Elke (1992), *Sprachwandel bei deutschen Verben: Flexionsklassenschwankungen starker und schwacher Verben*. Tübingen: Narr
- Thieroff, Rolf (1992), *Das finite Verb im Deutschen: Tempus – Modus – Distanz*. Tübingen: Narr (Studien zur deutschen Grammatik 40)
- Thieroff, Rolf & Tamrat, Matthias & Fuhrhop, Nanna & Teuber, Oliver (2000, Hrsg.), *Deutsche Grammatik in Theorie und Praxis*. Tübingen: Niemeyer.
- Toman, Jindřich (1987), *Wortsyntax: Eine Diskussion ausgewählter Probleme deutscher Wortbildung*. Tübingen: Niemeyer (Linguistische Arbeiten 137) [1983]
- Ulvestadt, Bjarne (1956), "The Strong Verb Conjugation System in German". *Word* 12, 91–105 [Übers.: "Das Konjugationssystem der starken Verben im Deutschen". In: Steger (1970, Hrsg.), 332–348]
- Urbaniak, Gertrud (1983), *Adjektive auf -voll*. Heidelberg: Winter (Monographien zur Sprachwissenschaft 13)
- Vater, Heinz (1975), "Werden als Modalverb". In: Calbert, Joseph H. & Vater, Heinz (Hrsg.), *Aspekte der Modalität*. Tübingen: Narr (Studien zur deutschen Grammatik 1), 71–148
- Vater, Heinz (1979), *Das System der Artikelformen im gegenwärtigen Deutsch*. Tübingen: Niemeyer (Linguistische Arbeiten 78) [1963]
- Veith, Werner (1984), "The Strong Verb Conjugation in Modern English Compared with Modern German". *Leuvense Bijdragen*, 73, 39–57
- Vogel, Petra Maria (1996), *Wortarten und Wortartenwechsel: Zu Konversion und verwandten Erscheinungen im Deutschen und in anderen Sprachen*. Berlin, New York: de Gruyter (Studia linguistica germanica 39)
- Wegener, Heide (1995), *Die Nominalflexion des Deutschen – verstanden als Lerngegenstand*. Tübingen: Niemeyer (Germanistische Linguistik 151)
- Wegener, Heide (2002), "Aufbau von markierten Pluralklassen im Deutschen – eine Herausforderung für die Markiertheitstheorie". *Folia linguistica*, 36, 261–295
- Weinrich, Harald (1971), *Tempus: Besprochene und erzählte Welt*. Stuttgart: Kohlhammer (Sprache und Literatur 16) [1964]
- Weisgerber, Leo (1964), "Vierstufige Wortbildungslehre". *Muttersprache* 74, 33–43 und 96 [Nachdruck in: Lipka & Günther (1981, Hrsg.), 36–54]
- Wellmann, Hans (1975), *Deutsche Wortbildung: Typen und Tendenzen in der Gegenwartssprache*, 2. Hauptteil: Das Substantiv. Düsseldorf: Schwann (Sprache der Gegenwart 32)
- Werner, Otmar (1965), "Vom Formalismus zum Strukturalismus in der historischen Morphologie: Ein Versuch, dargestellt an der Geschichte der Indikativ/Konjunktiv-Bildungen". *Zeitschrift für deutsche Philologie* 84, 100–127 [Nachdruck in: Steger (1970, Hrsg.), 349–384]
- Werner, Otmar (1969), "Das deutsche Pluralsystem: Strukturelle Diachronie". *Jahrbuch des Instituts für deutsche Sprache* 1968. Düsseldorf: Schwann (Sprache der Gegenwart 5), 92–125
- Werner, Otmar (1975), "Zum Genus im Deutschen". *Deutsche Sprache* 3, 35–58
- Werner, Otmar (1979), "Kongruenz wird zu Diskontinuität im Deutschen". In: Brogyanyi, Bela (Hrsg.), *Festschrift for Oswald Szemerényi on the Occasion of His 65th Birthday*, Bd. II. Amsterdam: Benjamins (Current Issues in Linguistic Theory 11), 959–988
- Werner, Otmar (1987), "The Aim of Morphological Change is a Good Mixture – Not a Uniform Language Type". In: Giacalone Ramat, Anna & Carruba, Onofrio & Bernini, Giuliano (Hrsg.), *Papers from the 7th International Conference on Historical Linguistics [Pavia 1985]*. Amsterdam: Benjamins (Current Issues in Linguistic Theory 48), 591–606
- Wiese, Bernd (1996), "Iconicity and Syncretism: On Pronominal Inflection in Modern German". In: Sackmann, Robin (Hrsg.), *Theoretical Linguistics and Grammatical Description*. Amsterdam & Philadelphia: Benjamins, 323–344
- Wiese, Richard (1996), *The Phonology of German*. Oxford: Clarendon Press (The Phonology of the World's Languages)
- Wiesinger, Peter (1983), "Die Einteilung der deutschen Dialekte". In: Besch et al. (Hrsg.), 807–900
- Wurzel, Wolfgang U. (1970), *Studien zur deutschen Lautstruktur*. Berlin/DDR: Akademie-Verlag (Studia grammatica 8)

Wurzel, Wolfgang U. (1984), *Flexionsmorphologie und Natürlichkeit: Ein Beitrag zur morphologischen Theoriebildung*. Berlin/DDR: Akademie-Verlag (Studia grammatica 21)

Wurzel, Wolfgang U. (1991), “‘Genitivverwirrung’ im Deutschen – wie regulär ist morphologischer Wandel?”. In: Boretzky, Norbert & Enninger, Werner & Jeßing, Benedikt & Stolz, Thomas (Hrsg.), *Sprachwandel und seine Prinzipien. Beiträge zum 8. Bochum-Essener Kolloquium über ‘Sprachwandel und seine Prinzipien’ vom 19. 10.–21. 10. 1990 an der Ruhruniversität Bochum*. Bochum: Brockmeyer

(Bochum-Essener Beiträge zur Sprachwandelforschung 14), 168–181

Wurzel, Wolfgang U. (1996), “Morphologischer Strukturwandel: Typologische Entwicklungen im Deutschen”. In: Lang & Zifonun (Hrsg.), 492–524

Wurzel, Wolfgang U. (1998), “Drei Ebenen der Struktur von Flexionsparadigmen”. In: Fabri et al. (Hrsg.), 225–243

*Elke Ronneberger-Sibold, Eichstätt (Deutschland)*

## 121. Français (Indo-européen: Roman)

1. Introduction
2. Types d’opérations de construction de mots: délimitation
3. Opérations dérivationnelles de construction de mots: conditions d’application
4. Structures et sens des mots dérivationnellement construits
5. Conclusion
6. Références

### 1. Introduction

Pour tenir la gageure consistant à exposer en peu de pages les principales propriétés du système morphologique du français, cet article sera consacré aux caractéristiques du système dérivationnel. Celles-ci sont en effet moins connues et ont donné lieu à moins de travaux que celles du système flexionnel. Pour ces dernières, on pourra se reporter à Nyrop (1924) pour un point de vue historique, et, pour des descriptions synchroniques menées de divers points de vue théoriques, à Dubois (1965), Mok (1968), Pinchon (1986: 3–10) pour la morphologie nominale, à Dubois (1967: 56–79), Foley (1979), Isaac (1985), Martinet (1969), Pinchon (1986: 162–174), Pinchon & Couté (1981), Plénat (1981), Schane (1968), Swiggers & Van den Eynde (1987, éds.), Van den Eynde & Blanche-Benveniste (1970) pour la morphologie verbale, à Félice (1950), Hunnius (1990), Mok (1990) pour un aperçu général du système flexionnel.

La description du système dérivationnel du français sera menée d’un point de vue **associatif**, c’est-à-dire que l’on admettra que les règles de construction des mots construisent

à la fois et de façon conjointe la structure et le sens des mots construits (une description et une justification détaillées de cette option théorique figurent dans Corbin 1987; 1991 a).

### 2. Types d’opérations de construction de mots: délimitation

Le domaine d’application des règles lexicales de construction de mots est variable suivant les théories. Sa délimitation reposera ici sur une définition contrainte des mots construits (2.1), sur la distinction entre la notion d’unité lexicale structurellement complexe et celle du mot lexicalement construit (2.2), et sur l’élimination d’un certain nombre d’opérations traditionnelles (2.3).

#### 2.1. Définition du mot construit

Je partirai de la définition suivante (cf. Corbin 1991 a: 16–18; 1991 b: 36–42): un mot est lexicalement construit s’il a une structure telle qu’à chacun de ses constituants soient associées des propriétés catégorielles et sémantiques stables, si son sens prédictible est compositionnel par rapport à sa structure morphologique sous-jacente, et si des mécanismes réguliers permettent d’expliquer les distorsions éventuelles entre son sens lexicalisé et le sens prédictible à partir de sa structure ou entre sa forme apparente et la structure conforme à son sens.

(a) La condition sur les propriétés des constituants permet de distinguer les mots construits des mots que leur forme pourrait inviter, à tort, à analyser comme tels. Par

exemple, même si *royaume* entretient avec *roi* une relation sémantique analogue à celle qui lie *duc* à *duc*, *comté* à *comte*, cette condition interdit de les relier dérivationnellement: *-aume* n'apparaît que dans *royaume* avec ces propriétés, et il n'est pas formellement reliable à *-é*. Même si la taille de ce à quoi réfère *amulette* semble inviter à voir dans ce mot le suffixe *-ette* (cf. *houvette*), la première condition l'élimine des mots construits, car le segment *amul-* n'est pas interprétable.

(b) Dans la mesure où la morphologie lexicale est l'un des domaines où les infractions superficielles aux relations biunivoques entre la forme des séquences et leur sens sont les plus nombreuses (cf. Art. 81 et 82), la condition de **compositionnalité** est actuellement au centre des débats. La deuxième condition impose que toutes les relations formelles et sémantiques ne soient pas analysées comme dérivationnelles, et situe le calcul de la **compositionnalité** au niveau de la structure sous-jacente et du sens prédictible (cf. Corbin 1987; 1989 a). Ainsi, même si *peuplier* et *pommier* semblent avoir la même structure, cette condition permet d'analyser le second comme construit, contrairement au premier: le sens de *pommier* se calcule par rapport à *pomme* et au suffixe *-ier* qui joue le même rôle dans ce mot que dans *abricotier*, *cerisier*, *poirier*, etc. (cf. Corbin & Corbin 1991); celui de *peuplier* est le même que celui de *peuple* – ancien mot conservé régionalement –, et *-ier* ne sert ici qu'à intégrer le nom dans la classe des noms d'arbres (ce n'est pas un suffixe mais ce qu'on appellera un intégrateur paradigmatique, cf. 2.3.2). La structure apparente de *dégraf(er)* s'analyse en *dé-graf(er)*, mais le sens du verbe, semblable à celui de *désagrafer* ('défaire ce qui est agrafé'), ne peut pas se calculer à partir des constituants ainsi isolés (\**graf* est ininterprétable); seule une structure construite sur la forme *agraf-* peut en rendre compte, ce qui suppose une **troncation** du *a*-initial (cf. 3.2): le choix de la base (*agrafe<sub>N</sub>*, *agraf(er)<sub>V</sub>* ou *agrafé<sub>A</sub>*) dépend ensuite de l'économie générale de la description du système dérivationnel (Corbin 1992 a fait l'hypothèse que la base de ce type de verbes en *dé-* est adjetivale). Au niveau sémantique, la démarche adoptée est la même: apparemment, le fait que *lunette* puisse référer, entre autres, à un instrument optique, à la vitre arrière d'une automobile ou à l'ouverture d'un siège d'aisance pourrait inciter à une position sémantiquement pessimiste; en fait, cette po-

lyréférentialité est expliquable, ainsi que l'écart entre ces acceptations et le sens 'petite lune' prédictible à partir de la structure, par des mécanismes sémantiques indépendamment motivés (cf. Corbin et al. 1993).

(c) Les exemples qui viennent d'être développés justifient la condition sur la régularité des mécanismes formels (*dégrafer*) et sémantiques (*lunette*) autorisant à trouver de la **compositionnalité** là où elle n'est pas évidente: seuls des mécanismes indépendamment motivés garantissent le caractère non ad hoc de la description.

## 2.2. Opérations non dérivationnelles de fabrication d'unités lexicales

Toutes les opérations servant à fabriquer des unités lexicales complexes n'appartiennent pas au domaine de la morphologie dérivationnelle:

(a) D'abord parce que beaucoup d'unités lexicales complexes ne répondent pas à la définition du mot construit développée ci-dessus. C'est le cas par exemple des mots-valises (*franglais*, de *français* + *(an)glais*), dans lesquels les segments tronqués ne sont pas nécessairement de nature morphologique (cf. Grésillon 1983), ou des mots résultant d'une suffixation parasitaire, d'une apocope et/ou d'un redoublement (*cradingue*, *crade*, *craqua*, formes "familierées" de *crasseux* (ex. de Plénat 1991 b: 63)), dont la forme suppose l'effacement d'un segment de nature non nécessairement morphologique de la source et dont le sens ne varie que stylistiquement par rapport à celle-ci. C'est aussi le cas des sigles (cf. Art. 92), dont le sens est le même que celui de leur source.

(b) Ensuite parce qu'une distinction doit être opérée entre le processus de lexicalisation et les opérations dérivationnelles de construction de mots. Par transcatégorisation (cf. Corbin 1991 a), des produits d'autres composants de la grammaire que le composant dérivationnel peuvent devenir des unités lexicales:

- Les unités citées en (a), qui sont les produits d'opérations du composant phonologique (cf. Plénat 1984; 1985; 1991 a).
- Les infinitifs et les participes, produits du composant flexionnel: les premiers deviennent des noms de procès (*le manger*, *le boire*, cf. Kerleroux 1990), et les seconds, des adjetifs à sens statif (*un fruit pourri*, cf. Riegel 1985: 190).

- Certains types de séquences syntaxiques nominalisées: *un boit-sans-soif*, *un hors-la-loi*, *un je ne sais quoi*, etc.
- Enfin, des séquences syntaxiques auxquelles se sont appliquées des règles sémantiques, qu'elles deviennent des unités verbales (*casser sa pipe*) ou des noms (*peau rouge*, *sang-froid*).

De la confusion entre les processus de création d'unités lexicales par transcatégorisation et par application de règles proprement lexicales découle la confusion générale entre mot composé et séquence complexe lexicalisée: les deux extrêmes sont représentés d'une part par les études menées dans la perspective du LADL (cf. par exemple Gross 1990) qui appellent mots composés toutes les séquences "figées", d'autre part par Di Sciullo & Williams (1987) qui soutiennent que le français ne dispose pas de règles lexicales de composition. La position adoptée ici (développée dans Corbin 1992 b) distingue (comme Di Sciullo & Williams 1987) entre lexicalisation et composition, mais, contrairement à eux, explique par des règles spécifiquement lexicales de composition les unités complexes dont les autres composants ne peuvent pas rendre compte (cf. 4.4 pour les structures et les sens ainsi construits).

### 2.3. Opérations traditionnelles inutiles

Un certain nombre de mécanismes traditionnellement répertoriés parmi les opérations dérivationnelles se révèlent inutiles soit parce qu'ils procèdent d'erreurs d'analyse, soit parce qu'ils peuvent être remplacés par des mécanismes plus adéquats.

#### 2.3.1. La dérivation "régressive"

La dérivation dite "régressive", qui prend pour base un mot plus long que le dérivé, tantôt calque le traitement linguistique sur la chronologie des attestations (c'est ce qui motive la dérivation d'*agress(er)* à partir d'*agression* proposée par Guilbert 1971: XXXIII), tantôt, quand elle sert à dénommer les relations entre des verbes (*vol(er)*) et des noms (*vol*) (cf. Béchade 1992: 135), confond le verbe et son infinitif, c'est-à-dire prend pour un constituant morphologique (en l'occurrence la base verbale: *vol<sub>V</sub>*) la dénomination métalinguistique de sa catégorie, qui, pour le français comme pour d'autres langues, se présente traditionnellement sous la forme de l'infinitif (le verbe *voler*). En fait, *agression* est suffixé sur *agressy* comme *divi-*

*sion* sur *divis<sub>V</sub>*, et aucun critère formel ne permet d'orienter la relation dérivationnelle entre *vol<sub>N</sub>* et *vol<sub>V</sub>*; seuls des arguments sémantiques peuvent inviter à dériver, par conversion, le nom du verbe (cf. Corbin 1987: 272–279).

#### 2.3.2. La dérivation "parasynthétique"

L'opération dite "parasynthèse", qui consiste à appliquer simultanément à une base un préfixe et un suffixe, procède d'une conception non hiérarchisée de la morphologie non exempte d'erreurs d'analyse (cf. Corbin 1980; 1987: 121–138). Selon les objets auxquels elle s'applique, elle recourt à trois types d'arguments:

(a) Des arguments d'attestation, liés à l'idée tenace que le lexique est un ensemble fini et non structuré de mots. Un mot de structure superficielle *préfixe-X-suffixe* est analysé comme parasynthétique si n'existent ni *préfixe-X* ni *X-suffixe* (cf. Darmesteter, par exemple <sup>2</sup>1893: 96–97, et bien d'autres après lui, comme Thiele 1987: 29): pour ce dernier, comme "ni *\*prisonnement* ni *\*emprison* n'existent avec le statut de lexème", *emprisonnement* est parasynthétique. Ce traitement, qui repose sur une analyse en constituants immédiats de la structure superficielle des mots, témoigne indirectement de la confusion dénoncée ci-dessus entre le constituant morphologique et le terme qui sert à le nommer: le nom *\*emprison* est impossible, mais le verbe *emprison<sub>V</sub>* (*emprisonner*) existe bel et bien, et peut servir de base à *emprisonnement* comme *ration<sub>V</sub>* à *rationnement*.

(b) L'argument selon lequel la préfixation ne pourrait pas déterminer la catégorie des mots construits (cf., après bien d'autres, Zwanenburg 1990: 72). Cet argument, associé au précédent et soutenu par la confusion entre le constituant verbal et sa dénomination, autorise par exemple Béchade (1992: 134) à analyser parasynthétiquement un verbe comme *enivr<sub>V</sub>*. Or, des mots comme *antigrippe<sub>A</sub>*, *apé-tale<sub>A</sub>* et, si l'on admet que l'affixe d'infinitif n'est pas un suffixe dérivationnel, *emprisonn(er)*, *aplat(ir)*, etc. ne peuvent qu'être les résultats d'une préfixation à pouvoir catégorisateur (respectivement sur *grippe<sub>N</sub>*, *pétale<sub>N</sub>*, *prison<sub>N</sub>*, *plat<sub>A</sub>*). Il en va de même pour *enivr(er)*, *enrich(ir)*, etc.

(c) Des arguments sémantiques motivent enfin l'analyse parasynthétique d'un adjectif comme *sous-marin* (Darmesteter <sup>2</sup>1893: 102):

*sous-marin* signifiant ‘sous la mer’, son sens n'est pas calculable par rapport à *sous* et *marin*, mais par rapport à *sous* et *mer*. Mais l'analyse parasyntétique n'est pas la seule solution pour résoudre ce problème. Certains choisissent une alternative pessimiste, consistant à “éparer les règles de structure et les règles sémantiques” (Zwanenburg 1990: 75). Une autre solution, préférable à mon sens car elle conserve la **compositionnalité**, consiste à dire que la structure sous-jacente de *sous-marín* est  $[sous[mer]_N]_A$  et que la finale *-in* n'est pas un suffixe, mais un **intégrateur** au paradigme des adjectifs (cf. en 2.1 l'exemple de *peuplier*) copié sur le suffixe servant à construire un adjectif (*marin*) sur la même base nominale. Une telle solution se justifie (i) par la tendance du français à rendre signifiante la finale des mots (cf. Nyrop 1936: 37 et Guiraud 1967: 93–105); (ii) par les observations empiriques suivantes:

- Les adjectifs préfixés sur des bases nominales peuvent avoir les deux structures apparentes  $[préfixe-X_N]_A$  (*trans-alaska*) et  $[préfixe-X_N\text{-suffixe}]_A$  (*circum-pol-aire*). Il est remarquable que la distribution de ces structures soit indexée, pour les adjectifs à préfixe localisateur, sur l'existence ou l'inexistence d'adjectifs relationnels suffixés sur les noms de base: sur *Alaska* n'est attesté en français aucun adjectif relationnel, l'adjectif signifiant ‘qui traverse l'Alaska’ a donc la forme *transalaska*; si l'adjectif non préfixé est attesté, la finale suffixoïde apparaît, quel que soit le préfixe: par exemple *diadermique* ‘qui traverse le derme’, *endocrânién* ‘qui se situe à l'intérieur du crâne’, *extracorporel* ‘qui se situe à l'extérieur du corps’, *interplanétaire* ‘qui se situe à l'intérieur des limites fixées par des planètes’, *périanal* ‘qui se situe dans la zone entourant l'anus’, *précolombien* ‘qui se situe avant (l'arrivée de) (Christophe) Colomb’, *supraterrestre* ‘qui se situe au-dessus de la terre’, etc.
- Certains adjectifs à préfixe non localisateur peuvent apparaître sous deux formes, l'une avec finale, l'autre sans (*antigrippal* et *antigrippe*, *acéphale* et *acéphalien*), avec un sens absolument identique (‘qui combat la grippe’, ‘dépourvu de tête’). Si ces formes recevaient des analyses différentes, leur synonymie resterait inexplicable (cf. Corbin 1991 a: 14).
- Comme le montrent les exemples cités, les finales suffixoïdes des adjectifs préfixés

sont toujours identiques aux suffixes des adjectifs dénominaux correspondants.

On conclura donc, sans préjuger des autres langues (cf. Booij 1977: 32, qui voit dans *ge-* et *-te* du néerlandais *geboomte* un affixe discontinu), que la parasyntthèse n'est pas un procédé de construction de mots en français (cf. Scalise 1984: 146–151 pour un point de vue proche sur l'italien).

#### 2.4. Bilan

Le domaine ainsi délimité, il reste quatre types structurellement différents d'opérations dérivationnelles de construction de mots en français: la suffixation (les produits ont la structure  $[[X]\text{suffixe}]$ ), la conversion ( $[[X]_C]_C'$ , avec  $C'$  différent de  $C$ ), la préfixation ( $[[préfixe]X]]$ ) et la composition ( $[[X][Y]]$ ). Avant d'exposer en 4 les types de sens associés à ces structures, je passerai rapidement en revue en 3 un certain nombre de questions préalables, relatives aux conditions d'application de ces opérations.

### 3. Opérations dérivationnelles de construction de mots: conditions d'application

#### 3.1. Types de bases

Les opérations dérivationnelles du français s'appliquent à des radicaux de mots français et, à l'exception de la conversion, de mots empruntés au latin ou au grec qui appartiennent de façon latente au français mais n'y apparaissent normalement pas à l'état autonome (en tant qu'unités lexicales potentielles, ils sont marqués ici du signe °): *rapide<sub>A</sub>* → *rapidité<sub>N</sub>* et °*céler<sub>A</sub>* → *célérité<sub>N</sub>*; *commode<sub>A</sub>* → *accommode(er)<sub>V</sub>* et °*céler<sub>A</sub>* → *accélér(er)<sub>V</sub>*; *croqu(er)<sub>V</sub>* + *monsieur<sub>N</sub>* → *croque-monsieur<sub>N</sub>* et °*anthrop(o)<sub>N</sub>* + °*phag<sub>V</sub>* → *anthropophage<sub>N</sub>*. On admet généralement, après Pichon (1942), que les constituants savants et les constituants non savants sont incombinables (cf. Dell & Selkirk 1978; Zwanenburg 1983), mais l'existence de deux systèmes dérivationnels hétérogènes est loin d'être assurée (cf. Corbin 1985).

Les bases des mots construits du français peuvent être non complexes ou complexes. Quand elles sont complexes, elles peuvent avoir été construites par des règles dérivationnelles (cf. 3.6), ou par des opérations d'autres composants de la grammaire (cf. 2.2): les bases peuvent être par exemple des

sigles (*O.N.U.* → *onusien*; *C.G.T.* → *cégétiste*, *anti-C.G.T.*, etc.), ou même, pour certains affixes particuliers, des séquences construites syntaxiquement (*une attitude je m'en foutiste*, *des mesures anti-destruction de la couche d'ozone*).

### 3.2. Modifications formelles liées à l'affixation

La forme observable des mots construits n'est pas nécessairement conforme à la concaténation des affixes et des bases. Elle peut avoir subi des modifications (**allomorphies**), des effacements (**troncations**) et des ajouts (**intégrateurs paradigmatisques** (cf. 2.3.2) et segments parasites).

(a) Les **allomorphies** sont fréquentes et variées. Elles résultent du caractère historiquement stratifié de la constitution du lexique français, et peuvent se traiter, en synchronie, comme des régularités contextuelles: ainsi, l'adjonction de certains suffixes à une base comportant une voyelle antérieure provoque la postériorisation de celle-ci (*mer* → *marin*, *nef* → *navette*, *fleur* → *floral*, *heure* → *horaire*, *seul* → *solitude*, etc.; cf. Dell & Selkirk 1978 et, pour une étude extensive du phénomène, Corbin 1987: 283–340, 737–775).

(b) Alors que les **allomorphies** ne peuvent accompagner qu'une suffixation ou une conversion, les **troncations** de segments basiques peuvent être d'origine suffixale ou préfixale. Elles sont de deux types: ou il s'agit d'une haploglie, liée à la succession dans certaines conditions de deux séquences identiques (la science des minéraux se dit *minéralogie*, et non *minéral-o-logie*, cf. Corbin & Plénat 1992), ou la **troncation** concerne des séquences non identiques: le verbe *désagraf(er)* a aussi la forme *dégraf(er)* (cf. 2.1); ‘rendre politique’ se dit *politis(er)*, et non *politicis(er)* (cf. Corbin 1987: 341–370).

(c) Les segments parasites sont le plus souvent des consonnes dentales qui apparaissent entre certaines bases et certains suffixes sans être formellement ou sémantiquement motivés: *café* → *caféier* et *cafetier(e)* (vs. *thé* → *théier*), *lion* → *lionceau* (vs. *pigeon* → *pigeonneau*), etc.

### 3.3. Contraintes

Si l'on met à part les **contraintes** catégorielles qui régissent la construction dérivationnelle des mots et qui, parce qu'elles conditionnent la structure de ceux-ci, seront étudiées en 4,

il reste toute une série d'autres **contraintes** dont on citera ici les types principaux (cf. Corbin 1989 b).

(a) **Contraintes sémantiques.** Les mots construits ne sont bien formés que si la base est compatible avec l'opération sémantique appliquée. Ainsi, des suffixes comme *-is(er)* ou *-ifi(er)* et la conversion verbale ne peuvent s'appliquer à des adjectifs que si ceux-ci réfèrent à un état non premier, résultatif, comme le montre l'impossibilité de verbes comme *\*impubérer* (\*rendre impubère), *\*puceler* (\*rendre puceau').

(b) **Contraintes structurelles.** Certaines opérations dérivationnelles ne sont compatibles qu'avec des bases non construites: il est remarquable par exemple qu'aucun préfixe entrant dans la construction de verbes de changement d'état (cf. 4.3 (e)) ne puisse s'appliquer à un adjectif construit, au contraire du suffixe *-is(er)* (*élargir*, *aplatir*, *enlaidir* vs. *insensibiliser*).

(c) **Contraintes sélectionnelles.** Les enchaînements de suffixes sont contraints en premier lieu par le type d'instruction sémantique dont chaque suffixe est porteur, ensuite par des attirances spécifiques ne semblant pas relever d'explications sémantiques. Par exemple, le suffixe formateur de noms de propriété *-ise* (*franchise*), indisponible par ailleurs, est le seul qui puisse servir à construire de tels noms sur des adjectifs suffixés ou terminés par *-ard* compatibles sémantiquement avec la formation de noms de propriété de ce type: *flémardise* (*flémard* est suffixé), *bâtardise* (*bâlard* ne l'est pas), *°blafardise*.

(d) **Contraintes phonologiques.** Certains suffixes ne s'appliquent qu'à certaines finales, par exemple *-if* ne peut s'appliquer qu'à une base terminée par une dentale ([d], [t], [s], [z]), et provoque l'effacement de ce qui suit celle-ci si elle n'est pas finale: face à *maladif*, *sportif*, *défensif*, *allusif* (sur *allusion*), il est impossible de construire *\*fiévrif*, *\*révif*, etc.

(e) **Contraintes de disponibilité.** La mesure de la disponibilité d'un procédé dérivationnel n'est pas simple, et ne doit pas être confondue avec celle de sa productivité (cf. Corbin 1987: 176–178). La notion n'est évidemment significative qu'à l'intérieur des contraintes associées à ce procédé. Dans cette mesure, il existe:

- des procédés disponibles pour certains contextes seulement (le suffixe *-ise* sur des adjectifs en *-ard*),

- des procédés disponibles seulement pour des bases très spécifiques (*archi-* ne peut s'appliquer qu'à des bases nominales renvoyant à des fonctions sociales susceptibles de comporter des degrés de dignité symbolique ou institutionnelle; cf. Corbin 1982: 127–131),
- des procédés dont la disponibilité a évolué avec le temps (*-iel-erie*; cf. Spitzer 1931) et d'autres qui n'ont jamais (*-ile*) ou qui ont toujours (*-(e)ment* adverbial) été disponibles dans la langue.

(f) **Contraintes historiques.** Il existe enfin de rares cas où l'origine historique de la base influe sur le choix d'un affixe, parmi plusieurs concurrents: ainsi le choix de *-ain(e)* ou *-ade* pour construire des noms d'ensembles sur des bases numérales est dicté par l'origine française (*quatraine*) ou grecque (*décade*) de la base (cf. Corbin 1987: 94–95).

### 3.4. Problèmes de frontières entre les opérations dérivationnelles

La délimitation proposée entre quatre opérations structurelles de construction de mots repose sur une distinction, pour l'instant non explicitée, entre les affixes et les constituants de mots composés. Les autres frontières ne sont pas problématiques: la conversion est structurellement identifiable dès lors que les affixes dérivationnels sont distingués des affixes flexionnels (cf. 2.3.1); suffixation et préfixation se différencient structurellement par la place des affixes à droite ou à gauche de la base. En revanche, la distinction entre un mot affixé et un mot composé dépend crucialement de la théorie. On posera ici que la différence entre un affixe et un constituant de mot composé tient au rôle sémantique que le morphème joue dans le mot: les constituants d'un mot composé sont des matériaux formels et sémantiques agencés par une règle, les affixes ont, en plus, un rôle propre d'**opérateurs de construction de sens**. Un constituant de mot composé a une référence, dans la mesure où il “présuppose l'existence d'un référent conceptuel construit à partir d'universels primitifs” (Kleiber 1981: 15); un affixe ne réfère pas, il sert à formuler des opérations sur et des relations entre des référents; son rôle sémantique consiste soit à sélectionner dans le sens des bases auxquelles il peut s'appliquer les propriétés conformes à son instruction sémantique, soit à spécifier le schéma d'interprétation défini par la règle à laquelle il est associé. Dans les deux cas, une règle

construit un agencement structurel et un sens lié à cet agencement. Par exemple, on peut décrire le sens prédictible du mot composé *poisson-chat* comme le résultat de l'insertion de deux noms référant respectivement à la classe des poissons et à la classe des chats dans une structure dont l'interprétation est prédéterminée par la règle de telle sorte que le premier constituant soit l'hyperonyme et le deuxième un caractérisant (cf. 4.4 (a)). Mais le sens d'un mot affixé est plus stratifié que celui d'un mot composé, dans la mesure où l'affixe joue dans le mot le double rôle de matériau formel utilisé par la règle et d'**opérateur secondaire de construction de sens** (cf. Corbin 1991 a; 1992 b). Pour expliquer que le préfixe *trans-* puisse intervenir dans des mots au sens compositionnel aussi différent que *transalpin* ('qui se situe au-delà de ou qui traverse les Alpes') et *transporter* ('déplacer d'un lieu à un autre en portant'), il faut recourir à l'idée qu'au préfixe est associée non une référence, mais une instruction sémantique impliquant le passage d'un pôle à l'autre dans un scénario à deux pôles; quand la base est nominale (*Alpes*), elle désigne l'espace séparant ces deux pôles, quand elle est verbale (*port(er)*), elle renvoie à la modalité du passage.

Dès lors que la différence entre un constituant de mot composé et un affixe est déplacée au niveau de leur rôle sémantique, on peut expliquer que, dans certains cas, un constituant de mot composé puisse jouer le rôle d'un affixe. Par exemple *micro-* renvoie à la propriété ‘très petit’. Sa combinaison avec le nom *organisme* en fait un prédicat de ce nom, et le sens de *microorganisme* peut se formuler par la paraphrase ‘type d’organisme très petit’: *microorganisme* est un mot composé. Par contre, la combinaison de *micro-* avec un nom comme *seconde* interdit une interprétation prédicative: dans *microseconde*, *micro-* prend un rôle conventionnel d'opérateur; indiquant que l'unité de mesure ‘seconde’ doit être divisée par un million, il devient un affixe.

### 3.5. Décomptes

En fonction de la théorie adoptée, l'inventaire et le dénombrement des procédés dérivationnels de construction de mots diffèrent considérablement d'une étude à l'autre. Pour ne prendre que les procédés morphématiques, les décomptes des suffixes et des préfixes varient selon la place et le statut accordés aux différences formelles, catégorielles et sémantiques, selon qu'on établit ou non une diffé-

rence entre les affixes disponibles et ceux qui ne le sont pas, selon que l'on distingue ou non les affixes et les constituants de mots composés et selon la façon dont on les distingue. Le Tableau 121.1 présente, à titre indicatif, le nombre des suffixes recensés dans Dubois (1962) et Morvan (1988), celui des suffixes et des préfixes recensés dans Arrivé et al. (1986), Thiele (1987), Béchade (1992) et de ceux que l'état actuel de la description permet d'identifier à partir des définitions proposées ci-dessus. Les chiffres ne donnent d'ailleurs que des indications très approximatives, dans la mesure où, même si les totaux sont parfois approchants, les inventaires sont loin d'être identiques (par exemple, l'intersection entre l'inventaire des suffixes figurant dans Arrivé et al. (131) et dans Morvan (132) n'est que de 30 unités de traitement, l'intersection entre celui de Morvan (132) et celui de Danielle Corbin (134) est de 26 unités, et encore les mêmes unités ne sont-elles pas nécessairement traitées partout de la même façon).

	suffixes	préfixes
Dubois 1962	215	
Morvan 1988	132	
Arrivé et al. 1986	131	60
Thiele 1987	127	69
Béchade 1992	98	59
Danielle Corbin	134	116

Tab. 121.1: Dénombrement des suffixes et des préfixes dans quelques études

### 3.6. Combinaisons des opérations

Les quatre opérations structurelles mentionnées peuvent se combiner entre elles selon des modalités complexes qui restent à étudier en détail:

(a) La réapplication de l'opération structurale de suffixation est possible (trois suffixations successives ont produit *institu-tion(n)-al-is(er)*), et, dans une moindre mesure, celle de la préfixation (*dés-em-bouteiller*), de la composition (*lave-pare-brise*, ex. de Wagner 1980: 116) et de la conversion (le nom *orange* renvoyant à la couleur orange est issu par conversion de l'adjectif *orange*, lui-même construit par conversion sur le nom désignant un fruit). Cette dernière paraît limitée à deux applications successives (cf. Corbin & Corbin 1991: 97–100).

(b) Quand la réapplication met en jeu le même affixe, elle est soumise à des contrain-

tes fortes (cf. Rainer 1986): en français, seuls certains préfixes sont récursifs sur eux-mêmes (*re-*), aucun suffixe ne l'est (cf. Corbin 1987: 496–501).

(c) Si les contraintes propres à chaque procédé sont respectées, la combinaison d'opérations structurelles différentes est possible. Voici des exemples illustrant l'application successive de deux types différents d'opérations:

- suffixation puis conversion: *baleine<sub>N</sub>* → *baleinier<sub>A</sub>* → *baleinier<sub>N</sub>* (cf. Corbin & Corbin 1991);
- suffixation puis préfixation: *discut(er)<sub>V</sub>* → *discutable<sub>A</sub>* → *indiscutable<sub>A</sub>*;
- suffixation puis composition: *cigare<sub>N</sub>* → *cigarette<sub>N</sub>*; *fum(er)<sub>V</sub>* + *cigarette<sub>N</sub>* → *fume-cigarette<sub>N</sub>*;
- préfixation puis suffixation: *démarr(er)<sub>V</sub>* → *redémarr(er)<sub>V</sub>* → *redémarrage<sub>N</sub>*;
- préfixation puis conversion: *connu<sub>A</sub>* → *inconnu<sub>A</sub>* → *inconnu<sub>N</sub>*;
- préfixation puis composition: *pluie<sub>N</sub>* → *parapluie<sub>N</sub>*; *protég(er)<sub>V</sub>* + *parapluie<sub>N</sub>* → *protège-parapluie<sub>N</sub>*;
- conversion puis suffixation: *valide<sub>A</sub>* → *valid(er)<sub>V</sub>* → *validable<sub>A</sub>*;
- conversion puis préfixation: *gros<sub>A</sub>* → *gross(ir)<sub>V</sub>* → *regross(ir)<sub>V</sub>*;
- conversion puis composition: *vide<sub>A</sub>* → *vid(er)<sub>V</sub>*; *vid(er)<sub>V</sub>* + *poche<sub>N</sub>* → *vide-poche<sub>N</sub>*;
- composition puis suffixation: <sup>°</sup>*gastr(e)<sub>N</sub>* + <sup>°</sup>*algie<sub>N</sub>* → *gastralgie<sub>N</sub>* → *gastralgique<sub>A</sub>*;
- composition puis préfixation: *port(er)<sub>V</sub>* + *avion<sub>N</sub>* → *porte-avion<sub>N</sub>* → *anti-porte-avion<sub>A</sub>*;
- composition puis conversion: <sup>°</sup>*brachy<sub>A</sub>* + <sup>°</sup>*céphale<sub>N</sub>* → *brachycéphale<sub>A</sub>* → *brachycéphale<sub>N</sub>*.

### 4. Structures et sens des mots dérivationnellement construits

Les mots dérivationnellement construits appartiennent aux catégories lexicales majeures, tout comme les bases des mots dérivés et les constituants des mots composés. Mais d'une part toutes les combinaisons catégorielles ne sont pas possibles, d'autre part à chaque structure peuvent correspondre un ou plusieurs sens, si bien que, dans la perspective adoptée ici, les règles de construction de mots se définissent par l'association d'une structure et d'une opération sémantique, elle-même spécifiable en fonction de l'affixe et du type de base.

On tentera ici un premier recensement, que les contraintes de place rendent nécessairement schématique et non argumenté, des types de sens associés à ces structures.

#### 4.1. Suffixation

Les opérations de suffixation produisent des noms, des verbes, des adjectifs et des adverbes à partir de bases catégorisables, dans leur très grande majorité, comme nominales, verbales, adjetivales; sauf les adverbes qui ne sont constructibles que sur des bases adjetivales, toutes les autres catégories peuvent servir de bases et accueillir les dérivés. Des bases numérales (cardinales) peuvent par ailleurs être suffixées pour donner des noms et des adjectifs ordinaires, et deux pronoms (*tu* et *vous*) servent de bases à des verbes (*tutoyer*) et *vousoyer*/vouvoy(er).

Le Tableau 121.2 résume les combinaisons catégorielles possibles; la lettre en exposant renvoie à la section où sera commentée sémantiquement la combinaison signalée:

mot construit	N	A	V	ADV
base				
N	+ <sup>(a), (b)</sup>	+ <sup>(c)</sup>	+ <sup>(d)</sup>	-
A	+ <sup>(e)</sup>	+ <sup>(a)</sup>	+ <sup>(f)</sup>	+
V	+ <sup>(g)</sup>	+ <sup>(h)</sup>	+ <sup>(a)</sup>	-
NUM	+	+	-	-
PRO	-	-	+	-

Tab. 121.2: Combinations catégorielles autorisées par la suffixation en français

Les opérations de suffixation construisent des sens abstraits: sens liés au passage d'une catégorie à une autre (typiquement, la formation d'adverbes en *-(e)ment* sur des bases adjetivales, mais aussi la construction de noms dés-adjectivaux et déverbaux); sens relationnels (typiquement, la construction d'adjectifs dénominaux); sens évaluatifs. Chaque suffixe est en principe affecté à une seule structure catégorielle, sauf les évaluatifs, en raison précisément de l'instruction sémantique qui leur est associée. On traitera donc à part l'opération évaluative, avant d'aborder les autres types.

(a) Le français possède une vingtaine de suffixes qui ont en commun de servir à construire des mots impliquant une évaluation (quantitative et/ou qualitative) de ce qui est désigné par la base (entité, propriété ou pro-

cès). Ces suffixes construisent donc des mots de même catégorie que leur base. C'est d'ailleurs la seule opération sémantique associée aux structures  $[[X]_A \text{ suffixe}]_A$  et  $[[X]_V \text{ suffixe}]_V$ . Les manifestations les plus courantes (non exclusives) de l'évaluation sont la diminution et la péjoration (cf. Dal 1991): par exemple, *-et(te)* construit préférentiellement un sens 'diminutif' qui s'adapte aux diverses catégories lexicales et sémantiques (*maison*<sub>N</sub> → *maisonnette*<sub>N</sub>, *propre*<sub>A</sub> → *propreté*<sub>A</sub>, *voil(er)*<sub>V</sub> → *volet(er)*<sub>V</sub>); *-asse* sert à exprimer l'idée que les entités, propriétés ou procès désignés sont évalués négativement par rapport à la norme associée à ce que désignent les bases des mots construits (*vin*<sub>N</sub> → *vinasse*<sub>N</sub>, *fade*<sub>A</sub> → *fadasse*<sub>A</sub>, *rêv(er)*<sub>V</sub> → *rêvass(er)*<sub>V</sub>).

(b) A la structure  $[[X]_N \text{ suffixe}]_N$  est également associé un sens dit "collectif", qui correspond en fait, semble-t-il, à une représentation de la base comme constituant ou principe de constitution de ce que désigne le dérivé. Ce sens est également associé à la structure  $[[X]_{NUM} \text{ suffixe}]_{NUM}$ . Il est véhiculé, avec des nuances différentes, par plus d'une dizaine de suffixes (*colonne*<sub>N</sub> → *colonnade*<sub>N</sub>, *citron*<sub>N</sub> → *citronnade*<sub>N</sub>, *cheveu*<sub>N</sub> → *chevelure*<sub>N</sub>, *feuille*<sub>N</sub> → *feuillage*<sub>N</sub>, *quatre*<sub>NUM</sub> → *quatraine*<sub>N</sub> etc.).

D'autres sens, très marginaux, sont associés à cette structure: quelques suffixes construisent des noms renvoyant à des professionnels de ce que désigne la base (*ciné(ma)*<sub>N</sub> → *cinéaste*<sub>N</sub>) et des noms locatifs (*chien*<sub>N</sub> → *chenil*<sub>N</sub>). Enfin, cette structure est privilégiée par les suffixations techniques et scientifiques, dont on ne traitera pas ici (cf. Kocourek 1991).

(c) La structure  $[[X]_A \text{ suffixe}]_A$  est sémantiquement relationnelle: l'adjectif conserve les propriétés de sa base que sélectionne le suffixe et peut ainsi exprimer un certain type de relation entre ce qui est désigné par le nom de base et ce qui est désigné par son nom recteur (cf. Méliès-Puchulu 1991). Une trentaine de suffixes sont associés à la règle ainsi définie. Par exemple, *-ier* sélectionne les propriétés pragmatiques de sa base (cf. Corbin & Corbin 1991), *-esque* les propriétés pouvant servir à une évaluation stéréotypique (cf. Corbin et al. 1993; Temple 1993).

(d) La structure  $[[X]_V \text{ suffixe}]_V$  se caractérise sémantiquement par le fait que le nom de base peut renvoyer à la plupart des actants

associés au verbe construit: l'agent (ce qui est retenu par le verbe est le comportement typique de celui-ci: *vampiriser*), le patient (le nom désigne l'état résultant du procès: *stariser*, *momifier*), l'instrument (*javelliser*, *revolvérer*), la forme du procès lui-même (*guerroyer*, *satiriser*). Cette diversité se retrouve dans la conversion (cf. 4.2 (b)).

(e) La structure  $[[X]_A \text{ suffixe}]_N$  sert à transposer un adjectif en nom, c'est-à-dire à construire un nom de propriété dont le sens peut ensuite être soumis à dérivation sémantique. Une quinzaine de suffixes, porteurs d'une spécification et de contraintes propres plus ou moins connues (cf. Temple 1993), sont affectés à ce rôle: *-ité* (*loyauté*) et *-isme* (*loyalisme*) sont les plus disponibles; alors que *-ité* paraît avoir un spectre large, *-isme* oriente la propriété désignée vers l'extrinsèque et l'institutionnel.

(f) A la formation d'un verbe à partir d'un adjectif ( $[[X]_A \text{ suffixe}]_V$ ) est associé un sens de changement d'état ( $pur_A \rightarrow purifi(er)_V$ ). L'influence de la combinaison catégorielle sur le sens du dérivé est mise en évidence par le fait que le même sens est associé aux structures converties ( $rouge_A \rightarrow roug(ir)_V$ ) et préfixées ( $plat_A \rightarrow aplat(ir)_V$ ), même si, dans ce dernier cas, le changement d'état n'est que la forme que prend, appliquée à un adjectif, une instruction plus fondamentale de nature aspectuelle.

(g) Les noms suffixés sur une base verbale ( $[[X]_V \text{ suffixe}]_N$ ) désignent soit le procès lui-même, soit ce qui favorise sa réalisation ( $arros(er)_V \rightarrow arrosoir_N$ ); le sens abstrait est souvent soumis à dérivation sémantique, soit qu'il se concrétise (*blessure\_N*), soit qu'il renvoie à l'agent collectif (*gouvernement*) ou à l'extension spatiale (*baignade*) ou temporelle (*cueillaison*) du procès. Une quinzaine de suffixes, souvent porteurs de nuances sémantiques et liés à des contraintes spécifiques, servent à la construction de noms de procès.

(h) Enfin, les adjectifs construits sur des verbes ( $[[X]_V \text{ suffixe}]_A$ ) expriment soit la possibilité pour le procès de se réaliser ( $concev(oir)_V \rightarrow conceivable_A$ ), soit la présence stable de la propriété conditionnant le procès, d'où leur très grande facilité à devenir par conversion des noms "agentifs" ( $rêv(er)_V \rightarrow rêveur_A \rightarrow rêveur_N$ ).

#### 4.2. Conversion

La conversion consiste structurellement en une transposition catégorielle sans changement de forme. La notion de suffixe zéro,

souvent utilisée, paraît un artifice destiné à préserver l'association d'un signifiant à un signifié. Au contraire, dans la théorie adoptée ici, la justification de règles lexicales de conversion repose sur la présence dans la langue d'opérations lexicales analogues, à la structure près: la conversion partage avec la suffixation ou la composition toutes les opérations sémantiques qu'elle construit. Parmi les combinaisons catégorielles qu'elle peut autoriser (Tableau 121.3), la seule impossibilité est la construction d'adjectifs sur des bases verbales.

	mot construit	N	A	V
base				
N		—	+ <sup>(a)</sup>	+ <sup>(b)</sup>
A		+ <sup>(c), (f)</sup>	—	+ <sup>(d)</sup>
V		+ <sup>(e)</sup>	—	—

Tab. 121.3: Combinaisons catégorielles autorisées par la conversion en français

Comme la suffixation, la conversion sert à construire:

- (a) des adjectifs dénominaux (cf. 4.1 (c)), qui ne retiennent souvent que les propriétés typiques du sens de leur base ( $orange_N \rightarrow orange_A$  'qui a la couleur typique d'une orange');
- (b) des verbes dénominaux, dans lesquels les noms de base occupent des rôles sémantiques variés (cf. 4.1 (d)):  $singe_N \rightarrow sing(er)_V$ ,  $scie_N \rightarrow sci(er)_V$ ,  $coffre_N \rightarrow coffr(er)_V$ ,  $frisson_N \rightarrow frissonn(er)_V$ , etc.;
- (c) des noms désadjectivaux de propriété:  $orange_A \rightarrow orange_N$  'couleur orange' (cf. 4.1 (e));
- (d) des verbes désadjectivaux exprimant un changement d'état:  $rouge_A \rightarrow roug(ir)_V$  (cf. 4.1 (f));
- (e) La conversion partage avec la suffixation et la composition la construction de noms déverbaux processifs: masculins ( $vol(er)_V \rightarrow vol_N$ ) ou féminins (genre actuellement seul disponible:  $chass(er)_V \rightarrow chasse_N$ ) et avec la composition celle de noms agentifs ( $gard(er)_V \rightarrow garde_N$ ).
- (f) Enfin, la conversion partage avec la composition la possibilité de construire des noms dénommant des entités qui incarnent, focalisent la propriété (le prédicat) exprimé par la (les) base(s). Alors que la composition exploite cette possibilité en construisant des

noms à partir de prédicats formés par la combinaison de constituants verbaux et nominaux (cf. 4.4 (d)), la conversion “de focalisation” (cf. Corbin & Corbin 1991) le fait sur des bases adjetivales: *commode<sub>A</sub>* → *commode<sub>N</sub>* ‘(meuble) dont la propriété vue comme saillante est d’être commode’, *laitier<sub>A</sub>* → *laitier(e)<sub>N</sub>* ‘(camion, personne, vache, etc.) dont la propriété vue comme saillante est de (transporter, vendre, être utilisé pour produire, etc.) du lait’ (on trouvera dans Corbin 1991 b une représentation lexicographique expérimentale de tous les mots de forme *laitier(e)<sub>N</sub>*).

#### 4.3. Préfixation

La préfixation est conçue ici comme un processus à pouvoir catégorisateur (cf. 2.3.2). Les seules lacunes observables dans les combinaisons catégorielles autorisées (Tableau 121.4) concernent la formation de noms sur des bases adjetivales et verbales. Elles s’expliquent sémantiquement: en effet, la préfixation construit des sens plus ‘concrets’ que la suffixation ou la conversion; cela tient à l’histoire des préfixes, majoritairement issus de prépositions; or les noms désadjectivaux et déverbaux ont des sens abstraits (ils dénomment une propriété ou un procès).

mot construit	N	A	V
base			
N	+ <sup>(a)</sup>	+ <sup>(b)</sup>	+ <sup>(c)</sup>
A	—	+ <sup>(d)</sup>	+ <sup>(e)</sup>
V	—	+ <sup>(f)</sup>	+ <sup>(g)</sup>

Tab. 121.4: Combinations catégorielles autorisées par la préfixation en français

Sémantiquement, aux préfixes du français est associée la notion de repérage, qui prend les formes suivantes: localisation spatiale, temporelle, aspectuelle (et, de façon dérivée, valuation), comparaison, quantification (y compris négation), etc. De là vient que l’instruction sémantique fondamentale dont est porteur un préfixe peut se réaliser de façon différente selon la catégorie et le type sémantique de la base. Contrairement à ce qu’affirme Béchade (1992: 121), la diversité des valeurs contextuelles ne doit pas être prise pour de l’homonymie.

(a) A la structure [préfixe[X]<sub>N</sub>]<sub>N</sub> est surtout associée la localisation: par exemple l’antériorité peut être exprimée, avec des nuances

différentes, par *avant-* (*avant-scène*), *ant(éli)-* (*antichambre*), *pré-* (*préhistoire*), *pro-* (*proto-rax*), *proto-* (*protohistoire*). Mais on y rencontre aussi des préfixes à sens quantitatif (*polyclinique*, *inconfort*) ou aspectuel (*autocritique*).

(b) La structure [préfixe[X]<sub>N</sub>]<sub>A</sub>, caractérisée par la possibilité que soit ajouté à la fin du mot un **intégrateur paradigmique** (cf. 2.3.2), peut exprimer une localisation (*mer<sub>N</sub>* → *sous-marin<sub>A</sub>*), une comparaison (*morphe<sub>N</sub>* → *isomorphe<sub>A</sub>*), une quantification (*langue<sub>N</sub>* → *multilingue<sub>A</sub>*).

Bien que les mêmes préfixes puissent y figurer, les structures [préfixe[X]<sub>N</sub>]<sub>A</sub> et [préfixe[X]<sub>N</sub>]<sub>N</sub> sont sémantiquement très différentes: dans le premier cas, le sens de la base sert de repère par rapport auquel le préfixe sert à instaurer un type précis de relation (dans *surrénal*, le préfixe sert à localiser ce que désigne le nom recteur d’une façon spécifique (‘au-dessus’) par rapport à ce que dénomme *rein*); dans le second, la base dénomme la classe superordonnée dans laquelle est inclus le référent du nom préfixé, et le préfixe précise quel type de relation d’inclusion relie les deux classes (dans *survêtement*, le préfixe sert à construire le nom d’une sous-classe de vêtements caractérisés par leur localisation (‘au-dessus’) par rapport aux vêtements ‘ordinaires’). Par ailleurs, l’adjectif de structure [préfixe[X]<sub>N</sub>]<sub>A</sub> peut être converti en nom par la conversion de focalisation (cf. 4.2 (f)). La différence de structure sous-jacente des noms de structure apparente préfixe-N permet donc d’expliquer leur différence d’interprétation: la structure [préfixe[X]<sub>N</sub>]<sub>N</sub> implique que le référent du nom construit et celui du nom de base appartiennent à la même classe (ex. *antistress<sub>1</sub>*, dans le sens ‘état opposé au stress’); la structure [[préfixe[X]<sub>N</sub>]<sub>A</sub>]<sub>N</sub> qu’ils appartiennent à des catégories disjointes (ex. *antistress<sub>2</sub>*, dans le sens ‘médicament destiné à combattre le stress’).

(c) Les préfixes des verbes de structure [préfixe[X]<sub>N</sub>]<sub>V</sub> servent tous à localiser le procès par rapport à ce que dénomme la base (ex. *terre<sub>N</sub>* → *enterr(er)<sub>V</sub>* ‘mettre dans la terre’, *atterr(ir)<sub>V</sub>* ‘arriver jusqu’à la terre’).

(d) A la structure [préfixe[X]<sub>A</sub>]<sub>A</sub> est toujours associé un sens de comparaison (*équidistant*) ou de quantification (*impur*). Même les préfixes à sens localisateur prennent, dans cette structure, un sens valutatif: *aigu<sub>A</sub>* → *suraigu<sub>A</sub>*.

(e) Les verbes préfixés sur des bases adjectivales expriment, comme cela a déjà été signalé, un changement d'état. Ici encore, il s'agit souvent d'une interprétation induite par la structure, beaucoup de préfixes entrant dans cette structure ayant une instruction sémantique d'ordre aspectuel. Comme dans la suffixation et la conversion, l'adjectif de base désigne l'état résultatif ( $plat_A \rightarrow aplat(ir)_V$ ,  $large_A \rightarrow élarg(ir)_V$ ); mais le préfixe *dé(s)-* implique qu'il renvoie à l'état initial ( $niais_A \rightarrow déniais(er)_V$ ) (cf. Corbin 1992 a).

(f) La structure [préfixe[X]v]<sub>A</sub> ne peut accueillir que des bases verbales empruntées au grec ou au latin; les préfixes y ont un sens localisateur (ex. *exogène* ‘qui est issu de l’extérieur’) ou quantificateur (*multimètre* ‘qui mesure de plusieurs façons’). A la différence de référentialité du premier constituant près, cette structure est proche de la structure de mot composé [[Y]<sub>N</sub>[X]v]<sub>A</sub> (*carnivore* ‘qui mange de la viande’).

(g) Enfin, les verbes préfixés sur des verbes expriment soit un procès localisé par rapport à un autre procès (*prdire* ‘dire (un événement) avant qu’il ne se produise’), soit un procès aspectualisé (inchoatif: *s’endormir*; itératif: *redire*; perfectif: *parfaire*).

#### 4.4. Composition

Délimitées selon les critères proposés (cf. 2.2 et 3.4), les opérations de composition ne construisent en français moderne que des noms et des adjectifs: des verbes comme *porter*, *maintenir* sont des résidus (cf. Benveniste 1966). Les combinaisons autorisées sont exposées dans le Tableau 121.5.

mot construit base	N	A
N + N	+ (a)	+ (c)
A + N	+ (c)	+ (c)
A + A	—	+ (b)
V + N/N + V	+ (d)	+ (d)

Tab. 121.5: Combinaisons catégorielles autorisées par la composition en français

Le sens associé à ces structures est déterminé par l'ordre et la nature des constituants. Appelons X le constituant sémantiquement et syntaxiquement recteur du composé, Y le constituant régi. L'ordre XY correspond à une composition “non savante”: ex. *poisson-chat*, *gris-bleu*, *aide-mémoire*; l'ordre inverse

YX à une composition “savante”: ex. *anthropomorphe*, *microcéphale*, *afro-cubain*, *électro-aimant*, *calorifère*, *anthropophage*. La distinction entre les deux types de composition ne repose en effet que secondairement sur l'origine des constituants: sauf exception héritée de la langue source (*misogyne*, *philosophe*, cf. Benveniste 1967), les constituants empruntés au latin et au grec ne peuvent se composer que dans l'ordre YX, mais les constituants français peuvent se combiner dans les deux ordres (*gris-bleu* vs. *afro-cubain*). La composition “savante” s’accompagne fréquemment de la **troncation** d'un segment final du premier constituant quand celui-ci est français (*afro-cubain*) et, si le premier constituant se termine par une consonne ou [σ], de la présence d'une voyelle médiane ([i] si les constituants sont empruntés au latin, [o] dans les autres cas). Certaines structures sont limitées à la composition savante: [[Y]<sub>A</sub>[X]<sub>N</sub>]<sub>N</sub> (ex. *mégalithe* ‘grande pierre’, *électro-aimant* ‘aimant électrique’), [[Y]<sub>A</sub>[X]<sub>N</sub>]<sub>A</sub> (ex. *microcéphale* ‘qui a une petite tête’), [[Y]<sub>N</sub>[X]<sub>N</sub>]<sub>A</sub> (ex. *anthropomorphe* ‘qui a la forme d’un homme’). Les autres sont disponibles pour les deux types de composition.

Les rapides observations qui précèdent invalident la notion de “tête” telle qu’elle est définie et utilisée dans les travaux de grammaire générative. En effet, une définition “syntaxique” de la tête, qui assigne à celle-ci la responsabilité des propriétés syntaxiques du composé (catégorie, genre, nombre, arguments) (cf. Di Sciullo & Williams 1987: 24), ne s’applique qu’aux cas où le composé a la même catégorie que l’un de ses constituants: or les combinaisons N + N et V + N/N + V peuvent former des adjectifs (“singe *cynocéphale*” ‘qui a une tête de chien’, “baguette *porte-bonheur*” ‘qui porte bonheur’, “animal *carnivore*” ‘qui mange la chair’). (L’hypothèse de Di Sciullo & Williams 1987: 81 d’une tête dont les traits ne seraient pas percolés est ad hoc). Le traitement consistant à donner cette responsabilité à la règle, et non à l’un des constituants, est donc plus adéquat et peut également rendre compte des généralisations d’ordre argumental que Di Sciullo & Williams (1987: 30) attribuent à la “tête”. Quand la définition de la “tête” s’accompagne d’une contrainte de localisation à droite (Di Sciullo & Williams 1987), elle aboutit tout naturellement à vider de son contenu la catégorie des mots lexicalement composés en français, et à générer syntaxiquement la plupart des mots considérés ici comme des com-

posés lexicaux (cf. Corbin 1992 b pour une critique d'un traitement syntaxique de ces mots). Une localisation à gauche (cf. Selkirk 1982) ne serait à l'évidence pas plus adéquate.

Ces précisions étant données, passons à l'exposé des sens prédictibles associés aux structures résumées dans le Tableau 121.5.

(a) L'interprétation canonique des noms composés de type  $[N + N]_N$  correspond à une relation hyponymique, dont la spécificité est exprimée par Y, entre le nom composé et le constituant recteur X (cf. Kleiber & Tamba 1990), quel que soit l'ordre dans lequel apparaissent X et Y. Le mot composé est traditionnellement dit "endocentrique". La relation sémantique observable entre les deux constituants est très variable: Y peut renvoyer à ce à quoi s'applique X (*gastralgie* 'douleur de l'estomac'), au lieu/instrument ayant permis à X d'exister (*bébé-éprouvette* 'bébé conçu dans une éprouvette'), à la finalité de X (*pause-café* 'pause destinée au café', *hippodrome* 'piste destinée aux chevaux'), à n'importe quel objet que peut évoquer une propriété de X (*poisson-chat* 'poisson dont [les barbillons] évoquent [les moustaches] d'un chat'), etc. La forme que prend cette relation dépend en fait largement de facteurs pragmatiques, en l'occurrence des rapports qui peuvent associer, dans notre représentation du monde, les catégories référentielles auxquelles renvoient les deux noms (cf. Noailly 1990: 33 pour les interprétations liées au nom *conférence-sandwich*).

Même quand les classes référentielles des deux constituants jouent un rôle équivalent dans la représentation à laquelle renvoie le nom composé, les deux constituants ne sont pas sémantiquement sur le même plan. Que le composé soit de type savant (ex. *sociolinguistique*) ou non (ex. *robe-manteau*), l'un des deux constituants (resp. *linguistique* et *robe*) est interprété comme étant l'hyperonyme du nom composé: même si le vêtement appelé *robe-manteau* sert à la fois de robe et de manteau (cf. Noailly 1990: 80 qui le traite comme un composé par "coordination"), le genre féminin du mot et le fait que *manteau-robe* ne corresponde pas à une description identique montrent que l'ordre des constituants oriente l'interprétation.

(b) L'interprétation du type  $[A + A]_A$  est parallèle à celle du type précédent, à la catégorie près. Le constituant situé en position rectrice en fonction du type de composition est inter-

prétré comme l'hyperonyme du mot composé: *gris clair* désigne un type de gris; une musique *afro-cubaine* est une musique cubaine d'inspiration africaine; *sourd-muet* est interprété comme renvoyant à un sous-type de sourd, même si l'infirmité désignée comporte à la fois de la surdité et de la mutité.

(c) Selon qu'elle donne lieu à des noms ou à des adjectifs, la structure  $[[Y]_A[X]_N]$ , limitée à des constituants savants, reçoit des interprétations différentes. Dans la structure  $[[Y]_A[X]_N]_N$  (*mégalithe*), Y prédique X, et le nom composé est un hyponyme de X. Dans la structure  $[[Y][X]_N]_A$ , que le premier constituant soit un adjectif (*macrocéphale*<sub>A</sub>, *quadrupède*<sub>A</sub>) ou un nom (*filiforme*, *anthropomorphe*), l'adjectif composé renvoie à une propriété exprimée de façon méronymique: X représente un attribut, une partie, dont la particularité est spécifiée par Y, d'une entité représentée dans la phrase par le nom recteur de l'adjectif. Ceci explique que ces adjectifs puissent se convertir très facilement en noms, quand la propriété qu'ils expriment se focalise dans une entité au point de lui donner son nom (cf. 4.2 (f)). On retrouve donc dans un couple comme *macrocéphale*<sub>N</sub> et *macromolécule*<sub>N</sub> la différence sémantique, traitée en termes d'homonymie structurelle, déjà signalée (4.3 (b)) à propos d'*antistress*<sub>N</sub>: le nom *macromolécule* renvoie à une molécule, et a donc la structure endocentrique  $[[macro]_A[molécule]_N]_N$ ; dans le nom *macrocéphale*<sub>N</sub>, "*céphale* renvoie à une partie de l'être ainsi dénommé, et le nom (exocentrique) a alors la structure  $[[[macro]_A[céphale]]_N]_A$ .

(d) A la structure  $[[X]_V[Y]_N]_N / [[Y]_N[X]_V]_N$  sont associées deux interprétations différentes. La première, fossilisée, fait du nom composé la dénomination d'un procès particulier décrit par le prédicat XY: *lèche-vitrine* (*faire du lèche-vitrine*) renvoie au procès 'lécher les vitrines', *homicide* au procès 'tuer un homme'. La deuxième interprétation est productive et disponible; le nom composé dénomme l'agent ou l'instrument dans lequel se focalise la propriété prédiquée par XY: *porte-clé* dénomme ainsi les entités (gardiens de prison ou objets utilitaires dans le lexique codé) que peut servir à décrire et classer la propriété 'porter des clés', *anthropophage* les personnes dont la propriété vue comme descriptivement saillante est exprimable par 'manger des hommes'. On voit la parenté sémantique qu'il y a entre ce type de composition et la conver-

sion de focalisation (4.2 (f)). Il faut noter l'ambiguité de certains mots comme homicide, qui renvoient au procès et à son agent.

## 5. Conclusion

Le domaine de la morphologie lexicale est loin d'être unifié. La théorie choisie influe autant sur la délimitation du domaine que sur sa description. J'ai tenté ici, de l'intérieur d'une théorie dont l'objectif fondamental est la recherche des relations prédictibles entre la forme et le sens des mots construits, de tracer à grands traits les caractéristiques actuellement mises au jour du système dérivationnel du français. Ce travail n'est qu'une esquisse, que devrait contribuer à préciser le programme de recherche de l'U.R.A. 382 du C.N.R.S. (SILEX), dont l'un des objectifs est l'élaboration conjointe d'une grammaire et d'un dictionnaire dérivationnels du français. Je voudrais pour finir en dégager deux traits:

- J'ai tenté de fonder la délimitation du domaine sur des critères explicites, et, notamment, d'introduire deux moyens de dessiner plus finement les frontières traditionnellement floues de la composition: en différenciant la notion d'unité lexicale complexe et celle d'unité lexicale construite dérivationnellement (2.2) et en dissociant l'identité lexicale d'un morphème, définie formellement et sémantiquement, de la fonction qu'il peut occuper dans un mot (3.4).
- La typologie des sens associés aux structures de mots construits proposée en 4 est, à ma connaissance, la première tentative de ce genre. Dans cette mesure, elle reste provisoire et imparfaite. Elle met pourtant en évidence deux faits essentiels: (i) les sens des mots construits dépendent autant de leur structure que des matériaux qui y figurent; (ii) certains sens des mots construits mettent en jeu des opérations cognitives à l'œuvre également hors du domaine de la dérivation: par exemple, le procédé de dénomination méronymique, qui apparaît dans la conversion de focalisation (4.2 (f)) et dans la composition combinant un constituant verbal et un constituant nominal (4.4 (d)), est aussi un procédé sémantique indépendant servant à dériver, par exemple, le sens 'objet doté d'une extrémité pointue' du sens 'extrémité pointue' de *pointe*. Il y a toutefois des différences, par exemple dans le type de

propriétés sélectionnables, dans la mise en œuvre de ces opérations cognitives selon les domaines linguistiques. Un champ spécifique de recherche se trouve ainsi cerné: la sémantique des mots construits, au confluent de la morphologie et de la sémantique lexicale.

## 6. Références

- Arrivé, Michel & Gadet, Françoise & Galmiche, Michel (1986), *La grammaire d'aujourd'hui: Guide alphabétique de linguistique française*. Paris: Flammarion
- Béchade, Hervé D. (1992), *Phonétique et morphologie du français moderne et contemporain*. Paris: Presses Universitaires de France
- Benveniste, Emile (1966), "Convergences typologiques". *L'Homme* 6.2, 5–12
- Benveniste, Emile (1967), "Fondements syntaxiques de la composition nominale". *Bulletin de la Société de Linguistique de Paris* 62.1, 15–31
- Booij, Geert E. (1977), *Dutch Morphology: A Study of Word Formation in Generative Grammar*. Lisse: The Peter de Ridder Press
- Corbin, Danielle (1980), "Contradictions et inadéquations de l'analyse parasyntétique en morphologie dérivationnelle". In: Dessaux-Berthonneau, Anne-Marie (éd.), *Théories linguistiques et traditions grammaticales*. Villeneuve d'Ascq: Presses Universitaires de Lille, 181–224
- Corbin, Danielle (1985), "Les bases non autonomes en français ou comment intégrer l'exception dans un modèle lexical". *Langue française* 66, 54–76
- Corbin, Danielle (1987), *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer (Linguistische Arbeiten 193–194) [= 1991 Villeneuve d'Ascq: Presses Universitaires de Lille]
- Corbin, Danielle (1989 a), "Form, Structure and Meaning of Constructed Words in an Associative and Stratified Lexical Component". *Yearbook of Morphology* 2, 31–54
- Corbin, Danielle (1989 b), "Contraintes et création lexicales en français". *L'information grammaticale* 42, 35–43
- Corbin, Danielle (1991 a), "Introduction: La formation des mots, structures et interprétations". *Lexique* 10, 7–30
- Corbin, Danielle (1991 b), "La morphologie lexicale, bilan et perspective". *Travaux de linguistique* 23 [Actes du colloque de lexicologie française, Louvain-la-Neuve, 21–23 mars 1991], 33–56

- Corbin, Danielle (1992 a), "Sens et définition: De la compositionnalité du sens des mots construits (Réponse à Claire Vanderhoeft)". *Lingvisticae Investigationes* 16.1, 189–218
- Corbin, Danielle (1992 b), "Hypothèses sur les frontières de la composition nominale". *Cahiers de grammaire* 17, 26–55
- Corbin, Danielle & Corbin, Pierre (1991), "Un traitement unifié du suffixe -ier(e)". *Lexique* 10, 61–145
- Corbin, Danielle & Dal, Georgette & Mélis-Puchulu, Agnès & Temple, Martine (1993), "D'où viennent les sens *a priori* figurés des mots construits? Variations sur *lunette(s)*, *ébéniste* et les adjectifs en *-esque*". *Verbum* 1–2–3
- Corbin, Danielle & Plénat, Marc (1992), "Note sur l'haplogologie des mots construits". *Langue française* 96, 101–112
- Corbin, Pierre (1982), "Le monde étrange des dictionnaires (3): La faisselle et autres contes, scolies sur le changement lexical". *Modèles linguistiques* IV.1, 125–184
- Dal, Georgette (1991), "Hyponymie et prototypie: Les noms en *-asse* et *-et(te)* du français". *Lexique* 10, 211–239
- Darmesteter, Arsène (?1893), *Traité de la formation des mots composés dans la langue française comparée aux autres langues romanes et au latin*, Revu, corrigé et en partie refondu par Gaston Paris. Paris: Champion
- Dell, François & Selkirk, Elisabeth O. (1978), "On a Morphologically Governed Vowel Alternation in French". In: Keyser, Samuel J. (éd.), *Recent Transformational Studies in European Linguistics*. Cambridge/MA: The MIT Press, 1–51
- Dubois, Jean (1962), *Etude sur la dérivation suffixale en français moderne et contemporain*. Paris: Larousse
- Dubois, Jean (1965), *Grammaire structurale du français: Nom et pronom*. Paris: Larousse
- Dubois, Jean (1967), *Grammaire structurale du français: Le verbe*. Paris: Larousse
- Eynde, Karel van den & Blanche-Benveniste, Claire (1970), "Essai d'analyse de la morphologie du verbe français : Présentation d'hypothèses de travail". *Orbis* 1970.2, 404–429
- Félice, Théodore de (1950), *Eléments de grammaire morphologique*. Paris: Didier
- Foley, James (1979), *Theoretical Morphology of the French Verb*. Amsterdam: Benjamins (Lingvisticae Investigationes: Supplementa 1)
- Grésillon, Almuth (1983), "Mi-fugue, mi-raison. Dévaliser des mots-valises". *DRLAV* 29, 83–107
- Gross, Gaston (1990), "Définition des noms composés dans un lexique-grammaire". *Langue française* 87, 84–97
- Guilbert, Louis (1971), "De la formation des unités lexicales". In: *Grand Larousse de la langue française en sept volumes (1971–1978)*. Paris: Larousse, IX–LXXXI
- Guiraud, Pierre (1967), *Structures étymologiques du lexique français*. Paris: Larousse
- Holtus, Günter & Metzeltin, Michael & Schmitt, Christian (1990, éds.), *Lexikon der Romanistischen Linguistik V, 1. Französisch. Le français*. Tübingen: Niemeyer
- Hunnius, Klaus (1990), "Französisch: Flexionslehre. Flexion". In: Holtus et al. (éds.), 59–71
- Isaac, Luc (1985), *Calcul de la flexion verbale en français contemporain*. Genève: Droz
- Kerleroux, Françoise (1990), "Du mode d'existence de l'infinitif substantivé en français contemporain". *Cahiers de grammaire* 15, 55–99
- Kleiber, Georges (1981), *Problèmes de référence: Descriptions définies et noms propres*. Paris: Klincksieck (Recherches linguistiques: Etudes publiées par le Centre d'Analyse Syntaxique de l'Université de Metz VI)
- Kleiber, Georges & Tamba, Irène (1990), "L'hyponymie revisitée: Inclusion et hiérarchie". *Langages* 98, 7–32
- Kocourek, Rostislav (?1991), *La langue française de la technique et de la science: Vers une linguistique de la langue savante* [2e éd. augmentée, refondue et mise à jour avec une nouvelle bibliographie]. Wiesbaden: Brandstetter [?1982]
- Martinet, André (1969), "De l'économie des formes du verbe en français parlé". In: Martinet, André, *Le français sans fard*. Paris: Presses Universitaires de France, 91–120
- Mélis-Puchulu, Agnès (1991), "Les adjectifs démoninaux: Des adjectifs de 'relation'". *Lexique* 10, 33–60
- Mok, Q. I. M. (1968), *Contribution à l'étude des catégories morphologiques du genre et du nombre dans le français parlé actuel*. The Hague: Mouton
- Mok, Q. I. M. (1990), "Französisch: Morphosyntax. Morphosyntax". In: Holtus et al. (éds.), 112–125
- Morvan, Danièle (1988), "Petit dictionnaire des suffixes du français". In: *Le Micro Robert: Dictionnaire d'apprentissage de la langue française*, nouv. éd. entièrement revue et mise à jour. Paris: Dictionnaires Le Robert, V–XIX
- Noailly, Michèle (1990), *Le substantif épithète*. Paris: Presses Universitaires de France

- Nyrop, Kristoffer (²1924), *Grammaire historique de la langue française, Vol. II: Morphologie*. Copenhague: Gyldendal
- Nyrop, Kristoffer (²1936), *Grammaire historique de la langue française, Vol. III: Formation des mots* [deuxième édition revue par Kristian Sandfeld]. Copenhague: Gyldendalske Boghandel Nordisk Forlag [¹1908]
- Pichon, Edouard (²1942), *L'enrichissement lexical dans le français d'aujourd'hui: Les principes de la suffixation en français*. Paris: D'Artrey [= ¹1935–1940, *Le français moderne* 1935.3, 209–223; 1935.4, 322–344; 1938.3, 299–304; 1939.1, 7–14; 1939.2, 133–149; 1939.3, 229–238; 1939.4, 317–328; 1940.1, 27–35; 1940.2, 121–130]
- Pinchon, Jacqueline (1986), *Morphosyntaxe du français: Etude de cas*. Paris: Hachette
- Pinchon, Jacqueline & Couté, Bernard (1981), *Le système verbal du français: Description et applications pédagogiques*. Paris: Nathan
- Plénat, Marc (1981), *L'“autre” conjugaison (ou: De la régularité des verbes irréguliers)*. Toulouse: Université de Toulouse-Le Mirail (Cahiers de grammaire 3)
- Plénat, Marc (1984), “Toto, Fanfa, Totor et même Guiguitte sont des anars”. In: Dell, François & Hirst, Daniel & Vergnaud, Jean-Roger (1984, éds.), *Forme sonore du langage: Structure des représentations en phonologie*. Paris: Hermann, 161–181
- Plénat, Marc (1985), “Morphologie du Largonji des loucherbems”. *Langages* 78, 73–122
- Plénat, Marc (1991 a), “Le javanais: concurrence et haplogologie”. *Langages* 101, 95–117
- Plénat, Marc (1991 b), “Commentaire sur la contribution de Danielle Corbin”. *Travaux de linguistique* 23 [Actes du colloque de lexicologie française, Louvain-la-Neuve, 21–23 mars 1991], 63–65
- Rainer, Franz (1986), “Recursiveness in Word-Formation, with Special Regard to Spanish”. *Acta linguistica Academiae Scientiarum Hungaricae* 36, 197–209
- Riegel, Martin (1985), *L'adjectif attribut*. Paris: Presses Universitaires de France
- Scalise, Sergio (1984), *Generative Morphology*. Dordrecht, Cinnaminson/NJ: Foris (Studies in Generative Grammar 18)
- Schane, Sanford A. (1968), *French Phonology and Morphology*. Cambridge/MA: The MIT Press
- Sciullo, Anna-Maria Di & Williams, Edwin (1987), *On the Definition of Word*. Cambridge/MA: The MIT Press
- Selkirk, Elisabeth O. (1982), *The Syntax of Words*. Cambridge/MA: The MIT Press
- Spitzer, Leo (1931), “Warum ersetzt frz. -erie (dtsch. -erei) das alte -ie (-ei)?”. *Zeitschrift für romanische Philologie* 51, 70–75
- Swiggers, Pierre & Eynde, Karel van den (1987, éds.), *La morphologie du verbe français*. Louvain (ITL. Review of Applied Linguistics 77–78)
- Temple, Martine (1993), *Le sens des mots construits: Pour un traitement dérivationnel associatif*. Thèse de doctorat, Université de Lille III
- Thiele, Johannes (1987), *La formation des mots en français moderne* [traduction et adaptation par André Clas]. Montréal: Presses de l'Université de Montréal
- Wagner, Robert-Léon (1980), “Les mots construits en français”. In: Wagner, Robert-Léon, *Essais de linguistique française*. Paris: Nathan, 111–122 [= ¹1969, *Bulletin de la Société de Linguistique de Paris* LXIII.I, 63–82]
- Zwanenburg, Wiecher (1983), *Productivité morphologique et emprunt*. Amsterdam, Philadelphia: Benjamins (Lingvisticae Investigationes: Supplementa 10)
- Zwanenburg, Wiecher (1990), “Französisch: Wortbildungsllehre. Formation des mots”. In: Holtus et al. (éds.), 72–77

*Danielle Corbin †, Lille (France)*

## 122. Russisch (Indogermanisch: Slawisch)

1. Allgemeines zur russischen Sprache
2. Der morphologische Grundcharakter des Russischen
3. Die Wortarten als morphologisch ausgewiesene Größen
4. Grammatische Morphologie
5. Wortbildungsmorphologie
6. Grammatische Morphologie und Wortbildung in der Literatur
7. Zitierte Literatur

### 1. Allgemeines zur russischen Sprache

Die russische Sprache wird in den europäischen und asiatischen Teilen der GUS von ca. 273 Millionen Sprechern als Muttersprache und/oder Koiné benutzt. Darüber hinaus besitzt das Russische nach wie vor einen hohen funktionalen Wert als Zweitsprache in den baltischen Republiken, Aserbaidschan und Georgien. Damit gehört diese Sprache zu den wenigen, deren Sprecherkontingent die 250 Millionen-Grenze überschreitet. Die russische Ethnospalte umfaßt die Dialektgruppen Nord-, Süd- und Mittelgroßrussisch. Letzteres ist ein Übergangssidiom der beiden ersten und bildet die Basis für die *Standardsprache*. Außer der überregionalen Standardsprache, dem *literaturnyj jazyk*, findet deren umgangssprachliche Form, die *razgovornaja reč'*, und eine Unterschichtenvariante, das *prostorečie*, breite Verwendung. Mit weiteren dia- und soziolakalen Varianten ist in den Gebieten mit russischer Minorität zu rechnen.

Der Beginn des russischen Schrifttums fällt zusammen mit der Christianisierung Rußlands durch Byzanz Ende des 10. Jahrhunderts. Dieses ist über Jahrhunderte stark vom *Kirchenslawischen*, einer südslawischen, von Byzanz zu Missionierungszwecken verwendeten Sprache, geprägt worden. Erst seit dem 18. Jahrhundert kam es zur Entwicklung einer russischen Schriftsprache im eigentlichen Sinn. Die ersten kodifizierenden Grammatiken (Adodurov, Lomonosov) des Russischen erschienen um die Mitte des 18. Jahrhunderts. In ihnen sind auch die ersten (präskriptiven) morphologischen Beschreibungen enthalten (vgl. 6). Von einer voll ausgebildeten und systematisch distribuierten russischen Standardsprache kann seit der 1. Hälfte des 20. Jahrhunderts gesprochen werden. Auf sie beziehen sich die folgenden

Ausführungen zur Morphologie. Dabei wird entsprechend der hier verfolgten Mitteilungsintention von der phonemisch-graphemischen Erscheinungsform des Russischen, nicht von der phonetischen, ausgegangen.

### 2. Der morphologische Grundcharakter des Russischen

Das Russische als Vertreter der indogermanischen Untergruppe des Slawischen hat ungeachtet bestimmter Einzelveränderungen in seiner morphologischen Struktur deren Grundcharakter seit ältester Zeit bewahrt und entspricht in vielen Hinsichten den ausgeprägt flektierenden Strukturen altslawischer Sprachzustände. Das bedeutet, grammatische, grammato-syntaktische und lexikalische Informationen werden in starkem Maße mit Hilfe synthetischer Funktionsmorpheme übermittelt. Sowohl im Flexions- als auch im Wortbildungsbereich finden Präfixe, Suffixe, Zirkumfixe, Infixe und Postfixe intensive Verwendung. Die Affigierungsprozesse sind häufig von morphonologischen Veränderungen des Stammes (russ. *čeredovanie*) begleitet (*peč'* 'backen' / *pěk-Ø* 'er backte', <ë> = [̄]; *gna-t'* 'jagen' / *gon-jú* 'ich jage'; *bog-Ø* 'Gott' / *bóž-ij* 'göttlich'). Ebenso können sie mit einem Wechsel der Akzentstelle gekoppelt sein (*okn-ó* 'das Fenster' / *ókn-a* 'die Fenster'; *ži-t'* 'leben' / *žit'-é* 'Existenz, Lebensart').

Bei der grammatischen Beschreibung des Russischen wird meist der **synkretistische** Charakter der grammatischen Morpheme herausgestellt. Es gibt jedoch auch zahlreiche **agglutinative** Funktionsmorpheme, insbesondere im Bereich bestimmter Kategorien des Verbs und des Adjektivs (s. 4.1, 4.4). Einige von ihnen fusionieren mit dem Stamm (*dórog-o* 'teuer' (Adverb), *doróž-e* 'teurer'). Die Menge der **morphologisch determinierten** Allomorphe überwiegt. Sie wird freilich durch eine erhebliche Anzahl von **phonologisch determinierten** ergänzt. So gehören beispielsweise die Substantive *učenik-Ø* 'Schüler' und *stol-Ø* 'Tisch' der gleichen Deklinationsklasse an (s. 4.2), dennoch treten im Nominativ Plural unterschiedliche Formen auf, da die auslautenden Konsonantenphoneme /-k/ bzw. /-l/ divergierenden phonotaktischen Regeln unterworfen sind: *učenik-i* : *stol-y*. Die

Akzentstelle kann ebenfalls allomorphiestiftend sein. Die Verben *peč'* ‘backen’ und *choté-t'* ‘wollen’ gehören der gleichen Konjugationsklasse an (s. 4.1), dennoch treten in mehreren Positionen des Präsensparadigmas unterschiedliche Allomorphe in Abhängigkeit von deren Akzentuierung/Nichtakzentuierung auf: *peč-eš'* ‘du bäckst’ : *chóč-eš'* ‘du willst’.

Hinsichtlich der Repräsentation syntaktischer Relationen muß das Russische als **synthetisch-analytische Sprache** eingeordnet werden, da diese sowohl durch Flexive als auch durch Präpositionen ausgedrückt werden: *Nín-a čítá-et gazét-u i p'ět čaj-∅ s limón-om* ‘Nina liest Zeitung und trinkt Tee mit Zitrone’ (*s limón-om* = präpositionaler Ausdruck von ‘mit Zitrone’). Wird eine semantosyntaktische Relation durch eine Präposition ausgedrückt, so tritt diese in Verbindung mit einem durch Rektionsregeln determinierten (meist redundanten) Flexiv auf (*iz komnat-y* ‘aus dem Zimmer’ (mit Genitivflexiv)).

### 3. Die Wortarten als morphologisch ausgewiesene Größen

Die **Wortarten** sind im Russischen eindeutig als solche durch morphologisierte grammatische Bedeutungen gekennzeichnet. Mit dem finiten **Verbum** verbinden sich die Kategorien Aspekt, Tempus, Modus und Genus verbi. Ferner reflektieren sich am Verbum semantische und grammatische Eigenschaften des Subjektaktaanten in Gestalt von Numerus, Person und Genus. Eine absolut kategorienneutrale Form besitzt das russische Verbum nicht. Der Infinitiv schließt stets die Kategorie des Aspektes ein (s. 4.1).

Das **Substantiv** trägt die Kategorien Numerus, Kasus und Genus. Bei letzterem handelt es sich um eine Art von immanentem Substantivklassifikator, der keinesfalls mit der lexikalischen Kategorie des Sexus gleichgesetzt werden darf (vgl. *muščin-a* ‘Mann’ (wie femininum dekliniert)). Die sog. “Kategorie der Belebtheit” drückt sich in der Verwendung des Genitiv- statt des Nominativflexivs in der Position des Objektkasus bei Substantiven mit dem Sem [+ belebt] aus. Diese “Kategorie” erfaßt nicht alle Deklinationsklassen und ist meist funktionslos, da mit ihr die bereits im Basismorphem gegebene Information [+ belebt] resp. [- belebt] lediglich wiederholt wird (*Anna ljúb-it brát-a* (GEN) ‘Anna liebt ihren Bruder’ / *Anna ljúb-it*

*teátr-∅* (NOM) ‘Anna liebt das Theater’; s. aber auch 4).

Die **Pronomen** als substitutive und/oder deiktische Wörter sind je nach Zahl und Art der mit ihnen verbundenen Kategorien in unterschiedliche Subtypen einzuteilen. Sie können bis zu vier echte Kategorien tragen und somit in hohem Maße synkretistisch sein. Beispielsweise verfügen die anaphorischen Personalpronomina im Singular über die Kategorien Person, Numerus, Genus und Kasus (*on-á* ‘3.SG.F.NOM (sie)’). Andere hingegen verfügen über lediglich eine Kategorie wie das nur den Kasus tragende Reflexivpronomen *seb'á* ‘mich, dich, sich, uns ...’ (AKK). Neben diesen Extremgruppen gibt es solche mit Zwischenwerten. Dieses Bild wird noch dadurch verkompliziert, daß sich mehrere Pronomengruppen ähnlich wie attributive Adjektive (s. 4.4) verhalten und durch Kongruenzregeln gesteuerte Kategorialausdrücke aufweisen. Beispielsweise reflektieren die Possessivpronomina in der 1. und 2. Person die kategorialen Eigenschaften des mit ihnen attribuierten Substantivs (*mo-já šljáp-a* ‘mein Hut’ (NOM SG F)). Mit dem Pronomen begegnet uns die russische Wortart, die sich hinsichtlich ihres Kategorienapparates am heterogensten verhält. Ihre Definition als Wortart kann folglich nicht konsequent auf dem Hintergrund der sich mit ihr verbindenden Kategorien erfolgen (die zudem weitgehend mit denen des Substantivs übereinstimmen). Vielmehr muß als grundlegendes Definitions criterium ihre oben genannte sprachliche Funktion dienen.

Das **Adjektiv** reflektiert als **attributive Einheit** alle kategorialen Eigenschaften des Substantivs und verfügt dazu über eigenständige Formenparadigmen. Dem Adjektiv als **prädikativer Einheit** fehlt logischerweise der Kasus – nicht aber das Numerus-Genus-Paradigma. Als echte adjektivische Kategorie besitzt das Russische nur die Komparation, mit der jedoch aus logisch-semantischen Gründen nicht alle Adjektive zu verbinden sind.

Das **Adverb** ist – wenn von einem Adjektiv abgeleitet – als solches morphologisch markiert. Einem wortartenimmanenteren Paradigma unterwerfen sich nur jene Adverbien, die aus semantischen Gründen der Komparation unterzogen werden können. Eine große Menge von Adverbien ist formal nicht eigens als Adverb gekennzeichnet (*óčen'* ‘sehr’, *počtí* ‘fast’, *závtra* ‘morgen’) und morphologisch stabil. Diese Stabilität erklärt sich aus der Kategorienlosigkeit dieser Adverbien.

#### 4. Grammatische Morphologie

Die **grammatischen Morpheme** im Russischen können je nach ihrer Funktion in drei Klassen eingeteilt werden:

- (a) in Morpheme, mit deren Hilfe auf 1. referentielle und 2. referentiell-relationale Gegebenheiten Bezug genommen wird (*Proféssor-Ø čítá-l-Ø lékci-ju* ‘Der Professor hält eine Vorlesung’). In diesem Satz signalisiert das Numerus-Kasus-Morphem -Ø an *proféssor*, daß mit dem gegebenen Lexem auf einen einzelnen Referenten abgehoben wird, und daß dieses für eine Subjekt-Agens-Funktion steht. Das Morphem -l- am Verb *čítá-* informiert über die Vergangenheit der Handlung, und das Morphem -ju in *lékciju* wiederum über die Singularität des Referenten und die Objekt-Patiens-Funktion des Lexems);
- (b) in Morpheme, die im phrasischen und/oder transphrasischen Rahmen formale Kohärenz erzeugen (*Lež-át na stol-é al'bóm-Ø i knížk-a. Oná sovsem pýl'n-aja* ‘Auf dem Tisch liegen ein Album und ein Buch. Es (das Buch) ist ganz staubig’). Im zweiten Satz wird mit dem Genusmorphem -a am Pronomen *oná* ein eindeutiger Bezug auf das Lexem *knižka* (beide Male femininum) und zugleich eine formale Kohärenz zwischen beiden Sätzen hergestellt);
- (c) in Morpheme, die weder Information tragen noch Kohärenz herstellen, also redundant sind (vgl. zur “Kategorie der Belebtheit” unter 3 und zur Verwendung von Kasusmorphemen bei gleichzeitigem Einsatz von Präpositionen unter 2. Ob ein grammatisches Morphem redundant ist oder nicht, entscheidet der Kontext. So kann z. B. die “Kategorie der Belebtheit” (s. 3) dann Informativität gewinnen, wenn sie in einem Satz mit zwei Akzenten, die von Lexemen mit dem Sem [+ männlich] und dem Genusmerkmal ‘M’ besetzt sind, die Agens- bzw. Patiensfunktion anzeigt: *Car'-Ø nenavíd-it sýn-a* ‘Der Zar haßt seinen Sohn’).

Das in vielen Grammatiken als eigene Wortart geführte **Numerals** kann unter dem Aspekt seines grammatisch-morphologischen Verhaltens kaum als eine solche gelten. Alle Zahlwörter verbinden sich mit den gleichen Kategorien wie die Nomen und entsprechen in formaler Hinsicht entweder dem Substantiv (s. 4.2; *pjat'-Ø* (F) ‘fünf’, *týsjač-a* (F) ‘tausend’) oder dem attributiven Adjektiv (s. 4.4; *odin-Ø* (NOM.M) ‘einer’, *odn-ogó* (GEN), *odn-ómu* (DAT)

etc.; teils mit modifizierten Flexiven: *tri* ‘drei’, *tréch* (GEN), *trém* (DAT) etc.). Einige Zahlwörter haben ein substantivisches Sonderparadigma mit extremer Flexivhomonymie (*st-o* ‘hundert’, alle anderen Kasus *st-a*).

Die Paradigmen aller Wortarten ergeben sich aus den Interrelationen zwischen den Flexiven (wobei Homonymie nicht selten ist). Grammatischer **Suppletivismus** kommt nur vereinzelt vor (z. B. Aspekttoppositionen: *pojmá-t'* (PFV) / *loví-t'* (IPFV) ‘fangen’; Komparationsoppositionen: *plóch-o* (Positiv) ‘schlecht’ / *chúze* (Komparativ) ‘schlechter’ u. ä.).

Es werden nun die morphologischen Strukturen der unterschiedlichen Wortarten näher erläutert. Dabei muß auf die Darstellung der funktionalen Eigenschaften der Kategorien wegen deren Komplexität weitgehend verzichtet werden. Die Skizze konzentriert sich also auf deren formale Seite. Die Beschreibungen werden nach synchronstrukturalistischen Prinzipien durchgeführt.

##### 4.1. Die Morphologie des Verbums

Die dem russischen Verb eigenen grammatischen Kategorien wurden bereits unter 3. angeführt. Unter ihnen ist die binäre Kategorie des **Aspektes** besonders markant. Sie dichotomisiert in funktionaler Hinsicht das gesamte Verbalparadigma. Mit dem Aspekt werden Prozesse und Resultate der Verbalinhalte näher charakterisiert. Der semantischen Opposition **perfektives : imperfektives Aspektglied** entsprechen die formalen Aspektkorrelate. Diese werden im wesentlichen durch folgende Verfahren abgebildet: durch suffixalische Opposition (*reš-i-t'* (PFV) / *reš-á-t'* (IPFV) ‘entscheiden’), durch präfixalische Opposition (*s-déla-t'* (PFV) / *Ø-déla-t'* (PFV) ‘tun’), durch infixalische Opposition (*pri-nes-tí* (PFV) / *pri-nosi-t'* (IPFV) ‘bringen’) und durch Suppletivopposition (*vzja-t'* (PFV) / *bra-t'* (IPFV) ‘nehmen’). Die Komplexität der formalen Grundverfahren wird durch die Zahl der individuellen Formative noch stark erhöht. Auch unterscheiden sich die Korrelate häufig durch morphonologische Varianz der Stämme (*zašít-i-t'* (PFV) / *zašítš-á-t'* (IPFV) ‘verteidigen’). Der Aspekt hat von allen Verbkatoren die formal heterogenste Struktur. Er erfaßt ca. 70 Prozent des Verbkorpus. Verben, die nur in einer Aspektbedeutung auftreten, gelten als aspektdefektiv.

Das russische **Tempus** verfügt über eine dreigliedrige Struktur, es umfaßt die Grammeme Präsens, Präteritum und Futur. Wäh-

rend das Präsens aus semantischen Gründen nur mit dem imperfektiven Aspekt kompatibel ist, können sich die beiden letzteren Tempus-Grammeme mit beiden Aspektkorrelaten verbinden. Die formale Repräsentation der Teiltempora erfolgt mit unterschiedlichen Mitteln. Das **Präsens** wird nach synkretistischem Prinzip ausgedrückt, d. h. das es repräsentierende Morphem trägt gleichzeitig die Grammeme von Person und Numerus (s. 4.2). Da im Russischen drei Personen und zwei Numeri unterschieden werden, ergeben sich sechs konkrete Möglichkeiten, das Präsens auszudrücken:

<i>ja čítá-ju</i> ‘ich lese’	<i>my čítá-em</i> ‘wir lesen’
<i>ty čítá-eš</i> ‘du liest’	<i>vy čítá-ete</i> ‘ihr lest’
<i>on čítá-et</i> ‘er liest’	<i>oní čítá-jut</i> ‘sie lesen’

Tab. 122.1: Russisches Präsensparadigma

Weitere Ausdrucksmöglichkeiten des Präsens liegen in Form von Partizipien und des Adverbialpartizips vor (s. 5.2, 4.5). Letzteres repräsentiert auch die Bedeutung der Gleichzeitigkeit (*On razgovárivaet s nej, pominútno pogládyva-ja na časy* ‘Er unterhält sich mit ihr, ständig auf die Uhr blickend’; *pogládyvaja* = Adverbialpartizip ‘blickend’).

Das Präteritum wird nach agglutinierendem Prinzip gebildet, d. h. das entsprechende Morphem trägt exklusiv die Bedeutung ‘PRÄT’. Es wird überwiegend durch das Morph *-l-*, in bestimmten Fällen auch durch *-o* realisiert (*oní čítá-l-i* ‘sie lasen’, *oná čítá-l-a* ‘sie las’; *on pogib-∅* ‘er kam um’ (zu *pogibnu-t*), aber: *oná pogib-l-a* ‘sie kam um’). Das Präteritum verbündet sich mit dem Numerus und dem Genus (s. 4.2). Weitere Ausdrucksformen präteritaler Zeiteinordnungen liegen in Form von Partizipien und des Adverbialpartizips vor (s. 5.2, 4.5). Letzteres repräsentiert auch die Bedeutung der Vorzeitigkeit (*Okónči-v konservatóriju, oná uéchala v Berlín* ‘Nachdem sie das Konservatorium abgeschlossen hatte, reiste sie nach Berlin’; *okónči-v* = Adverbialpartizip ‘nachdem (sie) abgeschlossen hatte’).

Das **Futur** wird in zweierlei Weise gebildet. Zur Formation des **perfektiven Futurs** wird das perfektive Aspektkorrelat benutzt, seine Konjugation entspricht der des Präsens (s. Tab. 122.1). Zur Formation des **imperfektiven Futurs** wird das imperfektive Aspektkorrelat benutzt. Es verbündet sich in Form des Infinitivs mit dem analytischen Futurindikator

*bud-*, der in der Art finiter Verben der präsens-tischen Konjugation unterworfen ist (*Závtra ja búd-u čítát* ‘Morgen werde ich lesen’; *Izvinjájt'sja ne bud-ut* ‘Sie werden sich nicht entschuldigen’).

Die russischen Verben können aufgrund unterschiedlichen formalen Verhaltens zwei Konjugationsklassen zugeordnet werden, der *e-* (s. Tab. 122.1) und der *i*-Konjugation. Diese umfassen ihrerseits weitere Subklassen, die sich auf der Basis der formalen Relation von Infinitiv- und Präsensstamm definieren (s. Švedova 1982, Hrsg.: 646–664; Kempgen 1989: 495–514). Da diese Relation durch eine hohe Variabilität gekennzeichnet ist, besitzt das russische Verb einen äußerst umfangreichen allomorphen Apparat, der besonders evident im Bereich der Stämme wird: *pe-č* ‘backen’, *pek-ú* ‘ich backe’, *peč-óš* ‘du bäägst’; *povys-í-t'* ‘erhöhen’, *povýš-u* ‘ich erhöhe’, *povýš-it* ‘er erhöht’. Allerdings ist nur eine Teilmenge dieser Subklassen produktiv. Zur Vielfalt des Formenapparates trägt auch der Infinitiv bei, der außer durch das dominierende Suffix *-t'* durch die Suffixe *-tí* und *-č'* ausgewiesen sein kann (*kur-í-t'* ‘rauchen’; *nes-tí* ‘tragen’; *mo-č'* ‘können’) (s. auch Isačenko 1968: 220–271).

Nach Auffassung der traditionellen formenorientierten Sprachbeschreibung verfügt die Kategorie des **Modus** im Russischen über drei formal markierte Grammeme: den Indikativ, den Konjunktiv und den Imperativ. Der **Indikativ** wurde bereits unter der Kategorie Tempus implizit beschrieben. Die dort angeführten Formen sind zugleich indikative Formen. Im Unterschied zu ihm bildet der **Konjunktiv** keine Zeitstufenformen. Es gibt nur ein zeitindifferentes Paradigma, das wie das des Präteritums gebildet wird unter jeweiliger Hinzufügung der Modalpartikel *by* bzw. *b.* Der Zeitbezug muß kontextuell definiert werden: *Iván čítá-l-∅ by segódňa* ‘Ivan würde heute lesen’; *Studénty otdochnú-l-i by včerá* ‘Die Studenten hätten sich gestern erholt’.

Der Imperativ kann sich auf die 2. Person Singular und Plural, auf die 2. Person Plural mit Ego-Inklusion und auf die 3. Person Singular und Plural beziehen. In der 2. Person erfolgt die Markierung durch agglutinative Suffixe, im geringeren Maße durch imperativische Partikeln (*piš-!* ‘schreib!’, *piš-i-te!* ‘schreibt!’, *napiš-em!* ‘wollen wir schreiben!’ (Intimform zwischen zwei Personen), *dav-áj búd-em pisát!*! (dasselbe mit dem imperfektiven Aspekt), *napiš-em-te!* ‘wollen wir schrei-

ben!' (Offizialform oder für mehr als einen Adressaten). In der 3. Person werden ausschließlich Partikeln in Verbindung mit dem präsenskonjugierten Verb verwendet (*pust' on napiš-et!* 'soll er schreiben!', *pust' oni piš-ut!* 'sollen sie schreiben!')

Das **Genus verbi** unterscheidet die Grammeme Aktiv, Passiv und Reflexivmedium, eine konsequente funktionale Trennung der beiden letzteren ist allerdings kaum möglich. Die bisher behandelten Formen waren ausnahmslos **aktivischer** Natur. Die **passivische** Verwendung der (in der Regel transitiven) Verben vollzieht sich im gleichen Aspekt-Tempus-Modus-Rahmen wie die aktivische, jedoch ist sie im Tempusbereich etwas differenzierter. Die passivischen Formen werden entweder mit speziellen Suffixen gebildet (**eigentliches Passiv**) oder mit dem agglutinativen Postfix *-sja* bzw. *-s'* (**Reflexivmedium**). Das eigentliche Passiv verbindet sich meist mit dem perfektiven Aspekt, das mediale mit dem imperfektiven Aspekt (*skáza-n-o* 'sag-PFV.PASS-SG.N (gesagt)'; *otkry-t-y* 'öffnen-PFV.PASS-PL (geöffnet)'; *čítá-et-sja* 'les-IPFV.PASS-3.SG (wird gelesen)'). Das Passiv-Medium-Paradigma des Verbums (*pročítat* 'lesen') hat in der Gestalt der 3. Person Plural folgende Form:

	perfektiv	imperfektiv
Präsens		<i>čítá-jut-sja</i>
Futur	<i>búd-ut pročítá-n-y</i>	<i>búd-ut čítá'-sja</i>
Präteritum	<i>bý-l-i pročítá-n-y</i>	<i>čítá-l-i-s'</i>
Perfekt	<i>pročítá-n-y</i>	
Konjunktiv	<i>bý-l-i by pročítá-n-y</i>	<i>čítá-l-i-s' by</i>
Imperativ	<i>búd'-te pročítá-n-y</i>	<i>čítá-j-te-s'</i>

Tab. 122.2: Russische Passivbildung

Es muß angemerkt werden, daß nicht jede mit dem Postfix *-sja* bzw. *-s'* ausgestattete Verbform ein Medium darstellt. Das genannte Postfix drückt auch verschiedene Qualitäten echter Reflexivität oder Deriva-

tionsbedeutungen (s. 5.2) aus: *brit'* 'rasieren' / *brit'sja* 'sich selbst rasieren'; *celovát'* 'küsselfen' / *celovát'-sja* 'sich gegenseitig küsselfen'; *lovít'* 'fangen' / *lovít'sja* 'sich fangen lassen'.

Die Kategorien der **Person**, des **Numerus** und des **Genus** sind bezüglich des russ. Verbums *Pseudokategorien*, da sie nicht die lexikalische Bedeutung des Verbums modifizieren, sondern formale Reflexe von Qualitäten des Subjekts. Tatsächlich sind (*ja čítá-ju* 'ich lese' (zugleich mit dem Indikativ Präsens wird die 1. Person Singular ausgedrückt); *Nína čítá-l-a* 'Nina las' (zugleich mit dem Indikativ Präteritum werden die 3. Person Singular und das Femininum ausgedrückt)). Diese Pseudokategorien finden im finiten Paradigma unterschiedliche Berücksichtigung (s. u. Tab. 122.3).

#### 4.2. Die Morphologie des Substantivs

Die mit dem russischen Substantiv verbundenen Kategorien werden synkretistisch ausgedrückt, d. h. Numerus, Kasus und Genus werden von ein und demselben Flexiv getragen. Insgesamt umfaßt ihr Paradigma  $2 \times 6$  Konstituenten, die folgenden Grammemen von **Numerus** und **Kasus** entsprechen: 'SG/PL' und 'NOM', 'GEN', 'DAT', 'AKK', 'INSTR', 'LOK'. Zur Illustration wird das Paradigma des Substantivs *sad* 'Garten' (M) gegeben:

	Singular	Plural
NOM	<i>sad-∅</i>	<i>sad-ý</i>
GEN	<i>sád-a</i>	<i>sad-óv</i>
DAT	<i>sád-u</i>	<i>sad-ám</i>
AKK	<i>sád-∅</i>	<i>sad-ý</i>
INSTR	<i>sád-om</i>	<i>sad-ámi</i>
LOK	(v) <i>sád-e</i> (o) <i>sád-e</i>	(v, o) <i>sad-ách</i>

Tab. 122.4: Russisches Deklinationsbeispiel

Das **Genus** als konstanter grammatischer Klassifikator (s. 4.) unterliegt in Verbindung

	Indikativ			Konjunktiv	Imperativ
	Präsens	Präteritum	Futur		
Numerus	+	+	+	+	+
Person	+	-	+	-	+
Genus	-	SG	-	SG	-

Tab. 122.3: Paradigmatische Distribution von Pseudokategorien des Verbums

mit ein und demselben Lexem in der Regel keiner Veränderung. Von einer paradigmatischen Entfaltung ist nur bezüglich unterschiedlicher Lexeme zu sprechen. In diesem Sinne werden Substantive mit **maskulinem**, **femininem** und **neutralem** Genus unterschieden. Als Klassifikationsgrundlage dienen die formale Qualität der Flexive im Nominativ Singular und/oder das Kongruenzverhalten genusabhängiger Kontexte (wie adjektivische Attribute, Präteritalprädikate, genusmarkierte Pronomen u. ä.). Marker des Maskulins ist das Flexiv -∅ nach Konsonant und -j (*sad-∅* ‘Garten’, *rubl'-∅* ‘Rubel’, *gerój-∅* ‘Held’). Marker des Femininums sind die Flexive -a nach Konsonant, -ja nach Vokal und -∅ nach palatalem Konsonant (*kníg-a* ‘Buch’, *nedél'-a* ‘Woche’, *zme-já* ‘Schlange’, *kost'-∅* ‘Knochen’). Die formale Koinzidenz zwischen dem Typ *rubl'-∅* (M) und dem Typ *kost'-∅* (F) wird in den Formen der obliquen Kasus zum Teil aufgehoben; eindeutiger ist der Kontext für die Disambiguierung: *zolot-ój* (M) *rubl'-∅* ‘Goldruble’ / *bol'š-ája* (F) *kost'-∅* ‘großer Knochen’. Marker des Neutrums sind die Flexive -o nach velarem Konsonant sowie -e und -ó nach palatalem Konsonant und /i/. Daneben gibt es einige wenige, aber sehr frequente Neutra mit dem Flexiv -a nach palatalem Konsonant sowie einer Stammerweiterung in den obliquen Kasus (*mést-o* ‘Platz’, *pól-e* ‘Feld’, *znani-e* ‘Wissen’, *pit'-ó* ‘Getränk’; *vrém'-a* ‘Zeit’, *vrémen-i* (GEN)). Von den angeführten Flexiv-Genus-Zuordnungen gibt es einige Abweichungen (*muščin-a* (M) ‘Mann’, *gorodišk-o* (M) ‘Stadt (peior.)’). Die genannten Genera dienen der deskriptiven Grammatik als wichtigstes Klassifikationskriterium der Substantive. Zusammen mit dem formalen Verhalten der Lexeme im Kasusparadigma bilden sie die Grundlage für ihre Einordnung in drei **Hauptdeklinationsklassen**. Neben diesen Standardtypen gibt es eine (nichtproduktive) Klasse von Substantiven, die sich in ihren Deklinationsformen an unterschiedlichen Standardtypen orientieren und folglich eine Mischklasse darstellen (s. Švedova 1982, Hrsg.: 483–530; Isačenko 1968: 90–129; Barnetová et al. 1979: 447–467). In und zwischen den Deklinationsklassen finden sich zahlreiche Fälle von **Flexivhomonymie** sowohl im intra- als auch im interparadigmatischen Sinn (zur intraparadigmatischen Homonymie: *kníg-a* (F) ‘Buch’, *knig-i* (GEN.SG und NOM.PL); *mys'-∅* (F) ‘Maus’, *mýš-i* (GEN/DAT/LOK.SG und NOM.PL); zur interparadigmatischen Homonymie: *pálec-∅* (M) ‘Finger’,

*pál'c-am* (DAT.PL); *kníg-a* (F) ‘Buch’, *knig-am* (DAT.PL); *tetrád'-∅* (F) ‘Heft’, *tetrád'-am* (DAT.PL). Interparadigmatische Homonymie ist besonders häufig im Bereich der Pluralparadigmen. Die Homonymie kann in bestimmten Fällen durch Akzentoppositionen kompensiert werden (*kol'c-ó* (N) ‘Ring’, *kol'c-á* (GEN.SG), *kól'c-a* (NOM.PL)).

Flexionsbedingte Stammveränderung ist beim Substantiv seltener als beim Verb. Zu den typischen allomorphen Veränderungen gehören die Erweiterungen des Stammes bestimmter Substantive in den obliquen Kasus sowie phonologische Veränderungen des auslautenden Konsonanten (*mat'-∅* (F) ‘Mutter’, *máter-i* (GEN etc. SG, NOM.PL), *sestrl-á* (F) ‘Schwester’, *sestr'l-é* (DAT.SG)). Eine häufig auftretende Stammveränderung ist das sogen. ‘flüchtige’ e bzw. o in bestimmten Positionen des Kasusparadigmas (*otéc-∅* ‘Vater’, *otc-á* (GEN.SG), *okn-ó* ‘Fenster’, *ókon* (GEN.PL), *pésn'-a* ‘Lied’, *pésen* (GEN.PL)).

Abschließend ist die Sondergruppe der **In-deklinabilia** zu erwähnen, die derzeit ca. 400 Lexeme umfaßt. Es handelt sich um Fremdwörter, Abbreviaturwörter (s. 5.2) und einige andere Typen, die aufgrund ihrer Form einen erschwerten Zugang zur Deklination haben (*interv'jú* ‘Interview’, *taksi* ‘Taxi’, MGU Abkürzung für ‘Moskauer Staatsuniversität’, aber auch *metró* ‘U-Bahn’, *rádio* ‘Radio’).

#### 4.3. Die Morphologie der Pronomen

Wie unter 3 ausgeführt, verhalten sich die **Pronomen** in ihrer Eigenschaft als Kategorienträger recht unterschiedlich. Dem entspricht die formale Heterogenität ihres grammatischen Apparates. Es sind ein pronominales, ein substantivisches und ein adjektivisches Deklinationsmuster zu unterscheiden. Zwischen diesen gibt es innerhalb der Einzelparadigmen zahlreiche Überschneidungen (s. Švedova 1982, Hrsg.: 531–539; Isačenko 1968: 469–520). In das pronominale Muster fließen sowohl substantivische als auch adjektivische Formen ein (*t-e* ‘jene’ (NOM.PL, pronominale Form), *t-a* ‘jene’ (NOM.SG.F, substantivische Form), *t-omu* (DAT.SG.M/N, adjektivische Form). Das substantivische Muster zeichnet sich durch Affinität zu substantivischen Flexionsformen aus, wobei Nominativ und oblique Kasus von unterschiedlichen Stämmen gebildet werden (*ja* ‘ich’, *mn-e* (DAT, LOK); *my* ‘wir’, *n-am* (DAT), *n-ámi* (INSTR). Im Falle der anaphorischen Personalpronomen finden sich neben dem NOM nach substantivischem Muster die obliquen Kasus mit adjek-

tivischem Muster (*on-i* ‘sie’ (NOM-PL), *i-m* (DAT), *i-mi* (INSTR)). Pronomen mit adjektivischem Muster verhalten sich formal wie Adjektive (s. 4.4). Besonders zu erwähnen sind die indeklinablen Possessivpronomen der 3. Person (*ich dom* ‘ihr (PL) Haus’, *ich straná* ‘ihr Land’, *s ich vračámi* ‘mit ihren Ärzten’; zu Funktion und Form der Pronomen s. auch Barnetová et al. 1979: 515–521).

#### 4.4. Die Morphologie des Adjektivs

Das russische **Adjektiv** besitzt zwei formal und funktional unterschiedliche Paradigmen, die durch die syntaktische Distribution ihrer jeweiligen Konstituenten bedingt sind (s. 3). Die Langform wird in attributiver und prädikativer, die Kurzform nur in prädikativer Funktion benutzt, d.h. letztere erfordert keine obliquen Kasus. Die Langform wird als die “Normalform” des Adjektivs verstanden, für die Bildung der Kurzform gelten bestimmte formale und semantische Restriktionen. Beide Formen reflektieren lediglich die kategorialen Eigenschaften der substantivischen Bezugsgröße in ihren Flexiven (Kongruenz). Diese werden zu pseudokategorialen Eigenschaften des Adjektivs. Sowohl an der Lang- als auch an der Kurzform werden eine harte und eine weiche Flexionsvariante unterschieden, die sich in der Velarität/Palatalität des stammauslautenden Konsonanten, teils auch im Flexivvokal manifestieren (*nól-vl-oe* ‘neues’ (NOM.SG.N, hart)/*zímln'-ee* ‘winterliches’ (NOM.SG.N, weich). Ein besonderes Paradigma hat das Possessivadjektiv: Es vereint in sich substantivische und adjektivische Formen (*otc-óv-∅* ‘dem Vater gehöriger’ (NOM.SG.M, substantivische Form), *otc-óv-ymi* (INSTR.PL, adjektivische Form). Auch in den Adjektivparadigmen ist Flexivhomonymie nicht selten (*nóv-oj* ‘neu’ (GEN/DAT/INSTR/LOK.SG.F)).

Langform	Kurzform
<i>nóv-yj</i> (NOM.SG.M)	<i>nóv-∅</i> (SG.M)
<i>nóv-aja</i> (NOM.SG.F)	<i>nov-á</i> (SG.F)
<i>nóv-oe</i> (NOM.SG.N)	<i>nóv-o</i> (SG.N)
<i>nóv-ye</i> (NOM.PL)	<i>nóv-y</i> (PL)
<i>nóv-ogo</i> (GEN.SG.M/N)	kein Kasusparadigma etc.

Tab. 122.5: Deklinationsformen des Adjektivs (hart)

Eine geringe Menge von entlehnten Adjektiven ist indeklinabel (*júbka bez* ‘ein Rock in

Beige’, *chíndi literatúra* ‘Literatur in Hindi’). Morphonologische Veränderungen des Adjektivstamms sind selten. Eine wichtige Ausnahme stellt allerdings die Kurzform dar, in der es zu Stammerweiterungen und Vokalalternationen kommen kann (*bol'-noj* ‘kranke’ / *bólen-∅*, *úmn-yj* ‘klug’ / *umén-∅*, *tépl-yj* ‘warm’ / *tepl-á*).

Die Kategorie der **Komparation** von Qualitätsadjektiven umfaßt die Grammeme **Positiv**, **Komparativ** und **Superlativ**. Hinsichtlich der Steigerungsformen ist zwischen einer **analytischen** und einer **synthetischen** Bildungsweise zu differenzieren (s. Art. 114). Im analytischen Paradigma entspricht der Positiv der einfachen Langform (*vážn-yj* ‘(ein) wichtiger’). Der Komparativ wird durch die Gradationspartikel *bólee* gebildet (*bólee vážn-yj* ‘(ein) wichtigerer’) und der Superlativ durch das mit dem Adjektiv kongruierende Steigerungswort *sám-* (*sám-yj vážn-yj* ‘(der) wichtigste’) oder die Gradationspartikel *naibólee* (*naibólee vážn-yj* dasselbe). Weniger regelhaft ist die synthetische Komparationsbildung. Für den Komparativ werden die agglutinativen (indeklinablen) Allomorphe *-ee*, *-e* und *-še* in geregelter Distribution verwendet. Die beiden letzteren können erhebliche morphonologische Stammveränderungen erzeugen (*kra-sív-* ‘schön’/*krasív-ee* ‘schöner’; *dórog-* ‘teuer’/*doróž-e* ‘teurer’; *ránn-* ‘früh’/*rán'-še* ‘früher’). Der (lexikalischen Restriktionen unterworfene) synthetische Superlativ wird mit dem Suffix *-ajš-/ejš-* gebildet, dem sich das adjektivische Flexiv anschließt (*vysók-* ‘hoch’/*vysoc-ájš-ij* ‘höchster’; *bogát-* ‘reich’/*bogat-éjš-ij* ‘reichster’). Außer den angeführten gibt es weitere, weniger frequente Bildungsweisen im Bereich der Komparation. Generell gilt, daß das formale Paradigma dieser Kategorie das am wenigsten stringente ist. Auf diesem Hintergrund ist auch der Umstand zu sehen, daß der synthetische Superlativ in der Funktion eines **Elativs** ohne direkte paradigmatische Bezugsgröße verwendet werden kann (*krépk-* ‘stark’/*krepl-ájš-aja* *vódka* ‘ein äußerst starker Vodka’).

#### 4.5. Die Morphologie des Adverbs

Das Adverb als Modifikator von Satzgliedbedeutungen verfügt im Russischen über sehr unterschiedliche Erscheinungsweisen wie Kasus- und Gerundialformen, Präpositionalgefüge, nichtaffigierte Wörter u. a. (s. Švedova 1982, Hrsg.: 703–705; Isačenko 1968: 176–193). Morphologisch eindeutig als Adverbien ausgewiesen sind neben den Gerunden (s.

4.1) nur jene Adverbien, die von Adjektiven mit Hilfe der Suffixe *-ol-e* und *-i* oder der Präpositionalstruktur *po + Adj + -omulemuli* abgeleitet sind (*býstr-yj* ‘schnell’ (Adjektiv, NOM.SG.M)/*býstr-o* (Adverb); *ískrenn-ij* ‘aufrichtig’ (Adjektiv)/*ískrenne* (Adverb); *brátsk-ij* ‘brüderlich’ (Adjektiv)/*brátsk-i* (Adverb); *nóv-yj* ‘neu’ (Adjektiv)/*po-nóv-omu* (Adverb)). Von Qualitätsadjektiven abgeleitete Adverbien verbinden sich mit der Komparationskategorie. Der **Komparativ** fällt in formaler Hinsicht mit dem synthetischen Komparativ des Adjektivs (s. 4.3) zusammen. Der selten verwendete **Superlativ** wird durch die Verbindung des Komparativs mit einer Konstruktion des Pronomens *vsë* ‘alles’ gebildet: *Oná skor-ée vsech napisála stat'jú* ‘Sie hat den Artikel am schnellsten (*skor-ée* vsech = schneller als alle) geschrieben.’

#### 4.6. Resümee

In Abschnitt 4 wurden die wesentlichsten grammatischen Strukturen des Russischen skizziert. Die formale Komplexität dieser Sprache wurde damit in keiner Weise erfaßt. Das Russische zählt – wie die meisten slawischen Sprachen – in grammatisch-morphologischer Hinsicht zu den komplexesten indogermanischen Sprachen. Für Zweitspracherwerber ergeben sich daraus in der Regel erhebliche Schwierigkeiten.

### 5. Wortbildungsmorphologie

Das russische **Wortbildungssystem** entspricht in seiner Grundstruktur der hochgradig affixalischen Struktur seines grammatischen Systems. Obgleich **analytische Wortbildungen** heute häufig sind und zunehmen (Typus: *sotrjasénie mózga* ‘Gehirnerschütterung’, wörtlich: Erschütterung des Gehirns), basiert die Bildung des zentralen Wortschatzes auf **synthetischen Prozessen**, an denen vornehmlich Suffixe, Präfixe und Zirkumfixe teilnehmen. Eine gewichtige Rolle spielen auch **Kompositabildungen**.

#### 5.1. Allgemeine Wortbildungsprinzipien

Das Russische verfügt über eine **identifikatorische** und eine **mutatorische Wortbildung**. Erstere vollzieht sich im Rahmen ein und derselben Wortart, letztere führt zur Produktion von Lexemen der Wortart Y aus Lexemen der Wortart X. Während die identifikatorische Wortbildung immer zu einer neuen lexikalischen Bedeutung führt (*nestí* ‘tragen’ >

*v-nestí* ‘hineinragen’), kann die mutatorische Wortbildung auch in der bloßen Überführung in eine andere Wortart bestehen (*znat* ‘wissen’ > *zná-n-ie* ‘das Wissen’). Identifikatorische Wortbildungen finden sich bei *allen* Wortarten des Russischen (*čaj* ‘Tee’ > *čaj-nik* ‘Teekanne’; *bélyj* ‘weiß’ > *bel-ovát-yj* ‘weißlich’ etc.). Mutatorische Wortbildungen finden zwischen folgenden Wortarten statt: Verb > Substantiv, Verb > Adjektiv, Substantiv > Verb, Substantiv > Adjektiv, Adjektiv > Verb, Adjektiv > Substantiv, Adjektiv > Adverb, Adverb > Adjektiv, Pronomen > Verb (*čitat'* ‘lesen’ > *čítá-tel'* ‘Leser’; *závtrak* ‘Frühstück’ > *závtrak-at'* ‘frühstück’ etc.).

Die **Derivationstiefe** russischer Wörter überschreitet selten fünf Takte. Ableitungen bis zu dieser Tiefe sind aber häufig (das Beispiel *u-prošč-énč-estv-o* ‘Simplifizierung’ inkorporiert vier Takte folgender Art: Adjektiv > Verb > Substantiv > Substantiv). Die Oberflächenstruktur derartiger Bildungen können sehr komplexe semanto-syntaktische Tiefenstrukturen vertreten.

#### 5.2. Wortbildungsstrukturen

Aufgrund des ausgeprägten affixalischen Wortbildungsmodus verfügt das Russische über ein umfangreiches Wortbildungsnetzsystem. Tichonov (1985, I:440) hat 12.542 Nester unterschiedlichen Umfangs mit insgesamt 126.034 derivierten Wörtern festgestellt.

Wie die Flexion können auch Derivationen zu erheblichen morphonologischen Veränderungen (der Ableitungsbasisen) führen. Solche treten in Verbindung mit Suffigierungen und Zirkumfigierungen auf (*ruká* ‘Hand’ > *rúč-k-a* ‘Händchen’; *reká* ‘Fluß’ > *za-réč*-e ‘Land hinter dem Fluß’).

Das Russische verfügt über ein beträchtliches Korpus von Wortbildungsmorphemen (s. Vinogradov 1960, Hrsg.: 202–271, 330–363, 576–603, 614–626). Die Vielfältigkeit der Mittel schließt jedoch Homonymie und Synonymie einer Reihe von Affixen nicht aus (Homonymie: *uči-tel'* ‘Lehrer’ (-tel) ‘männliche Person’)/*vklučá-tel'* ‘Schalter’ (-tel) ‘technische Vorrichtung’), Synonymie: *chvast-lív-yj* ‘prahlerisch’ (-liv- ‘geneigt zu’) / *zadúm-čív-yj* ‘nachdenklich’ (-čiv- ‘geneigt zu’). Im folgenden werden die einzelnen Bildungsverfahren vorgestellt.

(a) **Suffigierung:** Die Suffixe nehmen als Wortbildungsmittel den ersten Platz ein. Sie sind vornehmlich im Bereich des Nomens produktiv, bei der Verbbildung spielen sie

eine nachgeordnete Rolle (*drug* ‘Freund’ > *drúžba* ‘Freundschaft’). Suffixe können miteinander Kombinationen eingehen (*ideal-ist-ka* ‘Idealistin’), aber auch in Ø-Form erscheinen (*osmotréť* ‘besichtigen’ > *osmótr-o* ‘Besichtigung’). Als suffixalische Adjektivderivation muß auch die Bildung der unterschiedlichen Partizipien gelten.

(b) **Präfigierung:** Diese ist typisch für die Bildung von Verben, tritt aber auch im Bereich anderer Wortarten auf (*kléit* ‘kleben’ > *nakléit* ‘aufkleben’; *móščnyj* ‘leistungsstark’ > *sverch-móščnyj* ‘äußerst leistungsstark’). Auch Präfixe sind kombinierbar, allerdings wird dieses Verfahren selten praktiziert (*tol-káť* ‘stoßen’ > *vy-tálkivat* ‘herausstoßen’ > *po-vy-tálkivat* ‘nacheinander herausstoßen’).

(c) **Zirkumfigierung:** Dieses Bildungsverfahren stellt eine Kombination von (a) und (b) dar. Statt des Suffixes kann auch das Postfix *-sja* in Finalposition stehen (*luná* ‘Mond’ > *pri-lun-it'-sja* ‘auf dem Mond landen’; *den* ‘Tag’ > *po-děn-ščik* ‘Tagelöhner’).

(d) **Kompositbildung:** Wortbildungen, an denen mindestens zwei lexikalische Morpheme teilhaben, sind heute im Gegensatz zu älteren russischen Sprachzuständen recht produktiv. Mit ihrer Hilfe kann vermieden werden, daß ein Grad von Affixhomonymie entsteht, der eine kommunikative Dysfunktionalität zur Folge hätte. Kompositbildungsforschungen finden sich in allen Wortartenbereichen. Sie vollziehen sich im wesentlichen ebenfalls nach den unter 5.1 umrissenen Prinzipien und repräsentieren eine Vielzahl von semanto-syntaktischen Tiefenstrukturen. Nach ihrer Oberflächenstruktur können **Juxtaposita** und **Komposita** unterschieden werden (*górod* ‘Stadt’/*gerój* ‘Held’ > *górod-gerój* ‘Heldenstadt’; *sínij* ‘dunkelblau’ > *sínij-sínij* ‘tiefblau’; *neft* ‘Erdöl’/*dobýča* ‘Förderung’ > *neft-e-dobýča* ‘Erdölförderung’).

(e) **Abbreviaturbildung:** Eine Sonderform der Komposita machen die Abbreviationen aus, die nach dem Prinzip der Komposition von Silben- und Buchstabenkürzungen formiert werden. Die wichtigsten Typen sind die Silbenabbreviatur (*kollektívnoe chozjájstvo* ‘Kollektivwirtschaft’ > *kolchóz*), die Initialabbreviatur (*výsše učébnoe zavedénie* ‘Hochschule’ > *vuz*) und die Kombination von Abbreviaturen und Lexemen (*finánsovyy otdél* ‘Finanzabteilung’ > *finotdél*). Abbreviaturen können ihrerseits als Derivationsbasen dienen (*vuz* ‘Hochschule’ > *vúz-ovec* ‘Hochschüler’).

## 6. Grammatische Morphologie und Wortbildung in der Literatur

Die ersten Beschreibungen der formalen Strukturen des Russischen lieferten die seit Mitte des 18. Jahrhunderts aufkommenden präskriptiven Grammatiken dieser Sprache. Sie folgten in ihren Darstellungsverfahren den von den kirchenlawischen Grammatiken vermittelten Mustern (s. auch 1). Es handelte sich hier um einfache, zu Lehr- und Lernzwecken aufgelistete Flexionstabellen, deren innere Ordnung nach klassischem Vorbild den Gegebenheiten des Russischen oftmals nicht entsprach. Im 19. Jahrhundert entwickelte sich zusammen mit der historisch-vergleichenden Sprachwissenschaft auch das Interesse an der Erforschung der alten Strukturen des Russischen (s. Berezin 1979: 62–83). Die ersten historischen Grammatiken sind folglich in hohem Maße der Formen- und Wortbildung gewidmet. Noch in dieser Tradition entstanden um die Jahrhundertwende Konzeptionen einer synchronen Betrachtung der russischen Morphologie. Diese können durchaus als frühstrukturalistisch gelten. Hier ist vor allem die Moskauer Schule um und nach F. F. Fortunatov zu nennen (s. Lehfeldt & Kempgen 1984: 43–47). In Fortführung dieser Konzeptionen, aber auch in der Auseinandersetzung mit ihnen ist im Laufe des 20. Jahrhunderts in der UdSSR und im Ausland eine Fülle von Literatur entstanden, die jede Nennung von Namen und Titeln willkürlich macht.

Als breitangelegte aber wenig konsistente Gesamtdarstellungen der grammatischen Morphologie können die entsprechenden Kapitel in den Akademiegrammatiken angeführt werden (Vinogradov 1960, Hrsg.; Švedova 1982, Hrsg.). Ihr Wert liegt eher in der Materialkomilation als in einer ausreichend heuristischen Abbildung der sich hinter den Einzelphänomenen verborgenden Strukturen. Ähnlich einzuschätzen ist die an der Prager Akademie erarbeitete russische Grammatik (Barnetová et al. 1979). Allerdings ist die analytische Aufarbeitung des Materials hier erheblich effektiver. Seit den 60er Jahren haben die Arbeiten von A. A. Zaliznjak Aufmerksamkeit erregt. In ihnen wird versucht, die russische Formenbildung einer konsequenten Klassifikation zu unterziehen und algorithmische Syntheseregeln für die Teilparadigmen zu entwickeln. Als bedeutendste Arbeit ist der alle Wortarten berücksichtigende *Grammaticeskij slovar' russkogo jazyka* (Za-

liznjak 1977) zu nennen. Hier ist außer der systematisierenden Beschreibung der gesamten grammatischen Morphologie ein Wörterbuch *a tergo* (10.000 Einträge) enthalten. Im deutschen Sprachraum fehlen ähnlich erschöpfende und methodisch fortgeschrittene Gesamtdarstellungen. Allerdings kommt der Ende der 50er Jahre fertiggestellten Monographie von Isačenko (Isačenko 1968) besondere Bedeutung zu, da in ihr der rezente russische Formenhaushalt erstmals unter formalen und inhaltlichen Gesichtspunkten einer strukturalistischen Betrachtung unterzogen wird. Den gleichen Ansatz, obwohl ohne größere Würdigung der Inhaltsseite, zeigt Ďzurovič (1964). Erwähnenswert sind auch einige Spezialarbeiten zum russischen Verb von Daum & Schenk (¹1978) und von Kempgen (1989). Letzterer will mit seinem Modellentwurf Antwort auf die Frage geben, „wie eine beliebige Verbform eines beliebigen Verbs gebildet wird, wenn der Infinitiv gegeben ist“ (Kempgen 1989: 5).

Für Unterrichtszwecke kann weiterhin die informative und sprachlich leicht zugängliche, wenn auch nicht fehlerfreie Darstellung von Mulisch (¹1985) genannt werden.

Auch die russische Wortbildung wurde bereits im Schoße der historischen Sprachwissenschaft zum wissenschaftlichen Gegenstand. Nach gängiger Auffassung avancierte die Wortbildungslehre jedoch erst seit den 40er Jahren des 20. Jahrhunderts zu einer eigenständigen linguistischen Teildisziplin (Vinogradov 1949), die inzwischen bemerkenswerte Resultate vorlegen kann. Es wurden imposante Materialkorpora zusammengestellt, wie beispielsweise das 145.000 Wörter umfassende Wortbildungswörterbuch von Tichonov (1985) oder die Kompilationen der Akademiegrammatiken (Vinogradov 1960, Hrsg.; Barnetová et al. 1979; Svedova 1982, Hrsg.). An einer Theorie der Wortbildung unter semantischen und formalen Aspekten wird ebenfalls seit den 40er Jahren gearbeitet (vgl. Mel'čuk 1967; Lopatin 1977; Uluchanov 1977; zusammenfassend zur Theoriebildung Raecke 1984). Schließlich wurden komplexe Wortbildungsmodelle entwickelt, von denen sich das im Rahmen des Applikativen Generativen Modells erarbeitete und seine Modifikationen bewährt haben (Šaumjan & Soboleva 1968; Jachnow et al. 1980). Als ältere aber bewährte Arbeitsbücher sind Šanskij (1968) und Zemskaja (1973) zu erwähnen.

Neuere Darstellungen und Titel zur grammatischen Morphologie und zur Wortbildung findet der Leser bei Kempgen (1999), Lehfeldt & Kempgen (1999) und Raecke (1999).

## 7. Zitierte Literatur

- Barnetová, Vilma & Běličová-Křížková, Helena & Leška, Oldřich & Skoumalová, Zdena & Straková, Vlasta (1979), *Russkaja grammatika*, Bd. I. Praha: Academia
- Berezin, Fedor M. (1979), *Istorija russkogo jazyka*. Moskva: Vysšaja škola
- Daum, Edmund & Schenk, Werner (¹1978), *Die russischen Verben*. Leipzig: Enzyklopädie [¹1954]
- Ďzurovič, Ľubomir (1964), *Paradigmatika spisovnej ruštiny*. Bratislava: Slovenské pedagogické nakladatelstvo
- Isačenko, Alexander (1968), *Die russische Sprache der Gegenwart*, I: *Formenlehre*. Halle/S.: VEB Max Niemeyer
- Jachnow, Helmut (1984, Hrsg.), *Handbuch des Russischen*. Wiesbaden: Otto Harrassowitz
- Jachnow, Helmut (1999, Hrsg.), *Handbuch der sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen*. Wiesbaden: Otto Harrassowitz
- Jachnow, Helmut & Jäpel, Rita & Kemper, Hajo & Knoop, Iris & Koch, Peter & Marsall, Arkadij & Matzullat, Gerda & Özbel, Tekin & Zekan, Elisabeth (1980), *Zur Erklärung und Modellierung diachroner Wortbildungsprozesse anhand russischer substantivischer Neologismen*. München: Otto Sagner
- Kempgen, Sebastian (1989), *Grammatik der russischen Verben*. Wiesbaden: Otto Harrassowitz
- Kempgen, Sebastian (1999), „Morphemik“. In: Jachnow (Hrsg.), 87–108
- Lehfeldt, Werner & Kempgen, Sebastian (1984), „Die Formenbildung des Russischen und sowjetische Konzeptionen ihrer Beschreibung“. In: Jachnow (Hrsg.), 32–66
- Lehfeldt, Werner & Kempgen, Sebastian (1999), „Formenbildung“. In: Jachnow (Hrsg.), 109–149
- Lopatin, Vladimir V. (1977), *Russkaja slovoobrazovatel'naja morfemika*. Moskva: Nauka
- Mel'čuk, Igor' A. (1967), „K ponjatiyu slovoobrazovanija“. *Izvestija Akademii nauk* 4, 352–362
- Mulisch, Herbert (¹1985), *Morphologie (Die russische Sprache der Gegenwart)*. Leipzig: Enzyklopädie
- Raecke, Jochen (1984), „Zur neueren sowjetischen Wortbildungslehre“. In: Jachnow (Hrsg.), 103–142

- Raecke, Jochen (1999), "Wortbildung". In: Jachnow (Hrsg.), 150–181
- Šanskij, Nikolaj M. (1968), *Očerki po russkomu slovoobrazovaniju*. Moskva: Moskovskij universitet
- Šaumjan, Sebast'jan K. & Soboleva, Polina A. (1968), *Osnovaniya poroždajuščej grammatiki russkogo jazyka*. Moskva: Nauka
- Švedova, Natal'ja J. (1982, Hrsg.), *Russkaja grammatika, Bd. I.* Moskva: Nauka
- Tichonov, Aleksandr N.V. (1985), *Slovoobrazovatel'nyj slovar' russkogo jazyka, Bd. 1–2*. Moskva: Russkij jazyk
- Uluchanov, Igor' S. (1977), *Slovoobrazovatel'naja semantika v russkom jazyke i principy ee opisnija*. Moskva: Nauka
- Vinogradov, Viktor V. (1949), *Russkij jazyk*. Moskva: Vysšaja škola
- Vinogradov, Viktor V. (1960, Hrsg.), *Grammatika russkogo jazyka, Bd. I.* Moskva: Akademija
- Zaliznjak, Andrej A. (1977), *Gramatičeskij slovar' russkogo jazyka*. Moskva: Russkij jazyk
- Zemskaja, Elena A. (1973), *Sovremennyj russkij jazyk: Slovoobrazovanie*. Moskva: Nauka

*Helmut Jachnow, Bochum (Deutschland)*

## 123. Altgriechisch (Indogermanisch)

1. Einleitung
2. Der morphologische Typ des Altgriechischen
3. Wortarten
4. Flexivisch ausgedrückte Kategorien des Nomens und Verbs
5. Morphologische Prozesse; Homonymie
6. Flexion des Nomens
7. Verbalflexion
8. Wortbildung
9. Illustrativer Text
10. Unübliche Abkürzungen
11. Zitierte Literatur

### 1. Einleitung

Altgriechisch wurde im Altertum schon in der klassischen, vorhellenistischen Periode (bis 4. Jh. v. Chr.) nicht nur im eigentlichen Griechenland, d. h. auf der Balkanhalbinsel mit benachbarten Inseln, dem westlichen Kleinasien, kleinen Inseln des Ägäischen Meeres, Kreta und Zypern gesprochen, sondern durch Kolonisation weiter verbreitet, besonders nach Sizilien und Unteritalien und zur Schwarzmeerküste. In der hellenistischen Zeit kam es zu einer weiteren Ausbreitung des Griechischen. Die Bedeutung des Griechischen liegt jedoch in seinem Einfluß als Kultursprache. Griechisch gehört zur indogermanischen Sprachfamilie, ohne direkt unter eine ihrer Gruppen zu fallen (Latein mit fragmentarisch bezeugten italischen Sprachen und seinen romanischen Tochtersprachen, Keltisch, Germanisch, Slawisch, Baltisch, Indo-Iranisch, Hethitisch mit anderen Sprachen Anatoliens sind die Hauptgruppen,

dazu kommen als isolierte Sprachen noch das Armenische und das Albanische, die fragmentarisch bezeugten Sprachen Thrakisch, Illyrisch und Phrygisch, ferner das Tocharische). In dem dialektal stark gegliederten Sprachgebiet werden vier Dialektgruppen unterschieden, nämlich Arkadisch-Kyprisch, Ionisch-Attisch, Aiolisches und Westgriechisch (Dorisch und Nordwestgriechisch). Das Ionisch-Attische ist zur Sprache der griechischen Literaturprosa geworden, die in ihrer attischen Variante als klassische Literatursprache (5.–4. Jh. v. Chr.) Hauptrepräsentant des Altgriechischen ist und schon im Altertum Objekt der grammatischen Beschreibung (vgl. Art. 6) war. Als die für die synchronische Beschreibung am besten geeignete Sprachvariante stellt sie auch hier das Objekt der morphologischen Beschreibung dar.

Bis Anfang der fünfziger Jahre galt, neben den seit dem 8. Jh. v. Chr. belegten, meist kurzen Dialektinschriften, als älteste überlieferte Variante des Altgriechischen das Homerische, die dialektal stark gemischte Sprache der Epen Ilias und Odyssee, die in Ionien im 8. oder 7. Jh. auf der Grundlage einer älteren epischen Tradition verfaßt wurden. Dieser Überlieferung in alphabetischer Schrift geht aber das in Silbenschrift, besonders durch Tontafeln aus Knossos, Pylos und Mykene seit dem 14. Jh. bis 1200 überlieferte sog. mykenische Griechisch voraus, das im Jahre 1951 entziffert wurde und dialektal dem Arkadisch-Kyprischen am nächsten steht.

Für die nachklassische Entwicklung des Altgriechischen ist die Entstehung der sog.

Koiné kennzeichnend, einer Standardsprache, die die Dialekte verdrängte und in der hellenistischen Periode (3. Jh. v.–1. Jh. n. Chr.) auch als Schriftsprache verwendet wurde. Die klassische attische Literaturprosa blieb als Norm- und Schulsprache für die griechische Schriftsprache nicht nur des späten Altertums, sondern auch im Mittelalter und in der Neuzeit maßgebend.

## 2. Der morphologische Typ des Altgriechischen

Das Altgriechische repräsentiert den indogermanischen flexivischen Typ, und zwar dessen Variante mit besonders reichem Inventar an morphologischen Formen. Zu den charakteristischen Zügen der flexivischen Morphologie gehört die Kombination der basisverändernden (vgl. Art. 57–60), symbolischen (vgl. Sapir 1949: 126) Prozesse mit der Suffigierung, die aber von der Suffigierung in den agglutinierenden Sprachen (vgl. Kap. XV) verschieden ist. Die flexivische Endung, d. h. das auslautende Segment des Wortes, zeichnet sich aus durch Kumulation der morphosyntaktischen Merkmale (z. B. Kasus- und Numerusmerkmal in der Nominalendung, vgl. 4.1) und Fusion (Verschmelzung) ihrer Formantien miteinander und mit dem stamm-auslautenden Vokal (z. B. in den Endungen der 1. und 2. Deklination, vgl. 4.1). Der Ausdruck der grammatischen Kategorien wird durch komplexe interne Modifikation des Wortes, nicht durch bloße Affigierung wie in den agglutinierenden Sprachen verwirklicht, die einzelnen semantischen Merkmale werden durch Kombination mehrerer morphologischer Prozesse ausgedrückt. Weitere Züge sind einerseits starke Allomorphie (2.1) und andererseits starke Polyfunktionalität (2.2) der Formantien.

Die Polyfunktionalität der flexivischen morphologischen Formen hängt damit zusammen, daß ihre auf den Oppositionen innerhalb eines Paradigmas beruhende Bedeutung ziemlich breit ist (und vage, so daß bei einigen linguistischen Zugängen auf ihre Definition oder sogar auf ihre Anerkennung verzichtet wird).

### 2.1. Allomorphie

**Allomorphie** (vgl. Kap. VII) ist im flexivischen Typ sehr verbreitet und entweder formal morphologisch und phonologisch (2.1.1) oder semantisch (2.1.2) bedingt.

#### 2.1.1. Formal morphologisch und phonologisch bedingte Allomorphie

Charakteristisch für den flexivischen Typ ist die formal morphologische Allomorphie, die durch Zugehörigkeit des Wortes zu einer Deklinations- oder Konjugationsklasse bzw. -unterklasse konditioniert ist (s. 6 und 7). In einzelnen Fällen ist die allomorphe Variation innerhalb einer Deklinations- oder Konjugationsklasse phonologisch konditioniert. Z. B. wird der Akkusativ Singular durch folgende Allomorphe ausgedrückt: *-a* (3. Deklination, Stämme mit konsonantischem Auslaut), *-n* (3. Deklination, Stämme mit vokalischem Auslaut), *-on* (2. Deklination), *-ēn*, *-ān*, *-ān* (1. Deklination). (Abweichend von den Transkriptionsregeln wird hier und weiterhin auch die Länge von *ā* markiert, Kürze nur bei Hervorhebung des Gegensatzes.) Die Allomorphe innerhalb der 3. Deklination sind phonologisch durch konsonantischen und vokalischen Auslaut konditioniert. Die Variation innerhalb der 1. Deklination hat im Falle der Selektion zwischen langem *ē* und *ā* phonologische Bedingungen, die aber nicht so natürlich und intuitiv durchsichtig sind wie der Fall des vokalischen bzw. konsonantischen Auslauts (*ā* kommt nach *-r*, *-i* und *-e* vor, wo im Attischen der ionisch-attische Übergang *ā* > *ē* nicht stattgefunden hat oder ist hier *ā* wieder hergestellt worden). Die Selektion des kurzen *-ān* ist rein morphologisch durch Zugehörigkeit zu einer bestimmten Unterklasse konditioniert. Eine durch Assimilation motivierte Allomorphie des Stammes finden wir im Verbalparadigma in Fällen wie *ké-krum-mai*, ‘RDP:PF-verberg-MEDPASS.1.SG’ *ké-krup-tai* ‘RDP:PF-verberg:PF-MEDPASS.3.SG’ (vgl. Rix 1976: 102).

#### 2.1.2. Semantische Variation und semantisch bedingte Allomorphie

Im Altgriechischen ist die sog. Formenvariation (Lewy 1961: 205, 215) nicht nur auf die reine allomorphe Variation bei voller Synonymie begrenzt, sondern es bestehen innerhalb eines Morphems Ausdrucksvarianten, die sich durch zusätzliche semantische Merkmale voneinander unterscheiden. Beim aktiven Aorist begegnet uns eine Variation, die in den Bereich der Diathese im weiten Sinne, d. h. der mit der Argumentenstruktur zusammengehörigen Verbalmorphologie (vgl. Art. 107 und 108), fällt. Der semantische Unterschied zwischen dem sigmatischen Aorist auf der einen und dem athematischen Wurzelao-

rist oder thematischen Aorist auf der anderen Seite äußert sich meistens nur durch ihre Distribution auf semantisch begründete Verbklassen (vgl. Art. 48). Der sigmatische Aorist kommt bei aktiven, zum Agens bezogenen Verben vor, der Wurzel- und thematische Aorist dagegen bei inaktiven Verben, deren Subjekt eine andere semantische Rolle als die des Agens hat, z. B. *eidon* ‘ich habe gesehen’, wo dem Subjekt die Rolle des sog. Experiencer (vgl. Kap. XIII, Art. 94) zukommt. Bei einer Gruppe von Verben treten aber diese Aoriste zueinander in Opposition (vgl. besonders Kølln 1969: 15 f.). Der sigmatische Aorist ist transitiv/kausativ, der Wurzelaorist intransitiv: z. B. *é-phu-se* ‘PRÄT-wachs-AOR.AKT.3.SG (er ließ wachsen, entstehen)’ gegenüber *é-phu* ‘PRÄT-wachs:AOR.AKT.3.SG (er entstand, wuchs)’. Es handelt sich aber nicht um eine grammatisierte, sondern um eine nur auf einige wenige Verba beschränkte Opposition.

Auch bei den Präsensbildungen (d. h. Formationen des Präsensstammes – vgl. 7.1) finden wir semantische Distinktionen, die vor allem in den Bereich der Aktionsart (vgl. Art. 109) fallen, ohne daß jedoch die Oppositionen bei denselben Verben vertreten sind (vgl. Kujore 1973).

Semantische Variante ist auch die Pluralmarkierung des Nominativs/Akkusativs der Neutra. Fälle, in denen bei einem Substantiv zwei Pluralformen mit Bedeutungsunterschied vorkommen, z. B. maskulin *méroí* ‘einzelne Schenkelstücke’ gegenüber neutral *méra* ‘deren Gesamtheit’ (vgl. Schwyzer 1968: 581), ebenso wie der Singular des Verbs beim Plural des Neutrums (Typ *tà zô-a tréch-ei* ‘ART.NOM/AKK.PL.N Tier-NOM/AKK.PL.N laufen’ 3.SG (die Tiere laufen), im Attischen regelmäßig) zeigen, daß es sich beim maskulinen/femininen vs. neutralen Plural um keine bloße Allomorphie handelt. Beim Nominativ/Akkusativ des Neutrums handelt es sich in Bezug auf Nominativ und Akkusativ des Maskulinums oder Femininums um Allomorphie, die aber, anders als die durch die Deklinationsklassen bedingte Allomorphie, semantisch motiviert ist, da das Genus eine besondere semantische Kategorie des Nomens darstellt, auch wenn das Genusmerkmal in hohem Maße willkürlich verteilt ist. Eine semantisch bedingte Allomorphie kommt bei der Selektion des Kasus durch den Numerus vor, die aber als Kumulation des Ausdrucks der Morpheme des Nominativs Singular usw. zu verstehen ist, ebenso wie die selektiven Beziehungen in der Verbalendung (7.2).

## 2.2. Flexionsparadigma

Parallel zur Semantik, die durch Oppositionen innerhalb des morphologischen Paradigmas identifizierbar ist, ist die flexivische Form durch Oppositionen innerhalb des flexivischen Paradigmas identifizierbar. In einer agglutinierenden Sprache wie dem Ungarischen wird bei allen Nomina durch Suffigierung von *-k* der Plural, durch Suffigierung von *-t* das direkte Objekt (der Tradition nach als Akkusativ bezeichnet) signalisiert, z. B. *atyá-t* ‘Vater(SG)-AKK’, *atyá-k-at* ‘Vater-PL-AKK’, *barát-ot* ‘Freund(SG)-AKK’, *barát-ok-at* ‘Freund-PL-AKK’. In den entsprechenden griechischen Formen *patér-a* ‘Vater-AKK.SG’, *patér-as* ‘Vater-AKK.PL’, *phil-on* ‘Freund-AKK.SG’, *phil-ous* ‘Freund-AKK.PL’ ist nur in der Opposition zum Nominativ und zu den anderen Kasus des Singularparadigmas *-a* oder *-on* als Akkusativwendung und nur in der Opposition zum Akkusativ Singular *-as* oder *-ous* als Pluralendung definiert. Wir können bei den griechischen Formen des Akkusativs Plural nicht bestimmen, welcher Teil der Endung den Akkusativ und welcher Teil den Plural bezeichnet (Kumulation). Kennzeichnend für die eng an das Wort gebundene flexivische Morphologie ist die Tatsache, daß auch Adjektiv (3.1) und Pronomen (3.2) flektiert werden, und zwar nach der Deklinationsklasse oder dem Deklinationstyp, zu dem sie paradigmatisch gehören, also unabhängig vom Substantiv, auf das sie sich syntaktisch beziehen.

Das Verbalparadigma weist semantisch und formell beträchtliche Mannigfaltigkeit und Uneinheitlichkeit auf, die in 7.1 charakterisiert wird.

## 3. Wortarten

Die Hauptwortarten, Nomen und Verb, sind im Altgriechischen, ebenso wie in allen flexivischen indogermanischen Sprachen, morphologisch deutlich charakterisiert und von einander differenziert (s. 4–7). In 3.1–3.4 werden die übrigen Wortarten morphologisch charakterisiert.

### 3.1. Adjektiv

Das **Adjektiv** hat im Altgriechischen und Indogermanischen mit dem Substantiv gemeinsame morphologische und syntaktische Charakteristika (besonders die Funktion des Prädikats im Nominalzusatz ohne Kopula), so daß beide Wortarten in der grammatischen Tradi-

tion unter dem Begriff des Nomens zusammengefaßt werden. Auch wenn es besondere Adjektivsuffixe gibt (s. 8), werden Adjektiva von denselben grundlegenden Stämmen wie Substantiva gebildet und demgemäß nach denselben Deklinationsklassen und -unterklassen (s. 6) flektiert. Ebenso wie beim Substantiv werden beim Adjektiv Kasus und Numerus kumulativ in einer Endung ausgedrückt (vgl. *agath-ós* ‘gut-NOM.SG.M’). Dazu kommt, daß in der adjektivischen Endung noch das Genus (4.1) markiert wird, also *agath-ós* ‘gut-NOM.SG.M’, *agath-ē* ‘gut-NOM.SG.F’, *agath-ón* ‘gut-NOM.SG.N’. Die Femininformen des Adjektivs gehören immer der 1. Deklination an und werden wie Substantiva auf -ā oder -ē dekliniert, wenn sie zu den maskulinen Adjektiven der 2. Deklination gebildet werden, vgl. *agath-ós*, *agath-ē*; *dikai-os* ‘gerecht-NOM.SG.M’, *dikai-ā* ‘gerecht-NOM.SG.F’. Wenn sie zu den Adjektiven der 3. Deklination gebildet werden, deklinieren sie wie die Substantiva auf -ā, vgl. *mélās* ‘schwarz:NOM.SG.M’, *mélaina* ‘schwarz:NOM.SG.F’, *mélan* ‘schwarz:NOM.SG.N’ oder *glukús* ‘süß:NOM.SG.M’, *glukeá* ‘süß:NOM.SG.F’, *glukú* ‘süß:NOM.SG.N’. Es gibt aber auch Adjektiva zweier Endungen mit gemeinsamer Form für Maskulin und Feminin, und zwar regelmäßig zusammengesetzte und bestimmte abgeleitete Adjektiva, z. B. *katá-gei-os* ‘unter-irdisch-NOM.SG.M/F’ *eu-dáim-ōn* ‘glück-selig-NOM.SG.M/F’, *pátri-os* ‘überkommen-NOM.SG.M/F’. In einigen Fällen handelt es sich um semantisch motivierte Restriktionen, z. B. bei *árrhen* ‘männlich:NOM.SG.M/F’, wo die Femininform im Unterschied zu anderen Adjektiven mit demselben Suffix nicht belegt ist, z. B. *térēn* ‘zart:NOM.SG.M’, *téreina* ‘zart:NOM.SG.F’. Die Adjektiva einer Endung, mit nichtbelegtem Neutrum, bezeugen gut die fließenden Grenzen zwischen Substantiv und Adjektiv im Griechischen, z. B. *mákar* ‘selig:NOM.SG.M/F/N’, *phugás* ‘flüchtig:NOM.SG.M/F/N’. Das Altgriechische hat zwei allomorphe Komparativ- und Superlativbildungen: -iōn, -istos, z. B. *hēdiōn*, *hēdistos* ‘süßer, am süßesten’, und -teros, -tatos, z. B. *pistóteros*, *pistótatos* ‘treuer, am treuesten’. Das erste Suffix stellt das sog. primäre Suffix dar, das von der Wurzel gebildet wird, also nach der Beseitigung eines den Positiv charakterisierenden derivativen Suffixes angefügt wird, vgl. *hēd-ús* ‘süß-NOM.SG.M’, *hēd-iōn* ‘süß-KMPR.NOM.SG.M’, *hēd-istos* ‘süß-SUP.NOM.SG.M’. Im Gegensatz dazu sind *pistóteros*, *pistótatos* zu *pistós* ‘treu:NOM.SG.M’ (vgl. Adjektiv

mit demselben Stammsuffix -o wie *kakós* ‘häßlich:NOM.SG.M’, das die Komparationsformen *kak-iōn*, *kák-istos* aufweist) oder *chariés-teros*, *chariés-tatos* zu *charieis* ‘begabt:NOM.SG.M’ vom Stamm mit Suffix gebildet. Das Suffix -iōn ist gemeinsam für Maskulin und Feminin, -teros bildet wie alle Adjektiva auf -os die Femininform -téra, z. B. *pistotérā* ‘treuere’. Die Suffixe -teros, -tatos stellen produktive Steigerungsformen dar, -iōn, -istos kommen nur bei einer begrenzten Gruppe von Adjektiva vor.

Die Möglichkeit der Komparativbildung von Substantiven, wie *basileúteros* zu *basileús* ‘König’ oder *doulóteros* zu *dóulos* ‘Sklave’ (vlg. Schwyzer 1968: 536) ist wiederum Ausdruck der fließenden Grenze zwischen Substantiv und Adjektiv.

Auf dem Ausdruck der nominalen morphologischen Kategorien (Kasus, Numerus, Genus) beim Adjektiv beruht die sog. Kongruenz, die für die indogermanische Nominalphrase charakteristisch ist.

Das Adjektiv kann mit Hilfe des Artikels substantiviert werden, ohne Anwendung eines Derivationssuffixes. Auf diese Weise werden vor allem Neutra substantiviert, um abstrakte Begriffe wie *tò kalón* ‘ART.NOM/AKK.SG.N schön:NOM/AKK.SG.N (die Schönheit)’ zu liefern; *tà kalà* ‘ART.NOM/AKK.PL.N schön:NOM/AKK.PL.N (die schönen Sachen)’.

### 3.2. Pronomen

**Pronomina** drücken die nominalen Kategorien Kasus, Numerus und teilweise Genus aus, ihre Flexion ist aber von der der Substantiva und Adjektiva verschieden oder weist gewisse Besonderheiten auf.

Die Personalpronomina der ersten und zweiten Person und das in Tab. 123.1 angeführte Pronomen der dritten Person, das im Attischen nur als Reflexiv und in einfacher Form ziemlich selten verwendet wird, drücken keinen Genusunterschied aus, ähnlich wie es auch in anderen indogermanischen Sprachen der Fall ist.

Kennzeichnend ist die Unterscheidung betonter und unbetonter Formen und der besondere Nominativstamm, im Griechischen nur im Singular, deutlich bei *egō* ‘1.SG.NOM’. Die Segmentierung der Formen ist, wie so oft in flexivischen Sprachen, uneindeutig. Beim Nominativ und Akkusativ sind die Formen als Stammallomorphe mit -Ø-Endung zu verstehen (*su* gegenüber *se*), der Genitiv mit seines der nominalen 2. Deklination entsprechenden Endung -ou (6) unterstützt wieder

		1. Person	2. Person	Reflexiv
Sg.	Nom.	<i>egō</i>	<i>su</i>	—
	Gen.	<i>emoū, mou</i>	<i>soū, sou</i>	<i>hoū, hou</i>
	Dat.	<i>emoī, moi</i>	<i>soī, soi</i>	<i>hoī, hoi</i>
	Akk.	<i>emé, me</i>	<i>sé, se</i>	<i>hé, he</i>
Pl.	Nom.	<i>hēmeīs</i>	<i>humēīs</i>	<i>sphēīs</i>
	Gen.	<i>hēmōn</i>	<i>humōn</i>	<i>sphōn</i>
	Dat.	<i>hēmīn</i>	<i>humīn</i>	<i>sphīsi(n)</i>
	Akk.	<i>hēmās</i>	<i>humās</i>	<i>sphās</i>
Du.	Nom./Akk.	<i>nō</i>	<i>sphō</i>	
	Gen./Dat.	<i>nōjn</i>	<i>sphōjn</i>	

Tab. 123.1: Personalpronomina und Reflexiva

die Segmentierung (*e*)*m-ou* ‘1.SG-GEN’, *s-ou* ‘2.SG-GEN’ usw.

Numerusunterschiede sind bei den Personalpronomina nicht nur durch Endungen, sondern durch besondere Stämme markiert: *hēmeīs* ‘wir (Nom. Pl.)’, *humēīs* ‘ihr (Nom. Pl.)’, *hēmās* ‘uns (Akk. Pl.)’ *humās* ‘euch (Akk. Pl.)’, *nō* ‘uns beide (Nom./Akk. Pl.)’, *sphō* ‘ihr beide (Gen./Dat. Du.)’ usw. Darin äußert sich der Charakter der Numerusunterschiede innerhalb der Kategorie der Person (4.2.1). In der dritten Person sind die alten, genus-indifferenten Pronomina wie *he* ‘sich (Akk.)’ oder homerisch *min* ‘ihn/sie/es (Akk.)’ durch das genusunterscheidende Pronomen mit der Bedeutung ‘selbst’ ersetzt: *aut-ón* ‘ihn-AKK.SG.M’, *aut-én* ‘sie-AKK.SG.F’, *aut-ó* ‘es-AKK.SG.N’. Als Reflexivum fungiert dieses Pronomen in einer mit vorangehendem Stamm des Personalpronomens (für 1. und 2. Person Singular, vgl. Tab 123.1) oder Reflexivpronomens (vgl. Tab 123.1, für 3. Person Singular und Plural) zusammengesetzten Form, vgl. z. B. Dativ *em-autōj* ‘mir (selbst)’, *se-autōj* ‘dir (selbst)’, *he-autōj* ‘sich (selbst)’ (Dat. Sg.), *he-autoīs* ‘sich (selbst)’ (Dat. Pl.)’ usw.; im Plural der 1. und 2. Person sind die Formen juxtaponiert: *hēmīn autoīs* ‘uns (selbst)’ (Dat.)’ usw., auch in der 3. Person *sphīsin autoīs* ‘ihnen (selbst)’ (Dat.). Nominativ Plural *sphēīs* hat die Bedeutung ‘sie (selbst)’, lat. ‘ipsi’.

Alle außer den in Tab 123.1 angeführten Pronomina sind genusunterscheidend (vgl. 4.1). Sie sind entweder -o- (im Maskulin und Neutrum) und -ā- (im Feminin) Stämme und unterscheiden alle drei Genera, so Possessiva wie *emós* ‘mein’, *emé* ‘meine’, *emón* ‘mein’ usw., Demonstrativa, Artikel, *ho*, *hé*, *tó* ‘der,

die, das’, Relativum *hós*, *hé*, *hó* ‘welcher, welche, welches’, oder sie haben den Stammvokal -i und unterscheiden nur personale (Maskulin und Feminin) und nichtpersonale Formen (Neuter), so das Interrogativ- (betont) und Indefinitpronomen (enklitisch; s. 5) *tís* ‘wer’, *tis* ‘jemand’, *tí* ‘was’, *ti* ‘etwas’.

Die Flexion der pronominalen -o- und -ā- Stämme ist im Griechischen mit der nominalen 2. und 1. Deklination identisch, was teilweise dadurch gegeben ist, daß die nominalen -o- und -ā- Stämme einige Formen von der Pronominaldeklination übernommen haben. Abweichungen von der Nominaldeklination finden wir bei den Nominativformen, wo bei den Demonstrativa und dem Artikel sog. Stammsuppletivismus herrscht (vgl. bei Personalpronomina): Maskulin- und Femininformen des Singulars und Plurals haben einen vom übrigen Paradigma abweichenden Stamm, vgl. die Artikelformen: *ho* ‘ART. NOM.SG.M (der)’, *hé* ‘ART.NOM.SG.F (die)’, *hoī* ‘ART.NOM.PL.M (die)’, *hai* ‘ART.NOM.PL.F (die)’ gegenüber *toú* ‘ART.GEN.SG.M/N (des)’, *tōj* ‘ART.DAT.SG.M/N (dem)’. Weiter fehlt bei einigen -o-stämmigen Pronomina im Nominativ/Akkusativ des Neutrums die Endung -n: Artikel *tó* ‘ART.NOM/AKK.SG.N (das)’, *aut-ó* ‘es-NOM/AKK.SG.N’. Beim Artikel (und dem von ihm gebildeten Demonstrativ *hó-de* ‘dieser’) fehlt dem Nominativ des Maskulins die Endung -s. Die nichtnominativischen Formen des Interrogativums/Indefinitums *tís* ‘wer’/*tis* ‘jemand’ sind vom Stamm mit auslautendem -n gebildet und mit den Endungen der konsonantischen 3. Deklination versehen: Gen. Sg. *tín-os*, Dat. Sg. *tín-i*, Akk. Sg. *tín-a*. Das entspricht dem Verhältnis des Nominativs zu anderen Kasus des Singularparadigmas bei der

3. nominalen Deklination: Nominativ ohne Stammkonsonant und mit -s-Endung, übrige Kasus mit Stammkonsonant, vgl. *elpí-s* ‘Hoffnung-NOM.SG’, *elpid-os* ‘Hoffnung-GEN.SG’ oder *mélā-s* ‘schwarz-NOM.SG’, *mélanos* ‘schwarz-GEN.SG’. Die Form *tin* ist ursprünglicher Akkusativ (vgl. homerisches Personalpronomen *min* ‘ihn’), zu dem die Akkusativendung -a der konsonantischen Stämme hinzugefügt wurde (wahrscheinlich analogisch nach *hén-a* ‘ein-AKK.SG.M’, *oudén-a* ‘niemand-AKK.SG.M’), so daß das vorangehende Segment als Stamm interpretiert und verallgemeinert wurde. Das homerische Griechische besitzt noch die Genitiv- und Dativformen *teo*, *teōj* mit Stamm *te-*.

Der griechische **bestimmte Artikel** (beim Fehlen eines unbestimmten) dient der Substantivierung des Adjektivs (s. 3.1), des Infinitivs, des Adverbs (z. B. *hoi nún* ‘ART.NOM.PL jetzt (die heutigen Menschen)’), der präpositionalen Verbindung, des ganzen Satzes. Weiter können Adverbien oder präpositionale Verbindungen mit Hilfe des Artikels als Attribute verwendet werden (sog. Gelenk- oder Verbindungsfunction des Artikels, z. B. *hoi nún ánthrōpoi* ‘ART.NOM.PL jetzt Mensch: NOM.PL (die heutigen Menschen)'). Bei genügender morphologischer Bestimmung des Substantivs kommt es nicht zur Morphologisierung des Artikels. Über Disambiguierung mit Hilfe des Artikels vgl. unter 5.

Die Possessivität wird beim griechischen Pronomen auf zweifache Weise ausgedrückt. Possessivpronomen mit oder ohne Artikel, z. B. (*ho*) *emós patér* ‘mein Vater’, (*ho*) *huméteros patér* ‘euer Vater’, wird bei Emphasis gebraucht. Als grundlegender Ausdruck dient der Genitiv des Personalpronomens: (*ho*) *patér mou*, (*ho*) *patér humón*.

Für die pronominale Komposition sind außer der bei den Nominalkomposita üblichen Form mit erstem unflektierten und zweitem flektierten Glied (vgl. *he-autón*) zwei Formen charakteristisch. Es wird ein Partikel an das flektierte Pronomen angehängt, z. B. *hó-de* ‘ART.NOM.SG.M-PK’, oder es werden zwei flektierte Formen zusammengesetzt, z. B. *hóstis* ‘REL.NOM.SG.M-jemand.NOM.SG.M’. Als derivatisches Suffix ist das Suffix -*teros* zu erwähnen, das als Komparativsuffix bei den Adjektiven vorkommt. Bei den Pronomina ist die wohl ursprüngliche kontrastive Bedeutung (vgl. Wittwer 1969) erkennbar: *héteros* (abgeleitet von *heis* ‘einer’; s. 3.3) ‘ein anderer’, *ho héteros* ‘der eine von beiden, der andere von beiden’, *oudéteros* (von *oudeis* ‘kein,

keiner’) ‘keiner von beiden’, *póteros* (vom Interrogativstamm *po-*) ‘wer oder welcher von beiden’.

### 3.3. Zahlwort

Nur die **Kardinalia** 1–4 und ab 200 werden flektiert: *heis* (Mask.), *mía* (Fem.), *hén* (Neut.) ‘einer, eine, eines’ unterscheidet drei Genera und wird im Maskulin und Neutrum nach der 3. Deklination (*hen-ós* ‘einer-GEN.SG’ usw.), im Feminin nach der 1. Deklination flektiert; *dúo* ‘zwei’ ist genusindifferent und hat Dualflexion: (*dúo* (Nom./Akk.), *duoīn* (Gen./Dat.); *treís* (Mask./Fem.), *tría* (Neut.) ‘drei’ und *téttares* (Mask./Fem.), *tétara* (Neut.) ‘vier’ sind Plurale mit zwei Genusformen und flektieren nach der 3. Deklination. Die Kardinalia *diákósioi* ‘200’ und weiter flektieren als pluralische Adjektiva dreier Endungen (s. 3.1) nach der 1. und 2. Deklination.

Von den Kardinalia werden, mit gewissen Stammänderungen, mit Hilfe des Suffixes -*tos* **Ordinalia** abgeleitet: z. B. von *pénte* ‘fünf’ *pémp-tos*, *-tē*, *-ton* ‘fünfter, -te, -tes’, usw., die Adjektiva der 2. und 1. Deklination darstellen. Weitere zum Zahlwortzählende Bildungen sind adjektivische Multiplikativa wie *triplóus* ‘dreifach, dreifältig’, Proportionalia wie *tri-plásios* ‘dreimal soviel’, Adverbia wie *trís* ‘dreimal’, *pentákis* ‘fünfmal’, usw.

### 3.4. Unflektierte Wörter

Durch einen regelmäßigen, wortartverändernden Wortbildungsprozeß werden **Adverbien** zu Adjektiven gebildet, und zwar mit Hilfe des Suffixes -*os* ‘ADVR’, z. B. von *chalepós* ‘schwer-NOM.SG.M’ *chalep-ós* ‘schwer-ADVR’. Von Pronominalstämmen werden Adverbien durch verschiedenartige Suffixe oder Partikeln gebildet (so zu *állō-* ‘ander-’ *állō-thi* ‘anderswo’, *állō-then* ‘anderswoher’, *állō-te* ‘ein andermal, sonst’ usw.; vgl. auch 3.2). Als Steigerungsformen werden die Neutra der adjektivischen Steigerungsformen verwendet, und zwar im Singular für den Komparativ, z. B. *saphésteron* ‘deutlich:KMPR(ADVR) (deutlicher)’, im Plural für den Superlativ z. B. *saphéstata* ‘deutlich:SUP(ADVR) (am deutlichsten)’ zum Adverb *saph-ós* ‘deutlich-ADVR’, (vgl. Adjektiv *saphés* ‘deutlich:NOM.SG.M/F’, *saphés* ‘deutlich:NOM.SG.N’).

Die grammatischen Wörter (und zwar unflektierte ebenso wie flektierte: Artikel, einige Pronomina) weisen bestimmte morphonologische Charakteristika auf: Einsilbigkeit, Proklise und Enklise; vgl. 5). Das Griechische

ist reich an **Partikeln**: enklitische Partikeln wie *dé*, *gé*, *té* stehen nach dem ersten Worte oder der Wortgruppe im Satze oder im Kólon und dienen als Satzgliederungssignale, die Modalpartikel *án* spezifiziert die Bedeutung des Konjunktivs und Optativs (4.2.3) und fungiert als Marker der Irrealität beim Indikativ Präterits (besonders in Konditionalsätzen). Eine Sonderstellung der **Interjektionen** als Wortart ist dadurch gegeben, daß sie eine Satzaussage bilden können, und daß sie ferner als expressive Elemente in ihrer Lautgestalt von den übrigen Wörtern abweichen (Wiederholungssilben wie *eleleú* Kriegsruf bei Aeschylus, *attalattatá* Staunenruf bei Aristophanes) und oft onomatopoetisch sind, z. B. *pheú* ‘pfui’ (vgl. Schwyzer & Debrunner 1966: 601).

#### 4. Flexivisch ausgedrückte Kategorien des Nomens und Verbs

Die Semantik der Kategorien wird mit Rückicht auf die in 6 und 7 behandelten Ausdrucksstrukturen analysiert, und es können nur die für die indogermanische flexivische Struktur und das Altgriechische selbst (Optativ, Medium, Perfekt) besonders kennzeichnenden Tatsachen besprochen werden.

##### 4.1. Nominale Kategorien: Kasus, Numerus, Genus

Mit dem reichen Verbalsystem des Altgriechischen kontrastiert das **Kasussystem** mit nur 5 Kasus: Nominativ, Genitiv, Dativ, Akkusativ und Vokativ (gegen 8 Kasus im Altindischen und 6 im Lateinischen). Der Nominativ ist syntaktisch definiert als Subjektkasus. Alle drei übrigen Kasus (wenn wir den Vokativ als Kasus des Anrufs beiseite lassen) können Objekt- und Adverbialfunktionen ausüben, ihr gegenseitiges Verhältnis ist durch die strukturellen Merkmale – partiell, – marginal beim Akkusativ, + partiell beim Genitiv, + marginal beim Dativ gegeben. Das Griechische kennt keinen ausgesprochen adverbialen Kasus wie den lateinischen Ablativ, dessen Objektfunktion ganz marginal ist.

Beim **Numerus** treffen wir ein dreigliedriges System mit Singular, Dual und Plural, wobei der Dual im klassischen Attischen beim Nomen wie bei den kongruierenden Verbalendungen besonders regelmäßig verwendet wird. Er bezeichnet vor allem paarweise zusammengehörige Dinge (*tò ósse* ‘ART. NOM/AKK.DU Auge:NOM/AKK.DU (die beiden

Augen)’, *tò cheíre* ‘ART.NOM/AKK.DU Hand: NOM/AKK.DU (die beiden Hände)’) oder Personen (s. Beispiele in 9).

Das **Genus** bildet flexivische Oppositionen beim Adjektiv und Pronomen. Wenn sich das Pronomen oder Adjektiv (in substantivischer Stellung) nicht auf ein bestimmtes Substantiv, sondern auf eine Person oder Sache im allgemeinen bezieht, weist das Genus auf die natürliche Klasse der Personen gegenüber Sachen und innerhalb der Personen auf weibliche und männliche Personen hin (*hóde* ‘dieser’, *héde* ‘diese’, *tóde* ‘dieses’ – 3.3.; (*tò*) *agathón* ‘ART.NOM/AKK.SG.N gut:NOM/AKK.SG.N (das Gute, etwas gutes)’, (*hoi*) *agathoi* ‘ART. NOM.SG.M gut:NOM.SG.M (die guten Menschen)'). Sonst bezieht sich die lexikalische Klassifizierung durch Genusmerkmale auf die Substantiva, mit denen die Adjektiva und Pronomina kongruieren. Das substantivische Genus ist nur in begrenztem Umfang semantisch motiviert, besonders wieder bei den Personenbezeichnungen. Die Verteilung der Genera bei Sachbezeichnungen ist willkürlich, gewisse Regelmäßigkeiten und Präferenzen sind aber zu beobachten: Flüsse, Winde und Monate sind maskulin, Bäume, Länder, Inseln und Städte feminin.

Von den **Deklinationsklassen** ist nur die 1. oder *ā*-Deklination in ihrer Hauptunterklasse ohne *s*-Suffix im Nominativ Singular einem Genus reserviert, d. h. ausgesprochen femininisch. Nach ihr werden auch die Feminina zu den Adjektiven der 2. und 3. Deklination dekliniert (3.1) und auch die Feminina zu den Substantiven der 2. Deklination (wie *the-á* ‘Gott-NOM.SG.F (Göttin)’ zu *the-ós* ‘Gott-NOM.SG.M (Gott)’ usw.). Die zu der 1. Deklination zählenden Maskulina sind durch zwei unterscheidende Merkmale markiert, den Nominativ auf *-s* und den Genitiv auf *-ou*. Die 2. oder *o*-Deklination ist nicht so ausgesprochen den Maskulina vorbehalten, obwohl diese überwiegen, vgl. aber auch die *o*-Adjektiva zweier Endungen (3.1). Darin äußert sich der nichtmarkierte Charakter des Maskulins.

Numerus und Genus haben Kontrolle über die Kasuskategorie, indem sie ihre Markierung determinieren. Die Wahl des Kasusmarkers ist ikonisch in Bezug auf die Numerusmarkiertheit (s. 6). Das Genus determiniert den Ausdruck der grundlegenden Kasus in dem Sinne, daß die Neutra eine nichtdifferenzierte Nominativ/Akkusativform haben. Zugleich determiniert die Selektion des Neutrums die Form des Duals und Plurals, und

dies ist beim Plural teilweise eine semantische Formenvariation (2.1.2).

Eine Determinierungs- bzw. Kontrollbeziehung besteht also zwischen den der Wortkategorie inhärenten, auf lexikalischer Kategorisierung beruhenden grammatischen Kategorien und den relationellen, syntaktischen Kategorien. Das Genus ist ausgesprochen eine lexikalische, auf der Kategorisierung der Lexeme beruhende Kategorie. Der Numerus, der keine inhärenten Qualitätsmerkmale, sondern Quantität als externes Merkmal ausdrückt, nimmt in dieser Hinsicht eine Zwischenstellung zwischen Genus und Kasus ein.

#### 4.2. Verbale Kategorien

Auch beim Verb determiniert die auf lexikalischer Kategorisierung beruhende Kategorie des Aspekts den Ausdruck, bei Diathese und Tempus sogar das System der semantischen Merkmale. Das Tempus ist mit der Aspektkategorie so innerlich verbunden, daß wir im folgenden von aspekt-temporalen Flexionsstämmen sprechen. Tab. 123.2 zeigt das System der nach Flexionsstämmen geordneten flexivischen Kategorien und gibt zugleich über die in der Morphemübersetzung in diesem Artikel vorkommenden Bezeichnungen Auskunft (Perfekt und Imperfekt bezeichnen die Tempora Perfekt und Imperfekt, nicht die Aspekte perfektiv/imperfektiv; dem Aspekt bzw. Aktionsart nach ist Perfekt stativ/resultativ).

Die Kategorien Person und Modus sind an die Funktion des Verbs als Satzprädikat gebunden und werden nur beim **Verbum finitum**

ausgedrückt. Die Aspekte, das Futur und die Diathese werden auch beim **Verbum infinitum** ausgedrückt. Die so differenzierten Infinitive und Partizipien (vgl. 7.3) dienen in der griechischen Syntax in reichem Maße als subordinierende Mittel, die Nebensatzfunktionen vertreten können. Die Diathese ist in Tab. 123.2 nur beim Indikativ angegeben, andere Formen haben ein dem angegebenen Indikativ entsprechendes System; Konjunktiv Perfekt Mediopassiv und Optativ Perfekt Mediopassiv sind analytisch. Beim seltenen Futur Perfekt kommen Aktiv und Mediopassiv nur als Allomorphe, nicht oppositionell vor; siehe noch 7 über Futur und Diathese.

##### 4.2.1. Person

Das Morphem der **Person** wird einerseits beim Verb durch die Personalendung ausgedrückt, die im Griechischen und Indogermanischen die Subjektperson bezeichnet, d. h. es wird angegeben, wie das Subjekt des Verbs am Sprechakt beteiligt ist, andererseits durch die Personalpronomina (3.2), die nach Numerus und Kasus flektiert sind. Die Personalpronomina können neben der Personalendung auch in der Subjektsposition stehen, wenn das Merkmal Person hervorgehoben werden soll, in anderen Positionen stellen sie den einzigen Ausdruck der Person dar. Mit dieser besonderen Stellung des Subjekts des Personalpronomens hängt der Stammsuppletivismus zusammen. Der Ausdruck der Person ist im Altgriechischen nicht mit einer Unterscheidung des Geschlechts oder des Sozialstatus verbunden. Nur die Zahl der beteilig-

Aspekt- und Tempusmerkmale	Tempus (Ind.)	Diathese	Konj.	Opt.	Imp.	Inf.	Part.
+imperfektiv							
-präterital	Präsens	Akt./Medpass.	+	+	+	+	+
+präterital	Imperfekt	Akt./Medpass.	-	-	-	-	-
-imperfektiv							
futurisch	Futur	Akt./Med./Pass.	-	+	-	+	+
±imperfektiv							
-präterital	-	-	+	+	+	+	+
+präterital	Aorist	Akt./Med./Pass.	-	-	-	-	-
stativ/resultativ							
±präterital	Perfekt	Akt./Medpass.	+	+	+	+	+
+präterital	Plusquampr.	Akt./Medpass.	-	-	-	-	-
futurisch	Fut. Perf.	Akt./Medpass.	-	-	-	-	-

Tab. 123.2: Übersicht über das griechische Verbalsystem

ten Personen wird angegeben, und es wird wie beim Nomen nach Singular, Plural und Dual unterschieden. Inhaltlich ist aber der Nicht-Singular (Plural und Dual) innerhalb der Kategorie der Person vom Nicht-Singular der Nomina verschieden, da hier nicht eine Mehrzahl der durch das Singularlexem bezeichneten Individuen ausgedrückt wird, was besonders in der ersten Person deutlich ist (*hēmeis* ‘wir’ ist nicht eine Menge von *egō* ‘ich’). Damit in Einklang steht der Ausdruck des Plurals nicht nur durch Endungen, sondern durch besondere Lexeme (3.2).

### 4.2.2. Diathese

Die **Diathese** ist nicht auf die Opposition Aktiv : Passiv beschränkt, sondern umfaßt drei Glieder: Aktiv, Medium und Passiv. Das Medium wird direkt reflexivisch (*lou-omai* ‘wasch-PRÄS.MED.1.SG (ich wasche mich)’) und indirekt reflexivisch (*lou-omai tēn cheíra* ‘wasch-PRÄS.MEDPASS.1.SG ART.AKK.SG.F Hand:AKK.SG.F (ich wasche mir die Hand)’) gebraucht und semantisch dem Aktiv in verschiedenen Oppositionen gegenübergestellt (vgl. Perel'muter 1977: 137 ff.; 1995: 3 f., 66; Rijksbaron 1994: 131–160; Strunk 1980). Diese können unter dem Begriff der internen, zentripetalen Diathese zusammengefaßt werden, die dem äußeren Verhältnis des Agens zur Verbalaktion beim Aktiv gegenübersteht: (Der griechischen grammatischen Tradition entsprechend wird das Verb in der 1. Person Singular zitiert.) *gam-ō* ‘heirat-PRÄS.AKT.1.SG (heirate)’, *gam-oúmai* ‘heirat-PRÄS.MEDPASS.1.SG (mich verheirate)’; *phob-ō* ‘erschreck-PRÄS.AKT.1.SG (erschrecke)’, *phob-oúmai* ‘erschreck-PRÄS.MEDPASS.1.SG (fürchte)’; *poreú-ō* ‘in.Bewegung.setz-PRÄS.AKT.1.SG (setze in Bewegung)’, *poreú-omai* ‘in.Bewegung.setz-PRÄS.MEDPASS.1.SG (gehe)’.

Für die Merkmale 'intern, zentripetal' sind auch die zahlreichen nichtoppositionellen Media, sog. Media tantum (MT), markiert, z. B. *boúl-omai* 'woll-PRÄS.MT.1.SG (wolle)', *dérk-omai* 'seh-PRÄS.MT.1.SG' (blicke, sehe)' *mách-omai* 'kämpf-PRÄS.MT.1.SG (kämpfe)'.

Das Passiv, formell vom Medium nur im Aorist und teilweise im Futur unterschieden, ansonsten mit dem Medium Bestandteil des sog. Mediopassivs, kann semantisch mit dem Medium als interne Diathese aufgefaßt werden, syntaktisch hebt es sich aber vom Medium durch die Operation der Passivierung (vgl. Art. 108) ab. Passiviert werden können im Griechischen auch intransitive Verba mit Genitiv- (vgl. 1 a–1 b) oder Dativobjekt (vgl.

2 a–2 b), vereinzelt auch indirekte Dativobjekte in Konstruktionen wie in (3 a–b). Das impersonale Passiv von intransitiven Verben wie lat. *itur* ‘man geht’ gibt es im Griechischen nicht.

- (1) (a) *árch-omai*  
       hersch-PRÄS.MEDPASS.1.SG  
       ‘ich werde beherrscht’

(b) *árch-ō*                                      *tin-ós*  
       hersch-PRÄS.AKT.1.SG jemand-GEN.SG  
       ‘ich herrsche über jemanden’

(2) (a) *phthon-oúmai*  
       beneid-PRÄS.MEDPASS.1.SG  
       ‘ich werde beneidet’

(b) *phthon-ō*                                      *tin-i*  
       beneid-PRÄS.AKT.1.SG jemand-DAT.SG  
       ‘ich beneide jemanden’

(3) (a) *epitáss-omai*  
       auftrag-PRÄS.MEDPASS.1.SG  
       *ti*  
       etwas:NOM/AKK.SG  
       ‘ich bin mit etwas beauftragt’

(b) *epitáss-ō*  
       auftrag-PRÄS.AKT.1.SG  
       *tin-i*                                      *ti*  
       jemand-DAT.SG etwas:NOM/AKK.SG  
       ‘ich trage jdm. etwas auf’

Die unregelmäßige Verteilung der diathetischen Merkmale nach den einzelnen Flexionsstämmen, Tempora und lexikalischen Verbklassen ist ein typischer Zug des griechischen Verbalsystems mit seiner Selbständigkeit der Flexionsstämme und Formenvariation: Passiv als besondere Kategorie nur beim Aorist und Futur vertreten, semantische Variation von diathetischem Charakter beim Aorist (2.1.2), Media tantum (hier oben), mediale und passive Futura (7).

### 4.2.3. Modus

Für das Altgriechische ist charakteristisch, daß dem merkmallosen Indikativ neben dem Imperativ als Modus des Befehls zwei merkmalhafte **Modi**, Konjunktiv und Optativ, gegenüberstehen. Der Konjunktiv wird als Modus des Willens in Wunschsätzen adhortativ (*iōmen* ‘geh:PRÄS.KONJ.1.PL (läßt uns gehen!)’) und prohibitiv (mit Negation *mē*: *mē moi thorubēsēte* ‘NEG ich:DAT lärm:AOR. KONJ.2.PL (fangt mir nicht an zu lärmen!)’) und in Fragesätzen (deliberativ: *tí poiōmen?* ‘was:NOM.AKK.SG tu:PRÄS.KONJ.1.PL (was sollen wir tun?)’) gebraucht. Als Modus der Erwartung (sog. prospektiver Konjunktiv) und

Eventualität wird der Konjunktiv im Attischen (mit der Partikel *án* versehen) nicht mehr im selbständigen Aussagesatz, sondern im Satzgefüge gebraucht. Der Optativ ist Modus des Wunsches und der Potentialität (mit Partikel *án*). In bestimmten Nebensatztypen tritt der sog. *optativus obliquus* an die Stelle des Indikativs oder Konjunktivs, wenn das Hauptverb die Vergangenheit bezeichnet. Besonders im Konditionalsatz werden durch das griechische Modussystem, in dem mehrere Kombinationen der aspekto-temporalen und modalen Merkmale vertreten sind, feine Unterschiede ausgedrückt (vgl. Funk 1985). Obwohl wegen des Namens (und aufgrund sprachhistorischer Gesichtspunkte) der Optativ als Besonderheit des Griechischen betrachtet wird, weist gerade der Konjunktiv Verwendungen auf, die von "Standard Average European" abweichen: der sog. prospektive Konjunktiv zur Bezeichnung der Zukunft steht in Kontexten, wo moderne europäische Sprachen den Indikativ Futur verwenden, in den restriktiven Relativsätze erscheint der Konjunktiv, wenn es um die Bezeichnung des Idealen, Nichtfaktischen geht.

#### 4.2.4. Aspekt und Tempus

Präsens-, Aorist- und Perfektstamm sind **Aspektstämme** mit eigenen Konjunktiven, Optativen, Infinitiven und Partizipien. Durch die Kombination des Aspektsmerkmals mit modalen, temporalen (Vergangenheit beim Indikativ aller drei Aspektstämme, Futur beim neutralisierten präsento-aoristischen und beim Perfektstamm) und Diathesenmerkmalen ist die Semantik der morphologischen Formen des Verbalsystems determiniert. Präsens- und Aoriststamm drücken die Opposition imperfektiv : perfektiv (in der Tab. 123.3 als +imperfektiv : -imperfektiv bezeichnet, da der Aorist unmarkiertes Glied der Opposition darstellt – vgl. Ruipérez 1954), also grundlegende Aspektsmerkmale (vgl. Art. 109) aus. Das Perfekt als ursprüngliche Zustandskategorie wird im Attischen auch bei transitiven Verba verwendet, doch hier geht es auch um die Charakterisierung des Subjekts und seines Zustands (vgl. Rijksbaron 1994: 33–36 mit weiterer Literatur und Polemik gegen die traditionelle Auffassung des sog. resultativen Perfekts als auf den Zustand des Objekts bezogen). Von einigen Forschern werden die formalen Übereinstimmungen des Futurs mit den grundlegenden Aspektstämmen hervorgehoben, d. h. Bildung eines eigenen Partizips, Infinitivs und Optativs, so daß

das Futur den aspekto-temporalen Stämmen zugeordnet (so auch Rijksbaron 1994) oder sogar als Aspekt definiert wird (so Rix 1976: 196). Vielmehr steht das Futur semantisch wie formal dem Modus nahe (so Chanet 1985: 60), und sein Infinitiv und Optativ werden begrenzt als Formen der indirekten Rede gebraucht.

## 5. Morphologische Prozesse; Homonymie

Grundlegend in beiden Bestandteilen der Morphologie, der Flexion und der Wortbildung, ist die **Suffigierung**, die von Vokalwechsel, dem sog. Ablaut, und Akzentwechsel begleitet ist. Die vor der Endung stehenden Suffixe beteiligen sich mit ihrer Semantik entweder an der Wortbildung oder an der Flexion, oder sie sind semantisch leer und dienen als Marker der Deklinationsklassen bei Nomina und der Aspektstämme bei Verba. In der flexivischen Endung werden mehrere semantische Merkmale kumuliert (vgl. unter 2, 6 und 7). Der Ablaut ist quantitativ (Dehnstufe *ē* – Normalstufe *e* – Schwundstufe *Ø*, Normalstufe *ei* – Schwundstufe *i* usw.) und qualitativ (*e* – *o*, *ē* – *ō*) und kann in der Endung sowie im Stamm vorkommen. Meistens ist er an der Flexion und Derivation als Begleiterscheinung beteiligt: vgl. *patér* 'Vater(NOM.SG)', *patér-a* 'Vater-AKK.SG', *patrós* 'Vater-GEN.SG'. Nur marginal dient der Ablaut (bzw. auch Akzent) allein zur Unterscheidung der Wortformen: Nominativ gegenüber Vokativ, vgl. Nom. Sg. *patér*: Vok. Sg. *páter* (vgl. noch unter 6), oder Imperfekt gegenüber Aorist, vgl. *é-leip-on* 'PRÄT-lass: IMPF-1.SG': *é-lip-on* 'PRÄT-lass:AOR-1.SG'.

Auch der **Akzent** kann als einziger Unterscheidungsmerkmal erscheinen. Durch die Art des Akzents werden zufällige Homonymien der zu verschiedenen Paradigmen gehörenden Wortformen vermieden, so z. B. *tuchón* Genitiv Plural des Substantivs *túchē* 'Schicksal' und *tuchón* Partizip Aorist des Verbs *tun-chánō* 'treffe' (mit Nasalprefix im Präsens). Durch die Akzentstelle werden Infinitiv Präsens und Infinitiv Aorist unterschieden in Fällen wie *líteshai* gegenüber *litésthai* vom Verb *lísso-mail/lítomai* 'bitte'. In der Wortbildung unterscheiden sich Adjektiva von Substantiven durch die Akzentstelle, so *tómos* 'abgeschnittenes Stück' gegenüber *tomós* 'schneidend, scharf' oder durch Akzentstelle + Akzentart, so *leúkos* 'Weißfisch' gegenüber

*leukós* ‘weiß’ (vgl. Schwyzer 1968:420; Bornemann & Risch<sup>2</sup>1978: 8). Der morphonologische Akzentwechsel gilt für die Nominalparadigmata (einschließlich der infiniten Verbaformen), bei den Verba finita wird im Griechischen der Akzent so weit zum Anfang des Wortes zurückgezogen, wie es die griechischen Akzentregeln (nicht weiter als in der drittletzten Silbe) ermöglichen. Die grammatischen Wörter werden durch Proklisis (Artikel, einige Präpositionen und Konjunktionen, Negation *ou*) oder Enklisis (Personalpronomina, indefinite Pronomina und Adverbia, Partikeln usw.) gekennzeichnet, die entweder als Tonlosigkeit oder als Regelung innerhalb der gesamten Akzentgruppe erscheint.

Auch die **Reduplikation** zählt zu den wichtigen basisverändernden Prozessen, in der Form CI-C, mit dem Vokal *-i-*, ist sie der Marker des Präsensstammes (7.1.; Giannakis 1997), z. B. *dí-dō-mi* ‘RDP:PRÄS-geb-PRÄS.1.SG (ich gebe)’ zur Wurzel *dō-*, in der Form CE-C, mit Vokal *-e-* ist sie als Marker des Perfektstammes grammatisiert: *dé-dō-ka* ‘RDP:PF-geb-PF.1.SG (ich habe gegeben)’.

**Präfigierung** ist in der Flexion auf das Augment (s. 7) beschränkt, und auch in der Wortbildung treffen wir nur wenige Präfixe, besonders die Negationspräfixe *a-* und *dus-*. Dabei sind die Grenzen zwischen Präfigierung und Komposition fließend (Debrunner 1917 zählt auch die genannten Negationspräfixe zur Komposition), die Präverbierung ist aber als Bestandteil der Komposition zu verstehen. Infigierung kommt nur bei Präsensbildungen mit Nasal wie *tu< n>chán-ō* ‘zufällig.sein<PRÄS>-1.SG (ich bin zufällig)’ – *é-tuch-on* ‘PRÄT-zufällig.sein(AOR)-1.SG’ vor, vgl. 7.1.

Obwohl die Anzahl der Phoneme, die sich am Ausdruck der grammatischen Merkmale beteiligen, beschränkt ist, indem im Auslaut des griechischen Wortes von den Konsonanten nur *s*, *n* und *r* stehen können, ist die Homonymie der morphologischen Formen nicht groß, was durch die komplizierte Struktur des flexivischen Wortes gegeben ist. Die Homonymie der Formantien ist beträchtlich, wird aber als solche zum Teil nicht wahrgenommen, besonders wenn es sich um Formantien von unterschiedlichem Rang handelt, z. B. Endung und Suffix im Falle von *men* als Formans der 1. Person Plural *tí-the-men* ‘RDP:PRÄS-setz-PRÄS.AKT.1.PL’ und als Formans des mediopassiven Partizips *tithé-men-os* ‘setz-PART.PRÄS.MEDPASS-M.SG’.

Stärker fallen homonyme Endungen auf, auch wenn sie bei unterschiedlichen Wortarten und Wortklassen vorkommen: *s* als Marker des Nominativs Singular und der 2. Person Singular Indikativ und Imperativ (*hál-s* ‘Salz-NOM.SG’, *títhē-s* ‘RDP:PRÄS-setz-PRÄS.IND.2.SG’, *thé-s* ‘setz-AOR.IMP.2.SG’) oder *e* als Marker des Imperativs, des Vokativs und des Duals (*bál-e* ‘werf-AOR.IMP.2.SG’, *lúk-e* ‘Wolf-VOK.SG’, *daimon-e* ‘Dämon-NOM/AKK.DU’). Die Homonymie der Endungen innerhalb eines Deklinationsparadigmas führt in einigen Fällen zur Homonymie ganzer Formen, so *chór-ās* ‘Erde-GEN.SG/AKK.PL’, *pól-ei* ‘Stadt-DAT.SG/NOM/AKK.DU’. In diesen Fällen, ebenso wie in den unter 6 angeführten Homonymien *noús*, *neőj*, *neőn*, handelt es sich um Homonymien zwischen unterschiedlichen Kasus der verschiedenen Numeri. Störender ist die ziemlich verbreitete Homonymie des Nominativ/Vokativ Plural und Akkusativ Plural bei der 3. Deklination: *pól-eis* ‘Stadt-NOM/VOK/AKK.PL’, *dunám-eis* ‘Kraft-NOM/VOK/AKK.PL’.

Hier dient besonders der Artikel (Nominativ Plural *hoi* : Akkusativ Plural *toús*) zur Disambiguierung. Von anderer Art ist die systematische, semantisch motivierte Homonymie, die auf der Neutralisierung bestimmter Kasusmerkmale beruht, so die teilweise Homonymie des Nominativs und Vokativs und die vollständige des Nominativs/Akkusativs und Genitivs/Datifs beim Dual. Beim Verb ist die Homonymie der Endung durch Augment oder andere Mittel beseitigt: aspektotemporaler Flexionsstamm wie bei *é-leip-on* gegenüber *é-lip-on* (s. oben) usw., Augment wie bei *paideú-omen* ‘erzieh-PRÄS/IMPF.1.PL’, *e-paideú-omen* ‘PRÄT-erzieh-PRÄS/IMPF.1.PL’, *paideú-ete* ‘erzieh-PRÄS/IMPF.2.PL’, *e-paideú-ete* ‘PRÄT-erzieh-PRÄS/IMPF.2.PL’ oder Akzent (s. hier in 5 oben). Systematische Homonymie finden wir beim Imperativ und Indikativ der 2. Person Plural des Präsens: *paideú-ete* ‘erzieh-PRÄS.IND/IMP.2.PL’.

## 6. Flexion des Nomens

Durch die kumulativen Nominalendungen werden insgesamt 15 Kasus/Numerus-Kombinationen ausgedrückt, wobei die Morphemkombinationen Nominativ/Akkusativ/Vokativ Dual, Genitiv/Dativ Dual und Nominativ/Vokativ Plural systematisch homonym sind, so daß das Deklinationsparadigma maximal 11 Formen umfaßt. Die Neutra, wo

ferner der Nominativ mit dem Akkusativ sowie auch im Singular mit dem Vokativ identisch ist, unterscheiden 8 Paradigmaformen.

Man unterscheidet im Griechischen drei Deklinationen, nämlich die 1. oder *ā-*, die 2. oder *o-* und die 3. Deklination, welche die Konsonantenstämme sowie die *i-*, *u-* und Diphthongstämme umfaßt.

Deklination:	1.	2.	3.
Singular			
Nom.	-ēl/āl/ā	-os	-s/Ø
Gen.	-ēs/ās	-ou	-os
Dat.	-ēj/āj	-ōj	-i
Akk.	-ēn/ān/ān	-on	-a/n
Vok.	-ēl/āl/ā	-e	-Ø/s
Plural			
Nom.	-ai	-oi	-es
Gen.	-ōn	-ōn	-ōn
Dat.	-ais	-ois	-si
Akk.	-ās	-ous	-as/s
Vok.	-ai	-oi	-es
Dual			
Nom./Akk./Vok.	-ā	-ō	-e
Gen./Dat.	-ain	-oin	-oin

Tab. 123.3: Übersicht der Kasusendungen

Bei den ersten zwei Deklinationen unterscheidet man bei der synchronischen Beschreibung die Basis, die nicht mehr den kennzeichnenden Stammvokal enthält, und die Endung, die ein Produkt der Fusion des Stammvokals mit der eigentlichen Endung darstellt. Der vokalische Deklinationsmarker *-ē*, *-ā* (*-ā*) (zur Verteilung vgl. 2.1.1) ist in allen Formen der 1. Deklination außer dem Genitiv Plural *-ōn* als Bestandteil der Endungen vorhanden; der immer (abgesehen von der Akzentklasse) zirkumflektierte, auf Kontraktion hinweisende Genitiv, ist jedoch Zeichen der 1. Deklination. Über die maskuline Unterklasse der 1. Deklination s. 4.1. In der 2. Deklination steht *-e* im Vokativ (ursprünglich reine Stammform), *-o* (*-ō*) in den anderen Kasus, und zwar in unserer transliterierenden Notation in allen Kasus, doch war die Aussprache von *-ou* (Genitiv Singular) im Attischen des 5. Jh. bereits [u:]. Es ist nicht möglich, durch Zerlegung der Endungen der 1. und 2. Deklination Minimalendungen zu identifizieren, die mit der 3. Deklination gleich wären. Dies beruht nicht nur auf der Fusion des Stamm-

vokals mit vokalisch anlautenden Endungen (z. B. Dativ Singular *-ōj* < \**o-ei*), sondern auch darauf, daß diese Endungen vom Ursprung her verschieden sind. Beim Dativ Singular handelt es sich um Ablautvarianten desselben Kasussuffixes, nämlich Dativ/Lokativ *-ei* (2. Deklination)/*-i* (3. Deklination); im Genitiv Singular und Nominativ Plural sind *-ou*, *-oi*, *-ai* Allomorphe, die ursprünglich für die pronominale Deklination kennzeichnend waren; im Dativ Plural *-ois*, *-ais* gegenüber *-si* handelt es sich um allomorphe Varianten innerhalb der Nominaldeklination. Nur im Akkusativ Singular folgt nach dem Deklinationsmarker eine Endung *-n* (*-on*, *-ēn*, *-ān*, *-ān*), die in der 3. Deklination den Akkusativ Singular markiert, doch gilt diese Identität nur für die vokalisch auslautenden Stämme der 3. Deklination. Die Tatsache, daß die flexivische Morphologie im Unterschied zur Agglutination nicht auf der Hinzufügung eines bestimmten Segmentes beruht ist durch diese Gegebenheiten klarlegt. Deshalb kann die in Tab. 123.3 gegebene Übersicht der grundlegenden Endungen keine erschöpfende Information über die Kasusformen geben u. a. über die Verhältnisse im Akkusativ Plural der 3. Deklination, wo bei den vokalischen Stämmen zur Ersatzdehnung des der Endung *-s* (<*ns*) vorangehenden Vokals kommt und wo bei einigen Nomina und Adjektiva der Akkusativ Plural dem Nominativ Plural gleicht (vgl. in 5).

Die 3. Deklination umfaßt Stämme mit unterschiedlichem Stammauslaut, wobei der Akkusativ Singular nach Konsonant *-a*, nach Vokal *-n* lautet. Bei den vokalischen und *-s*-Stämmen sind durch Schwund des intervokalischen *i*, *u*, *s*, Vokalkontraktion und Metathese der Vokallängen (vgl. *basile-ōs* < *basile-os* ‘König-GEN.SG’ usw.) Formen entstanden, die eine Stammallomorphie und unklare Grenze zwischen Stamm und Endung (vgl. *pólei* ‘Stadt:DAT.SG’) aufweisen und mit ihren Endungen von den in Tab. 123.3 angeführten abweichen: *basileū-s* ‘König-NOM.SG’, *basile-ōs* ‘König-GEN.SG’, *basile-ā* ‘König-AKK.SG’; *póli-s* ‘Stadt-NOM.SG’, *póle-ōs* ‘Stadt-GEN.SG’, *pólei* ‘Stadt:DAT.SG’, *Dēmosthénēs* ‘Demosthenes:NOM.SG’, *Dēmosthénous* ‘Demosthenes:GEN.SG’, *Dēmosthénē* ‘Demosthenes:AKK.SG’.

Auch innerhalb der 2. Deklination haben wir eine kontrahierte Unterklasse (*noūs* ‘Sinn:NOM.SG’, *noū* ‘Sinn:GEN.SG’, *noūn* ‘Sinn:AKK.SG’) und eine Unterklasse mit durchgehendem *-ō*-Vokalismus der Endungen, die sog. attische Deklination (durch Metathese

der Vokallänge entstanden): *neōs* < *neōs* 'Tempel:NOM.SG', *neō* 'Tempel:GEN.SG', *neōn* 'Tempel:AKK.SG'. Die sich daraus ergebenden Homonymien betreffen in beiden Unterklassen immer verschiedene Singular- und Pluralkasus: *noūs* 'Sinn:NOM.SG/AKK.PL', *neōj* 'Tempel:DAT.SG/NOM.PL', *neōn* 'Tempel:AKK.SG/GEN.PL'.

Morphonologisch bestehen die Kasusendungen (s. Irigoin 1972) aus einem Phonem im Falle der Singularendungen der 3. Deklination (-s, -n, -a, -i) und weiter beim Vokativ Singular der 2. Deklination (-e), Nominativ/Vokativ Singular der 1. Deklination und Nominativ/Akkusativ Dual. Andere Endungen bestehen aus der Verbindung eines Vokals mit den Konsonanten -s, -n oder mit -i, -u als zweitem Element eines Diphthongs. Stärker markierte Form, nämlich die Verbindung des Diphthongs mit einem Konsonanten, treffen wir beim Akkusativ Plural der 2. Deklination -ous und beim Dativ Plural und Dual -ois, -ais, -oin, -ain; der Dativ Plural hat in der 3. Deklination als einziger Kasus eine sich vom übrigen Paradigma abhebende Form, nämlich die Verbindung eines Konsonanten mit folgendem Vokal (-si).

Als Regel gilt, daß die Pluralformen immer merkmalhaltiger sind als die Singularformen der betreffenden Kasus. Diese Ikonizität der formalen und semantischen Merkmalhaltigkeit kennt als einzige Ausnahme den Nominativ der 2. und der maskulinen Unterkasse der 1. Deklination. Hier ist der Nominativ Singular -os, -as merkmalhaltiger als der Nominativ Plural auf -oi, -ai oder steht auf der gleichen Stufe der Markiertheit, wenn wir das zweite Element des Diphthongs strukturell dem Konsonanten gleichsetzen. Diese Stellung des Nominativs Singular dürfte durch die Markiertheit der prototypischen Subjekte mit den Merkmalen 'belebt, individuell, maskulin' motiviert sein; die maskuline 1. Deklination ist den Personenbezeichnungen vorbehalten (vgl. *neanīas* 'Jüngling', *polītēs* 'Bürger'), bei der 2. Deklination sind sie zahlreich vertreten (vgl. *ánthrōpos* 'Mensch', *doułos* 'Sklave'). Bei der 3. Deklination ist die Stellung des Nominativs Singular durch bloßes -s bzw. -Ø nicht hinlänglich charakterisiert. Bei mehreren Unterklassen hat der Nominativ Singular einen von den übrigen Kasusformen verschiedenen Stamm, der durch Dehnung und/oder Fehlen des (der) stammauslautenden Konsonanten charakterisiert ist, vgl. Nominativ Singular mit -s-Endung *gigā-s* 'Riese-NOM.SG' gegenüber *gigant-os* 'Riese-GEN.SG',

*elpi-s* 'Hoffnung-NOM.SG' gegenüber *elpid-os* 'Hoffnung-GEN.SG', mit -Ø-Endung *geitōn* 'Nachbar(NOM.SG)' gegenüber *geiton-os* 'Nachbar-GEN.SG', *patér* 'Vater(NOM.SG)' gegenüber *patr-ós* 'Vater-GEN.SG' (mit Nullstufe des Stammes) und *patér-a* 'Vater-AKK.SG' (mit Grundstufe). Durch fehlende Dehnung (und durch Akzent), bzw. durch -s-Losigkeit ist der Vokativ vom Nominativ Singular unterschieden: *gigan* 'Riese (VOK.SG)', *geiton* 'Nachbar(VOK.SG)', *páter* 'Vater(vok.SG)'.

## 7. Verbalflexion

An der Flexion des Verbs sind die morphologischen Charakteristika der aspekto-temporalen Flexionsstämme (7.1), Endungen, in denen die eigentlichen Personmarker (auslautende Minimalendungen) mit den Markern anderer Kategorien kumuliert sind (7.2), und das sog. Augment, nämlich das Präfix *e-* als Marker des Präteritums, beteiligt. Die einzelnen semantischen Merkmale werden durch Kombination dieser Formantien ausgedrückt. Z. B. wird das Präteritum nicht nur durch das Augment, sondern auch durch den Typ der Endungen (7.2) ausgedrückt. Innerhalb der flexivischen Endungen lassen sich die Marker der einzelnen Merkmale nicht immer eindeutig trennen, und auch die Grenzen zwischen Stamm und Endung sind nicht eindeutig. Der Unterschied zwischen den Suffixen, die einen Flexionsstamm bilden, und denen, die als Bestandteil der Endung zu beschreiben sind, ist graduell. Die Suffixe können nämlich, nach bestimmten Regeln, zyklisch appliziert werden, so daß eine Suffixbildung eine Basis für die Applizierung der weiteren Suffixe darstellt. So werden *s* für den aktiven sigmatischen Aorist und *thē* bzw. *ē* für den passiven bzw. intransitiven Aorist als flexionsstammbildende Suffixe angesehen, da sie wie alle anderen unter 7.1 besprochenen Aoristmarker eine Basis für die Applizierung der Konjunktiv- und Optativmarker und, im Falle des passiven Aorists, auch des Futurmarkers bilden. Andererseits wird der Marker des Futurs zusammen mit denen des Konjunktivs und Optativs als Bestandteil der Endung (7.2.2) angesehen. Entscheidend dafür ist, außer semantischen Aspekten (4.2.4), daß das Futur zwar eine Basis für die Applizierung des Optativs (nicht aber des Konjunktivs) darstellt, aber ebenso wie Konjunktiv und Optativ von mehreren Flexionsstämmen gebildet wird (Futur Aktiv siehe 7.1; Futur Passiv, Perfekt Futur 7.2.2).

Die formale Komplexität und Mannigfaltigkeit des griechischen Verbs, die unter 7.1–7.3 nur angedeutet werden kann, ist zugleich mit einer starken Uneinheitlichkeit der Verbalparadigmata verbunden, die jeweils eine sehr verschiedene Zahl von Oppositionen umfassen. Das für die regelmäßigen Verba vom Typ *paideiō* ‘erziehe’ geltende morphologische Paradigma in Tab. 123.3 ist durch Restriktionen und Neutralisierungen auf der einen Seite, auf der anderen aber auch durch nicht voll paradigmatisierte Oppositionen (2.1.2) bei einzelnen semantischen Verbgruppen beeinträchtigt, und zwar besonders im Bereich der Diathese. Eine recht zahlreiche Gruppe von Verben mit mediopassivem Präsens (ob Media tantum oder oppositionell zu Aktiv) hat das Perfekt in der aktiven Form und/oder den Aorist auf *-ēl/thē*, bzw. einen starken thematischen Aorist, was durch den intransitiven Charakter dieser Perfekt- und Aoristbildungen zu erklären ist: vgl. (in 1. Person Singular) Präs. (Medium tantum) *dér-komai*, Pf. *dédorka*, Aor. *édrakon* ‘sehe’; Präs. (Mediopassiv) *phainomai*, Pf. *péphēna*, Aor. *ephánēn*. Die mediale Form des Futurs (*paideiū-so-mai* ‘erzieh-FUT.MED-1.SG’) bietet ein gutes Beispiel dafür, wie eine morphologische Form in verschiedenen semantischen Oppositionen stehen und unterschiedliche Geltung haben kann. Das mediale Futur steht oft in aktiver Bedeutung, also als Medium tantum auch in Fällen, wo das Präsens aktiv ist, z. B. *gelá-so-mai* ‘lach-FUT.MED-1.SG’ zu *gelō* ‘lach: PRÄS.AKT.1.SG’ (kontrahiert aus *gelá-ō*). Es ist regelmäßig dem Futur Passiv als mediale Form entgegengesetzt, bei einer Gruppe von Verben hat sie passivische Bedeutung wie bei *timō* ‘ehre’ (kontrahiert aus *timá-ō* ‘ehr-PRÄS.AKT.1.SG’), *timē-so-mai* ‘ehr-FUT-MEDPASS.1.SG (ich werde geehrt werden)’, ist also allomorphisch zum Futur Passiv, das bei *timáō* auch belegt ist: *timē-thē-so-mai* ‘ehr-FUT-PASS-1.SG’.

Die formalen und semantischen Unregelmäßigkeiten bzw. die nicht durchgeführte Paradigmatisierung macht das griechische Verb zu einer Domäne für die Zusammenarbeit von Linguistik und Didaktik (vgl. besonders Chanet (1985), Zeegers Vander Vorst (1987) und andere Beiträge in der Zeitschrift *Cratyle*).

### 7.1. Flexionsstämme

Die Opposition der grundlegenden Präsens- und Aoriststämme wird entweder durch morphologische Charakteristika auf beiden Seiten ausgedrückt, oder einer der Stämme ist

merkmallos. Der Präsensstamm ist besonders durch Suffix *-io* (Köstling 1998), nasale Suffixe und Infix und durch die Reduplikation (Giannakis 1997) charakterisiert. Vgl. folgende Beispiele (in der 1. Person Singular): Präsens mit Reduplikation und Suffix *gi-gnō-skō* – merkmalloser Aorist *é-gnō-n* ‘erkenne’; suffixales Präsens *kám-nō* – merkmalloser Aorist *é-kam-on* ‘ermüde’; Präsens mit nasalem Suffix und Infix *la<m>bá-nō* – merkmalloser Aorist *é-lab-on*; sich durch Ablaut unterscheidende Flexionsstämme Präsens *leip-ō* – Aorist *é-lip-on* ‘lass’; suffixales Präsens *deik-nu-mi* – suffixaler sog. sigmatischer Aorist *é-deik-s-a* ‘zeige’; merkmalloses Präsens *paideú-ō* – sigmatischer Aorist *e-paideu-s-a* ‘erziehe’. Der Aorist hat ferner einen besonderen Passiv- bzw. Intransitiv/Passivstamm auf *thē* oder *ē* (*e-paideú-thē-n* ‘PRÄTERIEZIEH-AOR.PASS.1.SG’). Das ausgeprägteste Charakteristikum des Perfektstammes ist die Reduplikation (mit Ausnahme des Perfektuums tantum *oīda* ‘weiß’ in allen Perfekten vorhanden), dazu kommt beim sog. schwachen Perfekt ein besonderer Marker *-k-* im Aktiv (*pe-paideu-k-a* ‘RDP-erzieh-PF.AKT-PF.AKT:1.SG’), beim sog. starken Perfekt meistens Ablaut (*lé-loip-a* ‘RDP-lass:PF.AKT-1.SG’ zu *leip-ō*), oder/und Aspiration (*pé-präch-a* ‘RDP-tu:PF.AKT-1.SG’ zu *práttō*, *pé-pomph-a* ‘RDP-schick: PF.AKT-1.SG’ zu *pémpō*).

Das aktive Futur hat bei mehreren Verben einen von Präsens und Aorist verschiedenen Stamm, indem es die Präsenssuffixe nicht aufweist und/oder im Ablaut vom Aorist bzw. auch Präsens abweicht, vgl. z. B. *phaín-ō* ‘zeig-PRÄS.AKT.1.SG’, *é-phēn-a* ‘PRÄT-ZEIG-AOR.AKT.1.SG’, *phan-ō* ‘zeig-FUT.AKT.1.SG’. Überwiegend hat aber das sigmatische Futur, mit *-s* Marker, einen mit dem Präsens identischen Stamm.

### 7.2. Endungen

#### 7.2.1. Typen der Personalendungen

Das Griechische hat mehrere Typen von Personalendungen, deren Wahl teils semantisch, teils lediglich formal allomorphisch bedingt ist. Beim Indikativ Präsens haben wir etwa bei der produktiven und regelmäßigen Konjugation von *paideiō* ‘erziehe’ die sog. primäre thematische (“thematischen” Vokal *e/o* erhaltende) Endungsreihe: *paideiū-ō*, *-eis*, *-ei*, im Singular, *paideiū-omen*, *-ete*, *-ousi* im Plural. Bei einer nicht geringen Zahl von Verben, z. B. bei *deiknumi* ‘zeige’ kommt aber eine allomorphe athematische Endungsreihe

vor, wo auch die Minimalendungen nicht gleich sind: *deiknu-mi*, -*s*, -*si*, im Singular, *deik-nu-men*, -*te*, -*nú-āsi* im Plural. Bei dem unregelmäßigen (und enklitischen) Verb *eimi* ‘bin’ lautet die 3. Sg. *es-tí* (mit der ursprünglichen Endung -*ti*). Auch innerhalb der thematischen Konjugation haben die kontrahierten Verbklassen (mit Stammvokal -*a*, -*e*, und -*o*) von der *paideuō*-Klasse verschiedene Endungen, vgl. die Singularreihen: *tim-ō*, -*ājs*, -*āj* ‘ehre usw.’, *poi-ō*, -*eis*, -*eī* ‘mache usw.’, *doul-ō*, -*oīs*, -*oī* ‘knechte usw.’.

Die Reihe der Sekundärendungen bezeichnet, in Kooperation mit dem Augment *e*, die Präteritalformen. Sie ist ebenfalls von thematischem oder athematischem Typ, beide Endungsreihen unterscheiden sich nur durch das Vorhandensein bzw. Fehlen des Themavokals *elo* im Falle des Imperfekts (Singular des Imperfekts *e-paideu-on*, -*es*, -*e*; *e-deiknu-n*, -*s*, -*Ø*). Nach den Konsonanten entsteht lautgesetzlich eine Variante -*a* zur Endung -*n* in der 1. Person Singular und 3. Person Plural (ähnlich wie bei Akkusativ Singular der 3. Deklination, vgl. unter 6), wodurch beim sigmatischen Aorist die Endung 1. Person Singular -*sa* und 3. Person Plural -*sa-n* (mit sekundär zugefügtem -*n*) entsteht und der so entstandene Formant -*sa-* auch zur 2. Person Singular, 1. Person Plural und 2. Person Plural (-*sas*, -*samen*, -*sate*, aber 3. Person Singular -*se*) übertragen wird: *e-paideu-sa*, -*sas*, -*se*, *e-paideu-samen*, -*sate*, -*san*, wobei bei “Verba liquida” mit Stamm auf *r*, *l*, *n*, *m* und bei einigen anderen Verba die Endungen im Attischen (durch lautliche Prozesse) *s*-los sind: z. B. *é-phēn-a*, -*as*, -*e* zu *phaínō* ‘zeige’. Zu den Optativendungen vgl. 7.2.2. Die Opposition von Aktiv und Mediopassiv wird durch den Typ der Personalendungen ausgedrückt, vgl. neben der primären Aktivreihe von *paideuō* die mediopassive Singularreihe: *paideu-omai*, -*ēj*, -*etai* (Präsens), neben der sekundären *e-paideu-ómēn*, *e-paideu-ou*, -*eto* (Imperfekt). Durch Endungen wird ferner der Imperativ gekennzeichnet: *paideu-e* ‘erzieh-AKT.IMP.2.SG’, *paideu-ete* ‘erzieh-AKT.IMP.2.PL’, *paideu-ou* ‘erzieh-MEDPASS.IMP.2.SG’, *paideu-esthe* ‘erzieh-MEDPASS.IMP.2.PL’. Besondere Endungsreihen hat das Indikativ des Perfekts: *pe-paideu-ka*, -*kas*, -*ke*, *pe-paideu-kamen*, -*kate*, -*kāsi*.

Besondere Dualendungen sind nur in der 2. (Akt. -*ton*, Medpass. -*sthon*) und 3. Person (Primärreihe Akt. -*ton*, Medpass. -*sthon*, Sekundärreihe Akt. -*tēn*; Medpass. -*sthēn*) und für den Imperativ belegt.

Bei der Mannigfaltigkeit der Endungen lassen sich keine identischen Zeichen der einzelnen Personen identifizieren. Doch sind gewisse Regelmäßigkeiten in der Verteilung der phonematischen Marker zu beobachten: hintere Vokale und nasale Konsonanten in der 1. Person gegenüber vorderen Vokalen, Dentalen und Spiranten in der 2. Person und in der 3. Person Singular. Die Ikonizität in der Markierung der Singular- und Pluralsachen gilt ohne Ausnahme. Mehr über Nichtarbitrarität in den griechischen Personalendungen bietet Panhuis (1980).

### 7.2.2. Konjunktiv-, Optativ- und Futurmarker

Morpheme mit besonderem Marker zwischen Flexionsstamm und Minimalendung sind Konjunktive, Optative und Futura. Zeichen des Konjunktivs sind der lange Themavokal und Primärendungen: *paideu-ēj* ‘erzieh:PRÄSKONJ.AKT.3.SG’, *paideu-s-ē-tai* ‘erzieh:AOR-KONJ-MEDPASS.3.SG’. Der *s*-Marker mit primären thematischen Endungen kennzeichnet das Futur in allen seinen Diathesen (*paideu-s-ō* ‘erzieh-FUT-AKT.1.SG’, *paideu-s-omai* ‘erzieh-FUT-MEDPASS.1.SG’, *paideu-thē-s-omai* ‘erzieh-PASS-FUT-MEDPASS.1.SG’ (wo sich Passivmarker *thē* und Endungen der mediopassiven Reihe auf dem Ausdruck der passiven Diathese beteiligen), auch bei dem schwach belegten Perfektfutur: *ke-ktē-se-tai* ‘RDP:PF-be-sitz-PF.FUT-MT.3.SG’. Das sigmatische Fut. Akt. hat eine allomorphe Variante (zur Verteilung siehe Hauri 1975) im “kontrahierten” Futur vom Typ *phan-ō*, -*eis*, -*eī* zum Präsens *phainō* ‘zeige’. Der Marker des Optativs enthält den Vokal *i*, der entweder der sekundären athematischen Endungsreihe in der Form *iē* vorausgeht: *gnōiē-n*, -*s*, -*Ø* ‘erkenn: AOR.OPT-AKT.1.SG, -AKT.2.SG, -AKT.3.SG’ oder in der Form *i* dem Themavokal *o* im Präsens oder dem Vokal *a* im sigmatischen Aorist folgt und teilweise besondere Endungen mit -*mi* ‘1.SG’, -*en* ‘3.PL’ aufweist: *paideu-oi-mi*, -*oi-s*, -*oi*, *oi-en* ‘erzieh:PRÄS-OPT-AKT.1.SG, -OPT-AKT.2.SG, -OPT(AKT.3.SG), -OPT-AKT.3.PL’.

### 7.3. Infinite Verbalformen

An den aktiven Infinitiven und Partizipien läßt sich der komplexe, durch Modifikation der ganzen Wortform bewirkte Ausdruck der grammatischen Morpheme (siehe unter 2) besonders gut beobachten. Sie werden von den drei Aspektstämmen mit unterschiedlichen Suffixen abgeleitet, so daß sie die aspekto-temporalen Oppositionen durch Flexions-

stammmarker und zugleich durch Suffixe markieren: *paideú-ein* (Inf. Präs. Akt.), *paideú-s-ai* (Inf. Aor. Akt.), *pe-paideu-k-éhai* (Inf. Pf. Akt.); *paideú-ón* (Part. Präs. Akt.), *paideú-sás* (Part. Aor. Akt.), *pe-paideu-kós* (Part. Pf. Akt.); nur im Futur sind die Suffixe mit dem Präsens identisch: *paideú-s-ein* (Fut. Inf. Akt.), *paideú-s-ón* (Part. Fut. Akt.). Alle griechischen Partizipia, einschließlich des Partizips Präsens (im Lateinischen ohne besonderer Femininform) unterscheiden alle drei Genera: vgl. zum Part. Präs. Aktiv Mask. *paideú-ón*, Fem. *paideú-ousa*, Neut. *paideú-on*.

## 8. Wortbildung

Als Sprache mit reich differenziertem Lexikon hat das Griechische eine reich entwickelte **Wortbildung** mit Hilfe suffixaler Derivation und Komposition, wobei beide Prozesse mit Ablaut einhergehen können. In der griechischen Wortbildungslehre überwiegt die sprachgeschichtliche Untersuchung: Debrunner (1917), für das homerische Griechische Risch (2<sup>1978</sup>), Monographien zur Geschichte einzelner Suffixe bzw. Kompositionstypen oder etymologisch zusammengehöriger Gruppen von Suffixen wie Bader (1965; 1974), Mignot (1972), Perpillou (1973), Meier (1975), Lamberterie (1990).

Einige ererbte Bildungen sind sehr frequent und polysemantisch, aber nicht mehr produktiv, z. B. *-u* Adjektive (s. Lamberterie 1990), andere haben im Ionisch-Attischen Produktivität auf der Basis einer stark abstrakten grammatisierten Bedeutung erreicht, z. B. verbale Substantiva auf *-ma*. Zu den Adjektiven mit dem Suffix *-u* wie *tachús*, *bathús* usw., welche die elementaren physikalischen Eigenschaften ausdrücken, standen die Abstrakta mit *-es/los* Suffix als parallele Bildungen gegenüber, und zwar als Ausdruck der graduierenden Eigenschaft nur zu den positiven Gliedern der Oppositionen wie *tachús* ‘schnell’ – *táchos* ‘Schnelligkeit’, während zu *bradús* ‘langsam’ erst durch suffixale Ableitung *bradutés* ‘Langsamkeit’ gebildet wurde, dann auch *tachutés* neben (nach Chantraine 1933: 418 weniger abstraktem) *táchos* (vgl. Lamberterie 1990: besonders *Bd. II* 568 ff.; Mignot 1972: 64 f., 122; Porzig 1942: 246 f.; vgl. auch 3.1 und 3.2).

### 8.1. Derivation

Besonders zahlreich sind die substantivischen **Derivationen**, wobei innerhalb einer funktionalen Gruppe mehrere konkurrierende,

durch Produktivität und semantische Nuancen sich unterscheidende Bildungen vertreten sind: Motionsfeminina mit Suffixen auf *-á* (1. Deklination), z. B. *léaina* ‘Löwin’ zu *léon* ‘Löwe’ mit ursprünglichem *-iā* Suffix ([j]) hinter Konsonanten ist mit Nachwirkungen verschwunden) oder mit Suffix *-is*, Gen. *-idos*, wie *summachis* ‘Bundesgenossin’ zu *súmmachos* ‘Kampf-, Bundesgenosse’; denominale Abstrakta mit den Suffixen *-tēs*, *-súnē* (*sōphrosúnē* ‘Verstand’ zu *sóphrōn* ‘verständig’), *-iā* (*philosophiā*), *-ikē* (*rhētorikē* ‘Redekunst’); deverbale Abstrakta mit den Suffixen *-sis* (*krísis* ‘Entscheidung’ zu *krínō* ‘entscheide’), *-mós*, *-mē* (*gnómē* ‘Einsicht’ zu *gignóskō* ‘erkenne’); Berufsbezeichnungen und Nomina agentis mit den Suffixen *-eus*, *-tēs* (*krités* ‘Richter’ zu *krínō* ‘entscheide, urteile’), *-tér*, *-tōr* (*rhē-tōr* ‘Redner’ zum Verbalstamm *rhē-* ‘sage’) usw. Adjektive werden von den Substantiven durch die Suffixe *-ios*, *-ikos* (*rhētorikós* ‘rednerisch’ zu *rhētōr* ‘Redner’), *-inos* usw. gebildet. Grammatikalierte Verbaladjektiva sind die Partizipien (7.3). Auch die Bildung der Adverbien zu den Adjektiven ist grammatisiert (3.4). Bei den abgeleiteten Verba ist die Funktion der Suffixe wenig ausgeprägt, sie dienen zur Bildung des Präsensstammes (7.1).

### 8.2. Komposition

Die **Komposition** ist im Griechischen ein grundlegender Prozeß, durch den die aus zwei oder mehreren Lexemen bestehenden Benennungen ausgedrückt werden. Bei der echten Komposition handelt es sich um die Verbindung von zwei lexikalischen Stämmen (sog. Vorderglied und Hinterglied), die mit einer flexivischen Endung versehen sind. Im endozentrischen Typ des Nominalkompositums wie *akró-polis* ‘Hochstadt’ (Adj. *ákros*, Fem. *ákrā* ‘hoch’ + *pólis* ‘Stadt (Fem.)’) ist sie gewöhnlich der Endung des Hintergliedes gleich, zu dem das Vorderglied in determinativem Verhältnis steht (daher wird dieser Typ Determinationskompositum genannt). Das Vorderglied (obwohl es sich um ein Femininum handelt) hat den für die Komposition typischen Stammvokal *o*. Verbreiteter (umgekehrt zum Deutschen) ist im Griechischen der exozentrische Typ mit Substantivnukleus außerhalb des Kompositums (auch Possessivkompositum genannt) z. B. *kakó-nous* ‘übelgesinnt’ (Adj. *kakós* ‘schlecht’ + *noús* ‘Sinn’). Sehr zahlreich sind auch die sog. Rektionskomposita mit Verbalstamm im Vorderglied – auf *-e*, *-si*, meist aber wiederum auf *-o*:

*philó-sophos* (*philéō*, kontrahiert *philō* ‘liebe’ + *sóphos* ‘weise’) – oder im Hinterglied, mit unterschiedlichen Endungen versehen, z. B. *phōs-phóros* ‘Lichtbringer’ (*phōs* ‘Licht’ + *phérō* ‘trage’), *nomo-thétēs* ‘Gesetzgeber’ (*nómos* ‘Gesetz’ + *títhēmi* ‘setze’). Das aristotelische zusammenfassende *philo-toioútos* ‘dies und dies liebend’ zeigt die Produktivität der Bildung und zugleich den adjektivischen Charakter des exozentrischen Kompositums (*toioútos* ‘solcher’).

## 9. Illustrativer Text

Dareí-*ou*            *kai Parusátid-os*  
 Dareios-GEN.SG und Parysatis-GEN.SG  
*gí-gno-ntai*  
 RDP:PRÄS-geboren.werd-PRÄS.MT.3.PL  
*paíd-es*            *dúo,*  
 Kind/Sohn-NOM.PL zwei(NOM),  
*presbú-teros*        *mén*  
 alt-KMPR.NOM.SG.M PK  
*Artaksérksés*        *neő-teros*  
 Artakserkses-NOM.SG, jung-KMPR.NOM.SG.M  
*dè Kûr-os;*        *epeì dè*  
 PK Kyros-NOM.SG; wenn nun  
*ēsthénei*  
 PRÄT:krank.werd:IMPF.AKT.3.SG  
*Dareí-*os**            *kai hupóptee-*  
 Dareios-NOM.SG und ahn:PRÄT-IMPF.AKT.3.SG  
*teleut-én*            *toû*            *bí-ou,*  
 Ende-AKK.SG ART.GEN.SG.M Leben-GEN.SG  
*e-boúle-to*            *tò*  
 PRÄT-woll-IMPF.MT.3.SG ART.AKK.DU  
*paíd-e*              *amphotér-ō*  
 Kind/Sohn-AKK.DU beide-AKK.DU  
*pareîmai.*            *ho*            *mén oún*  
 anwesend.sei:INF.PRÄS ART.NOM.SG.M PK PK  
*presbú-teros*  
 alt-KMPR.NOM.SG.M  
*parón*  
 anwesend.sei:PART.PRÄS.NOM.M  
*e-túnchan-e;*        *Kûron*  
 PRÄT-sich.treff-IMPF.AKT.3.SG Kyros-AKK.SG  
*dè metapémpe-tai*    *apò*  
 PK herbeiruf-PRÄS.MEDPASS.3.SG aus  
*tés*                  *arch-ēs*  
 ART:GEN.SG.F Regierung(=Provinz)-GEN.SG  
*hês*                  *aut-òn*        *satráp-ēn*  
 REL.GEN.SG.F er-AKK.SG.M Statthalter-AKK.SG  
*e-poíē-se;*  
 PRÄT-mach-AOR.AKT.3.SG;

‘Dem Dareios und der Parysatis waren zwei Söhne geboren worden, der ältere Artaxerxes und der jüngere Kyros. Als nun Dareios krank wurde und das Ende seines Lebens ahnte, wollte er beide Söhne bei sich anwesend sehen. Der Ältere war gerade anwesend; den Kyros aber ließ er aus der Provinz herbeirufen, zu deren Statthalter er ihn gemacht hatte.’

Anfang der *Kúrou Anábasis* des Xenophon (430–354 v. Chr.), ein Bericht über den Feldzug gegen Artaxerxes und Rückzug der zehntausend griechischen Söldner nach Trapezunt, nachdem Kyros d.J. bei Kunaxa gefallen war. Dareios = Dareios II., König von Persien 423–404. Zur Syntax und morphologischen Semantik nota: sog. Genitiv der Abstammung *Dareí-ou kai Parusátid-os*; sog. Präsens historicum (in präteritaler Bedeutung) *gígnontai* und *metapémptai*; Dual *tò paíd-e* (gegenüber *paíd-es dúo*); Konstruktion *tunchánō* ‘(mich) treffe’ mit Partizip *parón etúnchan-e* im Sinne ‘war gerade anwesend’; das Medium *metapémpe-tai* zu *metapémptō* ‘schicke nach’ im Sinne ‘zu sich nachschnicken, herbeirufen lassen’; enklitische Partikeln (= ‘PK’) als Satzgliederungssignale mit vager Bedeutung (*mén – dé* adversativ gegenübergestellt). Zur formalen Morphologie nota: Komparativ mit gedeihntem Stammvokal *o* (durch Kürze der vorangehenden Silbe konditioniert): *ne-ő-teros* gegenüber *pist-ó-teros* (3.1), wo die Silbe ‘durch Position’ lang (geschlossen) ist; das sog. temporale, d. h. in der Dehnung des anlautenden Vokals bestehende Augment in *ēsthénei* (*ē* > *ā*) zu *asthe néō*, *hupópteue* zu *hupopteúō*.

## 10. Unübliche Abkürzungen

MT Medium tantum  
 PK Partikel

## 11. Zitierte Literatur

- Bader, Françoise (1965), *Les composés grecs du type de demiourgos*. Paris: Klincksieck  
 Bader, Françoise (1974), *Suffixes grecs en -m-: Recherches comparatives sur l'hétéroclisie nominale*. Genève: Droz & Paris: Minard/Champion  
 Bornemann, Eduard & Risch, Ernst (1978), *Griechische Grammatik*. Frankfurt am Main: Moritz Diesterweg  
 Chanet, Anne-Marie (1985), *Memento de morphologie verbale du grec: Attique classique*. [Paris]: C.N.A.R.E.L.A.

- Chantraine, Pierre (1933), *La formation des noms en grec ancien*. Paris: Champion
- Debrunner, Albert (1917), *Griechische Wortbildungsllehre*. Heidelberg: Winter
- Funk, Wolf-Peter (1985), “On a Semantic Typology of Conditional Sentences”. *Folia Linguistica* 19. 365–413
- Giannakis, Giorgios K. (1997), *Studies in the Syntax and Semantics of the Reduplicated Presents of Homeric Greek and Indo-European*. Innsbruck: Institut für Sprachwissenschaft (Innsbrucker Beiträge zur Sprachwissenschaft 90)
- Hauri, Hans Walter (1975), *Kontrahiertes und sigmatisches Futur: Einflüsse von Lautstruktur und Aktionsart auf die Bildung des griechischen Futurs*. Göttingen: Vandenhoeck & Ruprecht
- Irigoin, Jean (1972), “La déclinaison thématique en attique classique: Essai de description”. In: *Mélanges de linguistique et philologie grecques offerts à Pierre Chantraine*. Paris: Klincksieck, 55–73
- Köstling, Pamela (1998), *Die griechischen primären Jotpräsentien: Untersucht und dargestellt nach Formenbestand, Aktionsarten und Etymologie*. Hamburg: Dr. Kovač (Philologia. Sprachwissenschaftliche Forschungsergebnisse 27)
- Kølln, Herman (1969), *Oppositions of Voice in Greek, Slavic, and Baltic*. København: Munksgaard
- Kujore, Obafemi (1973), *Greek Polymorphic Presents: A Study of their Development and Functional Tendencies*. Amsterdam: Hakkert
- Lamberterie, Charles de (1990), *Les adjectifs grecs en -us: Sémantique et comparaison*, Bd. 1–2. Louvain-La Neuve: Institut de Linguistique de Louvain (Cahiers de l’Institut de Linguistique de Louvain 54–55)
- Lewy, Ernst (1961), *Kleine Schriften*. Berlin: Akademie-Verlag
- Meier, Michael (1975), *-id-*: Zur Geschichte eines griechischen Nominalsuffixes. Göttingen: Vandenhoeck & Ruprecht
- Mignot, Xavier (1972), *Recherches sur le suffixe -tēs, -tētos (-tās, -tātos), dès origines à la fin du IV<sup>e</sup> siècle avant J.-C.* Paris: Klincksieck
- Panhuis, Dirk (1980), “The Personal Endings of the Greek Verb: An Exploration in the Subjectivity and Non-arbitrariness of a Paradigm”. *Studies in Language* 4, 105–117
- Perel'muter, Ilja Aronovič (1977), *Obšeindojevropeskij i grečeskij glagol: Vido-vremennyje i zalogovyye kategorii*. Leningrad: Nauka
- Perel'muter, Ilja Aronovič (1995), *Zalog drevnegrečeskogo glagola: Teorija, genezis, istorija*. S.-Peterburg: Izdatel'stvo “Noach”
- Perpillou, Jean-Louis (1973), *Les substantifs grecs en -eūs*. Paris: Klincksieck
- Porzig, Walter (1942), *Die Namen der Satzinhalte im Griechischen und Indogermanischen*. Berlin: De Gruyter
- Rijksbaron, Albert (1994), *The Syntax and Semantics of the Verb in Classical Greek: An Introduction*. Amsterdam: Gieben
- Risch, Ernst (1978), *Wortbildung der homerischen Sprache* [2. völlig überarbeitete Auflage]. Berlin, New York: de Gruyter
- Rix, Helmut (1976), *Historische Grammatik des Griechischen: Laut und Formenlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft
- Ruipérez, Martín Sánchez (1954), *Estructura del sistema de aspectos y tiempos del verbo griego antiguo: Análisis funcional sincrónico*. Theses et studia philologica Salamanticensia 7. Salamanca: Colegio trilingüe de la Univ.
- Sapir, Edward (1949), *Language: An Introduction to the Study of Speech*. New York, London: Harcourt Brace Jovanovich
- Schwyzer, Eduard (1968), *Griechische Grammatik, Bd. I: Allgemeiner Teil. Lautlehre. Wortbildung. Flexion*. München: Beck
- Schwyzer, Eduard & Debrunner, Albert (1966), *Griechische Grammatik, Bd. II: Syntax*. München: Beck
- Strunk, Klaus (1980), “Zum indogermanischen Medium und konkurrierenden Kategorien”. In: Bretschneider, Gunter & Lehmann, Christian (Hrsg.), *Wege zur Universalienforschung*. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler. Tübingen: Narr, 321–337
- Wittwer, Michael (1969), “Über die kontrastierende Funktion des griechischen Suffixes *-teros*”. *Glotta* 47, 45–110
- Zeegers Vander Vorst, Nicole (1987), “Initiation à la conjugaison grecque”. *Cratyle* 6 (Février)

Helena Kurzová, Prag  
(Tschechische Republik)

## 124. Finnish (Finno-Ugric)

1. Background
2. General characteristics of Finnish word structure
3. Parts of speech
4. Morphotactic structure and basic inflectional categories of nominals
5. Morphotactic structure and basic inflectional categories of finite verb-forms
6. Morphotactic structure and basic inflectional categories of nonfinite verb-forms
7. Clitics
8. Morphophonological alternations
9. Nominal inflectional types
10. Verbal inflectional types
11. Derivational morphology
12. Compounding
13. Morphological productivity and diachronic tendencies in the morphological system
14. Processing and acquisition of Finnish morphology
15. Illustrative text
16. Uncommon abbreviations
17. References

### 1. Background

In 2002, the population of Finland was 5,206,000 people out of whom 4,797,000 (92.1 percent) spoke Finnish as their first language. The name of the language in Finnish is *suomi* or *suomen kieli*. Finnish is also spoken by some 300,000 immigrants especially in Sweden, USA, Canada, Australia, and Russia. Finnish belongs to the Baltic-Finnic sub-branch of the Finno-Ugric languages. Finno-Ugric is one of the two main branches of the Uralic language group, the other main branch of which is the Samoyedic languages. Finnish is the biggest Baltic-Finnic language, the others are Estonian (1 million speakers), Karelian (70,000, including Aunus), Ingrian (300), Veps (6,000), Lude (5,000), Vote (few), and Livonian (few) (cf. Salminen 1993 for details). Vote and Livonian are close to extinction. The “intelligibility distance” between Finnish and Estonian is roughly similar to that between Swedish and Danish, or German and Dutch. Finno-Ugric tribes, i.e. early Lapps and perhaps early Finns, have been in Finland at least as early as around 4,000 BC (cf. Häkkinen 1996 for an overview of Finnish prehistory from the viewpoint of linguistics). In fact, as archeological evidence shows, Finland has been inhabited at least

since 7,500 BC but the language(s) of this earliest population cannot be determined.

Some 2/3 of the genome of present-day Finns is common Western European. Some geneticists hold the view that this fact must be explained by assuming that present-day Finns derive from Indo-Europeans that switched their language to Finno-Ugric after having entered Finland and made contacts with the indigenous population. From the linguistic point of view, the most likely interpretation of this genetic fact is extensive and extended contacts between early Finno-Ugrians and Indo-Europeans in prehistoric times, probably both north and south of the Gulf of Finland, with the Indo-Europeans being assimilated to the indigenous Finno-Ugrians. Some of the earliest Indo-European loanwords in Finnish have come from a very early, perhaps even Proto-Indo-European language form (Koivulehto 1983; 1994).

This article describes modern standard written Finnish. Morphologically, standard spoken Finnish is very close to the written standard. Finnish was first written around 1540 in connection with Mikael Agricola's (1510?–1557) work on translating the Bible. In 1820–1850, there was an intensive discussion of how much eastern Finnish vocabulary the standard language should contain. In 1863, Finnish was declared to be an official language alongside Swedish (1809–1917 Finland was an autonomous Russian province, after having belonged to Sweden from the Middle Ages to 1809). Present-day standard Finnish norms were basically established 1860–1920. There are eight areas of traditional regional dialects. All of these are mutually fully understandable. Due to urbanization, education, migration, etc. especially after World War II, the dialect differences and borders are not all that clear anymore. Helsinki urban speech has acquired a prestigious position after the 1960s but there are other strong regional urban variants as well. Sociolinguistic differences are not sizeable.

### 2. General characteristics of Finnish word structure

The segmental phonemes of Finnish are /i e æ y ø u o a/, orthographically <i e ä y ö u o a>. The opposition long/short is in extensive

use, phonemically long sounds are written with double letters: *taakka* [ta:k:a] ‘burden’. Henceforth, the phonemes /æ/ and /ø/ will be symbolized by the letters <ä> and <ö>.

Primary word stress is always on the initial syllable. Finnish is a suffixing language with a fairly elaborate inflectional morphology. The central inflectional categories are case for nouns and adjectives, person for finite verbs. The word structures are based on a rich system of suffixal morphotactic positions. Derivational endings occur between the root and the inflectional endings. There are very few prefixes, the most important one being *epä-* ‘non-’.

All types of suffixes are subject to vowel harmony which means that endings containing one of the harmony vowels /ä y ö a o u/ always have a corresponding ending variant with the opposite vowel of the pairs *a–ä*, *o–ö*, and *u–y*. The back ending variant occurs if there is back vowel /a o u/ to the left, else the front ending variant occurs. Thus: *talo-ssa* ‘in (the) house’ vs. *kylä-ssä* ‘in (the) village’ (inessive case), *tule-vat* ‘come-3.PL’ vs. *mene-vät* ‘go-3.PL’. The vowels /i e/ are neutral in endings and occur after both front and back vowels. If a stem has neutral vowels only, it takes front ending harmony as shown by *mene-vät*. Derivational endings are also subject to vowel harmony: *kahvi* ‘coffee’ – *kahvi-la* ‘cafeteria’ – *kahvi-la-ssa* ‘in (the) cafeteria’, *kylpy* ‘bath’ – *kylpy-lä* ‘public bath’ – *kylpy-lä-ssä* ‘in (the) public bath’. In what follows, vowel harmonic pairs will occasionally be denoted by the morphophonemes {A}, {O}, and {U} (other morphophonemes will also be enclosed in curly braces { }).

A typical feature of Finnish morphophonology is the wealth of morphophonological alternations dependent upon phonological, morphological, or even lexical (idiosyncratic) conditioning. Thus, Finnish is not a particularly agglutinating language, if that term is taken to mean that affixes are added to invariant stems. Cf. the stem alternants of the noun *käsi* ‘hand’ in various case forms: *käte-en* ‘hand-ILL (into the hand)’, *käde-ssä* ‘hand-INNESS (in the hand)’, *kät-tä* ‘hand-PRTV’, *käs-i-ssä* ‘hand-PL-INNESS (in the hands)’.

The more morphotactic positions there are in the word structures of a language, the more synthetic that language is. Finnish is a fairly synthetic language, a property which is emphasized by extensive use of nominal com-

pounding such as *merkin-ant-o-nappi* ‘sign-giving-button (button for giving a sign)’.

Finnish has few monosyllabic roots, less than 100, compared to at least 7,000 for English. The canonical structure of Finnish lexical morphemes, roots, is disyllabic. This in conjunction with the extensive use of ending classes and compounding leads to fairly long word-form tokens. Niemikorpi (1991: 391) reports that the average length of Finnish words in running text is 7.42 graphemes whereas that of English is 4.74 graphemes (the Brown Corpus) and Swedish 5.39 graphemes (the Press-65 corpus). Niemikorpi’s Finnish corpus contained five text-types, more than 420,000 word-form tokens.

### 3. Parts of speech

On purely morphological grounds, three groups of words may be distinguished: nominals, verbs, and basically non-inflecting words, generally called particles in the Finnish grammar tradition.

The **verbs** are most clearly marked as a distinct morphologically based part of speech because they alone inflect for person, tense, mood, and non-finiteness (infinitives, participles). Thus, of the verb stem *sano-* ‘say’ one obtains e.g. the word-forms *sano-tte* ‘say-2.PL(PRES.IND) (you say)’, *sano-i-tte* ‘say-PAST-2.PL (you said)’, *sano-isi-tte* ‘say-COND-2.PL (you would say)’, *sano-a* ‘say-INF<sub>1</sub> (to say)’. The endings *-tte* ‘2.PL’, *-i* ‘PAST’, *-isi* ‘COND’, *-a* ‘INF<sub>1</sub>’ could never occur after nominals or non-inflecting words.

The **nominals** are normally subdivided into nouns, adjectives, pronouns, cardinal numerals. Nominals share the properties of being inflected in number and case. Of the noun *talo* ‘house’, adjective *iso* ‘big’, pronoun *joka* ‘who, which’, and cardinal numeral *neljä* ‘four’ one thus obtains e.g. the following singular and plural case forms (see Tab. 124.1 next page).

The concept “nominal” is to be understood only as a cover term for this joint morphological behaviour. Overall, it would not make sense to postulate nominals as a basic part of speech in Finnish because nouns, adjectives, pronouns, and numerals of course differ from one another syntactically in several ways.

Upon closer inspection, there are some morphological differences between nouns, adjectives, pronouns, and numerals. Adjec-

		noun	adjective	pronoun	numeral
nominative	singular	<i>talo</i>	<i>iso</i>	<i>joka</i>	<i>neljä</i>
	plural	<i>talo-t</i>	<i>iso-t</i>	<i>jo-t-ka</i>	<i>neljä-t</i>
genitive	singular	<i>talo-n</i>	<i>iso-n</i>	<i>jo-n-ka</i>	<i>neljä-n</i>
	plural	<i>talo-j-en</i>	<i>iso-j-en</i>	<i>jo-i-den</i>	<i>nelj-i-en</i>
inessive	singular	<i>talo-ssa</i>	<i>iso-ssa</i>	<i>jo-ssa</i>	<i>neljä-ssä</i>
	plural	<i>talo-i-ssa</i>	<i>iso-i-ssa</i>	<i>jo-i-ssa</i>	<i>nelj-i-ssä</i>

Tab.124.1: Number/case inflection

tives may undergo comparison: *iso* ‘big’ – *iso-mpi* ‘bigger’ – *iso-in* ‘biggest’. Pronouns and nouns are not subject to morphological comparison. However, some nouns may be compared, e.g. *ranta* ‘beach’ – *ranne-mpi* ‘closer to the beach’ – *rann-in* ‘closest to the beach’. On the other hand, nouns may be inflected for possessive, e.g. *talo* ‘house’ – *taloni* ‘my house’. Numerals and especially pronouns do not accept possessive endings. Except for the possessive adjective *oma* ‘own’, adjectives do not readily take possessive endings either, even if occasional instances may be encountered such as *rakkaa-ni* ‘my dear (singular or plural)’ (cf. the base form *rakas* ‘dear’). When further considering the fact that nominals such as *suomalainen* are ambiguous (or homonymous) with both an adjectival reading (here, ‘Finnish’) and a noun reading (here, ‘Finn’), we see that nouns and adjectives are not clearly distinct in Finnish from the morphological point of view.

As frequently encountered in languages, many **pronouns** display morphological peculiarities. Thus, some pronouns superficially seem to be inflected by **infixation** rather than by suffixation, e.g. *jokin* ‘some’ – *jo <ssa>kin* ‘some<INESS> (in some)’ – *jo <lta>kin* ‘some<ABL> (from some)’. The etymological interpretation is that the original stem is \**jo-* while *-kin* is a lexicalized clitic. The pronoun *mikä* ‘which’ has identical singular and plural forms almost through the whole paradigm, e.g. *mi-ssä* ‘in which’ (singular or plural), in contradistinction to words with ordinary differentiated inflection like typical nouns: *auto-ssa* ‘in (the) car’ – *auto-i-ssa* ‘in (the) cars’. Many pronouns lack some of the (less frequent) cases. Thus, *tämä* ‘this’ does not have a singular abessive case form such as \**tä-ttä* ‘without this’. The sub-classification of pronouns is a difficult issue that cannot be further pursued here (cf. Hakulinen & Karlsson<sup>3</sup> 1995 for details).

Case and number are not inflectional categories exclusive for nominals because verbal participles are also inflected in case and number, just as adjectives, cf. the so-called first participle (here, in active voice): *sano-va* ‘say-PART.ACT(NOM.SG)’, *sano-va-t* ‘say-PART.ACT-NOM.PL’, *sano-va-n* ‘say-PART.ACT-GEN(SG)’, *sano-v-i-en* ‘say-PART.ACT-PL-GEN’, *sano-va-ssa* ‘say-PART.ACT-INESS(SG)’, and *sano-v-i-ssa* ‘say-PART.ACT-PL-INESS’.

The notion **particle** in traditional Finnish grammar is not a natural homogeneous part of speech but a cover term for several subclasses of words with no inflection, or clearly defective inflection. These include adverbs (with many subclasses), prepositions, postpositions, conjunctions, subjunctions, and interjections. Several adverbs, prepositions, and postpositions can be defectively inflected in 2–3 cases (out of 15), especially in some of the local cases. Thus, one finds adverb triplets such as *koto-na* ‘home-ESS (at home)’, *koto-a* ‘home-PRTV (from home)’, *koti-in* ‘home-ILL (“towards” home)’, or postposition triplets like *keske-llä* ‘middle-ADESS (in the middle of)’, *keske-ltä* ‘middle-ABL (from the middle of)’, and *keske-lle* ‘middle-ALL (towards the middle of)’. There are many postpositions (a few hundred, including the inflectional variants) but few prepositions. This is an obvious reflection of the historical fact that Finnish used to be an SOV language. The distinctions between adverbs, prepositions, and postpositions are frequently unclear. Should e.g. *yli* ‘over’ be classified as an adverb when occurring without a complement, as postposition when occurring with a pre-complement in the genitive case (*joe-n yli* ‘over the river’; cf. *joki* ‘river’), and as preposition when occurring with a post-complement in the genitive (*yli joe-n* ‘over the river’)? Or is there just one lexeme *yli* ‘over’ with an optional complement that may occur on either side? If the latter interpretation is

preferred, no neat part of speech label is available for classifying the word *yli*, but rather an undifferentiated adverb-preposition-postposition. Due to such fluctuations, the notion “adposition” is sometimes used as a cover term for prepositions and postpositions.

#### 4. Morphotactic structure and basic inflectional categories of nominals

The morphotactic structure of nominals is illustrated in Tab. 124.2 (cf. also Karlsson 2002).

In this morphotactic scheme, DERIV stands for all possible sequences of derivational endings (cf. 11), and CLITIC for cliticized bound morphemes that may be appended to nominals and to all types of verb-forms. These will be treated in more detail in 7.

The inflectional categories proper (as opposed to clitics) of nominals are number, case, and possessive.

Finnish has two **numbers**, singular and plural. In conformance with ordinary markedness tendencies, the singular has no overt endings. The plural has two suppletive manifestations, *-t* in the nominative and *-i-* in all other cases. The nominative plural *-t* is a portmanteau morph that actually belongs both to the number and case position classes in the scheme above. Plural *-i-* is manifested as *-j-* between two vowels, e.g. *talo-j-a* ‘houses’ (partitive plural).

Finnish has 15 **cases**. The following morphophonemes are used in Tab. 124.3: {A} representing /a/ or /ä/ as regulated by vowel harmony, {V} representing any vowel, {I}

representing /i/ or /e/, and {T}, {N} representing /t/ and /n/ or their absense, respectively. The character “=” in the plural column indicates that the form of the respective ending is identical in singular and plural.

The illative ending vowel {V} is realized by reduplication as a copy of the preceding vowel: *talo-on* ‘into (the) house’, *talo-i-hin* ‘into (the) houses’, *kauppa-an* ‘into the shop’, *sää-hän* ‘into (the) weather’, *puu-hun* ‘into (the) tree’.

The morphophonemes {T} and {N} are realized as zero when a possessive suffix follows, else as /t/ and /n/. Thus: *auto* ‘car (NOM.SG)’, *auto-t* ‘car-NOM.PL’, *auto-n* ‘car-GEN.SG’, *auto-on* ‘car-ILL.SG’, *auto-i-hin* ‘car-PL-ILL’, *auto-mme* ‘car(NOM.SG/GEN.SG/NOM.PL)-POSS.1.PL’, *auto-o-mme* ‘car-ILL-POSS.1.PL’, *auto-i-hi-mme* ‘car-PL-ILL-POSS.1.PL’, *auto-j-en* ‘car-PL-GEN’, *auto-je-mme* ‘car-PL.GEN-POSS.1.PL’. Note the three-ways morphological ambiguity of a form such as *auto-mme* ‘our car; our cars; our car’s’. The final vowel {I} in the translative is realized as /e/ when a possessive suffix follows, else as /i/: *vaimo-ksi* ‘to wife’, *vaimo-kse-ni* ‘to my wife’.

The distribution of the genitive plural, partitive, and illative allomorph types or “quasi-morphemes” (Karlsson 1983) -*eN*, -*deN*, -*teN*, -*tteN*; -*A*, -*tA*, -*ttA*; -*VN*, -*hVN*, -*seeN*, -*siiN* is dependent upon a complex interplay between the inflectional type and/or phonological properties of the stem, e.g. *nalle* – *nalle-a* (partitive singular) vs. *kone* – *kone-tta* (partitive singular). Some aspects of this will be treated in 8.

Only the personal pronouns *minä* ‘1.SG’, *sinä* ‘2.SG’, *hän* ‘3.SG’, *me* ‘1.PL’, *te* ‘2.PL’, *he*

ROOT	DERIV	NUMBER	CASE	POSSESSIVE	CLITIC	gloss
<i>talo</i>						(the) house
<i>talo</i>		<i>n</i>				house-GEN ((the) house’s)
<i>talo</i>		<i>ssa</i>				house-INNESS (in (the) house)
<i>talo</i>		<i>ssa</i>	<i>mme</i>			house-INNESS-POSS.1.PL (in our house)
<i>talo</i>		<i>sta</i>				house-ELAT (from (the) house)
<i>talo</i>	<i>i</i>	<i>ssa</i>				house-PL-INNESS (in (the) houses)
<i>talo</i>	<i>i</i>	<i>sta</i>	<i>mme</i>			house-PL-ELAT-POSS.1.PL (from our houses)
<i>talo</i>	<i>t</i>					house-NOM.PL ((the) houses)
				<i>ko</i>		house-INT ((the) house?)
			<i>ssa</i>	<i>ko</i>		house-INNESS-INT (in (the) house?)
<i>talo</i>	<i>mainen</i>					house-like
<i>talo</i>	<i>maise</i>		<i>ssa</i>			house-like-INNESS (in (the) house-like)
<i>talo</i>	<i>mais</i>	<i>i</i>	<i>ssa</i>	<i>ko</i>		house-like-PL-INNESS-INT (in (the) house-like ones?)

Tab. 124.2: Morphotactic structure of nominals

case endings	singular	plural	basic functions
nominative	—	T	base form
genitive	N	eN/deN/teN/tteN	possession; perfective aspect (obj.)
partitive	A/tA/lttA	A/tA	indefinite quantity; imperfective aspect (obj.)
accusative	t	=	object form of personal pronouns
inessive	ssA	=	interior location
elative	stA	=	movement from the interior
illative	VN/hVN/seeN	hVN/siiN	movement into
adessive	llA	=	exterior location; possessor; instrument
ablative	ltA	=	movement from the exterior
allative	lle	=	benefactive; movement towards
essive	nA	=	time; state
translative	ksI	=	change
instructive	n	=	instrument, manner
abessive	ttA	=	lack of
comitative	=	ine	in the company of

Tab. 124.3: Cases

‘3.PL’, and the interrogative pronoun *kuka* have a distinct morphological accusative form in *-t*: *minut*, *sinut*, *hänet*, *meidät*, *teidät*, *heidät*, *kenet*. These forms occur as direct objects in affirmative clauses with perfective aspect. All other nominals lack this distinct morphological accusative. Depending upon the syntactic properties of the clause, their object form is either genitive singular, nominative singular, or nominative plural:

- (1) *Mies näk-i häne-t* ~  
man(NOM) see-PAST 3.SG-ACC  
*auto-n* ~ *auto-t*.  
car-GEN.SG car-NOM.PL  
'The man saw him/her ~ a/the car ~ the cars.'
- (2) *Häne-t* ~ *auto* ~ *auto-t*  
3.SG-ACC car(NOM.SG) car-NOM.PL  
*näh-t-i-in*.  
see-PASS-PAST-4  
'He ~ the car ~ the cars was (were) seen.'

In traditional Finnish grammar, all noun phrase objects in nominative singular, genitive singular and nominative plural are called “accusatives” which is a source of confusion.

The nominative, genitive, and partitive are the central grammatical cases of Finnish marking noun phrases that function as subjects, objects, and (subject) predicate complements. Another important subgroup of cases are the local cases inessive, elative, illative, adessive, ablative, allative. They constitute a clear morphological and semantic subsystem:

	interior	exterior
stationary location	inessive ssA	adessive llA
movement from	elative stA	ablative ltA
movement towards	illative VN etc.	allative lle

Tab. 124.4: Local cases

These meanings are transparent in the respective case forms of nouns denoting location, e.g. *talo-ssa* ‘in (the) house’, *talo-sta* ‘out of (the) house’, *talo-on* ‘into (the) house’, *talo-lla* ‘at (the) house’, *talo-lta* ‘from (the vicinity of the) house’, *talo-lle* ‘to (the vicinity of the) house’. The interior local cases contain a basic and historically original “s”-element denoting interiority. Due to Proto-Finnic sound changes (\*s > h > Ø), this “s” is not seen as such in all present-day illative allomorphs. The exterior local cases have an “l”-element meaning ‘exterior’. Historically, the ‘stationary location’ component derives from a locative \*-nA (the precursor of the present essive), ‘movement from’ from the original separative meaning of the partitive \*-tA, and ‘movement towards’ from a previous locative \*-n still visible as the final segment of the illative allomorphs. The reconstructed Proto-Finnic subsystem of local cases is thus \*-s-nA, \*-s-tA, \*-s-n, \*-l-nA, \*-l-tA, \*-l-n.

The six local cases have other functions as well. All express meanings relating to time. The *l*-cases express possession, a natural extension of their local meanings:

- (3) *Miehe-llä on auto.*  
man-ADES be.3.SG car(NOM.SG)  
'The man has a car (lit. by the man is a car.)'
- (4) *Sa-i-n kirja-n miehe-ltä.*  
get-PAST-1.SG book-GEN.SG man-ABL  
'I got a book from the man.'
- (5) *Anno-i-n kirja-n miehe-lle.*  
give-PAST-1.SG book-GEN.SG man-ALL  
'I gave a book to the man.'

The essive *-nA* and translative *-ksI* are called general local cases. The essive expresses states and time, the translative denotes change of state or its endpoint. The other cases are more marginal. The instructive is mostly used in fixed adverb-like expressions like *käs-i-n* (plural) 'by hand', *jala-n* (singular) 'by foot'. The comitative ending *-ine-* is untypical because it lacks a regular singular form and requires to be followed by a possessive suffix. A form such as *vaimo-ine-en* is ambiguous:

- (6) *Kalle tul-i vaimo-ine-en.*  
Kalle come-PAST wife-COM-POSS.3  
'Kalle came accompanied by his wife ~ wives.'
- (7) *Miehe-t tul-i-vat*  
man-NOM.PL come-PAST-3.PL  
*vaimo-ine-en.*  
wife-COM-POSS.3  
'The men came accompanied by their wives.'

Pajunen & Palomäki (1982) report the following text frequencies for the cases, based on a corpus of 20,000 word-form tokens taken from four different text types (see Tab. 124.5).

The "traditional accusative" refers to the large syntactic interpretation of the notion accusative referred to above. On the average, the unmarked nominative is the most frequent case, followed by the genitive and the partitive. These three grammatical cases account for more than 60 percent of all case occurrences. The interior local cases are clearly more frequent than the exterior ones. Close to 70 percent of the word-form tokens (13,919 out of 20,000) in Finnish running text contain an instance of a case morpheme (including nominative singular zero).

The third nominal inflectional category is **possessive**. Traditional Finnish grammar has not always been explicit in regard to whether possessive suffixes are classified as inflectional morphemes proper, or as a separate

case	frequency	percent
nominative	4,105	29.5
genitive	2,824	20.3
accusative <i>-t</i>	12	0.1
"trad." accusative	433	3.1
partitive	1,908	13.7
essive	367	2.6
translative	308	2.2
inessive	993	7.1
elative	613	4.4
illative	871	6.3
adessive	610	4.4
ablative	145	1.0
allative	315	2.3
abessive	30	0.2
comitative	18	0.1
instructive	250	1.8
special inst.	117	0.8
sum	13,919	99.9

Tab. 124.5: Case frequency

type of clitic-like morpheme. The possessive endings act like clitics morphotactically because they occur to the right of the case endings, but left of the clitics proper. On the other hand, the possessive endings trigger the same morphophonological stem alternations as case endings do, and this is an argument for classifying them as inflectional rather than clitic in nature: cf. *käsi* 'hand' – *käte-en* 'hand-ILL(SG) (into the hand)' – *käte-ni* 'hand-POSS.1.SG (my hand; my hands; my hand's)'. A possessive ending (here *-ni*) triggers the inflectional stem *käte-* just as inflectional endings like the illative do.

The paradigm of possessive endings is:

	singular	plural
1st person	<i>ni</i>	<i>mme</i>
2nd person	<i>si</i>	<i>nne</i>
3rd person	<i>nsA/Vn</i>	=

Tab. 124.6: Possessive endings

Thus: *kirja-ni* 'my book', *kirja-si* 'your (singular) book', *kirja-nsa* 'his/her/their book', *kirja-mme* 'our book', *kirja-nne* 'your (plural) book', *kirja-ssa-mme* 'in our book', *kirjo-illa-an* 'with his/her/their books' (these forms have two more grammatical readings, nominative plural and genitive singular, not listed

here). The 3rd person possessive suffix is identical in singular and plural. It has two variants, *-nsA* (further specified by vowel harmony) which is possible in all contexts where possessive suffixes may occur, and *-Vn* which is possible (and the dominating variant) after overt case endings ending in a sequence of consonant – vowel: *kirja-ssa-an* ‘in his/her/their book’, *kirjo-i-lla-an* ‘with his/her/their books’. Alternative forms such as *kirja-ssa-nsa* and *kirjo-i-lla-nsa* are possible but sound archaic.

Finally, we pose the question how many distinct inflectional forms a Finnish noun may have. Given two numbers, fifteen cases, and six distinct possessive suffixes,  $2 \times 15 \times 7$  would yield a theoretical maximum of 210 forms (note that zero is the 7th possessive alternative), under the assumption that all slots would be populated by a distinct form. However, an ordinary noun does not have forms like accusative singular, and plural, instructive singular and plural, or comitative singular, with or without possessives. Furthermore, as noted above, the possessive forms of the grammatical words instantiating the feature combinations nominative singular, genitive singular, and nominative plural are identical. Thus, an ordinary noun such as *kirja* ‘book’ with no additional free variation has 158 segmentally distinct word-forms.

## 5. Morphotactic structure and basic inflectional categories of finite verb-forms

**Finiteness** in Finnish grammar is linked to the inflectional categories tense/mood and (grammatical) person, especially the latter

one. The morphotactic structure of finite verb-forms is illustrated in Tab. 124.7.

DERIV stands for all possible combinations of deverbal derivational endings (cf. 11).

The inflectional category **voice** is either active (zero) or passive (*-tta-*, *-ta-*). The Finnish passive is not a precise equivalent of the Indo-European syntactic passive construction even if it is functionally related. The Finnish passive basically expresses that the (semantic) agent is indefinite and human, and grammatically it belongs to the subsystem of person. It is expressed in two morphotactic positions, **VOICE** and **PERSON** above, i.e. discontinuatively. In the **PERSON** position passive is represented by the ending *-Vn* on which more will be said below. A more appropriate name for the Finnish “passive” voice subcategory would be “indefinite” or “indefinite person”. As seen in the glosses above, it corresponds to such overtly agentless (i.e. *by-phrase-less*) Indo-European passive constructions as English “it is said, it was said”.

The morphophonology of the Finnish passive morpheme is complex, it is partly grammatically, partly phonologically conditioned. In the present indicative, its overt morph shapes are either *-tA-* or *-dA-*: *sano-ta-an* ‘it is said’, *syö-dä-än* ‘people eat’. In the past tense and in other moods, the overt manifestation in **VOICE** position is either *-tta-* or *-ta-*, depending upon the inflectional type of the verb stem, and with further absence of the *-A-* in front of past tense *-i*. Thus: *sano-tta-isi-in* ‘one would say’ (conditional mood), *sano-tta-ne-en* ‘one probably says’ (potential mood), *sano-tta-ko-on* ‘may one say’ (imperative/optative mood), *sano-tt-i-in* (past tense);

ROOT	DERIV	VOICE	TENSE/MOOD	PERSON	CLITIC	gloss
<i>sano</i>				<i>n</i>		‘say-1.SG ((I) say)’
<i>sano</i>				<i>o</i>		‘say-3.SG ((he/she) says)’
<i>sano</i>				<i>vat</i>		‘say-3.PL ((they) say)’
<i>sano</i>		<i>i</i>		<i>t</i>		‘say-PAST-2.SG ((you) said)’
<i>sano</i>		<i>isi</i>		<i>mme</i>		‘say-COND-1.PL ((we) would say)’
<i>sano</i>		<i>ne</i>		<i>e</i>		‘say-POT-3.SG ((he/she) probably says)’
<i>sano</i>		<i>i</i>				‘say-PAST(3.SG) ((he/she) said)’
<i>sano</i>		<i>ko</i>		<i>on</i>		‘say-IMP/OPT-3.SG (may (he/she) say)’
<i>sano</i>	<i>ta</i>			<i>an</i>		‘say-PASS-4 (it is said)’
<i>sano</i>	<i>tt</i>	<i>i</i>		<i>in</i>		‘say-PASS-PAST-4 (it was said)’
<i>sano</i>	<i>tta</i>	<i>isi</i>		<i>in</i>		‘say-PASS-COND-4 (it would be said)’
<i>sano</i>	<i>tta</i>	<i>ko</i>		<i>on</i>		‘say-PASS-IMP/OPT-4 (may it be said)’
<i>sano</i>		<i>i</i>		<i>ko</i>		‘say-PAST-1.SG-INT (did I say?)’

Tab. 124.7: Morphotactic structure of finite verb-forms

*syö-tä-isi-in* ‘one would eat’, *syö-tä-ne-en* ‘one probably eats’, *syö-tä-köön* ‘may one eat’, *syö-t-i-in* ‘one ate’.

Some Finnish grammarians have treated the passive morpheme in VOICE position as a derivational rather than as an inflectional element. As seen from the scheme above, the voice morpheme does indeed occur immediately after the derivational endings. Such borderline problems are difficult to solve conclusively. The main argument for treating voice as an inflectional category is its productivity. All verbs with an appropriate lexical meaning (i.e. able of having a human agent) have passive forms. For obvious reasons, verbs like *sata-* ‘rain’ do not have passive forms: \**sade-ta-an* ‘it is rained’.

**Tense** and **mood** belong to the same morphotactic position. Finnish has two simple tenses, present (zero) and past (-i or -si, depending upon verb type). There are four moods: indicative (zero), conditional -isi ‘should’, potential -ne ‘probably’, and imperative/optative with partly amalgamated endings in the different persons:

	singular	plural
1st person	non-existing	<i>kAA-mme</i>
2nd person	—	<i>kAA</i>
3rd person	<i>kO-On</i>	<i>kO-Ot</i>
4th person		<i>kO-On</i>

Tab. 124.8: Imperative/optative

Of the verb *sano-* ‘say’ we thus obtain the following tense and mood forms in the 1st person singular: *sano-n* ‘I say’ (present indicative), *sano-i-n* ‘I said’ (past), *sano-isi-n* ‘I would say’ (conditional), *sano-ne-n* ‘I will probably say’ (potential). Examples of imperative forms in the other persons: *sano* ‘say!’ (2 singular), *sano-ko-on* ‘may he/she say’ (3 singular), *sano-kaa-mme* ‘let us say’ (1 plural), *sano-kaa* ‘say!’ (2 plural), *sano-ko-ot* ‘may they say’ (3 plural), *sano-tta-ko-on* ‘may one say’ (4th person, i.e. passive).

Simple tenses and moods belong to the same morphotactic position class as there are no past tense conditionals, potentials, or imperatives. The unmarked status of the present indicative is obvious, its surface exponent is zero. The potential mood is rare and even archaic, its use is confined to certain variants of written formal style. The morphological relation between tense/mood endings and person endings is transparent (agglutinative)

except for the imperative the paradigm of which is seen in Tab. 124.8. A form such as *sano-kaa* ‘say (2nd plural)’ seems to have no overt person ending. The unmarked status of the 2nd person in the imperative paradigm is obvious, both in singular and plural. In Finnish as in many other languages, the shortest form in the whole paradigm of verbal inflected forms is the 2nd person imperative singular: *sano* ‘say!', *hyppää* ‘jump!', *lue* ‘read!‘.

The personal endings of finite verb forms are (excluding the imperative forms just mentioned):

	singular	plural
1st person	<i>n</i>	<i>mme</i>
2nd person	<i>t</i>	<i>tte</i>
3rd person	<i>V</i>	<i>vAt</i>
4th person		<i>Vn</i>

Tab. 124.9: Personal endings of finite verb forms

The 3rd person singular ending -V is realized as a copy of the preceding vowel, i.e. as vowel lengthening. This is a process of reduplication similar to those encountered in 4. If the verb stem ends in a long vowel or a diphthong, or is inflected in the past tense or the conditional mood, there is no overt 3rd person singular ending. Examples (all 3rd person singular) formed on the verb stems *sano-* ‘say’, *tule-* ‘come’, *syö-* ‘eat’, and *hyppää* ‘jump’: *sano-o* ‘says’, *tule-e* ‘comes’ *syö* ‘eats’, *hyppää* ‘jumps’, *sano-i* ‘said’, *tul-i* ‘came’, *sö-i* ‘ate’, *hyppä-si* ‘jumped’, *sano-isi* ‘would say’, *tul-isi* ‘would come’, *sö-isi* ‘would eat’, *hyppä-isi* ‘would jump’. The indefinite 4th person (passive) ending is indeterminate between singular and plural. The ambivalent nature of the passive in regard to number is seen i.a. in the fact that the passive imperative person ending -*kO-On* is identical to the active 3rd person singular imperative ending, not to the active 3rd person plural ending (active *sano-ko-on* vs. passive *sano-tta-ko-on* but \**sano-tta-ko-ot*). On the other hand, syntactically the passive takes its predicate complement in the plural, not in the singular (*Oi-la-an ilois-i-a* ‘Let us be happy’, with the predicate complement adjective *ilois-i-a* in the partitive plural).

The personal endings yield to a straightforward analysis in terms of markedness theory. The singular endings are shorter than the corresponding plural ones. The lightest

ending of all is the 3rd person singular. Thus, formal simplicity of ending composition reflects cognitive and semantic primacy.

The number of distinct inflectional finite word-forms of a normal Finnish verb is around 40.

## 6. Morphotactic structure and basic inflectional categories of nonfinite verb-forms

The morphotactic structure of nonfinite verbal word-forms is illustrated in Tab. 124.10.

DERIV again stands for all possible combinations of deverbal derivational endings. There are two subclasses of nonfinite verb-forms: infinitives (above the line in the scheme) and participles (below the line). The infinitives function syntactically as nouns, the participles function either as parts of verb chains to express complex tenses, or as pre-modifying adjectives. Morphologically, every infinitive and participle has a marker (functor morpheme) of its own in position class NONFIN in the scheme. The very notion "nonfinite verb-form" implies that these are forms that lack (finite) tense/mood and person inflection, but do have at least some nominal inflectional properties. The participles have full nominal inflection in number, case, and possessive. Because they are basic verbs, they also inflect for voice. These forms are thus

the most complex ones in the Finnish inflection system. As for the infinitives, there are heavy idiosyncratic restrictions on what nominal properties every single infinitive has. No infinitive inflects for number, and no infinitive takes even close to all case forms. As seen from the glosses above, the nonfinite word-forms frequently correspond to Indo-European phrases or even clauses. In such constructions the fairly synthetic nature of Finnish word structure emerges most clearly.

There are three infinitives, normally called the I, II, and III infinitive in the Finnish grammar tradition. Infinitive I has one of the markers *-A*, *-dA*, *-tA* and occurs in two cases, endingless ("nominative", historically a remnant of an older lative) and translative *-ksI*. The translative takes further possessive endings. The I infinitive occurs in the active only. Thus: *sano-a* '(to) say', *syö-dä* '(to) eat', *juosta* '(to) run', *sano-a-kse-ni* 'in order for me to say', *syö-dä-kse-si* 'for you to eat'.

The II infinitive expresses manner or time and occurs in either instructive or inessive case. The functor endings correspond closely to those of inf. I: *-e*, *-de*, *-te*. Thus: *sano-e-n* 'saying', *sano-e-ssa-an* 'when he/she is saying', *syö-de-ssä-mme* 'as we eat ~ were eating', *sano-tta-e-ssa* 'as one says (pass.)'.

Infinitive III has the marker *-mA* and the largest infinitival array of cases: *sano-ma-ssa* 'saying' (inessive), *sano-ma-sta* 'from saying' (elative), *sano-ma-an* 'to say' (illative), *sano-*

ROOT	DERIV	VOICE	NONFIN	NUMBER	CASE	POSS	CLITIC	gloss
<i>sano</i>		<i>a</i>						'say-INF <sub>1</sub> ((to) say)'
<i>sano</i>		<i>a</i>			<i>kse</i>	<i>ni</i>		'say-INF <sub>1</sub> -TRANS-POSS.1 (in order for me to say)'
<i>sano</i>		<i>e</i>			<i>n</i>			'say-INF <sub>2</sub> -INSTR (saying)'
<i>sano</i>		<i>e</i>			<i>ssa</i>	<i>si</i>		'say-INF <sub>2</sub> -INES-POSS.2 (when you are saying)'
<i>sano</i>	<i>tta</i>	<i>e</i>			<i>ssa</i>			'say-PASS-INF <sub>2</sub> -INES (when it is said)'
<i>sano</i>		<i>ma</i>			<i>an</i>			'say-INF <sub>3</sub> -ILL ((to) say)'
<i>sano</i>		<i>ma</i>			<i>ssa</i>			'say-INF <sub>3</sub> -INES ((it/was) saying)'
<i>sano</i>		<i>va</i>						'say-PART.PRES (saying)'
<i>sano</i>		<i>v</i>	<i>i</i>		<i>en</i>			'say-PART.PRES-PL-GEN (of those that are saying)'
<i>sano</i>	<i>tta</i>	<i>v</i>	<i>i</i>		<i>ssa</i>			'say-PASS-PART.PRES-PL-INES (in those that may be said)'
<i>sano</i>			<i>nut</i>					'say-PART.PAST (said)'
<i>sano</i>			<i>ne</i>	<i>i</i>		<i>ta</i>		'say-PART.PAST-PL-PRTV (those (indef.) that were said)'

Tab. 124.10: Morphotactic structure of non-finite verb-forms

*ma-lla* ‘by saying’ (adessive), *sano-ma-tta* ‘without saying’ (abessive).

The participles have full nominal inflection. There are five participles: the present participle in the active (-vA) and the passive (-tTA-vA), the past participle in the active (-nUt) and passive (-tTU), and finally the agent participle (-mA). Examples: *sano-va* ‘saying’, *sano-tta-va* ‘that must/can/should be said’, *sano-nut* ‘said’ (act.), *sano-ttu* ‘said’ (pass.), *sano-ma* ‘said (by somebody)’. The present passive participle expresses lexicalized modal meanings. The morphological components of the past passive participle form an unsegmentable portmanteau morph. The agent participle is a nominalized relative clause. The noun phrase *Kalle-n kerto-ma tarina* thus means ‘Kalle-GEN tell-AG.PART story (a story (that was) told by Kalle)’, where the agent participle *kerto-ma* takes normal concord inflection in case and number as determined by the head noun, e.g. *Kalle-n kerto-m-i-ssa tarino-i-ssa* ‘Kalle-GEN tell-AG.PART-PL-INNESS story-PL-INNESS (in the stories told by Kalle)’.

Because each of the five participles has all case, number, and possessive forms of the nouns (and many have further comparative and superlative forms that also take ordinary inflection), it is obvious that the number of distinct inflected nonfinite word-forms of every single verb is in the thousands.

## 7. Clitics

Cliticized bound morphemes may be appended at the end of nominals, verbs (finite and nonfinite alike), and even at the end of many particles. Clitics differ from the inflectional bound morphemes treated earlier in that they never exert any morphophonological influence upon the preceding word. Clitics are genuinely cliticized to ready-made word-forms, without any stem-selection constraints so typical of inflection proper. In general terms, one may state that the clitics have discourse-related meanings. Mostly these are hard to pinpoint and gloss.

The most important clitics are -*kO* (signalling yes/no questions), -*kin* ‘also’, -*kAA*n ‘even’ (occurring in negated clauses), -*hAn* ‘emphasis; I assume you know’, -*pA* ‘emphasis; surprise’. There are also some combinations of clitics like -*kO-hAn*, -*pA-hAn*, -*kin-kO*.

As an example of a genuinely Finnish type of construction, consider the postpositional phrase *talo-n viere-en-kin* ‘also close to the

house’. The postposition *viere-en* is inflected in the directional illative case and further contains the clitic -*kin*.

## 8. Morphophonological alternations

Typical agglutinating languages have next to no morphophonological alternations, especially not grammatically conditioned ones. Finnish has a wealth of such alternations, several tens of pervasive ones. Therefore Finnish is not a typical agglutinating language. Most of these alternations are due to Proto-Finnic sound changes (cf. Posti 1953 for details).

The most conspicuous alternation is called **consonant gradation** (Finn. *astevaihtelu*), originally a lenition process of the stops /p t k/. Consonant gradation applies across the Finnish vocabulary to all inflectable words that have stop consonants in certain positions, i.e. to nominals and verbs alike. Roughly 20 Prozent of the words in the Finnish vocabulary are subject to consonant gradation. The essence of consonant gradation is that the phonemically long (= double) stops /pp tt kk/ alternate with the corresponding simple (short) ones in certain partly morphologically and partly phonologically definable contexts, e.g. *kauppa* ‘shop’ – *kaupa-ssa* ‘in (the) shop’, *matto* ‘mat’ – *mato-lla* ‘on (the) mat’, *kukka* ‘flower’ – *kuka-t* ‘flowers’. In the same contexts, simple stops undergo various types of lenition including voicing, assimilation, spirantization, and loss, e.g. *mato* ‘worm’ – *mado-t* ‘worms’, *kulta* ‘gold’ – *kulla-ssa* ‘in (the) gold’, *luku* ‘chapter’ – *luvu-t* ‘chapters’, *koko* ‘size’ – *koot* ‘sizes’. Some examples with verbs: *tappa-a* ‘(to) kill’ – *tapa-te* ‘you kill’, *katta-a* ‘(to) lay the table’ – *kata-mme* ‘we lay the table’ – *kato-i-tte* ‘you laid the table’, *luke-a* ‘(to) read’ – *hue-n* ‘I read’. The underlying grade in these words is called the strong grade, the lenited counterpart is the weak grade.

The context of the strong/weak grade alternation is a complex one. It is morphological in the sense that weak grades occur only in front of certain morpheme or word boundaries, i.e. in connection with inflectional processes invoking morphemes that must be morphologically defined (e.g., for nouns case but not possessive morphemes). On the other hand, there is a strong partial phonological subcondition as well, that weak grades occur in closed syllables, except if the sonant is a long vowel. Thus, in several common nomi-

nal inflectional types, we find the following distribution of strong and weak grades (syllable boundaries are indicated by dots "."):  
*kuk.ka* 'flower' (nominative singular, strong grade, open syllable), *ku.ka-n* (genitive singular, weak, closed), *kuk.ka-an* (illative singular, strong, long sonant even if closed syllable), *ku.k-i-s.sa* (iness. plural, weak, closed), *kuk.ka-na* (ess. singular, strong, open), *ku.ka-k.si* (tra. singular, weak, closed), *kuk.ka-m.me* (1st person possessive plural, which is not a case morpheme and therefore the grade remains strong even if the syllable is closed). Consonant gradation in verbs is even more clearly morphologized and estranged from the closed syllable criterion of weak grades as proven by weak forms such as *en ta.pa* 'I don't kill' (negative stem, weak, open), *ta.pe.-taan* (present passive, weak, open).

Consonant gradation is made even more complex by other inflectional types where the weak grade occurs in the base form and the strong grade in most other forms, e.g. *kide* 'crystal' (nominative singular, weak, open) – *ki.tee-n* (genitive singular, strong, closed).

Another pervasive morphophonological alternation is the mutation of vowels in front of certain suffixes with an initial *i*-vowel. Thus, before plural *-i* in nominals and past tense *-i* in verbs, long stem vowels are shortened, diphthongs ending in *-i* drop their *-i*, *-e* (in some inflectional types) is dropped, *-i* changes to *-e*, *-a* changes to *-o* if there is one of the vowels /i e a/ in the first syllable, etc.: *maa* 'country' – *ma-i-ssa* 'in (the) countries', *saa-n* 'I get' – *sa-i-n* 'I got', *hai* 'shark' – *ha(-)i-ssa* 'in (the) shark(s)', *kiele-ssä* 'in (the) language' – *kiel-i-ssä* 'in (the) languages', *mene-n* 'I go' – *men-i-n* 'I went', *lasi* 'glass' – *lase-i-ssa* 'in (the) glasses', *pala* 'piece' – *palo-i-ssa* 'in (the) pieces', *anta-a* '(to) give' – *anno-i-n* 'I gave' (note vowel mutation and consonant gradation in the same word-form).

A much discussed alternation (e.g. Skousen 1989) is the assimilation of /t/ to /s/ in verb stems when occurring before a past tense *-i* and the stem vowel simultaneously alternates with zero, e.g. *tunte-* 'feel' – *tunsi-* 'felt', *kiertä-* 'circulate' – *kiers-i* 'circulated'.

Initial consonant doubling is a sandhi-type alternation taking place over certain morphological boundaries, especially over word boundaries when the previous word belongs e.g. to the nominal inflectional type *vene*, is an infinitive I in the base-form ("nomi-

native"), an imperative in the 2nd person singular, or a negative verb stem in the present indicative. Examples: <*vene tul-i*> [venet:uli] 'the boat came' (an *e*-noun), <*en halua tul-la tänne*> [enhaluat:ul:at:än:e] 'I don't want to come here' (negative verb stem *halua*, infinitive *tulla*), <*tule täinne*> [tulet:än:e] 'come here' (imperative). This is a "morphophonetic" alternation in the sense that native speakers do not intuitively feel that the preceding word triggering the consonant doubling would itself end in a consonant.

In connection with how inflectional stems are formed, there occur tens of additional inflection-type-specific morphophonological alternations, e.g. the *s/V*-alternation in *vieras* 'guest' – *viera-a-t* 'guests', or the *e/zero-* and *s/ks*-alternations in *tulos* 'result' – *tulokse-t* 'results'.

## 9. Nominal inflectional types

There is no well established interpretation of how many inflectional types Finnish nominals should be subdivided into. There are two extreme opinions. One states that there are no declensional (inflectional type) differences whatsoever between nominals, all are inflected in the same way and the complex morphophonological differences between some of them are supposed to be described by more or less abstract rules. This view dates back to the 1840s (Elias Lönnrot), and similar views were expressed by early generativists in the 1960s, especially Kalevi Wiik (1967). On the other hand, the classical comprehensive dictionary of Standard Finnish (*Nykysuomen sanakirja*, 1951–61) postulated no less than 82 nominal inflectional types, often distinguished by minute allomorphic differences. An intermediate view (Karlsson 1983) postulates 8 nominal inflectional types, the main criteria being the lexical markings necessary if the nominative singular is taken to be the basis of describing inflectional forms. These types are:

nominative singular	genitive singular	partitive singular	partitive plural
<i>risti</i>	<i>risti-n</i>	<i>risti-ä</i>	<i>riste-j-ää</i>
<i>ovi</i>	<i>ove-n</i>	<i>ove-a</i>	<i>ov-i-a</i>
<i>susi</i>	<i>sude-n</i>	<i>sut-ta</i>	<i>sus-i-a</i>
<i>nalle</i>	<i>nalle-n</i>	<i>nalle-a</i>	<i>nalle-j-a</i>
<i>hame</i>	<i>hamee-n</i>	<i>hame-tta</i>	<i>hame-i-ta</i>
<i>varis</i>	<i>varikse-n</i>	<i>varis-ta</i>	<i>variks-i-a</i>
<i>vieras</i>	<i>viera-a-n</i>	<i>vieras-ta</i>	<i>viera-i-ta</i>
<i>rakka-us</i>	<i>rakka-ude-n</i>	<i>rakka-ut-ta</i>	<i>rakka-uks-i-a</i>

Tab. 124.11: Inflectional types

infinitive I	PRES.INDIC-1.SG	PAST (3.SG)	past participle
<i>alka-a</i>	<i>ala-n</i>	<i>alko-i</i>	<i>alka-nut</i>
<i>haka-ta</i>	<i>hakkaa-n</i>	<i>hakka-si</i>	<i>hakan-nut</i>
<i>saa-da</i>	<i>saa-n</i>	<i>sa-i</i>	<i>saa-nut</i>
<i>nous-ta</i>	<i>nouse-n</i>	<i>nous-i</i>	<i>nous-sut</i>
<i>lämme-tä</i>	<i>lämpene-e</i> (3. pers.)	<i>lämpen-i</i>	<i>lämmen-nyt</i>

Tab. 124.12: Verbal inflectional classes

These types each have typical inflectional stems, typical configurations of ending allomorphs in many cases, and typical morphophonological alternations. E.g., the *varis*-words have an inflectional vowel stem that ends in *-kse*, the partitive singular ending *-tA*, and loss of the stem vowel *-e* before plural *-i*.

## 10. Verbal inflectional types

A similar situation obtains with regard to verbs. Elias Lönnrot's opinion was that there is just one conjugation in Finnish. On the other hand, the maximalistic interpretation presented in *Nykysuomen sanakirja* is that there are 45 verbal inflectional types. An intermediate interpretation in the framework of concrete surface-oriented morphology is that there are 5 basic ways of inflecting verbs (Tab. 124.12).

Each type is characterized by typical inflectional stems, typical configurations of ending allomorphs, and typical morphophonological alternations. Thus, the verb type *salata* has the infinitive ending *-tA* after a vowel, weak grade with respect to consonant gradation in the infinitive, a long vowel or two vowels before the finite person endings, strong grade in the finite forms, the past tense ending *-si* occurring after a strong grade, and the past participle ending *-nUt* occurring after a weak grade and final *-n* in the verb stem.

## 11. Derivational morphology

Derivation is here taken to mean formation of new lexemes from existing ones using bound morphemes that are not inflectional. There are some 150–200 derivational endings in Finnish, depending upon how certain unfrequent and lexicalized instances are interpreted. Many of the Finnish derivational

morphemes are highly productive. From a noun such as *kahvi* 'coffee', one can form new nouns (*kahvi-la* 'coffee-shop'), adjectives (*kahvi-mainen* 'coffee-like'), and verbs (*kahvi-tta-a* 'serve coffee', where the final ending *-A* is the first infinitive). Morphotactically the derivational endings are located between the (minimal) stem and the inflectional endings (cf. the morphotactic schemata given in 4–6).

Derivation of nouns from nouns: *-nen* 'small N', e.g. *kala* 'fish' – *kala-nen* 'small fish', *muna* 'egg' – *muna-nen* 'small egg'; *-tAr* 'feminine N', e.g. *Helsinki* – *helsingi-tär* 'woman from Helsinki', *orja* 'slave' – *orjar-tar* 'female slave'; *-lA* 'N-place', e.g. *pappi* 'clergyman' – *pappi-la* 'vicarage'.

Derivation of nouns from verbs: *-ja* (nomen agentis), e.g. *luke-* 'read' – *luki-ja* 'reader', *anta-* 'give' – *anta-ja* 'one who gives'; *-minen* (nomen actionis), e.g. *luke-* – *luke-minen* 'reading', *täyty-* 'must' – *täyty-minen* 'being forced to'; *-mO* 'place', e.g. *korja-* 'repair' – *korja-mo* 'repair shop', *kampaa-* '(to) comb' – *kampaa-mo* 'hair salon'.

Derivation of verbs from verbs: This is a large class of derivational morphemes typical of Finnish. The basic classes are formation of causative, reflexive, momentaneous, and frequentative verbs. Causative verb derivatives: *avaa-* 'open' – *ava-utta-* 'have opened', *julkais-* 'publish' – *julkais-utta-* 'have published'. Reflexive verbs: *avaa-* 'open' – *ava-utu-* 'open oneself', *jättää-* 'leave' – *jättää-yty-* 'leave oneself'. Momentaneous verbs: *istu-* 'sit' – *ist-ahta-* 'sit down fast', *kiehu-* 'boil' – *kieh-ahta-* 'start boiling all of a sudden'. Frequentative verbs: *ui-* 'swim' – *ui-skentele-* 'keep swimming around', *ampu-* 'shoot' – *ammu-skele-* 'shoot here and there', *sylke-* 'spit' – *sylje-skele-* 'keep spitting'.

Derivation of verbs from nouns: *kone* 'machine' – *kone-ista-* 'mechanize', *suomalainen*

'Finn, Finnish' – *suoma-laista-* 'fennicize, make Finnish'; *auto* 'car' – *auto-istu-* 'become motorized'.

Sequences of several derivational morphemes are common, e.g. *nukku-* 'sleep' – *nuk-aht-ele-* 'keep on falling asleep, continuously be about to fall asleep' (momentaneous – frequentative verb), *pieni* 'small' – *piene-nt-y-* 'get smaller' (causative – reflexive), *laula-* 'sing' – *laula-tt-ele-* 'keep on to make sing' (causative – frequentative). The word *horj-aht-ele-mattom-uus* contains four derivational morphemes. It is a nominalization (-UUs) of a caritive adjective (-mAtOn : -mAttOmA-) derived from a momentaneous frequentative derived verb stem the root of which is *horju-* 'stagger'. The meaning of *horj-aht-ele-mattom-uus* is roughly 'the property of not keeping on making small staggerings'. This word is not made up and does not make an overly complex impression. Derived words of this degree of complexity are quite ordinary, cf. *asu-nto-la-mais-uus* 'the property of being like a dormitory', the root of which is the verb *asu-* 'live (e.g. in a house)' from which *asu-nto* 'dwelling' is derived which in turn is the basis for *asu-nto-la* 'dormitory'.

## 12. Compounding

Another feature typical of Finnish is the frequent use of compounding, especially nominal compounding, as a means of word-formation. Compounding and derivation often cooccur in complex words such as *täyde-nnyys=koulu-t-us=kys-el-y* 'further training inquiry' (where “=” indicates compound boundaries). The minimal free root morphemes here are *täysi* : *täyde-* 'full', *koulu* 'school', and *kysy-* 'ask', *-nn-* and *-t-* are causatives, *-el-* is frequentative, *-Us* and *-U* are nominalizers. The word *pysä-hty-mis=merki-n=ant-o=nappi* means 'button giving the stop signal', *hallitukse-ssa=ol-o=aika* 'period ("being") in the government'. As the last example indicates, compounds may occasionally contain even inflected words like *hallitukse-ssa* 'in the government', here inflected in the inessive case. Compound words with parts inflected in typical adverbial cases resemble incorporation.

The fairly synthetic nature of Finnish is a joint consequence of inflectional morphology, derivational morphology, and compounding.

## 13. Morphological productivity and diachronic tendencies in the morphological system

Most of the complex morphophonological alternations characterizing the various inflectional types of Finnish are due to Proto-Finnic sound changes many of which have been described and phonologically interpreted by Posti (1953). Having lost their original phonetic or phonological motivations, and having been at least partially morphologized, many of the current morphophonological alternations are perceived as morphological complications. There are many simplificatory diachronic tendencies in contemporary Finnish morphology. The basic tendency is to eliminate the most opaque morphophonological alternations as part of the well-known strive to establish the "one meaning – one form" principle. Thus, inflectional types where the root has many stem alternants are mostly not fully productive, and diachronically part of a process of gradually becoming more and more unfrequent, finally (perhaps) being ousted.

Thus, there are three potential ways of inflecting a noun with a base-form (nominative singular) ending in *-si*, viz. *lasi* 'glass' – *lasi-n* – *lasi-a* – *lase-j-a*, *jousi* 'bow' – *jouse-n* – *jous-ta* – *jous-i-a*, and *susi* – *sude-n* – *sutta* – *sus-i-a*, with two, three, and four stem allomorphs, respectively. The productive type is *lasi*, with the least amount of stem allomorphy, and consequently the closest conformance to the one meaning – one form principle. Neologisms are thus inflected e.g. *Masi* – *Masi-n* – *Masi-a* (hypocoristic form of the male name *Matti*), rather than with a genitive singular \**Mase-n* or even less \*\**Made-n* pro *Masi-n*. For nouns with *-e* in the base-form, the productive inflectional type is *nalle* – *nalle-n* – *nalle-a* – *nalle-j-a* (one stem), rather than *hame* – *hamee-n* – *hame-tta* – *hame-i-ta* (two or even more stems).

The same universal morphological tendency explains why so-called contracted verbs like *kaapa-ta* 'hijack' – *kaappaa-n* – *kaappa-si* – *kaapan-nut* – *kaapa-ta-an* (present passive) are more productive than the verb type *tappa-a* 'kill' – *tapa-n* – *tappo-i* – *tappa-nut* – *tape-ta-an*. Even if the contracted verbs also display a lot of stem allomorphy, as a whole their first two syllables are more invariant than the first two syllables of the *tappa-a* type.

Many individual morphophonological alternations have become more or less unproductive in (many or all types of) new words. Thus, most new underived minimal noun stems are no longer subject to qualitative consonant gradation, e.g. *auto* ‘car’ – *auto-n* (not \**audio-n*), *laku* ‘licorice’ – *laku-n* (\**lau-n*). Quantitative consonant gradation is more resistant in this regard, but there are clear signs of its becoming partially unproductive as well, e.g. in names such as *Beatty* – *Beatty-n*, *Stokka* ‘Stockmann, a department store in Helsinki’ – *Stokka-lla*.

As for morphological categories, the use of possessive suffixes is fluctuating in colloquial Finnish in favour of omission especially in nouns. Thus, standard Finnish (*minun*) *auto-ni* ‘my car’ is normally expressed analytically with omission of the possessive ending, *minun auto* or *mun auto*, in colloquial speech. On the basis of a careful sociolinguistic analysis of data starting from the 1930s, Pauonen (1996) has made the prognosis that the last and final synthetic possessive construction in Helsinki colloquial Finnish would be encountered some time in the year 2040.

The gradual loss of possessives is part of a more general process where Finnish morphology is developing in an analytic direction. Other manifestations of this tendency are the more or less obsolete nature of the potential mood in verbs (*sano-ne-e* ‘probably says’ becoming analytic *taita-a sano-a*), and the construction *me men-nä-än* ‘we go’ pro (*me*) *mene-mme* in the 1st person plural of finite verbs (where the ordinary passive verb form is expansive).

#### 14. Processing and acquisition of Finnish morphology

On the basis of single-word experiments dealing with aphasics as well as lexical decision and eye movement registration tests performed on normals, Jussi Niemi, Matti Laine and their colleagues have designed a processing model for Finnish nouns. For the input and central mental lexicons, their Stem Alloform / Inflectional Decomposition (SAID) model suggests morphological decomposition of inflected but not of derived noun forms (the most frequent inflected forms are an exception). Furthermore, the model claims that stems are represented by their allomorphs and not by deriving the stems using morphophonological rules or the like. The SAID

model also predicts that both inflected and productive derived form have decomposed representations in the output lexicon (Niemi et al. 1994; Laine et al. 1995). This work falsifies two views that have been widely held in current psycholinguistics: that the “Full Listing Hypothesis” (= all word-forms lexically stored) would be a viable alternative for describing cognitive or mental storage of word-forms in morphologically complex languages, and that morphophonological stem alternations would be processed by analogues of directional rules.

Jorma Toivainen (1980) has studied the L1 acquisition of Finnish inflectional affixes. He showed i.a. that the starting point of acquiring the verbal inflectional system is an unmarked basic 3rd person singular verb-form. As differentiation begins, the child first acquires the 1st person singular form and a primitive past tense form. Noun acquisition normally starts from nominative singular, occasionally from partitive singular or plural.

Maisa Martin (1995) did an empirical study of how L2 learners of Finnish acquired nominal inflection. Her error analysis results supported the SAID model of lexical storage: nouns stems are mentally stored as such. Morphophonologically complex words were more difficult to process than simpler ones. No extant model of morphological description (rule-based, paradigm-based, analogical) as such was potent enough to describe the learners’ errors.

#### 15. Illustrative text

(8) <i>Afrikassa</i>	<i>eli</i>	<i>pystyssä</i>
<i>Afrika-ssa</i>	<i>el-i</i>	<i>pysty-ssä</i>
<i>Africa-INESS</i>	<i>live-PAST</i>	<i>upright-INESS</i>
<i>käveleviä</i>		
<i>kävele-v-i-ä</i>		
walk-PRES.PART-PL-PRTV		
<i>ihmisapinoita</i>	<i>viidestä</i>	<i>kahteen</i>
<i>ihmis=apino-i-ta</i>	<i>viide-stä</i>	<i>kahte-en</i>
primate-PL-PRTV	five-ELAT	two-ILL
<i>miljoonaan</i>	<i>vuotta</i>	<i>sitten.</i>
<i>miljoona-an</i>	<i>vuot-ta</i>	<i>sitten</i>
million-ILL	year-PRTV	ago
<i>Mistä</i>	<i>ne</i>	<i>tulivat?</i>
<i>Mi-stä</i>	<i>ne</i>	<i>tul-i-vat</i>
Where-ELAT	they(NOM)	come-PAST-3.PL

‘In Africa, there lived upright walking primates from five to two million years ago. Where did they come from?’

## 16. Uncommon abbreviations

4	impersonal
AG.PART	agens participle

## 17. References

- Häkkinen, Kaisa (1996), *Suomalaisen esihistoria kielitieteen valossa* [The Prehistory of Finns in the Light of Linguistics]. Helsinki: Suomalaisen Kirjallisuuden Seura
- Hakulinen, Auli & Karlsson, Fred (3<sup>rd</sup> 1995), *Nykysuomen lauseoppia* [The Syntax of Modern Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura [1<sup>st</sup> 1979]
- Karlsson, Fred (1983), *Suomen kielen äänne- ja muotordakenne* [The Phonology and Morphology of Finnish]. Helsinki: WSOY
- Karlsson, Fred (2002), *Finnish: An Essential Grammar*. London and New York: Routledge
- Koivulehto, Jorma (1983), “Seit wann leben die Urfinnen im Ostseeraum? Zur relativen und absoluten Chronologie der alten idg. Lehnwortschichten im Ostseefinnischen”. *Mémoires de la Société Finno-Ougrienne* 185, 135–157
- Koivulehto, Jorma (1994), “Indogermanisch – Uralisch: Lehnbeziehungen oder (auch) Urverwandtschaft?”. In: Sternemann, Reinhard (ed.), *Bopp-Symposium 1992 der Humboldt-Universität zu Berlin*. Heidelberg: Winter, 133–148
- Laine, Matti & Niemi, Jussi & Koivuselkä-Sallinen, Päivi & Hyönen, Jukka (1995), “Morphological Processing of Polymorphemic Nouns in a Highly Inflecting Language”. *Cognitive Neuropsychology* 12: 457–502
- Martin, Maisa (1995), *The Map and the Rope: Finnish Nominal Inflection as a Learning Target*. Jyväskylä: University of Jyväskylä (Studia Philologica Jyväskylänsia 38)
- Niemi, Jussi & Laine, Matti & Tuominen, Juhani (1994), “Cognitive Morphology in Finnish: Foundations of a New Model”. *Language and Cognitive Processes* 9, 423–446
- Niemikorpi, Antero (1991), *Suomen kielen sanaston dynamiikkaa* [Dynamics of the Finnish Vocabulary]. Vaasa: Universitas Wasaensis (Acta Wasaensia 26)
- Pajunen, Anneli & Palomäki, Ulla (1984), *Tilastotietoja suomen kielen rakenteesta 2* [Frequency analysis of spoken and written discourse in Finnish]. Helsinki: Finnish Research Center for Domestic Languages
- Paunonen, Heikki (1996), “Language Change in Apparent Time and in Real Time: Possessive Constructions in Helsinki Colloquial Finnish”. In: *Samspel och variation: Språkliga studier tillägnade Bengt Nordberg på 60-årsdagen*. Uppsala universitet: Institutionen för nordiska språk, 375–386
- Posti, Lauri (1953), “From Pre-Finnic to Late Proto-Finnic: Studies on the Development of the Consonant System”. *Finnisch-ugrische Forschungen* 31, 1–91
- Salminen, Tapani (1993), “Uralilaiset kielet maailman kielten joukossa” [The Uralic Languages among the Languages of the World]. In: Salminen, Tapani (ed.), *Uralilaiset kielet tänään* [Uralic Languages Today]. Kuopio: Snellman-instituutti (A-sarja 13), 24–30
- Skousen, Royal (1989), *Analogical Modelling of Language*. Dordrecht etc.: Kluwer Academic Publishers
- Toivainen, Jorma (1980), *Inflectional Affixes Used by Finnish-speaking Children Aged 1–3 Years*. Helsinki: Suomalaisen Kirjallisuuden Seura (Suomalaisen Kirjallisuuden Seuran toimituksia 359).
- Wiik, Kalevi (1967), *Suomen kielen morfonemiikka*. Turku: Turun yliopiston fonetiikan laitoksen julkaisuja 3

*Fred Karlsson, Helsinki (Finland)*

## 125. Hebrew (Semitic)

1. The Hebrew language
2. The Hebrew *binyan* system
3. The sign-oriented approach
4. The traditional approach
5. The analysis
6. The data
7. Summary and conclusions
8. Illustrative text
9. References

### 1. The Hebrew language

This chapter deals with Modern Hebrew (also known as Israeli Hebrew, Contemporary Hebrew and ‘Ivrit’) (cf. Berman 1978a:1–4; Glinert 1989:1–5; Rosén 1977: 15–24) the national language of the Jewish majority (approximately five million) of the State of Israel and as a second language for the Jews of the world (as well as the Arab minority residing in Israel). The Hebrew language is usually divided into three or four major historical periods: Classical or Biblical Hebrew (c. 1200 B.C. to c. 200–300 B.C.), Mishnaic or Rabbinical Hebrew (c. 300 B.C. to c. 400–500 A.D.), Medieval Hebrew (c. 500 A.D. to 1700 A.D.) and Modern Hebrew (including the period of the Enlightenment and the revival of Hebrew in Israel) (cf. Bar-Adon 1975; Fellman 1973) c. 1700 A.D. to the present. Despite the wide spread of Jews throughout the world and the ingathering of the exiles in Israel with the multitude of mother tongues they speak, Modern Hebrew is strikingly uniform in its dialects and varieties of usage including both ethnic dialects between Jews of Afro-Asian origin (known as *Sepharadim*) and European-American origin (known as *Ashkenazim*) as well as sociolinguistic and regional dialects. The generic term Hebrew includes all the historical variations of the language which form a fairly comprehensible continuum of a single language rather than what might be considered to be different languages such as Old English versus Modern English (cf. Tobin 1981; Art. 154).

Hebrew, like other Semitic languages, has a fundamentally different structure than the Indo-European languages. In the Semitic languages, the isomorphic connections between phonology, morphology, syntax and semantics are much more overt. The vast majority of the words of the language can be analyzed into consonantal roots signaling broad se-

mantic fields which are combined with fixed morphophonemic patterns for what are traditionally called nominal, verbal and adjectival forms (cf. Art. 73, 74, 77). More often than not, the connections and relationships between the root and these fixed morphophonemic patterns are transparent. This will be illustrated in Tab. 125.1 which presents the triconsonantal (CCC) root *K-T-V* ‘write’ in the basic inflectional and derivational morphological systems of the language (cf. Art. 37, 38). This root (with this transcription reflecting modern Israeli Hebrew pronunciation) has been chosen because it is the one most frequently used to illustrate how Hebrew and other Semitic languages work. The root consonants have been written in upper case and the various affixes in lower case.

It should be clear from Tab. 125.1 that many traditional and familiar concepts such as root, stem, word, part of speech, etc. that have become part and parcel of traditional and neotraditional grammars (derived from grammars of Latin and Greek) may be questionable as an adequate means of describing a language such as Hebrew (cf. Art. 11). Indeed, Tab. 125.1 illustrates that on the most fundamental level the Hebrew lexicon, morphology, and grammar are primarily based on triconsonantal roots which undergo systematic inflectional and derivational processes to form primary verb patterns called *binyanim* (verbal ‘building blocks’) from which secondary *miskalim* (nominal and adjectival ‘form classes’) may be derived. Both these *binyanim* and *miskalim* are signaled by an elaborate set of mostly vocalic affixes (prefixes, infixes, and suffixes) (cf. Art. 54, 55).

### 2. The Hebrew *binyan* system

Some of the most – if not the most – crucial questions in linguistic analyses of the Hebrew language concern themselves with the Hebrew verb conjugation (*binyan* system). These questions include:

- (a) What is the best way to identify and analyze the verb *binyanim*, which includes how many *binyanim* are there?
- (b) Should the *binyan* be considered today as part of the grammar or as part of the lexicon?
- (c) Should the *binyan* still be considered to be a system today?

<i>KaTaV</i>	'wrote' – 3.M.SG past tense
	'reporter' – M.SG noun
<i>KoTeV</i>	'writes' – M.SG present tense/participial/adjectival form
	'writer' – M.SG noun
<i>yiKToV</i>	'will write' – 3.M.SG future tense
<i>KToV!</i>	'write!' – M.SG imperative
<i>KaTaVa</i>	'news report' – F.SG noun
<i>KToVeT</i>	'address' – F.SG noun
<i>KiTUViT</i>	'sub-title' – F.SG noun
<i>KeTuVa</i>	'marriage contract' – F.SG noun
<i>KaTuV</i>	'written' – M.SG adjectival form
<i>KTaV</i>	'handwriting' – M.SG noun
<i>KTiV</i>	'spelling' – M.SG noun
<i>KaTVanit</i>	'typist' – F.SG noun
<i>KaTVanut</i>	'typing' – F.SG noun
<i>niKTaV</i>	'was written' – 3.M.SG passive
<i>KiTev</i>	'inscribed' – 3.M.SG past tense
<i>MiKaTeV</i>	'inscribe' – M.SG present tense/participial/adjectival form
	'inscriber' – M.SG noun
<i>yiKaTeV</i>	'will inscribe'/'will be inscribed' – 3.M.SG future tense
<i>KaTeV!</i>	'inscribe!' – M.SG imperative
<i>KuTaV</i>	'was inscribed' – 3.M.SG passive
<i>miKTaV</i>	'letter' – M.SG present tense/participial/adjectival form
<i>miKTaViot</i>	'stationery' – F.SG noun (children's language)
<i>miKTaVa</i>	'desk' – F.SG noun
<i>meKuTaV</i>	'written', 'inscribed' – M.SG adjectival form
<i>hiKTiV</i>	'dictated' – 3.M.SG past tense
<i>maKTiV</i>	'dictates' – M.SG present tense/participial/adjectival form
	'dictator' (of dictations not countries) – M.SG noun
<i>yaKTiV</i>	'will dictate' – 3.M.SG future tense
<i>haKTeV</i>	'dictate!' – M.SG imperative
<i>haKTaVa</i>	'dictation' – M.SG noun
<i>huKTaV</i>	'was dictated' – 3.M.SG passive
<i>muKTaV</i>	'dictated' – M.SG adjectival form
<i>hitKaTeV</i>	'corresponded' – 3.M.SG past tense
<i>mitKaTeV</i>	'corresponds' – M.SG pres. tense/participial/adjectival form
	'correspondent' ('letter writer') – M.SG noun
<i>yitKaTeV</i>	'will correspond' – 3.M.SG future tense
<i>hitKaTeV!</i>	'correspond!' – M.SG imperative
<i>hitKaTVut</i>	'correspondence' – F.SG noun
<i>nitKatyv</i>	'was corresponded' – 3.M.SG past tense
<i>fiKTev</i>	'rewrote/revised' – 3.M.SG past tense
<i>fuKTaV</i>	'was rewritten/revised' – 3.M.SG passive
<i>fiKTuV</i>	'rewriting/revision' – M.SG noun

Tab. 125.1: The CCC Root *K-T-V* 'write'

Most modern scholars who have dealt with these questions may be divided into two camps based on their different approaches to the problem. The two approaches may be referred to as syntactic and lexicalist based on how they view the *binyanim* and their place in the analysis of Hebrew.

The syntactic approach assumes that there is an underlying paradigmatic system (cf. Art. 24) for *binyanim* as part of the grammar of the language, albeit a system which has not been adequately defined, described and explained. The syntactic approach begins with the traditional grammarians who ana-

lyzed Biblical Hebrew (cf. Gesenius 1910; Kimhi 1847) and was followed along different lines for Modern Hebrew in generative grammar (cf. Berman 1975 a; 1975 b; 1978 a; 1978 b; Bolozky 1978; Glinert 1989), formal logic (cf. Ariel 1971), structural and distributional approaches (cf. Ben-Asher 1972; Rosén 1977; Art. 19, 20); functional grammar (cf. Junger 1988), and language acquisition studies (cf. Berman 1982; Schwarzwald 1981 a; 1981 b; Art. 165). The syntactic approach generally lists seven major *binyanim* classifying them according to function (simple, intensive, causative, reflexive), voice (active, passive, middle) (cf. Art. 108), or types of morphological patterns (types of affixation). The syntactic approach is by far the most prevalent and accepted one despite the fact that all the syntactic analyses admit to exceptions, irregularities, contradictions, and point out overlapping syntactic, semantic and pragmatic functions and features among the *binyanim*. The three most comprehensive grammars of Modern Hebrew in English (Rosén 1977; Berman 1978 a; Glinert 1989 reviewed in Tobin 1991 c) follow this syntactic approach.

The lexicalist approach assumes that the *binyanim* do not constitute a grammatical system but are more lexical in nature. The lexicalists attribute the same functions and features to the same seven *binyanim* to the lexicon that the advocates of the syntactic approach include in the grammar. The lexical approach is fundamentally a reaction to the fact that there are so many unpredictable exceptions, irregularities, contradictions and overlaps within and across the *binyanim*. The lexicalists liken the *binyan* system to nominal word formation (*miskalim*) (cf. Art. 42) and other lexical phonological and morphological processes (cf. Art. 35, 36). The Lexicalists maintain that the *binyanim* are not systematic in the grammatical sense (like inflectional number and gender morphology) but are more lexical in nature (like derivational morphology). The lexicalists (cf. Ornan 1971 a; 1971 b; 1971 c; 1978; 1979; Schwarzwald 1974; 1975; 1981 a–c) justify their approach for the theoretical and methodological reasons of simplicity. This dichotomy between the syntactic and lexicalist approaches is not a purely theoretical issue. These opposed points of view underlie a most important methodological and lexicological problem which directly affects the structure of Hebrew language dictionaries (cf. Art. 171),

studies of language acquisition and language teaching (cf. Art. 177). The lexicalists advocate that each nominal and verbal form appears separately and independently in its own dictionary entry while the advocates of the syntactic approach believe that all verbal (and many nominal) forms should appear together as part of a single dictionary entry under the rubric of the consonantal roots.

Both advocates of the syntactic and lexicalist approaches share certain theoretical and methodological principles:

- (a) They both have primarily devoted their work to the extensive listing and discussion of the various syntactic, semantic and pragmatic functions that may be attributed to each of the seven major *binyanim*.
- (b) They have all tried to deal with the fundamental problem that each *binyan* does not necessarily have a single function or a set of functions which is unique to it alone.
- (c) They have all discovered that similar functions may be assigned to more than one *binyan* and have grappled with the key question of whether these functional overlaps are systematic or arbitrary.
- (d) Despite their opposed syntactic or lexicalist conclusions, they all recognize and try to account for the fact that the *binyan* system appears – on the surface at least – to be asymmetrical and random: not all the consonantal roots appear in each of the seven major *binyan* and where there is an overlap in functions among the *binyanim*, it is not always possible to predict which root will appear in which *binyan* for which specific set of shared functions.

None of the syntactic or lexicalist analyses have gone beyond listing the syntactic, semantic, and pragmatic functions of the seven major *binyanim* and/or grouping the *binyanim* based on their listing of which roots appear for which functions in which *binyanim*. No one has seriously proposed an analysis of the Hebrew *binyanim* based on the linguistic sign (cf. Art. 23), i.e. the possibility that an invariant meaning or a set of invariant meanings may be postulated for each individual *binyan* as part of a larger, semantically motivated paradigmatic grammatical system (cf. Art. 24). Nor has it been considered a viable theoretical or methodological possibility that the principle of invariance (cf. Art. 25) might account for why each *binyan* has the functions it has, and why different *binyanim* may share some of these functions.

### 3. The sign-oriented approach

In this chapter we will follow certain basic theoretical and methodological tenets of sign-oriented linguistics: invariance, markedness (cf. Art. 29) and distinctive feature theory (cf. Art. 28) – to show how the Hebrew *binyanim* may be viewed as a complex paradigmatic system composed of (at least) three interlocked systems. According to the sign-oriented approach used here (cf. Tobin 1988, ed.: xii–xxix; 1989 a: 1–19; 1990: ch. 4; Aphek & Tobin 1988; 1989; 1993; 1994: chs. 1, 2, 8; 1997), the Hebrew *binyan* system may be viewed as a set of systems revolving around the semantic domains of:

- (a) the **objective** versus **subjective conceptualization** of actions, states, or events.
- (b) the **single** versus **multiple perceptions** of these actions, states, or events.
- (c) whether these actions, states, or events are **autonomous** or not, from the deictic point of view of the encoder at the ‘here-and-now’ point of encoding.

In this analysis, there is an eight member conjugation (*binyan*) system which will be reduced to various (usually paired) sets of: objective vs. subjective, single vs. multiple and autonomous vs. non-autonomous conceptualizations and perceptions of actions, states, or events. The classification of these paired sets and subsets is iconically reflected in the partially overlapping and/or phonetically related morphological signals they share. This iconic, or nonarbitrary relationship illustrates the integrality of the signal and meaning sides of the linguistic sign as it functions within larger paradigmatic systems. This iconicity (cf. Art. 30) also reflects the markedness relationship of these linguistic signs within interlocked systems based on distinctive semantic features which cut across the arbitrary distinction between grammar versus lexicon and thus views the grammar

and the lexicon not as two distinct entities but rather as a ‘continuum’ of alternative sign systems and relationships (Tobin 1990: chs. 2–3).

### 4. The traditional approach

#### 4.1. The seven binyanim: definitions

The names of the traditional seven major *binyanim* (PAAL, NIFAL, PIEL, PUAL, HIFIL, HUFAL, HITPAEL) are derived from the triconsonantal (CCC) root *P-‘L* ‘action’ and the fixed morphophonemic pattern of each *binyan*. For the sake of orthographical convenience and simplicity, the names of the *binyanim* will be transcribed in upper case and omit the phoneme /*l*/ ayin of the triconsonantal root *P-‘L*. The most commonly used names for each *binyan* are being used here (e.g. HUFAL as opposed to HOFAL as it is sometimes called). The traditional seven major *binyanim* usually have been defined according to their most frequent function: simple, intensive, causative, reflexive, and/or voice (active, passive). These terms, primary functions, and designations of active and passive voice, have served, over the centuries, as the traditional labels for the *binyanim* as shown in Tab. 125.2.

Although most roots are triconsonantal (CCC), all roots be they bi- (CC), tri- (CCC), quadri- (CCCC) consonantal, etc., become part of the *binyan* system. This ‘root + *binyan* = verb’ process will be illustrated in Tab. 125.3 (for the third person, masculine, singular past forms from which the traditional labels are taken) using the triconsonantal root *K-T-V* ‘write’:

From Tab. 125.3, one can see that – at least from the point of view of the morphophonemic signal side of the linguistic sign – the four active *binyanim* appear to be fairly symmetrical and can be divided into:

(1) PAAL	‘simple’ conjugation (also known as QAL / ‘basic/simple’)
(2) NIFAL	‘passive’ of ‘basic-simple’ conjugation (PAAL/QAL)
(3) PIEL	‘intensive’ conjugation
(4) PUAL	‘passive’ of ‘intensive’ conjugation (PIEL)
(5) HIFIL	‘causative’ conjugation
(6) HUFAL	‘passive’ of ‘causative’ conjugation
(7) HITPAEL	‘reflexive’ conjugation

Tab. 125.2: The *binyan* system: traditional view

root	<i>binyan</i>	pattern	resultative verb
(1) <i>K-T-V</i>	PAAL	<i>CaCaC</i>	<i>KaTaV</i> ‘wrote’ (‘basic/simple’)
(2) <i>K-T-V</i>	NIFAL	<i>niCCaC</i>	<i>niKTaV</i> ‘was written’ (‘b./s.-passive’)
(3) <i>K-T-V</i>	PIEL	<i>CiCeC</i>	<i>KiTev</i> ‘inscribed’ (‘intensive’)
(4) <i>K-T-V</i>	PUAL	<i>CuCaC</i>	<i>KuTaV</i> ‘was inscribed’ (‘intensive-passive’)
(5) <i>K-T-V</i>	HIFIL	<i>hiCCiC</i>	<i>hiKtiV</i> ‘dictated’ (‘causative’)
(6) <i>K-T-V</i>	HUFAL	<i>huCCaC</i>	<i>huKTaV</i> ‘was dictated’ (‘passive-causative’)
(7) <i>K-T-V</i>	HITPAEL	<i>hitCaCeC</i>	<i>hitKaTeV</i> ‘corresponded’ (‘reflexive’)

Tab. 125.3: The *binyan* system: root-pattern-verb

- (a) two simple (i. e. infix only) patterns: PAAL  
*CaCaC* / PIEL *CiCeC*.  
(b) two complex (i. e. prefix + infix) patterns: HIFIL *hiCCiC* / HITPAEL *hitCaCeC*.

The three passive conjugations, on the other hand, appear to be less symmetrical:

- (a) The morphologically simple active (infix only) PAAL *CaCaC* has a morphologically complex passive (prefix + infix) NIFAL *niCCaC*.  
(b) The morphologically simple active (infix only) PIEL *CiCeC*, on the other hand, has a morphologically simple passive (infix only) PUAL *CuCaC*.  
(c) The morphologically complex active (prefix + infix) HIFIL *hiCCiC* has a morphologically complex passive (prefix + infix) HUFAL *huCCaC*.  
(d) The vocalic infix of the PUAL (*u-a*) overlaps with the prefix-infix HUFAL (*u-a*).  
(e) The HITPAEL reflexive conjugation seems to be without a passive.

Thus, according to the traditional view, the active *binyanim* are fairly symmetrically divided morphologically into two simple or infix-ed *binyanim* versus two morphologically complex or prefixed-infixed *binyanim*. The passive conjugations, however, are asymmetrical: both in their morphology (simple vs. complex) as well as in their distribution: the ‘missing reflexive passive’. The case of the missing reflexive passive, of course, leads to the first basic question raised in this chapter: what is the best way to identify and analyze the *binyanim* which includes how many *binyanim* are there? Most traditional and neotraditional grammarians of both the syntactic as well as the lexicalist approaches mention the existence of at least one ‘passive-reflexive’ HITPAEL, as well as other possible *binyanim*

such as a HOTPAAL, PAUL, MEFOAL, MUFAL, JIFEL/JUFAL, HITTAFEL, HITPAUL. All of the possible *binyanim* are usually considered to be peripheral to the major system, or are not directly dealt with because of their relative rarity or infrequency. Many of these forms appear in Tab. 125.1. The traditional view of the asymmetry of the *binyan* system is not limited to the morphology or morphophonemic signal side of the linguistic sign alone, but may also be found in the meaning or content side of the linguistic sign as well. This is only true, however, when one does not make a clear distinction between the invariant meaning of a linguistic sign and the variant uses, functions, and messages (the dictionary, or contextual discourse meanings) which may be inferred from these invariant meanings. The recognition of the difference between meaning (invariant meaning) versus message (contextual meaning) is the basis of sign-oriented linguistics and has been discussed at great length (cf. Tobin 1990: chs. 2–4; 1993; 1994: chs. 1, 2, 8).

#### 4.2. The seven *binyanim*: functions

The traditional and neotraditional analyses have uncovered a rather large overlap between the various syntactic, semantic and pragmatic functions and features within and across the *binyanim*. Tab. 125.4 lists the functions originally presented in Gesenius (1910), Berman (1975 a; 1978 a; 1978 b), Glinert (1989), Ariel (1971), and Junger (1988) which are fairly complete and exhaustive.

#### 4.3. The seven *binyanim*: summary

A review of the literature and even a cursory glance at the main and secondary functions and features found in Tab. 125.4 reveal that there are certain points about the *binyanim* upon which most scholars seem to agree:

	<i>binyan</i>	main functions/features	secondary functions/features
(1)	PAAL	basic/simple (+/-) transitive	middle (Ariel 1971)
(2)	NIFAL	'passive' of PAAL	middle of HIFIL basic/simple (intransitive) reciprocal (Berman 1975 a; Glinert 1989) ingressive (Berman 1978 a, b; Glinert 1989) expresses emotions ('passive'/mental') reciprocal or mutual actions active passive of PIEL & HIFIL (Gesenius 1910) inchoative-resultative of HIFIL pseudo-passive (Junger 1988)
(3)	PIEL	basic (+ transitive)	transitive or specialized to PAAL (Berman 1975 a) intensification strengthening repetition/iterative causative (Gesenius 1910) no distinctive meaning (Glinert 1989) (Berman 1975 a; Glinert 1989) passive and participle of PIEL (Gesenius 1910) pseudo-passive (Junger 1988)
(4)	PUAL	'PASSIVE' of PIEL	inchoative transitive of verbaliser basic in NIFAL (Berman 1975 a) reflexive (Berman 1978 b; Glinert 1989) transitive of intransitive PAAL
(5)	HIFIL	'causative' of PAAL	inchoative incipience of a certain condition and its continuation action in some particular mental direction denominatives expressing the drawing out, the production of a thing (Gesenius 1910) (Berman 1975 a; Glinert 1989) sometimes passive of PAAL (Gesenius 1910) pseudo-passive (Junger 1988)
(6)	HUFAL	'passive' of HIFIL	reflexive reciprocal inchoative iterative basic (Berman 1975 a; Glinert 1989) middle (Berman 1978 a, b) equivalent of PAAL
(7)	HITPAEL	(-tr) middle of PIEL	reciprocal middle inchoative (Gesenius 1910) pretending (Kimhi 1847)
		reflexive of PIEL	

Tab. 125.4: The *binyanim*: main and secondary functions and features

(a) The *binyanim* express certain syntactic and semantic features which appear to be (at least partially) systematically related to each other.

(b) The first *binyan* PAAL or QAL is the basic *binyan* from which the other *binyanim* are derived. Other scholars dealing with the *binyan* system of Arabic (cf. McCarthy

- 1981) also use the counterpart of the QAL *binyan* as the base form from which all the other conjugations are derived.
- (c) Phonologically, the PAAL or QAL is not only an infix-only *binyan* but the vocalic infix (*a–a*) itself is composed of the same ‘neutral’ or ‘unmarked’ vowel while the other infix-only *binyanim* PIEL (*i–e*) and PUAL (*u–a*) have different and more ‘marked’ vowels in the Jakobsonian sense (cf. Jakobson 1971).
  - (d) The *binyanim* usually are labeled according to syntactic and semantic functions derived from the PAAL/QAL: ‘intensive’, ‘causative’, and ‘reflexive’ – the traditional labels of PIEL, HIFIL and HITPAEL – all refer to the ‘simple’ action introduced in the PAAL/QAL while the NIFAL is considered the ‘passive’ of the PAAL/QAL.
  - (e) The PAAL OR QAL is also labeled ‘+/- transitive’ while the other *binyanim* are often defined along ‘transitivity lines’ with respect to the PAAL/QAL (cf. Berman 1978 a; 1978 b; Glinert 1989): [+/- transitive] PAAL/QAL; [+ transitive] PIEL, HIFIL; [- transitive] NIFAL, PUAL, HUFAL, HITPAEL. Unfortunately, these ‘lines of transitivity’ do not always work and there are [- transitive] verbs in the PIEL, [+ transitive] verbs in the NIFAL, HITPAEL, etc.
  - (f) The concepts of active vs. passive as a means of classifying the *binyanim* have also been generally recognized and accepted: at least three out of the four active *binyanim* (PAAL/QAL, PIEL and HIFIL) are considered to have passive counterparts (NIFAL, PUAL, HUFAL) although their functions do not always fall under the rubric of passivity. However not every root appearing in an active *binyan* will automatically have a passive counterpart. Sometimes a root will have an active message even though it only appears in a passive *binyan* (*niKNaS* ‘enter’/ NIFAL) or the same root will appear in both the active and passive paired *binyanim* with an ‘active’ message: *Laħam* (PAAL/QAL) / *niLħam* (NIFAL), ‘fought’.
  - (g) Only a very small percentage of roots are productive or appear in all seven *binyanim*: Gesenius (1910: 101) states that very few of the roots occur in all seven *binyanim*. Schwarzwald (1975; 1981 b) arrives at the figure of 2,3 percent based on her counterchecking of Barkali’s (1980) 11 percent. Junger (1988) found 165 roots

out of a total of 2452 which occur in all seven *binyanim*.

Despite the fact that there may be general agreement on principles (a)–(g), the plethora of exceptions, irregularities, contradictions and counterexamples to the syntactic, semantic and pragmatic features and rules postulated for the *binyanim*, the numerous overlaps in their functions, and the asymmetric distribution of the roots within the *binyanim*, still make the *binyan* system a rather difficult ‘nut to crack’: i. e. analyze, describe, understand and explain in a systematic way.

## 5. The analysis

### 5.1. Theory and methodology

The sign-oriented theoretical and methodological principles underlying the analysis of the *binyan* system include:

- (a) The Hebrew root – the integral connection between a set of acoustic signals, a sound pattern (**signifiant**) and a concept or invariant meaning (**signifié**) – represents, for us, a classic example of the Saussurian linguistic sign.
- (b) Each *binyan* may also be viewed as a linguistic sign: i. e. an integral connection between a sound pattern (infix, prefix + infix) and a concept (invariant meaning) or set of concepts (a set of invariant meanings belonging to interlocked systems).
- (c) Therefore the *binyanim* represent the systematic interaction and relationship between linguistic signs for each of which an invariant meaning or a set of interlocked invariant meanings may be postulated.
- (d) The interaction between a root and a *binyan* serves the communicative function of identifying, conceptualizing and classifying different kinds of actions, states and events.
- (e) The root presents the general semantic field of the action, state, and event while the *binyan* deictically conceptualizes and categorizes the kind of action, state and event.
- (f) This categorization of actions, states, and events is based on the orientation of the encoder at the ‘here-and-now’ point of encoding.

### 5.2. The interlocked systems

The *binyan* system itself represents three different interlocked systems each representing the kind of cognitive conceptualization or perception needed to identify an action, state, or event which will be called its semantic domain:

- (a) Domain #1 represents the ‘objective’ versus ‘subjective’ viewing of actions, states, or events. This means that the encoder may either be giving an objective description of an action, state and event or making a more subjective assessment or comment on an action, state and event. In other words, he may be merely ‘presenting the facts’ or ‘appraising the facts’: i. e. revealing more of his own personal point of view, impressions, or opinions of the scene. The notion of ‘objective’ versus ‘subjective’ might best be viewed as a continuum or cline across the lines of denotation – connotation respectively.
- (b) Domain #2: the ‘single’ versus ‘multiple’ perceptions of these actions, states, or events. This means that the action, state, and event may either be identified at a single perception or that more than one perception is necessary to identify it. Within this domain actions which may be viewed as potentially ‘multiple’ perceptions may also be perceived within integral space and time: i. e. occupying single continuous space and time rather than discontinuous space and time. This feature has been called elsewhere **semantic integrality** (cf. Tobin 1988 a; 1990: chs. 5–8; 1994 ch. 3.)

- (c) Domain #3 represents the ‘autonomy’ of actions, states and events. This means that the action, state and event may be identified independently (‘autonomous’), or is dependent on other elements (either explicitly or implicitly) to be identified (‘non-autonomous’).

### 5.3. The *binyan* system

In this analysis, there is an eight member system (including the ‘missing passive reflexive’ NITPAEL *binyan*) which is divided and subdivided into:

- (a) two sets of four ‘autonomous’ vs. ‘non-autonomous’ forms: (PAAL/PIEL/HIFIL/HITPAEL) vs. (NIFAL/PUAL/HUFAL/NITPAEL);
- (b) four sets of ‘objective’ vs. ‘subjective’ (‘autonomous’/‘non-autonomous’) pairs: (PAAL/HITPAEL)/(NIFAL/NITPAEL) vs. (PIEL/HIFIL) (PUAL/HUFAL);
- (c) two sets of ‘single’/‘integral’ vs. ‘multiple’/‘non-integral’ perceptions of ‘autonomous’ (‘objective’/‘subjective’) pairs: (PAAL/PIEL) vs. (HIFIL/HITPAEL);
- (d) all the ‘non-autonomous’ forms, by definition, involve ‘multiple’ perceptions but the NIFAL and PUAL involve potentially multiple perceptions of original ‘single’/‘integral’ perceptions perceived in an ‘integral’ way while the NITPAEL and HUFAL involve ‘multiple’/‘non-integral’ perceptions of original ‘multiple’/‘non-integral’ perceptions.

This view of the *binyan* system is presented schematically in Fig. 125.1.

The *binyanim* can thus be classified in the following way according to their opposed interlocked invariant meanings:

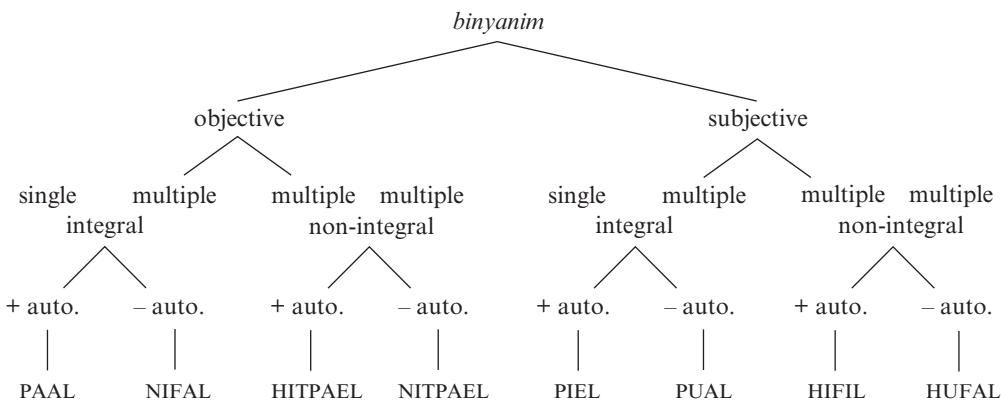


Fig. 125.1: The *binyan* system

	objective	subjective	integral/non-integral		autonomous	non-autonomous
			single	multiple		
PAAL	+		+		+	
NIFAL	+		+			+
HITPAEL	+			+	+	
NITPAEL	+			+		+
PIEL		+	+		+	
PUAL		+	+			+
HIFIL		+		+	+	
HUFAL		+		+		+

Tab. 125.5: The *binyan* system: forms and meanings

- (a) PAAL is marked for the three interlocked meanings: ‘objective’, ‘single’/‘integral perception’, ‘autonomous’;
- (b) HITPAEL is marked for the three interlocked meanings: ‘objective’, ‘multiple’/‘non-integral perception’, ‘autonomous’;
- (c) PIEL is marked for the three interlocked meanings: ‘subjective’, ‘single/integral perception’, ‘autonomous’;
- (d) HIFIL is marked for the three interlocked meanings: ‘subjective’, ‘multiple/non-integral perception’, ‘autonomous’;
- (e) NIFAL is marked for the three interlocked meanings: ‘objective’, ‘non-autonomous’, ‘multiple/integral perception of original single/integral perception’;
- (f) NITPAEL is marked for the three interlocked meanings: ‘objective’, ‘non-autonomous’, ‘multiple/non-integral perception of original multiple/non-integral perception’;
- (g) PUAL is marked for the three interlocked meanings: ‘subjective’, ‘non-autonomous’, ‘multiple/integral perception of original single/integral perception’;
- (h) HUFAL is marked for the three interlocked meanings: ‘subjective’, ‘non-autonomous’, ‘multiple’/‘non-integral’ perception of original ‘multiple’/‘non-integral’ perception.

This is illustrated schematically in Tab. 125.5.

#### 5.4. Reclassifying the *binyanim*

According to this sign-oriented classification, the traditional ‘active’ forms PAAL, PIEL, HIFIL and HITPAEL can now be reclassified according to their opposed interlocked invariant meanings in the following way:

- (a) The PAAL and HITPAEL share the meanings ‘objective’ and ‘autonomous’ but

- differ in the kind of perception – ‘single/integral’ vs. ‘multiple/non-integral’ – necessary to identify an action, state, and event.
- (b) The PIEL and HIFIL share the meanings ‘subjective’ and ‘autonomous’ but differ in the kind of perception – ‘single/integral’ vs. ‘multiple/non-integral’ – necessary to identify an action, state, and event.
- (c) The PAAL and PIEL share the meanings ‘single/integral’ and ‘autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state, and event.
- (d) The HITPAEL and HIFIL share the meanings ‘multiple/non-integral’ and ‘autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state, and event.

According to this sign-oriented classification, the traditional ‘passive’ forms NIFAL, PUAL, HUFAL and NITPAEL can now be reclassified according to their opposed interlocked invariant meanings in the following way:

- (a) The NIFAL and NITPAEL share the meanings ‘objective’, ‘non-autonomous’, ‘multiple’ but differ in the kind of perception – ‘single/integral’ vs. ‘multiple/non-integral’ – of the original identification of an action, state, and event.
- (b) The PUAL and HUFAL share the meanings ‘subjective’ and ‘non-autonomous’, ‘multiple’ but differ in the kind of perception ‘single/integral’ vs. ‘multiple/non-integral’ – of the original identification of an action, state, and event.
- (c) The NIFAL and PUAL share the meanings ‘multiple/integral’ and ‘non-autonomous’ perceptions but differ in their ‘objective’

- vs. ‘subjective’ perceptions of an action, state, and event.
- (d) The NITPAEL and HUFAL share the meanings ‘multiple/non-integral’ ‘non-autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state, and event.

### 5.5. The iconicity of the *binyan* system

From the point of view of iconicity, there is a connection between the signals and the meanings of the active forms: the simpler morphological forms share the simpler ‘single’ perception while the more complex forms share ‘multiple’ perceptions of actions, states, and events:

- (a) The PAAL and PIEL (the morphologically simple infix only conjugations) represent the basic ‘single/integral’ perceptions of ‘autonomous’, ‘objective’ vs. ‘subjective’ conceptualizations of actions, states, and events.
- (b) The HITPAEL and HIFIL (the morphologically complex *hi*-prefix + infix conjugations) represent the conceptually more complex ‘multiple/non-integral’ perceptions ‘autonomous’, ‘objective’ vs. ‘subjective’ conceptualizations of actions, states, and events.

From the point of view of iconicity, there is a connection between the signals and the meanings of the ‘passive’ forms sharing the meaning ‘non-autonomous’ with regard to their ‘objective’ vs. ‘subjective’ conceptualizations of an action, state, and event:

- (a) The ‘objective’, ‘non-autonomous’ *binyanim* NIFAL and NITPAEL share the prefix *ni*;
- (b) The ‘subjective’, ‘non-autonomous’ *binyanim* PUAL and HUFAL share the infixed vowels (*u-a*).

There still remains at least one other important question regarding the identification of the kind and number of linguistic signs found within the system. If we view the two sets of morphologically opposed ‘complex’ ‘autonomous’ and ‘non-autonomous’ forms: HIFIL/HITPAEL and NIFAL/NITPAEL we find an unaccounted for /t/ in the HITPAEL/NITPAEL forms. Therefore one must account for the occurrence of the /t/ in the HITPAEL/NITPAEL dyad. It is clear that:

- (a) the /t/ appears in only one set of *binyanim* with its own *P*-‘-*L* suffix (*pael*) thus

distinguishing it from all the other *binyanim*;

- (b) the /t/ only cooccurs with the *hi* and *ni* prefixes already found in the NIFAL and HIFIL binyanim thus creating a new opposition: *hi-ni-* versus *hit-lnit-* respectively.

Therefore it is possible that the /t/ may be there for phonological reasons: giving us two sets of allomorphs: *hi-ni-* preceding the historical spirantized allomorph [f] (HIFIL/NIFAL) vs. *hit-lnit-* preceding PAEL (with the occlusive allophone [p] (HITPAEL/NITPAEL). It is also possible to assume that the /t/ appears in these specific forms for reasons of ease of articulation. It should also be noted that the concept of ‘ease of articulation’ traditionally has been appealed to account for other sets of infixed consonants found in complex morphological forms in Biblical and Modern Hebrew. These phonological hypotheses are rejected for the following reasons:

- (a) There are no independent phonological reasons to postulate such a /t/ allomorph: neither for reasons of ‘ease of articulation’ nor because of the rules of the phonological system of the language.
- (b) The isomorphic and iconic oriented approach implies that differences in form should also include differences in meaning on all levels of language.

Other previously unaccounted for ‘epenthetic’ consonants found in the middle of Hebrew words have been analyzed as consonantal infixes which were found to be independent signs in their own right (cf. Tobin 1982; 1991 a). Therefore, it is possible that the /t/ might be a linguistic sign or part of a linguistic sign which will systematically allow us to oppose the NIFAL and HIFIL forms and the HITPAEL/NITPAEL dyad. There does seem to be a systematic and iconic connection between those paired forms containing this /t/ as opposed to those forms that do not:

- (a) The prefix *ni-* shows an iconic connection between those forms NIFAL/NITPAEL that signal the meaning: ‘objective’, ‘non-autonomous’, ‘multiple’ perceptions of actions, states, and events. The difference between NIFAL and NITPAEL, however, is that the former represents this conceptualization for actions, states, and events originally perceived in ‘single/integral’ perception while the latter – with the /t/ – is connected only with actions, states and events that were always per-

- ceived in ‘multiple/non-integral’ perceptions. Thus, a difference in form here reflects a difference in meaning as well. Iconically speaking, all the ‘objective’ ‘non-autonomous’ forms share the signal *ni*, but the /t/ is present for those ‘objective’ ‘non-autonomous’ forms which have always shared ‘multiple/non-integral’ perceptions.
- (b) Both the *hi-* forms HIFIL/HITPAEL share the meanings ‘autonomous’ and ‘multiple’ perceptions but differ in their ‘subjective’ vs. ‘objective’ conceptualizations of actions, states and events. Therefore it is the presence of the /t/ which serves to distinguish between the ‘objective’ vs. the ‘subjective’ conceptualizations of these ‘autonomous’ ‘multiple/non-integral’ actions, states and events. Once again, a difference in form here directly reflects a difference in meaning as well. In this case, it serves to distinguish between ‘objective’ vs. ‘subjective’ multiple/non-integral perceptions of actions, states, and events. Iconically speaking, we find that all the ‘objective’ ‘multiple/non-integral’ forms share the same signal /t/.
- (c) The /t/ may be viewed as a consonantal infix which can be analyzed as a linguistic sign or part of a linguistic sign: a signal which always may be linked to the two marked distinctive semantic features ‘objective’ ‘multiple/non-integral’ perceptions of actions, states and events. This particular sign, however, is unmarked or neutral with regard to the distinction between ‘autonomous’ and ‘non-autonomous’ conceptualizations of actions, states, and events making up the system.

Therefore, from the point of view of invariance, markedness, distinctive feature theory,

and iconicity, there is justification for the postulation of:

- each *binyan* as a linguistic sign;
- an eight member system (including the HITPAEL);
- a system based on the ‘objective’ vs. ‘subjective’ conceptualizations of actions, states and events;
- a system based on ‘single/integral’ vs. ‘multiple/non-integral’ perceptions of actions, states and events;
- a system based on the ‘autonomous’ vs. ‘non-autonomous’ perceptions of actions, states, and events.

## 6. The data

Let us return to the triconsonantal (CCC) root *K-T-V* ‘write’ and reanalyze it within the *binyan* system according to our analysis in Tab. 125.6. In table 125.6 the translations of the Hebrew forms are in the past tense: ‘he wrote’, ‘inscribed’, ‘dictated’, ‘corresponded’, ‘it was written’, ‘inscribed’, ‘written in a dictation’, ‘written in a correspondence’. Modern Hebrew does not use inflectional verbal forms for aspect (as did Biblical Hebrew) but fulfills this communicative need with adverbial forms or particles (cf. Tobin 1989 b). Therefore, the translations in Tab. 125.6 could also read: ‘he was writing’, ‘has/had written’, ‘did write’; ‘he was inscribing’, etc.:

The theoretical and methodological implications of the analysis for the active *binyanim* were:

- The PAAL and HITPAEL share the meanings ‘objective’ and ‘autonomous’ but differ in the kind of perception – ‘single/integral’ vs. ‘multiple/non-integral’ – necessary to identify an action, state, and event.

---

PAAL	<i>KaTaV</i>	‘he wrote’
NIFAL	<i>NiKTaV</i>	‘it was written’
HITPAEL	<i>hitKaTeV</i>	‘he corresponded’
NITPAEL	<i>nitKaTeV</i>	‘it was written in the correspondence’
PIEL	<i>KiTeV</i>	‘he inscribed’
PUAL	<i>KuTaV</i>	‘it was inscribed’
HIFIL	<i>hiKTiV</i>	‘he dictated’
HUFAL	<i>huKTaV</i>	‘it was written in the dictation’

---

Tab. 125.6: The *binyan* system: *K-T-V* ‘write’

- (b) The PIEL and HIFIL share the meanings ‘subjective’ and ‘autonomous’ but differ in the kind of perception – ‘single/integral’ vs. ‘multiple/non-integral’ – necessary to identify an action, state, and event.
- (c) The PAAL and PIEL share the meanings ‘single/integral’ and ‘autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state, and event.
- (d) The HITPAEL and HIFIL share the meanings ‘multiple/non-integral’ and ‘autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state, and event.

With regard to the distinction between the simple PAAL *KaTaV* ‘he wrote’ and the reflexive (in this case reciprocal) HITPAEL *hitKaTeV* ‘he corresponded’:

- (a) both of these forms describe basic, simple and independent actions: ('objective'/'autonomous');
- (b) the PAAL ‘he wrote’ needs no more than a single perception or an integral perception to identify the activity: ('single/integral');
- (c) but the reciprocal *hitKaTeV* ‘he corresponded’ implies an activity involving more than one person writing to each other at different moments in time: i. e. a reciprocal process which, by definition, involves more than a single perception viewed non-integrally in order to identify the activity: ('multiple/non-integral').

With regard to the distinction between the intensive PIEL *KitTeV* ‘he inscribed’ and the causative *hiKTiV* ‘he dictated’:

- (a) both of these forms describe a very specific, special, or particular conceptualization of a kind of ‘writing’ perceived as an independent action: ('subjective/autonomous');
- (b) the PIEL ‘he inscribed’ needs no more than a single perception or an integral perception to identify the activity: ('single/integral');
- (c) but the causative *hiKTiV* ‘he dictated’ (‘caused X to write Y’) implies an activity involving more than one person performing different roles reflected by different activities (‘speaking/reading’ vs. ‘writing’) at consecutive and/or successive moments of time, i. e. a reciprocal process which, by definition, involves more than a single

perception viewed non-integrally in order to identify the activity: ('multiple/non-integral').

The theoretical and methodological implications of the analysis for the passive *binyanim* were:

- (a) The NIFAL and NITPAEL share the meanings ‘objective’, ‘non-autonomous’, ‘multiple’ but differ in the kind of perception – ‘single/integral’ vs. ‘multiple/non-integral’ – of the original identification of an action, state, and event.
- (b) The PUAL and HUFAL share the meanings ‘subjective’ and ‘non-autonomous’, ‘multiple’ but differ in the kind of perception ‘single/integral’ vs. ‘multiple/non-integral’ – of the original identification of an action, state, and event.
- (c) The NIFAL and PUAL share the meanings ‘multiple/integral’ and ‘non-autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state, and event.
- (d) The NITPAEL and HUFAL share the meanings ‘multiple/non-integral’ and ‘non-autonomous’ perceptions but differ in their ‘objective’ vs. ‘subjective’ perceptions of an action, state and event.

With regard to the so-called passive forms:

- (a) The passive of the so-called simple PAAL QAL *KaTaV* ‘he wrote’; NIFAL *niKTAV* ‘it was written’ also implies an ‘objective’ action of what was originally perceived in a ‘single/integral’ perception, but it is now dependent on other elements not explicitly mentioned (‘X was written by someone/something’): i. e. (‘non-autonomous’) and, by definition, now implies a ‘multiple/integral’ perception.
- (b) The passive of the so-called reflexive (in this case reciprocal) HITPAEL *hitKaTeV* ‘he corresponded’: NITPAEL *nitKaTeV* ‘it was written in the correspondence’ also implies an ‘objective’ action perceived as part of the ‘multiple/non-integral’ perception inherent to the reciprocal process of ‘corresponding’, but it is dependent on other elements not explicitly mentioned (‘X was written by one of the correspondents in one of the letters’): (‘non-autonomous’).
- (c) The passive of the so-called intensive PIEL *KitTeV* ‘he inscribed’: PUAL *KuTaV* ‘it was inscribed’ also implies a ‘special kind of writing activity’ (‘subjective’) originally

- perceived in a ‘single/integral’ perception, but it is dependent on other elements not explicitly mentioned (‘X was inscribed by someone using a special kind of instrument on a special kind of surface for a particular purpose,’ etc.): (‘non-autonomous’) and, by definition, now implies a ‘multiple/integral’ perception.
- (d) The passive of the so-called causative HIFIL *hiKTiV* ‘he dictated’: HUFAL *huKTaV* ‘it was written in a dictation’ also implies a special kind of writing activity (‘subjective’) perceived as part of a process involving more than one person performing different roles and activities at different times (‘multiple/non-integral’), but it is dependent on other elements not explicitly mentioned (‘X was dictated by someone to someone else in a specific manner for a specific purpose’ etc.): (‘non-autonomous’).

The *binyanim* can thus be classified in the following way according to their opposed interlocked invariant meanings with regard to the triconsonantal (CCC) root *K-T-V* ‘writing’:

- (a) The PAAL *KaTaV* ‘he wrote’ is marked for the three interlocked meanings: ‘objective’, ‘single/integral perception’, ‘autonomous’.
- (b) The NIFAL *niKTaV* ‘it was written’ is marked for the three interlocked meanings: ‘objective’, ‘non-autonomous’, ‘multiple/integral perception of original single/integral perception’).
- (c) The HITPAEL *hitKaTeV* ‘he corresponded’ is marked for the three interlocked meanings: ‘objective’, ‘multiple/non-integral perception’, ‘autonomous’.
- (d) The NITPAEL *nitKaTeV* ‘it was written in the correspondence’ is marked for the three interlocked meanings: ‘objective’, ‘non-autonomous’, ‘multiple/non-integral perception of original multiple/non-integral perception’.
- (e) The PIEL *KiTéV* ‘he inscribed’ is marked for the three interlocked meanings: ‘subjective’, ‘single/integral perception’, ‘autonomous’.
- (f) The PUAL *KuTaV* ‘it was inscribed’ is marked for the three interlocked meanings: ‘subjective’, ‘non-autonomous’, ‘multiple/integral perception of original single/integral perception’.
- (g) The HIFIL *hiKTiV* ‘he dictated’ is marked for the three interlocked meanings: ‘sub-

jective’, ‘multiple/non-integral perception’, ‘autonomous’.

- (h) The HUFAL *huKTaV* ‘it was written in the dictation’ is marked for the three interlocked meanings: ‘subjective’, ‘non-autonomous’, ‘multiple/non-integral perception of original multiple/non-integral perception’.

It must be remembered that these interlocked invariant meanings based on marked distinctive features are not ‘absolute’, but rather ‘relative’ in nature, particularly as they function in language which is an open-ended system. Therefore, the choice of a specific feature is determined by how all these features function together as part of the larger eight member paradigmatic system. This means that whenever a root appears in all eight *binyanim*, all the opposed markedness values are being defined relatively to each other as they are being fully exploited within the system. The markedness values of these semantic features may appear to be ‘neutralized’ for those roots which are only partially distributed within the system. However, any root that appears in more than one *binyan* will maintain the opposed markedness values for the marked distinctive semantic features of the interlocked invariant meanings for those *binyanim* in which it appears.

## 7. Summary and conclusions

This chapter has presented a sign oriented view of the primary *binyan* system of Modern Hebrew based on the concept of invariant meaning serving as the basis for the postulation of paradigmatic systems. The following is a general view of the systems of Hebrew as an interlock of the root and various morphological and lexical systems:

- (a) The root system composed of consonantal signals presents a general semantic field which can be potentially conceptualized as an action, state, and event or as an entity by the encoder at the ‘here-and-now’ point of encoding.
- (b) The verbal *binyan* system composed of prefixes and infixes categorizes the kind of action, state, and event being conceptualized by the encoder as ‘objective/subjective’ and ‘autonomous/non-autonomous’ perceived in a ‘single/multiple’, ‘integral/non-integral’ perception.
- (c) The *binyan* system in Hebrew further interlocks with the systems of gender (M/F),

- number (SG/PL) and person (1, 2, 3) suffixes and tense which categorizes how actions, states, and events are encoded on a spatio-temporal-existential cline as being ‘proximate/remote’ and ‘experienced/non-experienced’ by prefixes and suffixes (cf. Tobin 1989 a; 1991 d; Art. 110).
- (d) The nominal *mīkāl* system composed of an alternate set of related affixes categorizes entities which may be derived from the roots as they appear (and often overlap) in the verbal *binyan* system (cf. Tab. 125.1).
  - (e) The nominal and adjectival systems further interlock with the systems of gender and number (including the dual) (cf. Tobin 1988 a; 1990: ch. 5; Tobin 1994: ch. 2).
  - (f) There are additional lexical and grammatical systems related to what is traditionally referred to as aspect and *Aktionsart* whose semantic domains involve the placement of actions, states, and events within various spatio-temporal-existential boundaries and/or as being parts of processes or results (cf. Art. 109). These systems usually involve lexical and morphological signals traditionally categorized as temporal adverbs, particles, and infinitives (cf. Tobin 1988 b; 1989 b; 1991 b).

## 8. Illustrative text

An example of how the Hebrew language works can be seen in the following Hebrew text (Ben-Amotz 1968: 11) and its English translation (Ben-Amotz 1979: 3):

... *le-?aħar Parba'-ah*  
 ... to-after four-M  
*yem-ei* *haflaq-ah*  
 days-M.PL.CONSTR.STATE sailing-F.SG  
*be-yam-Ø* *kaved-Ø, higa'-ti*  
 in-sea-M.SG heavy-M.SG arrive-1.SG.PAST  
*le-namal-Ø* *Genoa. mi-fa'-ah*  
 to-port-M.SG Genoa. from-hour-F.SG  
*fieva-Ø* *ba-bocker-Ø*  
 seven-F.SG in.the-morning-M.SG  
*'amade-ti* *darux-Ø ba-tartom-Ø*  
 stand-1.SG.PAST tense-M.SG in.the-prow-M.SG  
*ve-le-kol-Ø* *tseiħa-tam sel-Ø*  
 and-to-voice-M.SG screech-F.PL of-sea  
*seħaf-im hijate-ti* *?et 'eyn-ay*  
 gulls-M.PL extend-1.SG.PAST OBJ eyes-F.PL

*lifn-ei* *haʔoniy-ah,*  
 to.the.face-M.PL.CONSTR of.the.ship-F.SG  
*be-ħipus-Ø* *aħar 'Cape' ha-tikv-ah*  
 in-search-F.SG after Cape the-hope-F.SG  
*ha-tov-ah* *sel-i.*  
 the-good-F.SG of.mine

‘After four days at sea we arrived in Genoa. I had been standing in the prow since seven that morning amid a horde of screeching gulls, my eyes tensely scanning the horizon, seeking my own Cape of Good Hope.’

This example demonstrates the synthetic nature of Hebrew illustrated by (among others): the inflectional suffixes of nominal forms (including a distinct set of suffixes for construct state of nominal compounds) and their postposed adjectives which agree in gender and number: other suffixes inflected for gender and number which are attached to various parts of speech and serve various functions (including genitive or possessive markers) which agree with their nominal subjects for person, gender, and number: the fusion of prepositions and the definite article to distinguish between definite versus indefinite nominal forms in prepositional phrases, as well as verb affixation indicating person, gender, number, and tense. This text also further illustrates how the root and *binyan* systems – the primary systems of Hebrew – interact in the language and how the nominal *mīkāl* system may be derived from these primary root + *binyan* patterns.

## 9. References

- Aphek, Edna & Tobin, Yishai (1988), *Word Systems in Modern Hebrew: Implications and Applications*. Leiden, New York: E. J. Brill (The Sociology of Jewish Languages 3)
- Aphek, Edna & Tobin, Yishai (1989), *The Semiotics of Fortune-Telling*. Amsterdam, Philadelphia: John Benjamins (Foundations of Semiotics 22)
- Ben-Amotz, Dahn (1968), *lizkor ve-lifkoħ*. Tel Aviv: Amikam
- Ben-Amotz, Dahn (1979), *To Remember, To Forget*. Tel Aviv: Metziuth
- Ariel, Shlomo (1971), The System of Stem Forms in Colloquial Israeli Hebrew”. *Transactions of the Philological Society*, 192–272
- Bar-Adon, Aaron (1975), *The Rise and Decline of a Dialect*. The Hague: Mouton
- Barkali, Sha'ul (1980), *luħħ ha-peqalim ha-falem*. Jerusalem: Rubin Mas

- Ben Asher, M. (1972), "The Verb Patterns: A Matter for Grammar or the Lexicon". *Hauniversita* 17, 31–34
- Berman, Ruth A[ronson] (1975 a), "The Representation of the Verb-Roots and Patterns in the Lexicon". In: Ornan, Uzi & Fischler, Ben-Zion (eds.), *Rosen Memorial Volume*. Jerusalem: The Council for the Teaching of Hebrew, 25–36
- Berman, Ruth A. (1975 b), "The Morphological Realization of Syntactic Processes in the *binyanim* System". *Hebrew Computational Linguistics* 9, 25–39
- Berman, Ruth A. (1978 a), *Modern Hebrew Structure*. Tel-Aviv: University Publishing Projects
- Berman, Ruth A. (1978 b), "kifrei yahasa ve-hats-raxa ba-ivrit ha-hadafa". *Orxot* 10, 25–33
- Berman, Ruth A. (1982), "Verb Pattern Alteration: On the Interface of Morphology, Syntax and Semantics in Hebrew Child Language". *Journal of Child Language* 9, 169–191
- Bolozky, Shmuel (1978), "Word Formation Strategies in the Hebrew Verb System: Denominative Verbs". *Afroasiatic Linguistics* 5, 111–136
- Fellman, Jack (1973), *The Revival of a Classical Tongue: Eliezer Ben-Jehuda and the Modern Hebrew Language*. The Hague: Mouton
- Gesenius, William (1910), *Hebrew Grammar* ed. by E. Kauzch, rev. by A. E. Cowley. Oxford: Clarendon
- Glinert, Lewis (1989), *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press
- Jakobson, Roman (1971), *Studies on Child Language and Aphasia*. The Hague: Mouton
- Junger, Judith (1988), *Predicate Formation in the Verbal System of Modern Hebrew*. Dordrecht: Foris
- Kimhi, Rabbi David (1847), *sefer ha-forafim*. Berlin
- McCarthy, John J. (1981), "A Prosodic Theory of Non-Concatenative Morphology". *Linguistic Inquiry* 12, 373–417
- Ornan, Ozi (1971 a), "Verb patterns and Stems, Inflections and Declensions". *Hauniversita* 2, 15–22
- Ornan, Ozi (1971 b), "binyanim u-bsisim. netiyot u-gzarot". *Hauniversita* 16, 15–22
- Ornan, Ozi (1971 c), "New Approach to the Description of the Verb and Its Teaching, an Example of the Declension of the *benoni*". In: Kodesh Shmuel (ed.), *Kamrat Memorial Volume*. Jerusalem: The Council for the Teaching of Hebrew, 32–44
- Ornan, Ozi (1978), "netiyot ha-po'al keitsad?". *Orxot* 10, 47–68
- Ornan, Ozi (1979), "On the Meaning of the *binyanim*". In: Rabin, Chaim & Fischler, Ben-Zion (eds.), *minha le-kodef*. Jerusalem: The Council for the Teaching of Hebrew, 11–19
- Rosén, Haiim (1977), *Contemporary Hebrew*. The Hague: Mouton
- Schwarzwald, Ora (1974), "Roots, Stems and the Structure of Morphemes". *Lefonenu* 38, 131–136
- Schwarzwald, Ora (1975), "Comments on Root Pattern Relationships in the Hebrew Lexicon". *Hebrew Computational Linguistics* 9, 47–59
- Schwarzwald, Ora (1981 a), "Grammaticality in Modern Hebrew". *Journal of Middle East Studies* 13, 11–19
- Schwarzwald, Ora (1981 b), "Frequency Factors as Determinants in the *binyanim* Meaning". *Hebrew Studies* 22, 131–137
- Schwarzwald, Ora (1981 c), *Grammar and Reality in the Hebrew Verb*. Ramat-Gan: Bar-Ilan University Press
- Tobin, Yishai (1981), "The Biblical Connection: Sociolinguistic Problem of Modern Hebrew to English Translation". *Jewish Education* 49, 40–44
- Tobin, Yishai (1982) "Asserting One's Existence in Modern Hebrew: A Saussurian Based Analysis of the Domain of Attention in Selected 'Existentials'". *Lingua* 58, 341–368
- Tobin, Yishai (1988, ed.), *The Prague School and Its Legacy*. Amsterdam, Philadelphia: John Benjamins (Linguistic and Literary Studies in Eastern Europe 27)
- Tobin, Yishai (1988 a), "The Dual Number in Modern Hebrew: A Sign-Oriented Approach". *Journal of Literary Semantics* 12, 176–201
- Tobin, Yishai (1988 b), "Aspectual Markers in Modern Hebrew: A Sign-Oriented Approach". In: Morag, Shlomo (ed.), *Studies on Contemporary Hebrew, Vol. II*. Jerusalem: Academon, 584–592
- Tobin, Yishai (1989, ed.), *From Sign to Text: A Semiotic View of Communication*. Amsterdam, Philadelphia: John Benjamins (Foundations of Semiotics 20)
- Tobin, Yishai (1989 a), "Space, Time and Point of View in the Modern Hebrew Verb". In: Tobin (1989 a, ed.), 61–92
- Tobin, Yishai (1989 b) "The *ayin-dalet* System in Modern Hebrew: A Sign-Oriented Approach". In: Weydt, Harald (ed.), *Sprechen mit Partikeln*. Berlin, New York: Walter de Gruyter, 391–402
- Tobin, Yishai (1990), *Semiotics and Linguistics*. London, New York: Longman
- Tobin, Yishai (1991 a), "Existential Particles and Paradigms in Modern Hebrew". *Multilingua* 10.1–2, 93–108

Tobin, Yishai (1991 b), "Process and Result and the Hebrew Infinitive: A Study of Linguistic Isomorphism". In: Jongeling, Karel & van den Berg, H. L. & van Rompay, Lucas (eds.), *Studies on Hebrew and Aramaic Syntax*. Leiden, New York: E. J. Brill, 194–210

Tobin, Yishai (1991 c), "Review Article: Comparing and Contrasting Grammars". *Studies in Language* 15.2, 423–452

Tobin, Yishai (1991 d), "The Future Tense in Modern Hebrew: From Sign to Text". *Folia Linguistica* 24.3–4, 457–512

Tobin, Yishai (1993), *Aspect in the English Verb: Process and Result in Language*. London, New York: Longman

Tobin, Yishai (1994), *Invariance, Markedness and Distinctive Feature Analysis: A Contrastive Study at Sign Systems in English and Hebrew*. Amsterdam, Philadelphia: John Benjamins

Tobin, Yishai (1997), *Phonology as Human Behavior: Theoretical Implications and Clinical Applications*. Durham/NC, London: Duke University Press

*Yishai Tobin, Be'er Sheva (Israel)*

## 126. Türkisch (Turk)

1. Allgemeine Informationen
2. Agglutination und Allomorphie
3. Nominalflexion
4. Verbalmorphologie
5. Wortbildung
6. Illustrativer Text
7. Unübliche Abkürzungen
8. Zitierte Literatur

### 1. Allgemeine Informationen

Das Türkische ist Muttersprache von rund 60 Millionen Menschen, von denen der allergrößte Teil in der Republik Türkei lebt, doch zählen zur türkischen Sprachgemeinschaft auch bedeutende Minderheiten in anderen Ländern. So leben allein in der Bundesrepublik Deutschland über zwei Millionen Menschen türkischer Muttersprache. Das Türkische gehört zusammen mit Sprachen wie dem Aserbaidschanischen und dem Turkmenischen zu den südwestlichen Turksprachen, die wiederum in ihrer Gesamtheit zu den altaischen Sprachen zählen. Die türkische Sprachgemeinschaft ist unter allen Turksprachen die westlichste und zahlenmäßig größte.

Das moderne Türkisch unterscheidet sich vor allem im Vokabular, teilweise aber auch in Morphologie und Syntax beträchtlich von seinem Vorgänger Osmanisch, das von persischen und arabischen Elementen dominiert war. Es wird seit 1928 in einem lateinisch-basierten Alphabet geschrieben. Seit den 1930er Jahren wurde die Sprache im Rahmen der Verwestlichungspolitik massiv reformiert.

### 2. Agglutination und Allomorphie

Die Flexion im Türkischen erfolgt durch Suffixe. Da diese mit ganz wenigen Ausnahmen jeweils nur eine grammatische Kategorie symbolisieren, so daß zum Ausdruck einer Kombination mehrerer grammatischer Kategorien an einem Stamm mehrere Suffixe aneinandergefügt werden müssen, werden Sprachen dieses Typs als **agglutinierend** bezeichnet. Das Türkische gilt (und durchaus nicht zu unrecht) als einer der typischsten Vertreter des agglutinierenden Sprachbaus. Doch lassen sich regional in der Umgangssprache Ansätze von Suffixfusion hin zu Kumulativmorphemen beobachten. So sind im Standard Aspekt- und Personalsuffixe (siehe 4.4) klar segmentierbar, z. B. *gel-iyor-sun* 'komm-PROG-2.SG (du kommst)' oder *gel-ecek-sin* 'komm-FUT-2.SG (du wirst kommen)'. In der Umgangssprache wird hieraus regional gelegentlich *gel-iyon* und *gel-icen*. Eine Konsequenz der Fusion ist die, daß die Interrogativpartikel *mi* nicht mehr, wie im Standard, zwischen Aspekt- und Personalsuffix treten kann, und so wird *gel-iyor mu-sun?* 'komm-PROG INT-2.SG (kommst du?)', zu *geliyon mu?* (analog hierzu *gelecek misin? > gelicen mi?* 'wirst du kommen?') restrukturiert. Ansonsten ist die Flexion des Türkischen fast ausnahmslos regelmäßig.

Die Agglutination führt zusammen mit einem beachtlichen Synthesegrad dazu, daß in extremen Fällen relativ lange Wortformen entstehen können, die aus einer ein- oder zweisilbigen Wurzel und einer Abfolge von

Derivationssuffixen, stammbildenden Suffixen und Flexionssuffixen bestehen wie in (1):

- (1) *yapıştırılamadıklarındır*  
*yapış-tr-ı-ama-dık-lar-m-dan-dir*  
 mach:REZ(kleb)-KAUS-PASS-NPOT-  
 OBLPART-PL-POSS.3-ABL-PRÄD  
 '(das) gehört zu denen, die nicht angeklebt werden konnten.'

Andere Bildungstypen kommen nur in der Wortbildung und auch dort nur relativ vereinzelt vor, nämlich Komposition, markerlose Konversion und Reduplikation.

Phonologisch konditionierte Allomorphie im Türkischen ist hauptsächlich unter dem Begriff **Vokalharmonie** bekannt. Diese kann, grob gesagt, als eine Art progressiver Assimilation auf der Ebene der Vokale beschrieben werden (also über die dazwischenliegenden Konsonanten hinweg), und sie betrifft hauptsächlich die Vokale von Suffixen. Man kann das so interpretieren, daß einige der Merkmale der Vokale dieser Suffixe in ihrer Lexikoneintragung zunächst unspezifiziert bleiben, also z. B. für [hinten] oder [rund], und daß diese dann vom entsprechenden Merkmal des letzten Stammsuffixes bzw. des in der Suffixkette vorangehenden Suffixes spezifiziert werden (vgl. z. B. Bassarak 1996). Dabei sind die Suffixe, die überhaupt über mehrere Formen verfügen, mindestens hinsichtlich des Merkmals [hinten] differenziert, also *a* vs. *e* oder *i* und *u* vs. *i* und *ü*, während bei Suffixen, die hinsichtlich mehrerer Merkmale differenziert sind, zusätzlich noch zwischen nicht-runden (*i* und *i*) und runden Vokalen (*u* und *ü*) unterschieden werden muß. Wenn im folgenden ein Suffix mit dem Vokal *a* oder dem Vokal *i* angegeben wird, ist dies also so zu verstehen, daß dieses Suffix zwei bzw. vier sich lediglich im Vokal voneinander unterscheidende Formen hat.

Nicht nur der Vokal eines Suffixes kann an die vorhergehende Silbe angepaßt werden, sondern auch der Anlautkonsonant eines Suffixes. Dies betrifft im Türkischen vor allem die Konsonantenpaare *d/t* (2 a) und *c/ç* (2 b), seltener *g/k*.

- (2) (a) *gel-di, yap-tı*  
 komm-PRÄT(3.SG) mach-PRÄT(3.SG)  
 'er/sie kam, er/sie machte'  
 (b) *Alman-ca, Türk-ce*  
 deutsch-ADVR türkisch-ADVR  
 '(auf) deutsch, (auf) türkisch'

Wie das Beispiel *-di/-di/-du/-dü/-ti/-til/-tu/-tü* zeigt, können daher Suffixe, die sowohl der

Vokalharmonie als auch der konsonantischen Angleichung unterliegen, damit je nach ihrem phonetischen Vorgängerkontext bis zu acht verschiedene Formen annehmen.

Neben den umgebungsbedingten Veränderungen der phonologischen Form der türkischen Suffixe (Vokalharmonie und konsonantische Angleichung) gibt es noch weitere Veränderungen, die ebenfalls mit der phonologischen Form des Stammauslauts bzw. des Auslauts des Vorgängersuffixes zusammenhängen, und zwar damit, ob dieser dem Suffix vorangehende Auslaut vokalisch oder konsonantisch ist. Es ist ja klar, daß nicht jedes Suffix nach jedem beliebigen vorangehenden Auslaut gleich gut artikuliert werden kann. Hinzu kommt, daß die Silbenstrukturregeln des Türkischen Konsonantencluster nur in beschränktem Maße und Vokalkombinationen im Prinzip gar nicht erlauben. Deshalb haben einige türkische Suffixe zwei Anlautformen, deren Verwendung davon abhängt, ob sie auf einen Vokal oder auf einen Konsonanten folgen, wie z. B. *-(y)i* (3 a), oder *-(s)i* (3 b).

- (3) (a) *ev-i, kaptı*  
 Haus-AKK Tür-AKK  
 (b) *ev-i, kaptı-sı*  
 Haus-POSS.3 Tür-POSS.3

Um beim Anhängen des Progressivsuffixes *-yor* an einen auf Vokal auslautenden Stamm eine Sequenz aus zwei Vokalen zu vermeiden, wird der zum Stamm gehörige Vokal getilgt. Diese Regel erklärt, warum die Progressivformen zweier Verbstämme wie *yık-* 'zerstören' und *yıka-* 'waschen' identisch sind. In beiden Fällen erhält man *yık-yor*.

### 3. Nominalflexion

#### 3.1. Numerus, Definitheit und Kasussystem

Im Türkischen werden nicht (wie etwa in zahlreichen indo-europäischen Sprachen) die einzelnen Bestandteile eines Nominalsyntaxmas, sondern das Syntaxma als Ganzes flektiert. Es gibt somit keine Kongruenz zwischen Adjektiven, Pronomen und ihrem Bezugswort, Flexionssuffixe erscheinen lediglich am Ende des Syntaxmas.

Die Numerus-Opposition wird mit Hilfe des Pluralsuffixes *-lar/-ler* ausgedrückt, wobei das Fehlen dieses Suffixes in bestimmten Kontexten einer Neutralisierung besagter Opposition, also einer transnumeralen Verwendung entspricht (4).

- (4) *Kitap oku-dum.*  
 Buch les-PRÄT(3.SG)  
 'Ich habe (in) ein(em) Buch/(in) einige(n) Bücher(n) gelesen.'

Der Ausdruck der Numerus-Opposition ist also bei indefiniten direkten Objektkomplementen fakultativ, was mit der universellen Tendenz zur Koaleszenz von transitiven Verben mit schwach individuierten Objektkomplementen zusammenhängt (vgl. deutsch *Zeitung lesen*, *Fenster putzen* u. ä.). Je nach syntaktischer Rolle stehen verschiedene Mittel zum Ausdruck unterschiedlicher Definitheitsgrade zur Verfügung, z. B. Wortstellung, indefiniter Artikel oder Indefinit- und Demonstrativpronomen. Hier soll nur die Rolle des Akkusativsuffixes im Rahmen der differentiellen Objektmarkierung kurz beschrieben werden. Individuation ist eine Kombination verschiedener Eigenschaften wie Empathie, Referentialität und Definitheit. Daraus ergibt sich, daß der Sprecher stets am stärksten individuiert ist. Er ist somit prädestiniert für die Rolle des Agens, das wiederum typischerweise als Subjekt erscheint. Ist er Referent eines Objektkomplements, ist dies weniger typisch und wird daher in vielen Sprachen morphologisch signalisiert. Neben dem Sprecher können zunächst weitere menschliche oder allgemein belebte, schließlich auch sonstige definite oder lediglich referentielle Entitäten zur Kategorie der höher individuierten Protagonisten gezählt werden, wie das Nominalsyntagma *bir öğrenci-yi* in (5).

- (5) *Sevim bir öğrenci-yi ara-di.*  
 Sevim ein Schüler-AKK such-PRÄT(3.SG)  
 'Sevim suchte einen (bestimmten) Schülern.'

Geringer individuierte Entitäten sind besser als Patiens geeignet und werden daher in der Objektposition nicht besonders symbolisiert, wie in (6).

- (6) *Sevim çarşı-ya gid-ip (bir)*  
 Sevim Markt-ALL geh-ADVR ein  
*karpuz al-di.*  
 Melone kauf-PRÄT(3.SG)  
 'Sevim ging auf den Markt und kaufte (eine) Melone(n).'

Dies erklärt das Fehlen des Akkusativsuffixes in (6), wo auf eine oder mehrere beliebige (zudem anempathische) Melonen verwiesen wird, und sein Vorhandensein in (5), wo das Objektkomplement zwar indefinit, aber referentiell (und zudem empathisch) ist (vgl.

Dede 1986). Die Grenze zwischen markiert und unmarkiert ist je nach Sprache unterschiedlich definiert. Die universelle Funktion differentieller Objektmarkierung ist in jedem Falle die Signalisierung weniger typischer, also gewissermaßen "unerwarteter" Objektkomplemente.

Während also der Nominativ immer und der Akkusativ teilweise suffixlos sind (was in der Zentralität dieser beiden Fälle begründet ist), ist das Verhalten der restlichen vier Kasus weniger auffällig: Genitiv *(-(n)m)*, Lokativ *(-da)*, Ablativ *(-dan)*, sowie ein Allativ, der den Funktionsbereich des Dativs mit abdeckt, auf *(-y)a*. Der Instrumental oder Komitativ mit dem Suffix *(-y)la*, hervorgegangen aus der Postposition *ile* 'mit', weist infolge von Grammatikalisierung einige Eigenschaften von Kasus auf (z. B. Suffixstatus und damit Bildung von Allomorphen gemäß Vokalharmonie), ist aber noch nicht voll grammatikalisiert (vgl. auch Bassarak 1988).

### 3.2. Possessivkonstruktionen und Suffixreihenfolge

Den Possessivpronomina im Deutschen und zahlreichen anderen Sprachen entsprechen im Türkischen **Possessivsuffixe** (s. Art. 103).

	Numerus	Singular	Plural
Person		des Possessors	
1.	<i>-(i)m</i>	<i>-(i)miz</i>	
2.	<i>-(i)n</i>	<i>-(i)niz</i>	
3.	<i>-(s)i(n)</i>	<i>-[lar]i(n)</i>	

Tab. 126.1: Possessivsuffixe des Türkischen

So heißt also *baba-n* 'dein Vater' oder *ev-imiz* 'unser Haus'. Die Reihenfolge der Suffixe ist im Maximalfall Substantivstamm – Numerus – Possessiv – Kasus (+ postprädiktative Suffixe). Die Anordnung der Suffixe unterliegt also einer festen Reihenfolge.

Beim Possessivsuffix der dritten Person kann sowohl der Plural des durch das Substantiv bezeichneten Objekts als auch der Plural des Possessors durch das Pluralsuffix *-lar* gekennzeichnet werden, so daß ohne Kontext nicht immer entscheidbar ist, wessen Plural hier gemeint ist. Und da das Suffix *-lar* in dem Falle, daß sowohl das Objekt als auch der Possessor mehrzahlig sind, nicht verdoppelt werden kann, hat eine Form mit der Suffixkombination *-lari/-leri* unter Umständen sogar drei Lesungen:

	1. Sg.	1. Pl.	2. Sg.	2. Pl.	3.		
					jener	diese(r/s)	folgende(r/s)
<b>Singular</b>							
Nom./Ind.	<i>ben</i>	<i>biz</i>	<i>sen</i>	<i>siz</i>	<i>o</i>	<i>bu</i>	<i>şu</i>
Genitiv	<i>ben-im</i>	<i>biz-im</i>	<i>sen-in</i>	<i>siz-in</i>	<i>on-un</i>	<i>bun-un</i>	<i>şun-un</i>
Dativ	<i>ban-a</i>	<i>biz-e</i>	<i>san-a</i>	<i>siz-e</i>	<i>on-a</i>	<i>bun-a</i>	<i>şun-a</i>
Akkusativ	<i>ben-i</i>	<i>biz-i</i>	<i>sen-i</i>	<i>siz-i</i>	<i>on-u</i>	<i>bun-u</i>	<i>şun-u</i>
Ablativ	<i>ben-den</i>	<i>biz-den</i>	<i>sen-den</i>	<i>siz-den</i>	<i>on-dan</i>	<i>bun-dan</i>	<i>şun-dan</i>
Lokativ	<i>ben-de</i>	<i>biz-de</i>	<i>sen-de</i>	<i>siz-de</i>	<i>on-da</i>	<i>bun-da</i>	<i>şun-da</i>
<b>Plural</b>							
Nom./Ind.	—	<i>biz-ler</i>	—	<i>siz-ler</i>	<i>on-lar</i>	<i>bun-lar</i>	<i>şun-lar</i>
Genitiv		<i>biz-ler-in</i>		<i>siz-ler-in</i>	<i>on-lar-in</i>	<i>bun-lar-in</i>	<i>şun-lar-in</i>
Dativ		<i>biz-ler-e</i>		<i>siz-ler-e</i>	<i>on-lar-a</i>	<i>bun-lar-a</i>	<i>şun-lar-a</i>
Akkusativ		<i>biz-ler-i</i>		<i>siz-ler-i</i>	<i>on-lar-i</i>	<i>bun-lar-i</i>	<i>şun-lar-i</i>
Ablativ		<i>biz-ler-den</i>		<i>siz-ler-den</i>	<i>on-lar-dan</i>	<i>bun-lar-dan</i>	<i>şun-lar-dan</i>
Lokativ		<i>biz-ler-de</i>		<i>siz-ler-de</i>	<i>on-lar-da</i>	<i>bun-lar-da</i>	<i>şun-lar-da</i>

Tab. 126.2: Pronominalflexion

(7) *ev-ler-i*

Haus-PL-POSS.3  
'seine/ihre (Sg.) Häuser'/  
'ihr (Pl.) Haus'/  
'ihre (Pl.) Häuser'

Klarheit kann in solchen Fällen ein lexikalischer Possessor schaffen. Dieser steht im Genitiv vor dem Possessum-Nominal, das mit dem Possessivsuffix der dritten Person versehen wird. Die possessive Relation wird also zweifach ausgedrückt, die redundante Wiederholung des Pluralsuffixes entfällt jedoch, wenn die Pluralität schon an anderer Stelle ausgedrückt wird (8).

(8) *kadın-lar-m ev-i*

Frau-PL-GEN Haus-POSS.3  
'das Haus der Frauen'

Daneben existiert in der Umgangssprache eine Konstruktion aus Personalpronomen im Genitiv, was einem Possessivpronomen gleichkommt, und suffixlosem Possessor-Nominal (9).

(9) *biz-im ev = ev-imiz*

1.PL-GEN Haus Haus-POSS.1.PL  
'unser Haus'

## 3.3. Pronominalflexion

Im Unterschied zu den Substantiven können bei den Demonstrativpronoma (proximal *bu*, mesial *şu*, distal *o*) nur die Kategoriengefüge Numerus und Kasus und bei den Perso-

nalpronomina (*ben*, *sen*, *o*, Plural *biz*, *siz*, *on-lar*) nur der Kasus symbolisiert werden. Die Kasussuffixe sind im wesentlichen die gleichen wie bei den Substantiven. Nur im Genitiv gibt es eine Modifikation, die aber nur die 1. Person betrifft (*-im* statt *-(n)m*). Diese Formen können als schwach suppletiv gelten. Es ist aber kein Zufall, daß sich eine solche assimilatorisch bedingte Irregularität lediglich im Bereich der Pronomina durchsetzen und halten konnte, da diese ohnehin allgemein stark zur Suppletion tendieren. Außerdem gibt es bei den Personalpronomina der 1. und 2. Person Singular im Dativ eine Stammveränderung von *-e* zu *-a* (*ben/sen* vs. *banal/sana*), also einen der für eine agglutinierende Sprache seltenen Fälle von schwacher Suppletion.

*O* ist sowohl das Personalpronomen der 3. Person Singular und (zusammen mit dem Pluralsuffix, also in der Form *onlar*) Plural als auch das Demonstrativpronomen 'jene(r/s)'. Die Formen *bizler* und *sizler*, also Pluralformen von Personalpronomina, die bereits Pluralia sind, entsprechen ungefähr 'wir alle' bzw. 'ihr/Sie alle'.

## 4. Verbalmorphologie

## 4.1. Stammbildende Suffixe

Die Flexion der Verben erfolgt durch Suffixe, die an den Stamm angehängt werden. Dieser kann durch die reine Verbwurzel alleine re-

präsentiert sein oder durch Verbindungen von einer Verbwurzel mit solchen Suffixen, die vor dem Infinitivsuffix *-mak* stehen können. Die Reihenfolge dieser stammbildenden Suffixe liegt im Türkischen fest (s. auch Art. 118).

In der ersten Position unmittelbar nach der Verbwurzel kann das Reflexivsuffix *-(ı)n* oder das Reziproksuffix *-(ı)s* stehen. Hierbei handelt es sich um valenzreduzierende Operationen (s. Art. 107). In der zweiten Position kann ein valenzerweiterndes Kausativsuffix stehen. Das Kausativsuffix kann die Formen *-tir*, *-(ı)rl-ar* und *-(ı)t* haben. Die Verteilung zwischen diesen drei bzw. vier Suffixal-lomorphen ist nur teilweise phonologisch konditioniert und in vielen Fällen lexikalisch determiniert. Die dritte Position ist dem Kategoriengefüge Aktiv/Passiv vorbehalten. Das Aktiv ist suffixlos, das Passiv hat das Suffix *-ll-(ı)n*. Die Verteilung zwischen diesen beiden Suffixvarianten ist in diesem Falle strikt phonologisch geregelt. In der vierten Position nach der Verbwurzel kann das Negationssuffix *-ma* stehen. Zu beachten ist, daß das Negationssuffix *-ma* nicht betonbar ist, so daß in der Regel die Silbe vor diesem Suffix den Wortakzent trägt. Dadurch ist das Betonungsmuster eines negierten Verbs gegenüber dem eines positiven (affirmierten) Verbs stark distinkтив. Die fünfte Position wird von den Suffixen des Kategoriengefüges (Non-)Potential eingenommen. Das Potentialsuffix ist *-(y)abil*, wobei aus dieser Darstellung schon hervorgeht, daß der erste Teil des Suffixes *((y)a(y)e)* der Vokalharmonie unterliegt, der zweite Teil, der mit dem Stamm des Verbs *bilmek* ‘können, wissen’ identisch ist, hingegen nicht. Die negierte Form lautet *-(y)ama* (vokalharmonisch, wechselt also mit *-(y)eme*). Die Reihenfolge der Suffixe illustriert das Kontinuum zwischen Derivation (näher an der lexikalischen Wurzel) und Flexion (näher an den grammatischen Aspekt-/Tempus- und Personalsuffixen).

- (10) *Bun-u yap-abil-ir-sin.*  
dies-AKK mach-POT-DISP-2.SG  
'Du kannst das machen.'
- (11) *Bun-u yap-ama-z-sin.*  
dies-AKK mach-NPOT-DISP-2.SG  
'Du kannst das nicht machen.'

In der sechsten Position, also nach den stammbildenden Suffixen, können das Infinitivsuffix *-mak*, einige weitere umkategorisierende Suffixe sowie die Tempus- und Modus-

suffixe stehen. Der Infinitiv dient in Wörterbüchern als Zitierform des betreffenden Verbs.

## 4.2. Partizipialkonstruktionen

### 4.2.1. Partizipien mit Attributstatus

Von Verben können zahlreiche **Partizipien** abgeleitet werden, die die Funktion der Subordinatoren (Pronomen und Konjunktionen) indo-europäischer Sprachen erfüllen. Für Relativkonstruktionen werden die Suffixe *-(y)an* und *-dik* verwendet. Ist das Subjekt des Verbs der Relativkonstruktion mit dem Subjekt des Verbs der Matrixklause identisch, wird ein Partizip auf *-(y)anl-(y)en* gebildet.

- (12) *Yaklaş-an adam-tı ta*  
sich.näher-SBJPART Mann-AKK schon  
*uzak-tan gör-dü-m.*  
fern-ABL seh-PRÄT-1.SG  
'Den sich nähernden Mann sah ich schon von weitem.'

Hingegen kann ein Partizip auf *-dik* verschiedene oblique syntaktische Rollen repräsentieren, z. B. ein allatives Komplement in (13). Das Subjekt der Relativkonstruktion erscheint dann in Form eines Possessivsuffixes am Partizip, ein lexikalisches Subjekt steht entsprechend im Genitiv.

- (13) *Elif-in git-tığ-i*  
Elif-GEN geh-OBLPART-POSS.3.SG  
*yer-i bil-m-iyor-um.*  
Ort-AKK wiss-NEG-PROG-1.SG  
'Den Ort, zu dem Elif gegangen ist/geht, kenne ich nicht.'

Dieses Partizip kann sowohl Präsens- als auch Vergangenheitsbedeutung haben. Das System der Tempusoppositionen ist also in der Subordination stark reduziert. Soll eine Vergangenheitsbedeutung eindeutig hervorgehoben werden, wird ein zusammengesetztes Partizip verwendet.

- (14) *Elif-in git-mış*  
Elif-GEN geh-PART.PF  
*ol-dug-u yer-i*  
AUX.DETR-OBLPART-POSS.3.SG Ort-AKK  
*bil-m-iyor-um.*  
wiss-NEG-PROG-1.SG  
'Den Ort, zu dem Elif gegangen ist, kenne ich nicht.'

Nachzeitigkeit wird durch *-(y)acak* zum Ausdruck gebracht.

- (15) *Elif-in gid-eceğ-i*  
 Elif-GEN geh-PART.FUT-POSS.3.SG  
*yer-i bil-m-iyor-um.*  
 Ort-AKK wiss-NEG-PROG-1.SG  
 ‘Den Ort, zu dem Elif gehen wird, kenne ich nicht.’

#### 4.2.2. Partizipien mit Komplementstatus

Die Partizipien auf *-dik* und *-(y)acak* können auch als Komplemente eines Verbs erscheinen.

- (16) *Elif-in nere-ye*  
 Elif-GEN wo-ALL  
*git-tığ-in-i*  
 geh-OBLPART-POSS.3.SG-AKK  
*bil-m-iyor-um.*  
 wiss-NEG-PROG-1.SG  
 ‘Ich weiß nicht, wohin Elif geht/gegangen ist.’

In dieser Strukturposition trifft man auch den flektierbaren verkürzten Infinitiv auf *-mal-me* an.

- (17) *Elif-in gel-me-si-ni*  
 Elif-GEN komm-INF-POSS.3-AKK  
*iste-di-m.*  
 woll-PRÄT-1.SG  
 ‘Ich wollte, daß Elif kommt/kommen soll.’

Die Distribution von *-dik/-y)acak* vs. *-ma* hängt von der Semantik des regierenden Verbs ab. Tendenziell werden die zuerst genannten Formen nur gebraucht, wenn sich eine Handlung tatsächlich ereignet hat oder absehbar ist, während der verkürzte Infinitiv auch für eventuell eintretende Handlungen gebraucht werden kann. So wird der verkürzte Infinitiv vor allem bei Verben des Wünschens, Bittens und Befehlens gebraucht (Tekinay 1988: 296), während bei Kognitionsverben die Possessivpartizipien bevorzugt werden. Diese Einteilung erinnert an Modusoppositionen bei der Subordination in indoeuropäischen Sprachen.

#### 4.3. Verbaladverbien

Das Türkische verfügt über zahlreiche Suffixe zum Ausdruck adverbialer Subordination. Diese können zwar wie eine finite Verbform Subjekt, Objekt und sonstige Dependenten haben, weisen aber keine Kongruenz mit dem Subjekt auf.

Die Klasse der Verbaladverbien (auch Gerundien oder Konverbien genannt) ist nicht

im strengen Sinne abgeschlossen. Hier seien die wichtigsten Verbaladverbsuffixe und die Konjunktionen, die ihnen in etwa entsprechen, aufgeführt:

<i>-(y)ip</i>	‘und’
<i>-(y)arak</i>	‘indem’
<i>-dikça</i>	‘immer wenn’, ‘so oft/ solange als’, ‘in dem Maße wie’
<i>-(y)mca</i>	‘sobald’, ‘wenn’
<i>-(y)ali</i>	‘seit’
<i>-madan</i>	‘ohne zu’
<i>-maksızın</i>	‘ohne zu’
<i>-(y)acağına</i>	‘anstatt zu’
<i>-maktansa</i>	‘anstatt zu’
<i>-r + -maz</i>	‘sobald’
<i>-(y)a + -(y)a</i>	‘indem ... ständig’

Tab. 126.3: Verbaladverbsuffixe

#### 4.4. Tempus, Aspekt, Modus

Nach den in 4.1 aufgeführten stammbildenden Suffixen stehen Aspekt- und/oder Tempussuffixe, an die dann die Personalsuffixe treten. Die Aspektsuffixe *-iyor* {PROGRESSIV}, *-miş* {PERFEKTIV}, *-(y)acak* {PROSPEKTIV/FUTUR}, *-(i)r* {DISPOSITIV} können jeweils mit den Tempussuffixen *-Ø* {unmarkiertes PRÄSENS} oder *-di* {PRÄTERITUM} kombiniert werden. Je nach dem Virtualitätsgrad der Aspekt-Tempus-Konstellation ist der Übergang zu modalen Bedeutungen fließend. Zur Illustration enthält (18) einige Kombinationen:

- (18) (a) *yaz-iyor-Ø-um*  
 schreib-PROG-PRÄS-1.SG  
 ‘ich schreibe’  
 (b) *gel-mış-ti-k*  
 komm-PFV-PRÄT-1.PL  
 ‘wir waren gekommen’  
 (c) *gel-ecek-ti-n*  
 komm-FUT-PRÄT-2.SG  
 ‘du beabsichtigst zu/wolltest kommen’  
 (d) *oku-r-du-m*  
 les-DISP-PRÄT-1.SG  
 ‘ich hätte gelesen/las regelmäßig’

Während im Präteritum die Aspektposition leer bleiben kann (*oku-du-m* ‘les-PRÄT-1.SG’ entspricht einem punktuellen Geschehen in der Vergangenheit), ist im (unmarkierten) Präsens die Aspektmarkierung obligatorisch.

Possessiv-suffixe	Personalsuffixe		
	Normalfall	nach <i>-dil-sa</i>	beim Imp./Opt.
<b>Singular</b>			
1. -( <i>i</i> ) <i>m</i>	-( <i>y</i> ) <i>m</i>	- <i>m</i>	- <i>ayim</i>
2. -( <i>i</i> ) <i>n</i>	- <i>sin</i>	- <i>n</i>	- <i>o</i>
3. -( <i>s</i> ) <i>l(n)</i>	- <i>ø</i>	- <i>ø</i>	- <i>sin</i>
<b>Plural</b>			
1. -( <i>i</i> ) <i>miz</i>	-( <i>y</i> ) <i>iz</i>	- <i>k</i>	- <i>alim</i>
2. -( <i>i</i> ) <i>niz</i>	- <i>siniz</i>	- <i>niz</i>	- <i>in(tz)</i>
3. -[ <i>lar</i> ] <i>l(n)</i>	[ <i>lar</i> ]	[ <i>lar</i> ]	- <i>sin[lar]</i>

Tab. 126.4: Possessiv- vs. Personalsuffixe

Während der prospektive Aspekt im Präsens nichts anderes als ein Futur ergibt (allerdings ein objektiv vorhersehendes, denn für Absichtserklärungen wird der Dispositiv bevorzugt), ist die Abgrenzung zwischen Progressiv und Dispositiv von mehreren Faktoren abhängig. Der wichtigste dürfte der Aktualitäts- oder Realisierungsgrad der Situation sein, wobei für reale (wirklich stattfindende) Situationen der Progressiv, für virtuelle (verallgemeinerte, denkbare, beabsichtigte) Situationen der Dispositiv bevorzugt wird. Beim Progressiv wird also etwas über ein Ereignis ausgesagt, beim Dispositiv wird hingegen dem Referenten des Subjekts eine Disposition (Neigung, Eignung) zu der im Verb ausgedrückten Tätigkeit zugeschrieben (vgl. Savaşır 1986). Der perfektive Aspekt schließlich hat sich weiter grammatisiert, hat außer bei der Bildung des Plusquamperfekts (und als attributives Partizip) evidentielle Bedeutung und kann in dieser Funktion an die anderen Aspektsuffixe angehängt werden (vgl. Aksu-Koç 1986). Mit Einschränkungen können nicht-tempusmarkierte Aspektformen attributiv verwendet werden.

Konditionalgefüge werden mit Hilfe des Suffixes *-(y)sa* gebildet, das an die Aspektsuffixe angehängt wird.

- (19) *Sen yurt dis-m-dan*  
 2.SG Heimat Äußeres-POSS.3-ABL  
*gel-iyor-sa-n*                            *borç*  
 komm-PROG-KOND-2.SG Schulden  
*al-ama-z-sin*.  
 nehm-NPOT-DISP-2.SG  
 ‘Wenn du aus dem Ausland kommst,  
 kannst du kein Geld leihen.’

Die Kombinationsmöglichkeiten des Debitivsuffixes *-mali* innerhalb der Verbform sind stark eingeschränkt, so daß komplexere Konstellationen durch peripherastische Konstruktionen ausgedrückt werden (vgl. 20 a vs. 20 b). Derartige Konstruktionen existieren auch für den Potential (20 c) (nicht aber beim Volitional, der immer persönlich ausgedrückt wird).

- (20) (a) *Gel-meli-yim*.  
 komm-OBLG-1.SG  
 ‘Ich muß kommen.’  
 (b) *Gel-me-m*  
 komm-INF-POSS.1.SG  
*gerek-ecek*.  
 nötig.sein-FUT(3.SG)  
 ‘Ich werde kommen müssen.’  
 (c) *Gel-me-m*  
 komm-INF-POSS.1.SG  
*imkânsız-dı*.  
 unmöglich-PRÄT(3.SG)  
 ‘Ich konnte nicht kommen.’

Die Personalsuffixe sind den Possessivsuffixen sehr ähnlich, besetzen allerdings andere Strukturpositionen.

Die Personalsuffixe können zusammen mit einigen der Tempus-Aspekt-Suffixe auch wie eine Kopula an Prädikatsnominale angehängt werden, z. B. *ögrenci-ydi-m* ‘Student-PRÄT-1.SG (ich war Student)’.

Die Vielzahl der Suffixe, die einer Verbwurzel folgen können, ist eine hervorragende Illustration des agglutinierenden Sprachtypus. Der Übersicht halber dürfte es nützlich sein, hier noch einmal alle potentiell möglichen Suffixe mit ihren Positionen aufzulisten (s. Bassarak 1997):

Pos. 0:	Verbwurzel
1:	Reflexiv/Reziprok
2:	Kausativ
3:	Passiv
4:	Negation
5:	Potential/Non-Potential (0–5 = Verbstamm)
6:	Tempus- und Modussuffixe, umkategorisierende Suffixe (Infinitiv, Verbalnomina, Verbaladjektive, Verbaladverbien)
7:	Numerus
8:	Possessivsuffixe
9:	Kasus
10:	Fragepartikel
11:	postprädiktative Aspekt-Enklitika
12:	Personalsuffixe
10 a:	Alternativposition der Fragepartikel
13:	Prädikativitätssuffix <i>-dur</i>
7 a:	Alternativposition des Numerus

Tab. 126.5: Positionen der Suffixe

In Tab. 126.6 werden noch einmal einige mögliche Kombinationen von Suffixen ab Position 6 demonstriert:

## 5. Wortbildung

Wie man es von einer agglutinierenden Sprache erwartet, hält das Türkische eine ganze Reihe von verbalen und nominalen Derivationsmöglichkeiten bereit. (21) illustriert das bei der Bildung von Verben aus Nomina produktivste Suffix *-la*.

- (21) *imza-la-mak*  
Unterschrift-VR-INF ‘unterschreiben’  
*temiz-le-mek*  
sauber-VR-INF ‘reinigen’  
*şırça-la-mak*  
Bürste-VR-INF ‘bürsten’

Es gibt mindestens 17 Suffixe zur Bildung von Verbalnomina, darunter das sehr produktive *-lk*.

- (22) *güzel-lik*  
schön-NR ‘Schönheit’  
*göz-lük*  
Auge-NR ‘Brille’  
*söz-lük*  
Wort-NR ‘Wörterbuch’

Auch Komposition, wie in *büyükanne* ‘Groß-Mutter’, spielt eine gewisse Rolle. Abschließend noch einige Beispiele für Reduplikation. Es gibt im wesentlichen zwei Typen: vollständige und partielle Reduplikation.

- (23) *adım adım* ‘schrittweise, Schritt für Schritt’ (von *adım* ‘Schritt’)  
*elele* ‘Hand in Hand’ (von *el* ‘Hand’)  
*ara sira* ‘dann und wann’ (unechte Reduplikation aus: *ara* ‘Zwischenraum’ + *sira* ‘Zeitpunkt’)

Reduplikationen von Teilen der ersten Silbe von Adjektiven haben eine Art Steigerungseffekt (s. (24)), besonders bei Farben (s. (25)).

- (24) *epeski* ‘uralt’ (von *eski* ‘alt’)  
*besbelli* ‘völlig klar’ (von *belli* ‘klar’)

0–5	6	7	8	9	10	11	12	13	7 a
Stamm	TM/VN	NUM	POSS	KASUS	QU	PP	PS	<i>-dur</i>	NUM
<i>yap</i>	<i>ar</i>						<i>m</i>		
<i>yap</i>	<i>miş</i>					<i>ti</i>	<i>m</i>		
<i>yap</i>	<i>tyor</i>					<i>muş</i>	<i>sun</i>		
<i>yap</i>	<i>ar</i>				<i>mi</i>		<i>simz</i>		
<i>yap</i>	<i>acak</i>	<i>lar</i>			<i>mi</i>				
<i>yap</i>	<i>miş</i>	<i>lar</i>						<i>dir</i>	
<i>yap</i>	<i>miş</i>							<i>tir</i>	<i>lar</i>
<i>yap</i>	<i>ar</i>					<i>ken</i>			
<i>yap</i>	<i>miş</i>				<i>mi</i>	<i>ydi</i>	<i>niz</i>		
<i>yap</i>	<i>miş</i>						<i>sin</i>	<i>dir</i>	
<i>yap</i>	<i>tyor</i>	<i>lar</i>				<i>sa</i>			
<i>yap</i>	<i>mak</i>		<i>ta</i>					<i>dir</i>	<i>lar</i>
<i>yap</i>	<i>mak</i>		<i>ta</i>				<i>yken</i>		
<i>yap</i>	<i>tik</i>	<i>lar</i>	<i>umz</i>	<i>dan</i>	<i>mi</i>	<i>ymış</i>			

Tab. 126.6: Kombinationen von Suffixen ab Position 6

- (25) *bembeyaz* ‘reinweiß’ (von *beyaz* ‘weiß’)  
*kapkara* ‘kohlrabenschwarz’ (von *kara* ‘schwarz’)

Der einzige produktive Typ von Reduplikation ist der mit *m*-Anlaut, der bei allen Substantiven möglich ist, die nicht schon mit *m* beginnen, und der etwa ‘N und/oder so(was)’ bedeutet:

- (26) *kitap mitap* ‘Bücher und so’ (von *kitap* ‘Buch’).

## 6. Illustrativer Text

- (27) *Vatan-das-lar-in* çoğu-u,  
Heimat-Gefährte-PL-GEN viel-POSS.3  
*İngiliz-ler-in* yönetim  
englisch-PL-GEN Leitung  
*merkez-ler-in-i* uzun yıl-lar  
Zentrum-PL-POSS.3-AKK lang jahr-PL  
*kendi el-ler-in-de*  
eigen Hand-PL-POSS.3-LOK  
*tut-ma-lar-in-dan* ve Çin  
halt-INF-PL-POSS.3-ABL und chinesisch  
*köken-li-ler-i* bilinç-li  
Wurzel-ADJR-PL-AKK Bewußtsein-ADJR  
*ol-arak* politika-dan  
AUX.DETR-ADVR Politik-ABL  
*uzak-laş-tır-ma-lar-in-dan*  
fern-PROZ-KAUS-INF-PL-POSS.3-ABL  
*yakın-iyor-lar.* Özel-lik-le  
ktag-PROG(3)-PL besonders-NR-KOMIT  
*de İngiliz yönetim-i-nin*  
auch englisch Leitung-POSS.3-GEN  
*demokratik hak-lar-i* çok geç  
demokratisch Recht-PL-AKK sehr spät  
*uygula-ma-ya* koy-ma-sı  
anwend-INF-ALL stell-INF-POSS.3  
*eleştir-il-iyor.* Gerçek-ten  
kritisiер-PASS-PROG(3.SG) wirklich-ABL  
*de Hongkong-da ilk özgür*  
auch Hongkong-LOK erst frei  
*parlamento seçim-ler-i* bin  
Parlament Wahl-PL-POSS.3 tausend  
*dokuz yüz doksan beş-te*  
neun hundert neunzig fünf-LOK  
*yap-il-di.*  
mach-PASS-PRÄT(3.SG)

‘Die meisten Bürger klagen darüber, daß die Engländer die Schaltstellen viele Jahre lang in ihren eigenen Händen gehalten hätten und die Chinesischstämmigen bewußt von der Politik ferngehalten hätten. Besonders wird auch kritisiert, daß die englische Verwaltung demokratische Reformen sehr spät durchgeführt habe. Und in der Tat wurden die ersten freien Parlamentswahlen in Hongkong 1995 durchgeführt.’

## 7. Unübliche Abkürzungen

DISP	Dispositiv
NPOT	Non-Potential
OBLPART	Obliquegerichtetes Partizip
SBJPART	Subjektgerichtetes Partizip
PRÄD	Postprädikatives Suffix

## 8. Zitierte Literatur

- Aksu-Koç, Ayhan (1986), "The Acquisition of Past Reference in Turkish". In: Slobin & Zimmer (Hrsg.), 247–264

Bassarak, Armin (1988), "Morphologisierungstendenzen im Türkischen". In: 2. Jenaer Semantik-Syntax-Symposium, *Wissenschaftliche Beiträge der Friedrich-Schiller-Universität Jena*, 222–233

Bassarak, Armin (1996), "On Turkish Vowel Harmony". In: Rona, Bengisu (Hrsg.), *Current Issues in Turkish Linguistics. Proceedings of the Fifth International Conference on Turkish Linguistics. School of Oriental and African Studies, 15–17 August, 1990, Bd. I.* Ankara: Hitit Yayinevi, 23–27

Bassarak, Armin (1997), "On the Position of the Question Marker in Turkish (An Integrated Model of Turkish Inflection Revised)". In: İmer, Kâmile & Uzun, N. Engin (Hrsg.), *Proceedings of the VIIIth International Conference on Turkish Linguistics*, August 7–9, 1996, Ankara, 183–190

Dede, Müserref (1986), "Definiteness and Referentiality in Turkish Verbal Sentences". In: Slobin & Zimmer (Hrsg.), 147–163

Savaşır, İskender (1986), "Habits and Abilities in Turkish". In: Slobin & Zimmer (Hrsg.), 137–146

Slobin, Dan Isaac & Zimmer, Karl (1986, Hrsg.), *Studies in Turkish Linguistics*. Amsterdam, Philadelphia: Benjamins (Typological Studies in Language 8)

Tekinay, Alev (1988), *Günaydin: Einführung in die moderne Türkische Sprache, Teil 2*. Wiesbaden: Ludwig Reichert

*Armin Bassarak, Berlin (Deutschland)  
Gerd Jendraschek, Toulouse (Frankreich)*

## 127. Hunzib (North-East Caucasian)

1. Introduction
2. Nouns
3. Adjectives and pronouns
4. Verbs
5. Illustrative Text
6. Uncommon abbreviations
7. References

### 1. Introduction

Hunzib is a Tsezic language belonging to the Dagestanian branch of North-East Caucasian languages. The other Tsezic languages are: Tsez, Hinukh, Bezhta, and Khvarshi. At present there are ± 2,000 speakers of Hunzib. Approximately half of them live in three villages (Hunzib, Nakhada, Garbutli) in the mountainous Tsunta district; the majority, however, lives in three villages on the Dagestanian lowlands. A small number of Hunzib live in Georgia. There are minor differences between the language spoken in Hunzib and in Nakhada and Garbutli, most of a lexical kind. These differences persist on the lowlands.

The Tsezic languages are non-written. Avar and Russian are the medium of instruction in school and are also used in newspapers, broadcasting, and literature. The Tsezic languages are spoken at home, with neighbours and friends in the village. The Hunzib living in the mountains understand Bezhta as well, but they generally speak Avar with speakers of Bezhta and the other Tsezic languages. Speakers of the Tsezic languages living on the plains use Avar or Russian in communication with people of other nationalities. The Tsezics have a fair command of both Avar and Russian but Russian is often preferred. Men have a far better command of these languages than women.

Few Hunzib have finished higher education. The main occupation for men is livestock-breeding, construction works, and small trade; women work in and around the house. The Tsezics are sunni muslims. They are endogamous, marrying within their village and even, if possible, within the family. Nowadays, marriage to a person of another nationality is not uncommon; when a non-Tsezic woman marries into a Tsezic community she is expected to learn the language of her husband. As long as the Hunzib continue living in their closed communities, the Hunzib lan-

guage will continue to play a major role in their lives.

Lomtadze (1956) considers Bezhta and Hunzib to be dialects of one language. The first sketch of Hunzib separately is Bokarev (1959). Since 1976 the Dagestanian scholar Isakov has been working on Hunzib, which resulted in a Hunzib-Russian dictionary (Chalilov & Isakov, 2001), and an unpublished grammatical sketch (cf. Isakov, ms.). Hunzib nominal lexicon is also represented in Kibrik & Kodzasov (1990). Van den Berg (1995) is based on on-site fieldwork among the Hunzib in the early 1990s.

### 2. Nouns

#### 2.1. Noun classes

Hunzib has five **noun classes** (cf. Art. 98), which serve as the basis for agreement between the intransitive subject or transitive patient and co-referring verbs, pronouns, and some adjectives and adverbs. Almost half of the verbs and a small part of the adjectives show agreement by class prefixes; apophony is used to indicate agreement in pronouns, the auxiliary and a few verbs (cf. van den Berg 1995: 80).

Class 1 contains nouns denoting male persons, class 2 female persons, class 3 inanimate objects, class 4 animate and inanimate objects, class 5 inanimate objects and the word /q'əra/ 'child'. In the plural only the distinction 'human' vs. 'non-human' is made.

	noun class	prefix	'be'	'this'
singular	CL1	Ø-	lo	bəd
	CL2	j-	lo	bodu
	CL3	j-	li	bəd
	CL4	b-	lo	bodu
	CL5	r-	li	bəd
plural	human	b-	lo	bəd-ra
	non-human	r-	li	bəd-ra

Tab. 127.1: Agreement for noun class in Hunzib

Instead of the class prefixes /b-/ and /r-/ we find /m-/ and /n-/ respectively, before a root-initial nasalised vowel, which then loses its nasalisation. Examples:

- (1) (a) /bəd i?eru ože  
this:CL1 CL1:small boy(ABS)  
ãq'e/  
CL1:come(PRES.3)  
'this small boy comes'
- (b) /bodu j-i?eru kid  
this:CL2 CL2-small girl(ABS)  
j-ãq'e/  
CL2:come(PRES.3)  
'this small girl comes'
- (c) /bodu b-i?eru wə  
this:CL4 CL4-small dog(ABS)  
m-aq'e/  
CL4:come(PRES.3)  
'this small dog comes'
- (d) /bəd-ra b-i?er-ar ož-da  
this-PL HUM.PL-small-PL boy-PL(ABS)  
m-aq'e/  
HUM.PL:come(PRES.3)  
'these small boys come'

## 2.2. Basic stems

There are four stems to which the case suffixes are attached: the base stem, the oblique singular stem, the plural stem and the oblique plural stem (cf. also Kibrik 1992 a).

The plural stem is derived from the base stem by a plural suffix, most frequently by /-la/, /-wa/, /-ja/, /-r/, /-a/, or /-ba/. The distribution is lexically determined. E.g.: /bitʃ/ 'ewe', plural /bitʃ'-la/; /'ut'/ 'silkworm', /'ut'-la/; /maru/ 'mountain', /maru-wa/; /k'odo/ 'grapes', /k'odo-wa/; /nani/ 'doll', /nani-ja/; /more/ 'log', /more-ja/; /qala/ 'fortress', /qala-r/; /qp'a/ 'paw', /qp'a-t/; /bərus/ 'wooden plough', /bərus-a/; /iž/ 'sieve', /iž-a/; /tiga/ 'chamois', /tig-ba/; /roq/ 'load', /roq'-ba/.

A quarter of all nouns, mainly abstract and collective nouns and nouns denoting materials, do not have a plural form (cf. also Kibrik 1992 b).

The large majority of Hunzib nouns form oblique stems by adding an oblique marker, most frequently /-li/, /-al/, /-lo/, /-jo/, /-i/, or /-bo/ to the base stem. The distribution is lexically determined. E.g.: /žzo/ 'courtyard', oblique stem /žzo-li-/; /q'al/ 'belt', /q'al-li-/; /qâšu/ 'throat', /qâš-a-/; /sam/ 'mint', /sam-a-/; /žagi/ 'bushes', /žagi-lo-/; /vav/ 'crop', /vav-lo-/; /kudari/ 'jug', /kudari-jo-/; /ist'oli/ 'table', /ist'oli-jo-/; /štsu/ 'door', /šts-i-/; /mor/ 'eagle owl', /mor-i-/; /χõxe/ 'tree', /χõχ-i-/; /tʃ'inik/ 'scythe', /tʃ'inik'-bo-/; /nutsu/ 'honey', /nutsu-bo-/.

Some nouns in a vowel do not have an oblique marker (/abu/ 'father', oblique stem

/abu-/) or form their oblique singular stems by apophony (/botso/ 'moon', oblique stem /bitsə-/; /hâ/ 'blood', /hiža-/). The oblique plural stem is derived from the plural stem by the oblique morpheme /-la/.

## 2.3. Syntactic cases: form and function

The absolute ending -Ø follows the base and plural stems. All other endings, like ergative /-l(o)/, genitive /-s/, instrumental /-d(o)/, and the local endings, are added to the oblique stems. In final position endings with the form -C(o) are -C after vowel and -Co after consonant in non-final position they are -Co. The absolute typically denotes intransitive subject/transitive patient (cf. examples (1 a-d)), the ergative the transitive agent.

- (2) /abu-l ože hehe-r/  
father-ERG boy(ABS) hit-PAST  
'the father hit the boy'

The genitive typically marks attributive nouns, attributive adverbs, and present participles modifying nouns in the absolute, e.g. /sin-a-s miy/ 'fork-OBL-GEN tail(ABS) (the haft of a fork)'. In general, an attributive form precedes its head.

The instrumental does not only denote the instrument, but also attributive nouns, attributive adverbs, and present participles modifying nouns in an oblique case; therefore, the instrumental is sometimes called the 'second genitive'. E.g. /mesed-li-d sin-a-s miy/ 'gold-OBL-INSTR fork-OBL-GEN tail(ABS) (the haft of a golden fork)'.

## 2.4. Location and direction

The seven local endings, dative, adessive, superessive, subessive, contactive, comitative, and allative have both essive and lative meaning. Some endings also have a syntactic, non-local, use. The local endings can be accompanied by postpositions which reinforce their meaning. Two directional suffixes, elative and translative, can occur after the local endings or after the accompanying postpositions.

	elative	translative
dative	-V	-V-s
adessive	-g(o)	-go-s
superessive	-tʃ(o)	-tʃ'o-s
subessive	-tʃ	(-tʃ-sə)
contactive	-t	-t-sə
comitative	-yur	-yur-sə
allative	-dər	-dər-sə

Tab. 127.2: Location and direction in Hunzib

The dative ending is usually a copy vowel that yields a single long vowel with the copied vowel, e.g.: /han-a-a/ ‘forest-OBL-DAT (in(to) the forest)’, /han-a-a-s/ ‘forest-OBL-DAT-ELAT (out of the forest)’. The dative denotes the experiencer, the adjunct of some transitive verbs, and the local meaning ‘inside’. In the last case it often combines with the postposition /ēdu/ ‘inside’, e.g.: /χōx-i-i ēdu-s/ ‘tree-OBL-DAT inside-ELAT (out of the tree)’.

The adessive combines with animate nouns only and has a wide range of non-local uses: it denotes the adjunct of verbs of speech and vision, the causee with causative verbs derived from transitives, and the involuntary agent with intransitive verbs.

Superessive /-tɬ'(o)/ often combines with the postposition /tɬ'odo/ ‘above’, e.g.: /ist'-oli-jo-tɬ'/ ‘table-OBL-SUPESS’ or /ist'oli-jo-tɬ' tɬ'odo/ ‘table-OBL-SUPESS above’ both mean ‘on(to) the table’, and /t'irə-tɬ'o-tɬ'/ ‘bridge-SUPESS-TRNSL’ or /t'irə-tɬ' tɬ'odo-tɬ'/ ‘bridge-SUPESS above-TRNSL’, both ‘over the bridge’.

Subessive /-tɬ/ virtually always occurs with the postposition /tɬ'irə/ ‘under’, to which the directional suffixes are added. E.g.: /ist'oli-jo-tɬ' tɬ'irə-s/ ‘table-OBL-SUBESS under-ELAT (from under the table)’.

The contactive denotes ‘in complete contact with’, e.g.: /nəsu-4/ ‘wall-CNCT (on(to) the wall)’. The comitative denotes ‘next to an (in)animate object, together with an animate object’, e.g.: /halmay-li-vur/ ‘friend-OBL-COM (with a friend)’. The allative ending always involves motion ‘towards an animate object’, e.g.: /is-t'i-dər/ ‘brother-OBL-ALL-(to (my/ your/his) brother)’.

## 2.5. Nominal compounds and derivations

Nominal compounds are few. They consist almost always of two nouns with a related or opposite meaning, e.g.: /iju+abu/ ‘mother+father (parents)’, /hawa+boq/ ‘sky+sun (weather)’, /helu+sige/ ‘pea+buckwheat (hail)’, and, in the village of Nak-hada only, /ožə+kid/ ‘boy+girl (ladybird)’. Only the last member of these compounds inflects, whereas both terms agree for number. There are some attributive compounds, e.g.: /art'o+zaman/ ‘before+time (earlier times)’, /hara+mitʃ/ ‘cow+nettle (non-stinging nettle)’, /heresi+xabar/ ‘falsehood+story (untrue story)’. There is a special group of attributive compounds with the second element /ʒo/ ‘thing’, e.g.: /gudo+ʒo/ ‘hen+thing (something like a hen)’.

The suffix /-i/ derives abstract nouns from nouns or adjectives; e.g.: /is=4i/ ‘sibling=

ABSTR (brotherhood)’, /k'ot't'u=4i/ ‘good =ABSTR (well-being)’ (cf. also 4.5.3 for the derivation of deverbal nouns).

## 3. Adjectives and pronouns

### 3.1. Adjectives

Hunzib **adjectives** can be used attributively, substantivally, and predicatively. When used attributively, the absolute adjective ends in /-u/ (cf. examples (1 a–c)), which is usually absent in the oblique cases (cf. example (3)). A singular adjective used predicatively also lacks /-u/ (cf. example (4)). Substantivally used adjectives are fully inflected with the usual case-endings following the stem-extending oblique morpheme /-o/ (cf. example (5)). The plural adjective suffix is /-ar/ (cf. example (1d)). Some adjectives show agreement by means of class prefixes (cf. examples (1 a–d)).

- (3) /boɬu j-i?er kid-bo-s nani/  
this:OBL CL2-small girl-OBL-GEN doll(ABS)  
‘this small girl’s doll’
- (4) /kid j-i?er lo/  
girl(ABS) CL2-small be:CL2  
‘the girl is small’
- (5) /j-i?er-o-s nani/  
CL2-small-OBL-GEN doll(ABS)  
‘the doll of the small one (F)’

The adjective has no special comparative form. The comparative ending /-ja/ is added to the object of comparison, e.g.:

- (6) /əg i?er abu-ja/ lo/  
that:CL1 CL1:small father-CMPR be:CL1  
‘he is smaller than father’

### 3.2. Personal pronouns

**Personal pronouns** do not formally distinguish between the absolute and the ergative case. There are no separate pronouns for third person; demonstrative pronouns are used instead. Hunzib has no possessive pronouns; the genitive and instrumental cases of the personal pronouns are also used to indicate possession. The second person singular instrumental forms are in free variation.

	me	you	we	you(PL)
ABS/ERG	də	mə	ile	miže
GEN	dije	dibə	ilu-s	mižu-s
DAT	dī?i	dibi	ilu-u	mižu-u
INSTR	di-do	dibə-d/	il-do	miž-do
OBL	di-	du-	ilu-	mižu-

Tab. 127.3: Hunzib personal pronouns

### 3.3. Demonstrative pronouns

**Demonstrative pronouns** in the absolute singular distinguish between classes 1/3/5 and 2/4 (cf. 2.1), except for speakers from Garbutli, who only have the 2/4 series (cf. van den Berg 1995: 348). This distinction is absent in the oblique singular cases and in the plural. There are three degrees of distance: /bəd, bodu/ ‘near the speaker, this’, /bəl, bolu/ ‘near the hearer, this’, /əg, ogu/ ‘far from both speaker and hearer, that’.

	1/3/5	2/4	oblique	plural
‘this (near the speaker)’	bəd	bodu	boṭu	bəd-ra
‘this (near the hearer)’	bəl	bolu	boṭu	bəl-la
‘that’	əg	ogu	oṭu	əg-ra

Tab. 127.4: Hunzib demonstrative pronouns

When used substantively the demonstratives are fully inflected, compare example (3) with /boṭu-s nani/ ‘this-GEN doll(ABS) (his/her/ its doll)’.

The roots /bəd/, /bəl/, and /əg/ also occur in the adverbs /bə?i/ ‘here’, /bəli/ ‘here’, /əgi/ ‘there’, /bədaa/ ‘like this’, /bəlaa/ ‘like this’, /əgaa/ ‘like that’.

### 3.4. Interrogative pronouns

**Interrogative pronouns** are, e.g.: /suk’u/ ‘who?’, /ʃijo/ ‘what?’, /nijo/ ‘where?’, /hinaa/ ‘how?’, /sisər/ ‘which?’, /hidə/ ‘when?’, /hidil/ ‘how much?’, /surba/ ‘why?’. /suk’u/ ‘who?’ and /ʃijo/ ‘what?’ have the oblique stem /sa-/ and /sini-/ respectively.

The specific indefinite pronoun is formed by adding the indefinite particle /di/ to an interrogative pronoun, e.g. /suk’u di/ ‘someone’, /ʃijo di/ ‘something’, /nijo di/ ‘somewhere’. These indefinite pronouns only combine with affirmative verb forms.

- (7) /suk’u di əq’e-r-i/  
      who(ABS) INDEF CL1:come-PAST  
      ‘someone came’

The non-specific indefinite pronoun is formed by adding the coordination suffix /-n(o)/ to an interrogative pronoun. In sentences with an interrogative verb form it expresses indefiniteness, in combination with a negative verb form, negation.

- (8) (a) /suk’u-n əq’e-r-i/  
          who(ABS)-and CL1:come-PAST-INT  
          ‘did anyone come?’

- (b) /suk’u-n əq’-is/  
      who(ABS)-and CL1:come-PAST.NEG  
      ‘no one came’

### 3.5. Reflexive pronouns

The third person **reflexive pronoun** /žu/ ‘self’ has the oblique stem /žini-/, plural stem /žuru/, and plural oblique stem /žide-/. The pronouns are used for reflexivization and anaphorically for co-reference. In a reflexive expression the pronoun occurs twice, in the cases as required by the verb. First and second person use a reflexive element /-ž-/ between the two occurrences of the pronoun, each in the case as required by the verb. Compare example (2) with (9).

- (9) (a) /ož-di-l žin-lo žu  
      boy-OBL-ERG self:OBL-ERG self(ABS)  
      hehe-r/  
      hit-PAST  
      ‘the boy hit himself’  
      (b) /də-ž-də hehe-r/  
      me-REFL-me hit-PAST  
      ‘I hit myself’

## 4. Verbs

Hunzib **verb** forms consist of a stem, which can be preceded by a class prefix and followed by endings. Simple stems consist of a root only, complex stems of a root and one or more derivational suffixes. Almost half of the verbs have the possibility to indicate the class of their intransitive subject (cf. examples (1 a–d)) or their transitive patient (cf. example (13 a)). Except for the present tense, negation is indicated with port-manteau suffixes combining tense and negation (cf. 4.6).

### 4.1. Verbal derivation

The causative suffix /-k’(V)/ and verbalizing suffix /-l(V)/ have the allomorphs /-k’/ and /-l/ after vowel, /-k’V/ and /-lV/ after consonant. The vowel of the suffix is /ə/ when the root vowel is /ə, i/, in all other cases it is /e/. Examples: /hehe=k’/ ‘hit=CAUS’, /-ūq=k’e/ ‘eat=CAUS’, /-isə=k’/ ‘sell=CAUS’, /-əd=k’ə/ ‘manage=CAUS’.

#### 4.1.1. Causative

The **causative suffix** /-k’(V)/ combines with all adjectives and almost all verb roots. It derives transitive verbs from adjectives.

- (10) /ile qoqo haldu=k’-er/  
      we house(ABS) white=CAUS-PAST  
      ‘we made the house white’

With derivation from verbs, the valency increases: intransitives become transitives (example (11)), transitives become ditransitives (compare example (2) with (12)).

- (11) /iju-l            kid  
mother-ERG girl(ABS)  
j-artſe=k'-er/  
CL2-get.up=CAUS-PAST  
'mother woke up the daughter'
- (12) /obu-l        ož-di-g        wə  
father-ERG boy-OBL-ADESS dog(ABS)  
hehe=k'-er/  
hit=CAUS-PAST  
'father made the boy hit the dog'

The suffix /-k'(V)/ may occur twice in one and the same form: the first causative suffix derives a transitive verb, the second adds the instigation. In some cases, however the addition of a second suffix intensifies the meaning of 'make (someone do something) to 'force (someone to do something)'.

#### 4.1.2. Verbalizer

While derivation with the **verbalizing suffix** /-l(V)/ is not productive and its meaning is less transparent, stems with /-l(V)/ are presented as petrified derivations. /-l(V)/ derives intransitive, inchoative verbs from adjectives: /haldū/ 'white', /haldul/ 'become white', /-iq'q'u/ 'big', /-iq'lə/ 'grow (up)'.

As opposed to /-k'(V)/, the suffix /-l(V)/ does not always increase the valency when deriving from a verb. In some cases, it derives a transitive from an intransitive verb, e.g. /-axe/ 'tear, itr.', /-axel/ 'tear, tr.'; /-ek/ 'fall', /-ekle/ 'make fall'.

In other cases, it merely changes the meaning of the verb, e.g. /-iqə/ 'close', /-iqəl/ 'hide'; /nis(ə)/ 'say', /nisəl/ 'ask'; /itʃ'e/ 'dress oneself', /itʃ'el/ 'dress another person'.

#### 4.1.3. Detransitive

The **detransitive suffix** /-laa/ derives intransitive verbs from transitive verbs (cf. Art. 107). The original agent becomes subject (the action carried out is characteristic of subject), while the former patient disappears and cannot be expressed in an oblique case (cf. examples (13 a–b)). Some detransitivized verbs have a reflexive meaning (cf. example (14)). Though the suffix /-laa/ is rather frequent, detransitive verbs cannot be derived from every transitive verb. Besides /-laa/ the petrified affixes /-raa/ and /-daa/ also occur.

- (13) (a) /oču-l            χανу  
that:OBL-ERG onion(ABS)  
m-isə-r/  
CL4-sell-PAST  
'(s)he sold onions'  
(b) /əg        ̄isə=laa-r/  
that:CL1 CL1:sell=DETR-PAST  
'he traded'
- (14) /kid        q'uti-la-a  
girl(ABS) trunk-OBL-DAT  
j-ūtsu=laa-r/  
CL2-hide=DETR-PAST  
'the girl hid (herself) in the trunk'

#### 4.1.4. Inchoative

The **inchoative suffix** /-k(e)/ derives intransitive verbs from adverbs and other expressions of time, e.g. /az-du=k(e)/ 'summer-OBL=INCH (set in (of summer))', /ədu=k(e)/ 'inside=INCH (go inside)'.

#### 4.2. Number

Some verbs only express **number** by their class prefixes. In some forty percent of the verbs, however, plurality of the intransitive subject/transitive patient can (also) be expressed affixally, either by the suffix /-baa/, or by the infix /-a-, -å-, -jå-, -wå-/ . The suffix /-baa/ follows most stems in long /aa/, e.g. /hehedaa/ 'neigh', /hehedaa-baa/ 'neigh-PL', /koſe=laa/ 'mow=DETR', /koſe=la-baa/ 'mow=DETR-PL' (for shortening of long vowels, cf. van den Berg 1995: 22).

The plural infix /-a-, -å-, -jå-, -wå-/ always supplies the second vowel of the verb. The choice of the allomorph is determined as follows:

(a) If there is one vowel before the last consonant of the stem, the infix /-a-, -å-, -jå-, -wå-/ follows the vowel. Stem-final /o, u, ə/ change into /e/ when the infix is inserted. The infix has the following distribution:

- /-a-/ after /a/, e.g.  
/(m-)a<a>ts'ə/ 'see<PL>'  
/-a-/ after /a/, e.g.  
/-a<a>hu/ 'take<PL>'  
/-ja-/ after /i, e/, e.g.  
/e<ja>k/ 'fall<PL>',  
/-i<ja>q'e/ 'know<PL>'  
/-wa-/ after /i, ə, o, u/, e.g.  
/-u<wa>t/ 'sleep<PL>',  
/k'o<wa>k'e/ 'be.ill<PL>',  
/-ə<wa>q'e/ 'lead<PL>'.

(b) If there are two or more vowels before the last consonant of the stem, the infix /-a-/ replaces the second vowel, e.g. /-ežel/ 'take.away', /-ež<a>l/ 'take.away<PL>'; /hatfuk/ 'look', /hatf<a>k/ 'look<PL>'.

(c) If there is one vowel before a consonant group, the infix /-a-/ is inserted between the consonants of the group, e.g. /-ekle/ 'make.fall', /-ek<a>le/ 'make.fall<PL>', /-iχ=k'ə/ 'hot=CAUS (warm up)', /iχ-a=k'e/ 'hot-PL=CAUS (warm up)'.

#### 4.3. Finite verb endings

Stems end in a consonant, a short vowel or long /aa/. As sequences of non-identical vowels are avoided, a stem-final short vowel is deleted before a following dissimilar vowel, e.g.: /hehe/ 'hit', /heh-ar/ 'hit-FUT', /heh-is/ 'hit-PAST.NEG'. With stems ending in long /aa/, the vowel of the suffix is deleted, with the exception of /i/, which yields /j/, e.g.: /helaa/ 'boil', /helaa-r/ 'boil-FUT', /helaa-js/ 'boil-PAST.NEG'. A table of affirmative and negative finite endings is given in 4.6.

##### 4.3.1. Present

The Hunzib verb shows personal conjugation in the **present** tense. The first and second person intransitive subject/transitive agent have the present ending /-tʃ(o)/, whereas the third person has no overt present tense ending. The present ending /-tʃ(o)/ is /-tʃo/ after consonant, /-tʃ/ after vowel in final position; in non-final position it is /-tʃo/, e.g.: /ētʃ'e-tʃ/ 'CL1:go-PRES', /tʃax-tʃo/ 'write-PRES', /ētʃ'e-tʃo-s/ 'CL1:go-PRES-GEN (going)', /tʃax-tʃo-s/ 'write-PRES-GEN (writing)' (cf. 4.5.3).

##### 4.3.2. Preterite

The **preterite** ending /-(V)r/ is /-r/ after a vowel, and /-Vr/ after a consonant. The initial vowel of the ending is /e, u, o, ə/ according to the following distribution:

- /e/ after palatal consonants (/tʃ, tʃ̥, ſ, ſ̥/), e.g. /bož-er/ 'believe-PAST'.
- /e/ after a stem containing a plural infix /-ja-, -wa-/., e.g. /-u<wa>q-er/ 'eat<PL>-PAST'.
- /e/ after the suffixes /-k'(V)/ or /-l(V)/ with all other root vowels than /i, ə/, e.g. /hehe=k'-er/ 'hit=CAUS-PAST', /ʃo?o=k'-er/ 'forget=CAUS-PAST'.
- /e/ after a stem containing the root vowels /i, e, a, ə/, e.g. /-it'-er/ 'divide-PAST', /-ek-

er/ 'fall-PAST', /tʃ'al-er/ 'notice-PAST', /-ütʃ'aax-er/ 'slumber-PAST'.

/u/ after a stem containing the root vowel /u/, e.g. /-ūq-ur/ 'eat-PAST'.

/o/ after a stem containing the root vowel /o/, e.g. /-ots'-or/ 'chase-PAST'.

/ə/ after a stem containing the root vowels /i, ə/, also after the suffixes /-k'(V)/ or /-l(V)/, e.g. /-əd-ər/ 'manage-PAST', /nitʃ'-ər/ 'give-PAST', /-isə=k'-ər/ 'sell=CAUS-PAST'.

##### 4.3.3. Generic and future

The generic suffix /-a/ denotes an event in the non-past, which may be habitual or not. Some Hunzib do not distinguish between the generic tense and the future tense marked by the suffix /-ar/. The single corresponding negative ending has already completed this development (cf. 4.6).

##### 4.3.4. Imperative

Intransitive verbs have a zero **imperative** ending, whereas transitive verbs show the imperative ending /-o/. The imperative ending often lacks with transitive verbs derived by /-k'(V)/ and /-l(V)/, which are overtly marked as transitive. E.g.: /kax-o/ 'paint-IMP', /nis-o/ 'say-IMP', /tʃ'e/ 'CL1:go(IMP)', /r-atʃ'el-o/ 'CL5-strike-IMP' or /r-atʃ'el/ 'CL5-strike(IMP)'. Experiencer verbs do not have an imperative; they take the imperative of the derived causative verb (cf. 4.1), most of the time without an overt imperative ending, e.g. /ʃo?o=k'-o/ 'forget=CAUS-IMP' or /ʃo?o=k'/ 'forget=CAUS(IMP)', both 'forget it' (cf. van den Berg 1995: 123).

#### 4.4. The verb 'to be'

The paradigm of the verb 'to be' is suppletive and defective. It consists of the suppletive roots /li, lo/ in the present (cf. Tab. 127.1), negative /ga(j)/, and /zuq'u/ in other forms. Missing forms are taken from the roots /-aq(u)/ 'happen, usually be' or /zəv/ 'find, appear to be, probably be'.

#### 4.5. Non-finite verb endings

##### 4.5.1. Infinitive

The **infinitive**, marked by a suffix /-a/, is used in purpose clauses (cf. /hatfuk'-a j-ētʃ'e-n lo/ in example (24 a)) and complement clauses, e.g.:

- (15) /ož-di-i j-at'ə-r kid  
 boy-OBL-DAT CL2-want-PAST girl(ABS)  
 heh-a/  
 hit-INF  
 'the boy wanted to hit the girl'

The infinitive ending can be followed by the benefactive clitic /dija/ which explicitly denotes purpose (supine), e.g.: /bet'erbaqi b-uw-a dija/ 'household(ABS) CL4-do-INF BEN (to make a living)'.

#### 4.5.2. Gerunds

The present tense form in /-tʃ(o)/ is also used as a present **gerund**. Usually, this form is followed by benefactive /dija/ and denotes simultaneity. Without /dija/, the present gerund is only found in analytic predicates.

- (16) /nəsu r-uwo-tʃ dija də  
 wall(ABS) CL5-do-PRES BEN me  
 ek-er/  
 CL1:fall-PAST  
 'while making a wall I (masc.) fell'

The preterite gerund ending -(V)n has the same distribution of allomorphs and vowel quality as the preterite ending -(V)r (cf. 4.3.2). The preterite gerund denotes an event accomplished before the event referred to by the main verb, cf. /b-ek'etʃ-en wə-j-l m-uq-un lo ěʃ/ in example (24h) and the Hunzib proverb given in example (17), which contains the negative preterite gerund (cf. 4.6).

- (17) /sərjo-n b-uh-it'o  
 horse(ABS)-and CL4-die-GER.NEG  
 suk'u-n uh-it'o  
 man(ABS)-and CL1:die-GER.NEG  
 qala b-ah-ojs/  
 fortress(ABS) CL4-take-FUT.NEG  
 'unless horse and man die, you cannot take the fortress'

Both the present and the preterite gerund can be found in complex finite forms. Analytic predicates consist of a gerund or participle and a form of the verb 'to be'. Participle predicates, which focus on the participants of the event, are less frequent than predicates with gerunds, which focus on the event as a whole.

Preterite gerunds in combination with the auxiliary /lo, li/ render the perfect tense; a present or preterite gerund with /zuq'u-r/ 'was' forms the pluperfect. A present or preterite gerund with the perfect of 'be' /zuq'u-n lo, li/ renders the evidential, which denotes an event in the past that was not witnessed by the speaker.

There is an aspectual difference between complex predicates with the preterite or the present gerund: predicates with a present gerund depict events as uncompleted, repeated, or habitual, while predicates with a preterite gerund present them as completed.

#### 4.5.3. Participles

The present **participle** consists of present /-tʃ(o)/ and the attributive markers /-s, -d/ with the usual distribution (cf. 2.2), e.g. /j-uhu-tʃo-s kid/ '2-die-PRES-GEN girl(ABS) (the dying girl)' and /j-uhu-tʃo-d kid-bo-s hare/ '2-die-PRES-INSTR girl-OBL-GEN eye(ABS) (the eye of the dying girl)'.

For the preterite participle the adjectival ending /-u/ is added to the preterite, giving /-ru/ (cf. 3.1 for their attributive and substantival use).

- (18) /də ãts'ə-ru kid/  
 me CL1:see-PAST.PART girl(ABS)  
 'the girl who saw me (M)'  
 (19) /hĩja-d blood:OBL-INSTR  
 r-əts'-əru tʃ'it'/  
 CL5-be.covered-PAST.PART knife(ABS)  
 'the knife covered with blood'  
 (20) /yalq'  
 people(ABS)  
 b-u<wa>t'-eru zaban/  
 HUM.PL-sleep<PL>-PAST.PART time(ABS)  
 'the time when the people were sleeping'

A deverbal noun is derived from the preterite participle by the abstract suffix /-ti/, resulting in /-(V)r=ti/. The subordinate intransitive subject /transitive patient is either in the absolute, or in one of the two attributive cases (genitive, instrumental). Examples: /q'əra qoda-baa-r=ti/ 'child(ABS) study-PL-PAST. PART=ABSTR (children's study)', /abu uhu-r =ti/ 'father(ABS) CL1:die-PAST.PART=ABSTR' or /abu-s uhu-r=ti/ 'father-GEN CL1:die-PAST.PART=ABSTR' both meaning 'father's death'.

Both the preterite and the present participle can occur in complex predicates (cf. 4.5.2 and van den Berg 1995: 102–104).

#### 4.5.4. /-oʃ/ 'when'

The subordinating /-oʃ/ denotes an event anterior to the event of the main clause, often with a causal relationship. Forms in /-oʃ/ are preferred over the preterite gerund in subordinate clauses, when the transitive agent or experiencer of the subordinate clause does

not co-refer with an actant of the main clause.

- (21) /fi-t<sup>3</sup>'o j-ēt<sup>3</sup>-o<sup>4</sup>, ḡgi  
water-SUPESS CL2-go-when there  
vurdelo lo/  
mullah(ABS) be:CL1  
'when she went for water, the mullah  
was there (i.e. at the well)'
- (22) /eg tadaax-o<sup>4</sup>, t<sup>3</sup>irə  
that:CL1 be.tired-when under  
ek-er/  
CL1:fall-PAST  
'as he got tired, he fell'
- (23) /wadə n-aq'-i<sup>4</sup>,  
rain(ABS) CL5-come-when.NEG  
baχ qoqo-n li/  
grass(ABS) dry-GER be:CL5  
'because there was no rain, the grass  
has dried'

#### 4.6. Negation

Except in the present tense, **negation** is indicated with port-manteau suffixes combining tense and negation. Most suffixes share the element /i/. A negative suffix /-at'(o)/ follows the present tense suffixes only. The forms of the negative preterite gerund /-it'(o)/ and negative imperative /-aq'(o)/, with or without /o/, are in free variation.

	affirmative	negative
PRES.3	-Ø	-Ø-at'
PRES	-t <sup>3</sup> (o)	-t <sup>3</sup> -at'
PRES-GEN	-t <sup>3</sup> o-s	-t <sup>3</sup> -at'o-s
PAST	-(V)r	-is
PAST.PART	-(V)ru	-isu
PAST.PART=ABSTR	-(V)r=4i	-is=4i
FUT	-ar	-ojs
GNR	-a	-ojs
IMP	-Ø, -o	-aq'(o)
GER	-(V)n	-it'(o)
when	-o <sup>4</sup>	-i <sup>4</sup>

Tab. 127.5: Affirmative and negative verbal endings

#### 4.7. Interrogation

The **interrogative suffix** is /-j/ after vowel and /-i/ after consonant. It follows a verbal form or the nominal part of a predicate. Note that the combination of the interrogative and the present suffixes always yields /-t<sup>3</sup>o-j/, e.g. /hehe-t<sup>3</sup>o-j/ 'hit-PRES-INT (are you hitting?)' (cf. /hehe-t<sup>3</sup>/).

Instead of a suffix, one also finds an interrogative infix /-j-/ inserted before root-final consonant. Verb forms with an interrogative infix may have a negative nuance: the activity denoted by the verb should not have been done or, conversely, should have been done but has not, e.g. /tʃa<j>χ-is/ 'write<INT>-PAST.NEG (did (s)he (really) not write? ((s)he should already have written))'; /tʃa<j>χ-er/ 'write<INT>-PAST (did (s)he (really) write? (why did (s)he, (she) should not have))'. The infix is also used to check a message which is not completely understood, e.g.: /j-a<j>t<sup>3</sup>el-er/ '2-freeze<INT>-PAST ({what did you say?} did you (F) freeze?)'.

#### 4.8. Verbal compounds

Verbal compounds either consist of native Hunzib material only or contain a borrowed element. Native compounds are few: almost all are made up of two simple roots, e.g.: /-ef+χut<sup>3</sup>/ 'eat+drink (eat)', /-ahda+e<sup>3</sup>/ 'work+eat (live)', /-etʃe+nutsu/ 'stay+have.to (wait)', /-ūq'e+nuts'u/ 'bend+have.to (insult)'.

The Hunzib verbs /-aq(u)/ 'happen' and /-uwo/ 'do' can combine with Avar nouns or verbs in the infinitive. Compounds with /-aq(u)/ are intransitive, those with /-uwo/ transitive, e.g.: /beʃkize+aq(u)/ 'stop+happen (stop)', /hardizi+uwo/ 'ask+do (ask, beg)'. Compounds with nouns often have a fixed class prefix, e.g.: /gotʃi+b+aq(u)/ 'movement+happen (resettle)', /χal+b+uwo/ 'attention+do (pay attention)'.

#### 5. Illustrative Text

This text is the final part of a story relating the fate of a young girl who was chased away from home by her step mother. At this point in the story the step mother has discovered that the girl is alive and pays her a visit.

- (24) (a) /hat<sup>3</sup>uk'-a j-ēt<sup>3</sup>e-n lo,  
watch-INF CL2-go-GER be:CL2  
zahru-n b-eʒe-n  
poison(ABS)-and CL4-take-GER  
oɻu-u m-ij-a dija/  
that:OBL-DAT CL4-send-INF BEN
- (b) /əgi-s ēzo-li-i  
there-ELAT courtyard-OBL-DAT  
j-āq'-o<sup>4</sup>, suk'u di, itʃil  
CL2-come-when who INDEF old  
qadam-li-d q'aʃida-l-t<sup>3</sup>o  
person-OBL-INST outfit-OBL-SUPESS

- j-ëtʃ'e-n lo/  
CL2-go-GER be:CL2
- (c) /“waj, dije kid, mə  
Gee me:GEN girl(ABS) you  
bə?i-j lo, dibi  
here-INT be:CL2 you:DAT  
nukar  
servant(ABS)  
j-at’-at’-i”/  
CL2-want-(PRES.3)-NEG-INT
- (d) /bed ołu-l nisə-n li  
then that:OBL-ERG say-GER be:CL5  
“di?i nukar-no  
me:DAT servant(ABS)-and  
j-at’-at’, dije  
CL2-want(PRES.3)-NEG me:GEN  
bet’erhan lo”/  
husband(ABS) be:CL1
- (e) /bed-do-n nisə-n k’ot’t’u  
then-APPRX-and say-GER good  
tsəsə, չabar m-iye-n-no  
bad story(ABS) CL4-tell-GER-and  
j-ijaa-n lo, tʃe?e-n  
CL2-cry-GER be:CL2 laugh-GER  
lo j-etʃe-n lo əgi  
be:CL2 CL2-stay-GER be:CL2 there  
ogu/  
that:CL2
- (f) /j-etʃe-n-no zahru-n  
CL2-stay-GER-and poison(ABS)-and  
m-iye-n lo ołu-u/  
CL4-send-GER be:CL4 that:OBL-DAT
- (g) /zahru-n m-iye-n  
poison(ABS)-and CL4-send-GER  
ẽʃ gul-un lo  
apple(ABS) put:CL4-GER be:CL4  
ist’oli-jo-tʃ’/  
table-OBL-SUPESS
- (h) /ẽʃ gul-oɿ  
apple(ABS) put:CL4-when  
b-ek’etʃ’-en wə-j-l  
CL4-jump-GER dog-OBL-ERG  
m-uq-un lo ogu  
CL4-eat-GER be:CL4 that:CL4  
ẽʃ/  
apple(ABS)
- (i) /ẽʃ m-uq-oɿ wə  
apple(ABS) CL4-eat-when dog(ABS)  
b-uhu-n lo əgi  
CL4-die-GER be(ABS) there  
q’ere/  
downwards
- (j) /“bolu-j di?i ՚illa  
this:CL4-INT me:DAT trick(ABS)  
zuq’u-r bołu-s, bolu  
be-PAST this:OBL-GEN this:CL2

zuq’u-n j-ãts’ej di-do  
be-GER CL2-seem me:OBL-INST  
abu-s aqe” tʃ/e/  
father-GEN wife(ABS) QUOT

‘She took poison with her to give to her and went to look for her. When she came from there onto the courtyard, she went in the disguise of an old lady. ‘Gee, my darling, are you here, don’t you want a servant?’ She said: ‘I don’t want a servant, I have a husband.’ Then she said good and bad things, talked a lot, cried, laughed and in the end stayed there. As she stayed there, she gave her the poison. She gave the poison putting an apple onto the table. Having put the apple there, the dog jumped and ate the apple. After eating the apple, it fell down dead. ‘Was that her trick for me? Then it surely is my fathers wife,’ she said.’

## 6. Uncommon abbreviations

APPRX approximative  
CNCT contactive

## 7. References

- Berg, Helma E. van den (1995), *A Grammar of Hunzib (with Texts and Lexicon)*. München: Lincom
- Bokarev, E[vgenij] A. (1959), *Cezskie (didojskie) jazyki Dagestana*. Moskva: AN SSSR
- Chaililov, M[ajid] Sh. & Isakov, I[sak] A. (2001), *Gunzibsko-russkij slovar’*. Moskva: Nauka
- Isakov, I[sak] A. (ms.), *Očerki opisanija gunzibskogo jazyka: Fonetika, morfologija i leksika*. Machačkala
- Kibrik, A[leksandr] E. (1992 a), “Principy organizacii imennoj paradigmy v dagestanskikh jazykach (sopostavitel’no-tipologičeskie nabljudenija)”. *Očerki po obščim i prikladnym voprosam jazykoznanija*. Moskva: MGU, 80–101
- Kibrik, A[leksandr] E. (1992 b), “Čislovye formy nesčetnyx sušestvitel’nyx v dagestanskikh jazykach”. *Očerki po obščim i prikladnym voprosam jazykoznanija*. Moskva: MGU, 131–134
- Kibrik, A[leksandr] E. & Kodzasov, S[andro] V. (1990), *Sopostavitel’noe izučenie dagestanskich jazykov. Imja. Fonetika*. Moskva: MGU
- Lomtadze, È[lizbar] A. (1956), “Analiz kapučinsko-gunzibskogo jazyka”. *IKJa*. 8, 369–412

## 128. Ketisch (Jenisseisch)

1. Allgemeine Information
2. Zur Phonologie
3. Allgemeines zur morphologischen Typologie
4. Das Verb
5. Das Substantiv
6. Das Pronomen
7. Andere Wortarten
8. Zur Wortbildung
9. Illustrativer Text
10. Unübliche Abkürzungen
11. Zitierte Literatur

### 1. Allgemeine Information

Das Ketische wird heute von rund tausend Tajgajägern und Fischern gesprochen, die seit jeher am Jenissei und an seinen Nebenflüssen Dubtsches, Podkamennaja Tunguska, Baklanicha, Bachta, Jeloguj, Surguticha, Imbak, Turuchan und Kureika ansässig sind. Es ist das letzte noch lebendige Mitglied einer ehemals großen Sprachfamilie, deren mögliche genetische Beziehungen zu anderen Sprachen Asiens oder Nordamerikas ungeklärt bleiben. Die übrigen bekannten Jenissei-Sprachen – das Arinische, das Assanische, das Kottische und das Pumpokolische – sind bereits im 19. Jahrhundert ausgestorben; das dem Ketischen am nächsten verwandte Jugische (am Fluß Sym) ist in den 80er Jahren des vergangenen Jahrhunderts ebenfalls endgültig verstummt.

Die russische Pluralform des Wortes *Ket* (eigentlich ‘Mensch’) gilt als offizieller Volksname seit den 30er Jahren; daneben ist auch die alte Selbstbezeichnung *ostøyən* (‘Ostjaken’) gebräuchlich. In der älteren Literatur war der Name “Jeniszej-Ostjakisch” üblich.

Auf dem riesengroßen Territorium seiner Verbreitung zerfällt das Ketische in zahlreiche Mundarten, die allerdings nur relativ geringe Unterschiede aufweisen. Bei der Volkszählung 1989 gaben nur 538 von insgesamt 1113 Keten Ketisch als ihre Muttersprache an. Praktisch alle beherrschen sie Russisch, wenn auch in unterschiedlichem Maße. Daneben sprechen viele von ihnen das samojedische Selkupisch als die Sprache eines größeren Nachbarvolkes.

Ketisch existiert nur in mündlicher Form. Der in den 30er Jahren vorgelegte Schriftentwurf hat keine Verwendung gefunden. Seit den 80er Jahren werden Versuche unternommen, ein neues Schriftsystem zu entwickeln

und Ketisch als Muttersprache in den Schulunterricht einzuführen.

Den Grundstein zu einer wissenschaftlichen Beschreibung der Jenissei-Sprachen hat Mathias A. Castren Mitte des 19. Jahrhunderts gelegt (Castren 1858). Von den vielen Forschern, die einen bedeutenden Beitrag zum Studium des Ketischen geleistet haben, sind vor allem Kai Donner (Donner 1955), Nikolaj K. Karger (Karger 1934), Andrei P. Dul’zon (Dul’zon 1968) und Eruchim A. Krejnović (Krejnović 1968) zu nennen.

### 2. Zur Phonologie

Abgesehen von einigen irrelevanten Details erweist sich das phonologische System des Ketischen als relativ einfach und schließt nicht mehr als 20 Phoneme ein (Vall & Kankin 1990: 12–21).

Vokale	/a-æ/, /i/, /e-ɛ/, /u/, /o-ɔ/, /ɪ/, /ə-ʌ/
Konsonanten	/m/, /n/, /ŋ/, /l/, /ʃ/, /s/, /h/, /t/, /d-r/, /b-p-v/, /k-g-ɣ/, /q-x-R/, /ʔ/ (unter Berücksichtigung der wichtigsten Phonemvarianten)

Tab. 128.1: Phoneme

Die phonologische Bewertung des Stimmänderverschlusses ist umstritten, was aber für die folgende Beschreibung belanglos bleibt.

### 3. Allgemeines zur morphologischen Typologie

Der klassischen morphologischen Sprachtypologie, die voll ihre Bedeutung behält, liegt bekanntlich die Wortstruktur zugrunde. Für das Wesen der Isolierung, der Agglutination und der Flexion sind dabei vor allem zwei Momente entscheidend: die Verwendung von grammatischen Morphemen sowie der Grad ihrer Unentbehrlichkeit in der Wortform (vgl. Plotkin 1989: 14–17). Das Wort kann aus einem Stamm bzw. aus einer bloßen Wurzel bestehen oder neben dem Stamm ein oder mehrere Affixe enthalten. Diese sind entweder fakultativ und lassen sich je nach Bedarf mehr oder weniger frei und lose an den Stamm anschließen und genauso leicht ent-

fernen, oder sie sind im Prinzip unerlässliche Bestandteile einer Wortform, ohne die der Stamm in der Regel nicht vorkommt und nicht selbständige fungieren kann. Im letzten Fall kann ihre Verbindung mit dem Stamm so eng sein, daß es zu einer Fusion kommt, bei der die Morphemgrenzen verwischt und die Morpheme selbst oftmals stark deformiert werden. Eine Wortform kann also entweder ein lockeres Morphemgefüge sein, das vom Sprechenden nach einem Modell zusammengesetzt wird, oder vielmehr ein irregulärer und in seine Bestandteile nur schlecht zerlegbarer Morphemkomplex, der dem Sprechenden als eine fertige Einheit gegeben ist. Als fertige Formen treten natürlich auch die unberechenbaren Suppletivbildungen auf.

Es muß betont werden, daß verschiedene Worttypen in ein und derselben Sprache nebeneinander bestehen können, wobei sie an verschiedene Wortarten gebunden sind. In der ketischen Morphologie weisen das Verb und das Pronomen zahlreiche Fusions- und Suppletivverscheinungen auf, während für das Substantiv die Agglutination kennzeichnend ist und das Adjektiv, das Adverb und das Numerale praktisch unveränderlich bleiben.

#### 4. Das Verb

##### 4.1. Die Wortformstruktur

Das **Verb** ist die weitaus wichtigste Wortart im Ketischen. Eine finite Verbform bildet an und für sich schon stets eine grammatisch vollwertige Kernaussage, die eventuell nur erweitert wird. Das Verb hat das umfangreichste Paradigma, und seine Formen haben die komplizierteste Morphemstruktur.

Verhältnismäßig selten besteht der Verbalsstamm nur aus einer Wurzel. In den meisten Fällen handelt es sich um eine Vereinigung von zwei Wurzelmorphemen, die nicht immer unmittelbar nebeneinander stehen, mit einigen anderen Elementen – sogenannten Determinativen, Sproßvokalen und Bindekonsonanten sowie manchen Segmenten, deren Funktion sich nicht ohne weiteres bestimmen läßt. Sie bilden eine diskontinuierliche Verbalbasis, in die die grammatischen Affixe eingebaut werden. Sowohl die Basiselemente als auch die Affixe sind größtenteils kurz – einsilbig oder unsilbisch – und unterliegen gewissen Alternationen und phonetischen Deformierungen, was zu ihrer weitgehenden Homonymie führt; je nach einzelnen Ver-

klassen nehmen sie zum Teil verschiedene Stellen ein. Das alles macht die Wortformbildung äußerst unregelmäßig und unerwartet.

Von den grammatischen Bedeutungen des Verbs, die durch die Affixe ausgedrückt werden, beziehen sich die einen auf den Prozeß (Zustand, Vorgang bzw. Tätigkeit), die anderen charakterisieren seinen Träger bzw. seine Teilnehmer (vgl. Jakobson 1972: 99).

##### 4.2. Die Rollenformantien

Jede finite Verbform enthält die Angabe eines Zustands-, Vorgangs- bzw. Tätigkeitsträgers. Die transitiven oder die Handlungsverben haben außerdem noch besondere Affixe für das Ziel der Handlung, die nicht entfernt werden können. Alle Verben zerfallen also in zwei voneinander scharf abgegrenzte Klassen: die einen nehmen grundsätzlich kein Handlungszielformans an; die anderen können es grundsätzlich nicht aufgeben und sind somit ständig transitiv. Daher sind jegliche Diathesentransformationen ausgeschlossen, und es gibt keine Passivkonstruktion. Die Formantien beziehen sich unmittelbar und unumkehrbar auf die konstanten semantischen Größen (Rollen – Agens und Patiens) und nicht auf die variablen syntaktischen (Satzglieder – grammatisches Subjekt und grammatisches Objekt). Das Prinzip der direkten Rollenbezeichnung ist ein überaus wichtiges sprachtypologisches Charakteristikum (vgl. Van Valin & Foley 1980).

Die **Rollenformantien** sind vieldeutig. Außer den beiden genannten Rollenbedeutungen drücken sie die Belebtheit bzw. Unbelebtheit der Prozeßteilnehmer aus; bei den belebten auch noch drei Personen und die Zahl (Singular und Plural); in der 3. Person Singular wird das Geschlecht (Maskulinum, Femininum und Neutrumb) unterschieden.

Sowohl für das Agens als auch für das Patiens werden im allgemeinen dieselben Formantien verwendet; ihre jeweilige Bedeutung hängt teils vom Verbtyp, teils von ihrer Position im Wort ab. Gewöhnlich werden im ganzen sieben bis acht Reihen von Rollenformantien gezählt (Krejnović 1968: 22–25; Verner 1974; Verner & Živova 1981). Sie können aber eigentlich auch in vier Reihen gruppiert werden, die sich trotzdem nicht in allen Formen unterscheiden (Tab. 128.2).

Da sich zwischen den einzelnen Reihen offensichtlich keine semantischen und sonstigen funktionalen Unterschiede feststellen lassen, müssen die entsprechenden Formantien als Allomorphe und ihre Wahl in jedem kon-

	1.	2.	3.	4.
<b>Singular</b>				
1. Person	<i>di-</i>	<i>di-</i>	<i>ba-</i>	<i>bɔ-</i>
2. Person	<i>ku-</i>	<i>ku-</i>	<i>ku-</i>	<i>ku-</i>
3. Person				
Maskulinum	<i>du-</i>	<i>a-</i>	<i>a-/bu-</i>	<i>o-/bu-</i>
Femininum	<i>da-/da-</i>	<i>i-</i>	<i>i-/bu-</i>	<i>u-/bu-</i>
Neutrum/unbelebt	<i>da-/da-/b-/m-</i>	<i>i-/b-/m-</i>	<i>i-/bu-/lØ-</i>	<i>u-/bu-</i>
<b>Plural</b>				
1. Person	<i>di-... -n</i>	<i>daŋ-</i>	<i>daŋ-</i>	<i>dɔŋ-</i>
2. Person	<i>ku-... -n</i>	<i>kaŋ-</i>	<i>kaŋ-</i>	<i>koŋ-</i>
3. Person				
belebt	<i>du-... -n</i>	<i>aŋ-</i>	<i>aŋ-/bu-</i>	<i>oŋ-/bu-</i>
unbelebt	<i>b-/m-</i>	<i>i-/b-/m-</i>	<i>i-/lØ-</i>	<i>u-/bu-</i>

Tab. 128.2: Rollenformantien

kreten Fall einfach als Gegebenheit betrachtet werden. Dabei ist diese Wahl durchaus nicht frei; einzelne Gruppen von Verben nehmen nur ganz bestimmte Formantien an.

#### 4.2.1. Die Agensformantien der intransitiven Verben

Die für die intransitiven Verben einzig mögliche Rolle des Agens (im weiten Sinne, d. h. sowohl eines Tätigkeits- als auch eines Zustandsträgers) kann durch die Formantien aller vier Reihen bezeichnet werden, von denen die Reihe 1 für die meisten Verben kennzeichnend ist. Die Formantien der einzelnen Reihen unterscheiden sich voneinander durch ihre Position in zusammengesetzten Verben, die zwei Wurzelmorpheme enthalten.

Nur in der Reihe 1 werden Belebtheit, Person und Geschlecht immer vereint und präfixal, die Zahl aber immer abgesondert durch das Pluralsuffix *-n* ausgedrückt. Vor das Pluralsuffix kann ein Sproßvokal treten; die Präfixe (außer *da-* im Femininum) werden dagegen oft auf den bloßen Konsonanten reduziert, und zwar wenn der Stamm mit einem Vokal beginnt oder mehrsilbig ist. Dadurch fallen die 1. und die 3. Person formal zusammen, was für viele Verben die Norm ist. Gegebenenfalls kann das Präfix durch die Assimilation völlig mit dem Stamm verschmelzen; phonetische Erscheinungen solcher Art seien hier aber nur am Rande erwähnt. Für die regelmäßigen Formen genügen folgende Beispiele:

- (1) (a) *di-γ-a-daa*  
1.SG-DET-PRÄS-leb  
'ich lebe (wohne)'

- (b) *du-tal*  
3.SG.M-frier  
'er friert'  
(c) *da-t-a-Rɔt*  
3.SG.F-DET-PRÄS-schlaf  
'sie schläft'  
(d) *dɔŋ-l-d-in*  
1/3-PRÄT-IPFV-leb-PL  
'wir/sie lebten'

Nur das Formans der Unbelebtheit *b/v/m* befindet sich innerhalb der Verbbasis:

- (2) (a) (*don*) *t-a-b-o-Rɔt*  
(Messer) DET-PRÄS-UNBELEBT-Ø-lieg  
'(das Messer) liegt'  
(b) (*tis*) *k-i-b-a-tin*  
(Stein) DET-Ø-UNBELEBT-VER.3.SG-(hinab)roll  
'(der Stein) rollt hinab'  
(c) (*qonoks*) *d-a-v-Ran*  
(Morgen) DET-PRÄS-UNBELEBT-beginn  
'(der Morgen) bricht an'

So unterscheiden sich z. B. (3 a) und (3 b), (4 a) und (4 b):

- (3) (a) *di-t-a-d-daq*  
1.SG-DET-PRÄS-VER.1.SG-(-(hin)fall  
'ich falle hin'  
(b) (*tis*) *t-a-b-daq*  
(Stein) DET-PRÄS-UNBELEBT-(-(hin)fall  
'(der Stein) fällt hinunter'  
(4) (a) (*oks*) *d-a-j-a-tij*  
(Baum) 3.SG-PRÄS-Ø-VER.3.SG-wachs  
(der Baum: belebt)  
'(der Baum) wächst'

- (b) (dan) *a-b-a-tij*  
 (Gras) PRÄS-UNBELEBT-VER.3.SG-wachs  
 (das Gras: unbelebt)  
 '(das Gras) wächst'

Die unzerlegbaren Formantien der Reihe 2 kommen nur in der Mitte einiger zusammengesetzter Verben vor:

- (5) (a) *sit>di<ja*  
 1.SG<aufwach>  
 'ich wache auf'  
 (b) *sit>a<ja*  
 3.SG.M<aufwach>  
 'er wacht auf'  
 (c) *sit>day<a*  
 1.PL<aufwach>  
 'wir wachen auf'

Die Reihe 3 unterscheidet sich von der Reihe 2 nur in der 1. Person Singular sowie in der Form für das unbelebte Agens, die ausnahmsweise keinen Rollenformantien hat. Diese Formantien verbinden sich sowohl mit einfachen als auch mit zusammengesetzten Stämmen und nehmen dementsprechend entweder die Vorder- (6) oder die Mittelstellung (7) ein (vgl. auch (8) ohne Rollenformans):

- (6) (a) *ba-γ-i-s-sal*  
 1.SG-DET-Ø-PRÄS-übernacht  
 'ich übernachte'  
 (b) *a-γ-i-s-sal*  
 3.SG.M-DET-Ø-PRÄS-übernacht  
 'er übernachtet'  
 (c) *day-k-i-s-sal*  
 1.PL-DET-Ø-PRÄS-übernacht  
 'wir übernachten'  
 (7) (a) *us-ba-t-a-Ran*  
 warm-1.SG-DET-PRÄS-beginn  
 'ich beginne, warm zu werden'  
 (b) *us-a-t-a-Ran*  
 warm-3.SG.M-DET-PRÄS-beginn  
 'er beginnt, warm zu werden'  
 (c) *us-day-t-a-Ran*  
 warm-1.PL-DET-PRÄS-beginn  
 'wir beginnen, warm zu werden'  
 (8) (a) (*in*) *il-t-aq*  
 (Nadel) klein-DET-werd  
 '(die Nadel) bricht'  
 (b) (*kansa*) *qon-t-aq*  
 (Rauchpfeife) verlier-DET-geh  
 '(die Rauchpfeife) kommt abhanden'  
 (c) (*ayj*) *halaj-t-aq*  
 (Strick) verwickel-DET-werd  
 '(der Strick) verwickelt sich'

- (d) *ob-ba-t-aq*  
 Vater-1.SG-DET-werd  
 'ich werde Vater'

Im Prinzip genauso verhalten sich die Formantien der Reihe 4:

- (9) (a) *bɔ-γ-a-tn*  
 1.SG-DET-PRÄS-geh  
 'ich gehe'  
 (b) *o-γ-a-tn*  
 3.SG.M-DET-PRÄS-geh  
 'er geht'  
 (c) *u-γ-a-tn*  
 3.SG.F-DET-PRÄS-geh  
 'sie geht'  
 (10) (a) *da-sulej-o-k-s-a*  
 3.SG.N-röt-3.SG.M-DET-PRÄS-werd  
 'er wird rot (wörtl. es rötet ihn)'  
 (b) *da-sulej-u-k-s-a*  
 3.SG.N-röt-3.SG.F-DET-PRÄS-werd  
 'sie wird rot (wörtl. es rötet sie)'  
 (c) *da-sulej-op-k-s-a*  
 3.SG.N-röt-3.PL-DET-PRÄS-werd  
 'sie werden rot (wörtl. es rötet sie)'

Das Formans der Unbelebtheit fällt in dieser Reihe mit dem Feminin-Formans zusammen; es kann aber die Zahl nicht unterscheiden und ist deshalb gesondert zu behandeln:

- (11) (al) *u-γ-a-aRan*  
 (Suppe) 3.SG.N-DET-PRÄS-koch  
 '(die Suppe) kocht'

#### 4.2.2. Die Agens- und die Patiensformantien der transitiven Verben

Die Rolle des Agens wird nur durch die Formantien der Reihe 1 ausgedrückt; zur Bezeichnung des Patiens dienen die drei anderen Reihen je nach einzelnen Verbgruppen. Das Patiensformans schließt sich entweder unmittelbar dem Agenspräfix an oder ist durch ein Determinativ, ein Bindeelement oder ein Wurzelmorphem von ihm getrennt, was aus folgenden Beispielen ersichtlich ist:

- (12) Reihe 2  
 (a) *da-us-q-i-ku-(i)t*  
 AG.3.SG.F-warm-KAUS-Ø-PAT.2.SG-R  
 (warm + KAUS + R = wärmt)  
 'sie wärmt dich'  
 (b) *k-us-q-i-day-it*  
 AG.2.SG-warm-KAUS-Ø-PAT.1.PL-R  
 (warm + KAUS + R = wärmt)  
 'du wärmt uns'  
 (c) (*don*)  
 (Messer)

- k-et-t-a-b-sin*  
AG.2.SG-scharf-DET-PRÄS.UNBELEBT-R  
(scharf + R = scharf.mach)  
'du machst es (das Messer) scharf'
- (13) Reihe 3
- (a) *da-ba-t-uŋ*  
AG.3.SG.F-PAT.1.SG-DET-seh  
'sie sieht mich'
  - (b) *d-i-t-uŋ*  
AG.1.SG/3.SG.M-PAT.3.SG.F-DET-seh  
'ich/er sehe/sieht sie'
  - (c) *d-anij-daj-s-i-bet*  
AG.1.SG/3.SG.M-denk-PAT.1.PL-PRÄS-Ø-mach  
'ich/er denke/denkt an uns'
- (14) Reihe 4
- (a) *d-daj-k-s-aq*  
AG.1.SG/3.SG.M-PAT.1.PL-DET-PRÄS-führ  
'ich/er führe/führt uns weg'
  - (b) *da-kaj-k-s-aq*  
AG.3.SG.F-PAT.2.PL-DET-PRÄS-führ  
'sie führt euch weg'
  - (c) *k-et-bɔ-k-s-u-Ro*  
AG.2.SG-gesund/lebendig-PAT.1.SG-DET-PRÄS-Ø-R (gesund/lebendig + R = gesund/lebendig.mach)  
'du machst mich gesund'

Die Reflexivformen der 1. und der 2. Person weisen keinerlei Besonderheiten auf. Sie entstehen innerhalb eines üblichen transitiven Paradigmas:

- (15) (a) *k-i-di-tan*  
AG.2.SG-Ø-PAT.1.SG-gürt  
'du gürtest mich'
- (b) *k-a-j-tan*  
AG.2.SG-PAT.3.SG.M-Ø-gürt  
'du gürtest ihn'
- (c) *k-i-j-tan*  
AG.2.SG-PAT.3.SG.F-Ø-gürt  
'du gürtest sie (fem.)'
- (d) *ku-yu-tan*  
AG.2.SG-PAT.2.SG-gürt  
'du gürtest dich'
- (16) (a) *di-γu-s*  
AG.1.SG-PAT.2.SG-anzieh  
'ich ziehe dich an'
- (b) *di-di-s*  
AG.1.SG-PAT.1.SG-anzieh  
'ich ziehe mich an'

In der 3. Person sind die entsprechenden Formen (in transitiven Paradigmen) zweideutig, vgl.:

- (17) (a) *d-a-i-s*  
AG.3.SG.M-PAT.3.SG.M-Ø-anzieh  
'er zieht ihn/sich an'
- (b) *da-j-i-j-s*  
AG.3.SG.F-Ø-PAT.3.SG.F-Ø-anzieh  
'sie zieht sie (fem.)/sich an'

Um diese Zweideutigkeit zu überwinden, gebraucht man zusätzlich im Satz das Objektpronomen *bu* oder das Pronomen *bin* 'selbst/sich selbst':

- (18) (a) *bu d-a-i-s*  
OBJ AG.3.SG.M-PAT.3.SG.M-Ø-anzieh  
'er zieht ihn an'
- (b) *bu bindu*  
OBJ selbst:M  
*d-a-i-s*  
AG.3.SG.M-PAT.3.SG.M-Ø-anzieh  
'er zieht sich selbst an'
- (c) *bu binda*  
OBJ selbst:F  
*da-j-i-j-s*  
AG.3.SG.F-Ø-PAT.3.SG.F-Ø-anzieh  
'sie zieht sich selbst an'

Die reflexive Bedeutung kann auch in den Formen der Subjektversion enthalten sein, vgl. (18 b) vs. (19 a) und (18 c) vs. (19 b). In (19) ist die Reflexivität mit der Benefaktivität durch das Versionsaffix *a-* ausgedrückt.

- (19) (a) *bu du-j-a-s*  
OBJ AG.3.SG.M-Ø-VER.3.SG-anzieh  
'er zieht sich an'
- (b) *bu da-j-a-s*  
OBJ AG.3.SG.F-Ø-VER.3.SG-anzieh  
'sie zieht sich an'

#### 4.2.3. Besondere Fälle

Viele einfache und zusammengesetzte Verben zeigen Stammveränderungen im Plural:

- (20) (a) *d-il<si>bet*  
1.SG/3.SG.M-<PRÄS>atm  
'ich/er atme/atmet (wörtl. ich/er Atem mache/macht)'
- (b) *d-il<si>γet-n*  
1/3-<PRÄS>atm-PL  
'wir/sie atmen'

Einige transitive und intransitive Verben haben das Element *-b/v-*, das sich wie das Formans der Unbelebtheit verhält, aber mit ihm nicht identisch ist, weil es mit keinen anderen Formantien wechselt (vgl. Kreinović 1968: 37). Dieses Affix kann in diesem Fall nur als Instrument betrachtet werden:

- (21) (a) *d-ba-t-i-b-es*  
 AG.1.SG/3.SG.M-PAT.1.SG-DET-Ø-  
 INSTR-zeichn  
 'er zeichnet mich (wörtlich: er zeich-  
 net mich mit etwas)'  
 (b) *o-g-b-un*  
 AG.3.SG.M-DET-INSTR-ausrutsch  
 'er rutscht aus (wörtlich: er rutscht  
 durch etwas (z. B. Glatteis) aus)'

Viele Verben haben gleichzeitig zwei Agensformantien, von denen das erste der Reihe 1 angehört, während das zweite die Bedeutung der Person und der Zahl wiederholt und sich nicht entfernen lässt. Der funktionale Zweck solcher Verdoppelung ist unklar. Beispiele sind:

- (22) (a) *d-h-a-di-tas*  
 AG.1.SG/3.SG.M-DET-PRÄS-VER.1.SG-auf-  
 steh  
 'ich stehe auf'  
 (b) *d-d-a-j-a-dun*  
 AG.1.SG/3.SG.M-DET-PRÄS-Ø-  
 VER.3.SG-schrei  
 'er schreit'  
 (c) *d-bu-to γ-in*  
 AG.1/3-VER.3.PL-erschreck-PL  
 'sie erschrecken'

Das Agens einiger Schallverben wird mit Hilfe anderer Formantien bezeichnet, deren eigentliche Funktion der Ausdruck der Possessivität ist (vgl. 5.2), z. B.:

- (23) (a) *ab-kutolij-b-a-ta*  
 POSS.1.SG-Pfiff-UNBELEBT-VER.3.SG-R  
 (Pfiff + R = pfeif)  
 'ich pfeife (wörtl. mein Pfeifen ist  
 zu hören)'  
 (b) *uk-kutolij-b-a-ta*  
 POSS.2.SG-Pfiff-UNBELEBT-VER.3.SG-R  
 (Pfiff + R = pfeif)  
 'du pfeifst'  
 (c) *(bu)da-kutolij-b-a-ta*  
 POSS.3.SG.M-Pfiff-UNBELEBT-  
 VER.3.SG-R (Pfiff + R = pfeif)  
 'er pfeift'  
 (d) *(bu)d-kutolij-b-a-ta*  
 POSS.3.SG.F-Pfiff-UNBELEBT-  
 VER.3.SG-R (Pfiff + R = pfeif)  
 'sie pfeift'

#### 4.3. Aspektaffixe

Die **Aspektaffixe** *n-* (Perfektivität) und *l-* (Imperfektivität) kommen nur in Vergangenheitsformen vor, und wenn das Tempusaffix

Ø- 'PRÄT' fehlt, wird das Tempus (Vergangenheit) durch die Aspektaffixe kenntlich gemacht. Diese haben allerdings zwei verschiedene Bedeutungen – eine temporale und eine modale –, je nachdem, ob das Agensformans vorhanden ist oder fehlt. Ihre Hinzufügung zum finiten Verb bezeichnet das Vorangehen des Prozesses einem bestimmten Zeitpunkt (der nicht unbedingt mit dem Redemoment zusammenfallen muß). Durch die gleichzeitige Eliminierung des Agensformans wird der Imperativ gebildet. Da sich die Willensäußerung und die Bedeutung der Vergangenheit gegenseitig ausschließen, können beide durch ein und dasselbe Formans ausgedrückt werden. Andere temporale und modale Bedeutungen werden außerhalb des Verbparadigmas durch Partikeln ausgedrückt.

In den meisten Fällen befindet sich das Aspektaffix unmittelbar vor dem Wurzelmorphem bzw. vor dem zweiten Wurzelmorphem eines zusammengesetzten Verbs; von zwei zusammenstoßenden Vokalen bleibt nur einer erhalten:

- (24) (a) *da-i-t-ɔ-l-Rot*  
 3.SG.F-Ø-DET-PRÄT-IPFV-schlaf  
 'sie schlief'  
 (b) *t-a-l-Rot*  
 DET-PRÄT-IPFV-schlaf  
 'schlafe!'  
 (c) *d-si>t-ɔ-l<γit*  
 3.SG.M-DET-PRÄT-IPFV<fege>  
 'er fegte auf'  
 (d) *si>t-a-l<γit*  
 DET-PRÄS-IPFV<fege>  
 'fege auf!'  
 (e) *sit>day-n<a*  
 1.PL-IPFV<aufwach>  
 'wir wachten auf'  
 (f) *da-ba-t-ɔ-l-uŋ*  
 AG.3.SG.F-PAT.1.SG-DET-PRÄT-IPFV-  
 seh  
 'sie sah mich'

In der dritten Person betrifft es nur die Reflexivformen, während sonst die übliche Morphemfolge herrscht:

- (25) (a) *di-l-a-s*  
 AG.3.SG.M-IPFV-VER.3.SG.M-anzieh  
 'er zog sich an'  
 (b) *d-a-ɔ-l-s*  
 AG.3.SG.M-PAT.3.SG.M-PRÄT-IPFV-an-  
 zieh  
 'er zog ihn an' (normalerweise *d-ɔ-l-s*)

Die Basis vieler Verben unterliegt Modifikationen verschiedenen Grades bis zu Suppletivbildungen, z. B.:

- (26) (a) *t-a-b-Röt*  
DET-PRÄS-UNBELEBT-lieg  
'es liegt'
- (b) *t-ɔ-b-i-l-Röt*  
DET-PRÄT-UNBELEBT-Ø-IPFV-lieg  
'es lag'
- (c) *ba-γ-i-s-sal*  
1.SG-DET-Ø-PRÄS-übernacht  
'ich übernachte'
- (d) *ba-γ-i-n-sal*  
1.SG-DET-Ø-PFV-übernacht  
'ich übernachtete'
- (e) *bɔ-γ-a-tn*  
1.SG-DET-PRÄS-geh  
'ich gehe'
- (f) *bɔ-γ-ɔ-n-Ø*  
1.SG-DET-PRÄT-PFV-geh  
'ich ging'

In der Vergangenheitsform des Verbs *-tn* 'geh' in (26 f) verschwindet das Wurzelmorphem.

Auffallend ist das obligatorische Weglassen des Elements {b/v/m} im Imperativ, sei es das Formans der Reihe 2 für das unbelebte Patiens der transitiven Verben oder der formale Bestandteil des Verbstammes (s. 27 a aber 27 b; vgl. 4.2.3):

- (27) (a) *d-b-i-n-dit*  
AG.1.SG/3.SG.M-PAT.UNBELEBT-Ø-PFV-aufzähl  
'ich/er zählte es auf'
- (b) *i-n-dit*  
Ø-PFV-aufzähl  
'zähle (es) auf!'

## 5. Das Substantiv

Im Vergleich zum Verb hat das ketische Substantiv eine bedeutend einfachere Wortstruktur und kann affixlos auftreten. (Nach anderen Analysen ist das ketische Substantiv mit 11 Kasusaffixen versehen, vgl. Werner 1995; 1997.) Sein Stamm besteht aus einem oder zwei Wurzelmorphemen, von denen die meisten einsilbig sind:

- (28) *bok* 'Feuer'  
*tis* 'Stein'  
*qus* 'Zelt/Haus'  
*boktis* 'Feuerstein'  
*tisqus* 'Steinhaus'

### 5.1. Die Kategorie des Numerus

Das einzige gebundene grammatische Morphem des Substantivs ist das **Pluralsuffix**, das in den Allomorphen *-n* und *-ŋ* (eventuell mit einem vorangehenden Sproßvokal) auftritt. Deren Wahl ist teils phonetisch, teils semantisch bedingt (Dul'zon 1968: 68 f.; Porotova 1968: 14–16):

- (29) (a) *ogdə-n* 'Ohr-PL'  
*bolva-n* 'Pilz-PL'  
*bes-n* 'Hase-PL'  
*qim-n* 'Frau-PL'
- (b) *am-aŋ* 'Mutter-PL'  
*ob-aŋ* 'Vater-PL'  
*hun-aŋ* 'Tochter-PL'  
*aj-aŋ* 'Sack-PL'

Einige Dutzend Wörter haben unregelmäßige Fusions- und Suppletivformen:

- (30) (a) *qoj* 'Bär'; *qon* 'Bär.PL'
- (b) *qus* 'Haus'; *quŋ* 'Haus.PL'
- (c) *tib* 'Hund'; *tab* 'Hund.PL'
- (d) *ses* 'Fluß'; *sas* 'Fluß.PL'
- (e) *ket* 'Mensch'; *deŋ* 'Mensch.PL'

Eine Reihe von Substantiven unterscheidet keine Numerusformen. Es sind vorwiegend Stoff- und Sammelnamen, Abstrakta und Unikalien, z. B.:

- (31) *ul* 'Wasser'  
*dik* 'Teer'  
*to* 'Salz'  
*sul* 'Blut'  
*dan* 'Gras'  
*dk* 'Leben'  
*ulej* 'Nebel'  
*ulas* 'Regen'  
*bejas* 'Wind'

Dazu gehören aber auch die Bezeichnungen einiger Lebewesen:

- (32) *bən* 'Ente'  
*is* 'Fisch'

### 5.2. Die latenten Kategorien: die Belebtheit und das Geschlecht

Außer der zweigliedrigen Kategorie des Numerus hat das ketische Substantiv die Bedeutung der Belebtheit bzw. der Unbelebtheit mit Unterteilung der Belebtheit in das männliche und das weibliche Geschlecht. Diese Bedeutungen, die den benachbarten uralischen und altaischen Sprachen übrigens völlig fremd sind, finden ihren Ausdruck aber nicht im Substantiv selbst, sondern in anderen Wörtern, die mit ihm grammatisch verbun-

den sind. Zum einen ist dies das finite Verb, dessen Rollenformantien in der 3. Person mit dem Substantiv kongruieren. Zum anderen stellen die nominalen klitischen Partikeln *{di}*, *{da}*, *{na}* eine formale sowie auch eine inhaltliche Verbindung zwischen zwei Substantiven oder zwischen einem Substantiv und einer Postposition her. Ihr Gebrauch richtet sich nach den genannten latenten Bedeutungen des ihnen vorangehenden Substantivs und hängt von seinem Numerus ab. Nur mit dem Singular tritt die Partikel *{da}* auf, die das männliche Geschlecht ausdrückt; nur mit dem Plural tritt die Partikel *{na}* auf, die die Belebtheit bezeichnet; die Partikel *{d(i)}* begleitet beide Numerusformen und hat jeweils die negative Bedeutung: das nichtmännliche Geschlecht im Singular und die Unbelebtheit im Plural. Die Wahl der Partikeln wird also immer vom ersten der zwei Substantive bestimmt:

## (33) Singular

- (a) *hiy da bay*  
Mann KONN.SG.M Platz  
'des Mannes Platz'
- (b) *qim di bay*  
Frau KONN.SG.NM/PL.UNBELEBT Platz  
'der Frau Platz'
- (c) *qus di bay*  
Zelt KONN.SG.NM/PL.UNBELEBT Platz  
'des Zeltes Platz'

## Plural

- (d) *hon na bay*  
Mann.PL KONN.PL.BELEBT Platz  
'der Männer Platz'
- (e) *qim-n na bay*  
Frau-PL KONN.PL.BELEBT Platz  
'der Frauen Platz'
- (f) *quj di*  
Zelt:PL KONN.SG.NM/PL.UNBELEBT  
*bay*  
Platz  
'der Zelte Platz'

Genauso verhält es sich beim Gebrauch der meisten Postpositionen, die der Raum-, Richtungs- und Zeitangabe dienen oder Ziel, Ursache, Folge und ähnliche Bedeutungen ausdrücken. Etymologisch gehen sie oft auf Substantive wie *qo* 'Mund', *oyot* 'Rücken', *kup* 'Schnabel', *ugde* 'Länge' usw. zurück; z. B.:

- (34) (a) *sel da oyot*  
Ren KONN.SG.M Rücken  
'wir setzen uns auf das Ren'

- (b) *sen na oyot*  
Ren:PL KONN.PL.BELEBT Rücken  
'wir setzen uns auf die Rentiere'

Mit den drei gebräuchlichsten Postpositionen bilden die Partikeln unzerlegbare Einheiten – *{diga}*, *{danya}*, *{naya}* mit der Grundbedeutung 'hin'; *{digal}*, *{dapat}*, *{nayal}* oder *{dil}*, *{dal}*, *{nal}* mit der Grundbedeutung 'her'; *{digt(ən)}*, *{dant(ən)}*, *{najt(ən)}* mit der Grundbedeutung 'da':

- (35) (a) *bok dija aq*  
Feuer KONN.SG.NM:hin Holz  
*d-obilda*  
3.SG.M-leg  
'er legte ein Holzscheit ins Feuer'
- (b) *ad quj digal al'a*  
1.SG Zelt KONN.SG.NM:her hinaus  
*bɔ-γ-ɔt*  
1.SG-DET-geh  
'ich gehe aus dem Zelt hinaus'
- (c) *bu du-yodaq bayus*  
3.SG.M 3.SG.M-wohn Erdhütte  
*digitən*  
KONN.SG.NM:da  
'er wohnt in einer Erdhütte'

Dagegen ist die linksseitige Verbindung der Partikeln mit dem vorangehenden Substantiv niemals so fest; durchaus möglich ist z. B. (36 a) neben (36 b):

- (36) (a) *kida bay ultayin*  
hier Erde klein.Sumpf  
*dintə* *onay*  
KONN.SG.NM:da viel.sein
- (b) *kida bay digitə*  
hier Erde KONN.SG.NM:da  
*ultayin* *onay*  
klein.Sumpf viel.sein  
'hier gibt es viele Sumpfe'

Das Substantiv, auf das sich die Partikel bezieht, kann weggelassen werden:

- (37) (a) *ob da don*  
Vater KONN.SG.M Messer  
'Vater sein Messer'/Vaters Messer'
- (b) *da don*  
KONN.SG.M Messer  
'sein Messer'
- (c) *hib dayt*  
Sohn KONN.SG.M:da  
*d-anit-i-l-e-vet*  
AG.1.SG/3.SG.M-denkt-Ø-IPFV-Ø-mach  
'ich/er dachte an seinen Sohn'

- (d) *dajt*  
 KONN.SG.M:da  
*d-anij-i-l-e-vet*  
 AG.1.SG/3.SG.M-denk-Ø-IPFV-  
 Ø-mach  
 'ich'er dachte an ihn'

Offensichtlich hat das Substantiv keine eigenen Ausdrucksmittel für die Belebtheit und das Genus; dennoch handelt es sich zweifellos um nominale Kategorien, die berücksichtigt werden müssen.

Zu den belebten Substantiven gehören die Bezeichnungen aller Lebewesen, zu denen auch die Bäume gezählt werden:

- (38) *qaj* 'Elch'  
*tit* 'Mücke'  
*qoj* 'Bär'  
*il* 'Espe'  
*qud* 'Hecht'  
*us* 'Birke'  
*kil* 'Krähe'  
*din* 'Tanne' usw.

Als belebt gelten außerdem:

- (39) *i* 'Sonne'  
*ses* 'Fluß'  
*baŋ* 'Erde'  
*bok* 'Feuer'  
*eqŋ* 'Donner'  
*hu* 'Herz'

sowie die Benennungen einzelner wichtiger Teile des Rentierschlittens, für die es keine europäischen Entsprechungen gibt.

Die Grenze zwischen den belebten und den unbelebten Substantiven ist im ganzen ziemlich stabil. Von den wenigen Ausnahmen sind vor allem die Wörter {*is*} 'Fisch' und {*oks*} 'Baum' zu nennen, die beiden Klassen angehören, je nachdem ob es sich um einen lebendigen oder einen toten Fisch handelt bzw. um einen wachsenden oder einen gefällten Baum.

Dagegen ist die Grenze zwischen den Genusklassen der belebten Substantive sehr beweglich. Nur wenige Wörter haben ständig dasselbe Geschlecht; immer männlich sind z. B.:

- (40) *ket* 'Mensch'  
*eqŋ* 'Donner'  
*tet* 'Ehemann'  
*sul* 'Schlitten'  
*op* 'Vater'  
*ulas* 'Regen'

Immer weiblich sind (Dul'zon 1968: 63–69):

- (41) *qim* 'Frau'  
*i* 'Sonne'  
*am* 'Mutter'  
*baŋ* 'Erde'  
*hun* 'Tochter'  
*qo* 'Stern'  
*hu* 'Herz'  
*bok* 'Feuer'

Das eine oder das andere Genus überwiegt auch bei generalisierendem Gebrauch einzelner Tiernamen als Gattungsbezeichnungen (Dul'zon 1968: 63–69). Die meisten belebten Substantive wechseln aber ihr Genus frei:

- (42) (a) *dil biseb*  
 Junge Geschwister/Bruder  
*d-ket>a-γ-i-s<qut*  
 AG.3.SG.M-PAT.3.SG.M-DET-Ø-PRÄS  
 <beleidig>  
 'der Junge beleidigt seinen Bruder'  
 (b) *dil biseb*  
 Junge Geschwister/Schwester  
*d-ket>i-γ-i-s<qut*  
 AG.3.SG.M-PAT.3.SG.F-DET-Ø-PRÄS  
 <beleidig>  
 'der Junge beleidigt seine Schwester'

## 6. Das Pronomen

Den Kern dieser Wortart bilden die **Personalpronomen**. Im Unterschied zu den meisten indoeuropäischen Sprachen beziehen sie sich nur auf Lebewesen und können keine Dingbezeichnungen ersetzen.

Die ketischen Personalpronomen haben im ganzen zwei Verwendungsbereiche. Vor einem finiten Verb oder im selbständigen Gebrauch unterscheiden sie nur die Person und den Numerus:

	Singular	Plural
1. Person	<i>ad</i>	<i>ətn</i>
2. Person	<i>uk</i>	<i>əkŋ</i>
3. Person	<i>bu</i>	<i>buŋ</i>

Tab. 128.3: Personalpronomen

In adverbaler Stellung wiederholen sie die entsprechenden Bedeutungen der Rollenformantien und beseitigen die häufige Homonymie:

- (43) (a) *ad d-is>k-a<Rɔt*  
 1.SG 1.SG/3.SG.M-DET-PRÄS<dös>  
 'ich döse (wörtl. ich liege (im Schlummer))'

- (b) *bu d-is>k-a<Rot*  
 3.SG 3.SG.M-DET-PRÄS<dös>  
 ‘er döst’

Vor einem Substantiv oder einer Postposition werden die Pronomen (mit Ausnahme der Suppletivform der 1. Person und der unveränderten Form der 2. Person Singular) von den bereits erwähnten unbetonten Partikeln {di}, {da}, {na} begleitet, die in der 3. Person auch selbstständig auftreten können und im Singular das Genus unterscheiden:

	Singular	Plural
1. Person	(a) <i>b</i>	<i>ətn na</i>
2. Person	(u) <i>k</i>	<i>əkn na</i>
3. Person	mask. fem.	( <i>bu</i> ) <i>da</i> ( <i>bu</i> ) <i>di</i>
		( <i>bu</i> ) <i>na</i>

Tab. 128.4: Personalpronomen

In Verbindung mit Substantiven drücken sie die persönliche Zugehörigkeit aus; irgendwelche besonderen Possessivpronomen gibt es nicht:

- (44) (a) (*a*)*b op*  
 1.SG Vater  
 ‘mein Vater’  
 (b) (*u*)*k am*  
 2.SG Mutter  
 ‘deine Mutter’  
 (c) (*bu*)*da don*  
 3.SG.M Messer  
 ‘sein Messer’

Diese Formen werden auch mit den meisten Postpositionen gebraucht:

- (45) (a) *ab kupka*  
 1.SG vor  
 ‘vor mir’  
 (b) (*bu*) *dija*  
 (3.SG) KONN.SG.NM:hin  
 ‘ihr’  
 (c) (*bu*) *da konət*  
 (3.SG) KONN.SG.M zu  
 ‘zu ihm’  
 (d) *ətn na balga*  
 1.PL KONN.PL.BELEBT zwischen  
 ‘zwischen uns’

Außerdem werden in das Pronominalparadigma auch die verbalen Rollenformantien der Reihe 1 einbezogen, die mit dem Morphem {bin} emphatische Formen bilden:

- (46) (a) *ad bin-di sul*  
 1.SG selbst-1.SG Schlitten  
*d-b-i-l-e-vet*  
 1.SG/3.SG.M-UNBELEBT-Ø-IPFV-Ø-mach  
 ‘den Schlitten habe ich selbst gemacht’

- (b) *bu bin-du da*  
 3.SG selbst-3.SG.M KONN.SG.M  
*don d-kai>n<em*  
 Messer 3.SG.M-PFV<nehm>  
 ‘er nahm sein eigenes Messer’

Der für das Interrogativpronomen relevante Unterschied ‘belebt/unbelebt’ wird mit Hilfe von Suppletivstämmen ausgedrückt: {ana} oder {bisə} ‘wer?’; {akus} ‘was?’, z. B.:

- (47) (a) *bisə k-a-t-uŋ ?*  
 wer AG.2.SG-PAT.3.SG.M-DET-seh  
 ‘wen siehst du?’  
 (b) *akus ku-t-uŋ ?*  
 was AG.2.SG-DET-seh  
 ‘was siehst du?’

Sowohl die Demonstrativ- als auch das Interrogativpronomen treten mit Postpositionen auf:

- (48) (a) *bu kit-daya bən*  
 3.SG diesen-auf NEG  
*d-bu-ŋ-l-ɔ-Rot*  
 AG.3.SG.M-VER.3.SG-DET-IPFV-Ø-seh  
 ‘er sah nicht auf diesen’  
 (b) *bu kida-dija bən*  
 3.SG diese-auf NEG  
*d-bu-ŋ-l-ɔ-Rot*  
 AG.3.SG.M-VER.3.SG-DET-IPFV-Ø-seh  
 ‘er sah nicht auf diese’  
 (c) *ana dagtə*  
 wer bei  
*ku-γ-i-n-sal*  
 2.SG-DET-Ø-PFV-übernacht  
 ‘bei wem hast du übernachtet?’  
 (d) *akus diŋte ta?*  
 was in Salz  
*t-a-b-o-Rot*  
 DET-PRÄS-UNBELEBT-Ø-lieg  
 ‘worin liegt das Salz?’

## 7. Andere Wortarten

Alle übrigen Wörter kennen keine Formenbildung und lassen sich nur aufgrund ihrer Semantik und zum Teil aufgrund ihrer Distribution in Adjektive, Adverbien, Modalpartikeln und Numeralien einteilen. Zwar schließen sich die Rollenformantien der Reihe 2

z. B. auch an Adjektive an, doch handelt es sich dabei eigentlich um keine wortmorphologische Erscheinung, sondern um die Form des nominalen Prädikats, die auch den Substantiven und Pronomen eigen ist, jedoch nur, wenn diese von Postpositionen begleitet werden. Also:

- (49) (a) (*ad*) *oŋ-di*  
           (1.SG) gesund-1.SG  
           ‘ich bin gesund’
- (b) *ətn na op*  
           1.PL KONN.PL.BELEBT Vater  
           *qus-ka-du*  
           Zelt-LOK-3.SG.M  
           ‘unser Vater ist im Zelt’
- (c) (*bu*) *aslin-ka-da*  
           (3.SG) Boot-LOK-3.SG.F  
           ‘sie ist im Boot’

## 8. Zur Wortbildung

Die Bildung der Substantive beschränkt sich im wesentlichen auf die Zusammensetzung von Wurzelmorphemen, die von der Fusion begleitet werden kann (Dul'zon 1968: 91):

- (50) (a) *bul* ‘Fuß’ + *taq* ‘Finger’: *bultak* ‘Zehe’  
      (b) *bul* ‘Bein’ + *ol* ‘Hülle’: *bulol* ‘Strumpf’  
      (c) *tak* ‘Finger’ + *ol* ‘Hülle’: *tayol* ‘Fingerring’  
      (d) *bok* ‘Feuer’ + *ul* ‘Wasser’: *boγul* ‘Schnaps’  
      (e) *baŋ* ‘Erde’ + *qus* ‘Zelt’: *baŋus* ‘Erdhütte’

Unter den im ganzen schlecht ausgeprägten Adjektiven und Adverbien gibt es dennoch eine Anzahl von Ableitungen mit den produktiven Suffixen *{tu}* und *{an}* mit gegenseitig antonymischer Bedeutung:

- (51) (a) *anuŋ*      ‘Verstand’  
           *anuŋtu*     ‘klug’  
           *anuŋan*     ‘dumm’
- (b) *kul*          ‘Bart’  
           *kultu*        ‘bartig’  
           *kulan*        ‘bartlos’
- (c) *qim*          ‘Frau’  
           *qimtu*        ‘verheiratet’  
           *qiman*        ‘ledig’

Die Wortbildung der ketischen Verben ist äußerst kompliziert und lässt sich nicht kurz zusammenfassen. Viele Verbalstämme enthalten Elemente, deren Bedeutung und Funktion

nur ungefähr oder gar nicht zu bestimmen sind. Mit einem und demselben Wurzelmorphem *{aq}* werden z. B. folgende Verben gebildet:

- (52) (a) *du-j-aq*  
           AG.3.SG.M-Ø-komm  
           ‘er kommt (heraus)’
- (b) *d-t-i-j-aq*  
           AG.3.SG.M-nach-PAT.3.SG.F-Ø-komm  
           ‘er kommt ihr nach’
- (c) *d-suŋa-k-s-aq*  
           3.SG.M-zurück-DET-PRÄS-komm  
           ‘er kommt zurück’
- (d) *d-ɬ'a-k-s-aq*  
           3.SG.M-hinaus-DET-PRÄS-komm  
           ‘er geht mal hinaus (und kehrt wieder zurück)’
- (e) *d-si>j<aq*  
           3.SG.M-Ø<frag.nach>  
           ‘er bittet (wörtl. er fragt nach)’

Die Bedeutung der jeweiligen ersten Wurzel der zusammengesetzten Verben ist klar: *{suŋa}* ‘zurück’, *{ɬ'a}* ‘hinaus’, *{si}* ‘bitten’, ‘Bitte’. Aber die Funktion der sogenannten Determinative *{t}* und *{ks}* bleibt rätselhaft, um so mehr, als einige von ihnen in der Vergangenheitsform regelmäßig wegfallen. So heißt es z. B.:

- (53) *d-ɬ'a-l-aq*  
           3.SG.M-hinaus-IPFV-komm  
           ‘er ging mal hinaus (und kehrte wieder zurück)’

## 9. Illustrativer Text

- (54) *ab daʔq*  
           1.SG Leben
- (55) *ad baqta diŋtan ket*  
           1.SG Bachta KONN.SG.NM:da Mensch(SG)
- (56) *ad uj diŋtan*  
           1.SG Wiege(SG) KONN.SG.NM:da  
           *di-t-a-Rɔt, ab op*  
           1.SG-DET-PRÄS-lieg 1.SG Vater(SG)  
           *əŋa di-n-bes*  
           hierher 1.SG-PFV-komm  
           *qɔl dija*  
           steinig.Tunguska KONN.SG.NM:hin
- (57) *tamanun siq-n ab op siŋ*  
           einige Jahr-PL 1.SG Vater(SG) hier  
           *d-ɔ-l-daq*  
           1.SG/3.SG.M-PRÄT-IPFV-leb

- (58) *qajye tiya*  
dann flußabwärts  
*o-γ-ɔ-n-Ø*  
3.SG.M-DET-PRÄT-PFV-geh  
*dija, hanin ad*  
KONN.SG.NM:hin klein 1.SG  
*d-en>t<o*  
1.SG/3.SG.M-DET<erinner>
- (59) *ad ob-aŋ as*  
1.SG Vater-PL zusammen  
*bɔ-γ-ɔ-n-Ø, tiya*  
1.SG-DET-PRÄT-PFV-geh flußabwärts  
*dajg-ɔ-n-Ø*  
1.PL-DET-PRÄT-PFV-geh
- (60) *til øtn d-ɔ-l-d-in*  
dann 1.PL 1.PL-PRÄT-IPFV-leb-PL  
*ob-aŋ as*  
Vater-PL zusammen
- (61) *ab op du-n-o,*  
1.SG Vater(SG) 3.SG.M-PFV-sterb  
*am dɔ-n-o*  
Mutter(SG) 3.SG.F-PFV-sterb
- (62) *ab op da bisep*  
1.SG Vater KONN:SG.M Bruder(SG)  
*d-kas>di-n<os,*  
AG.1.SG/3.SG.M-PAT.1.SG-PFV<nehm>  
*ad di-l-di-tos*  
1.SG AG.3.SG.M-IPFV-PAT.1.SG-großzieh
- (63) *ab op da*  
1.SG.POSS Vater(SG) KONN.3.SG.M  
*bisep tamanun siq-ŋ*  
Bruder(SG) einige Jahr-PL  
*d-ɔ-l-daq, qajye øya*  
1.SG/3.SG.M-PRÄT-IPFV-leb dann hierher  
*di-n-bes qɔ?l*  
1.SG-PFV-komm steinig.Tunguska  
*dija*  
KONN.SG.NM:hin
- (64) *ab øyten ob-i-l-da*  
1.SG bei PRÄT-UNBELEBT-IPFV-sein  
*enayo aŋam am qo siq-ŋ*  
ungefähr sechs noch zehn Jahr-PL
- (65) *ad d-tet-n-am,*  
1.SG 1.SG/3.SG.M-Ehemann-PFV-R  
(Ehemann + R = heirat)  
*dej naya*  
Leute(PL) KONN.PL.BELEBT:hin  
*bɔ-γ-ɔ-n-Ø*  
1.SG-DET-PRÄT-PFV-geh
- (66) *ab op da bisep*  
1.SG Vater(SG) KONN.3.SG.M Bruder(SG)  
*haj tiya*  
wieder flußabwärts
- ɔ-γ-ɔ-n-Ø, baqta*  
3.SG.M-DET-PRÄT-PFV-geh Bachta  
*dijtən du-n-o*  
KONN.SG.NM:da 3.SG.M-PFV-sterb
- (67) *ab øytən a dep*  
1.SG bei sechs Leute(PL)  
*ɔb-i-l-da, qan*  
PRÄT-UNBELEBT-IPFV-sein fünf  
*du-n-o-n, hanin se qɔk*  
3.PL-PFV-sterb-PL klein SUP eins  
*da-t-ɔ-n-a-t*  
3.SG.F-DET-PRÄT-PFV-VER.3.SG-bleib
- (68) *ab tet armija*  
1.SG Mann(SG) Armee(SG)  
*dijtən tambiles*  
KONN.SG.NM:da irgendwo  
*ɔ-γ-ɔ-n-Ø*  
3.SG.M-DET-PRÄT-PFV-geh
- (69) *ab hun*  
1.SG Tochter(SG)  
*da-qɑ-ɔ-yon,*  
3.SG.F-groß-PRÄT-werd,  
*da-tet-n-am,*  
3.SG.F-Ehemann-PFV-R (Ehemann + R  
= heirat)  
*dej naya*  
Leute(PL) KONN.PL.BELEBT:hin  
*u-γ-ɔ-n-Ø*  
3.SG.F-DET-PRÄT-PFV-geh
- (70) *ad qɔk ket*  
1.SG ein Mensch  
*d-igd-ɔ-Ron*  
1.SG/3.SG.M-Verbleib-PRÄT-werd
- (71) *qɔk ket ad bən in*  
ein Mensch 1.SG nicht lang  
*d-ɔ-l-daq, qajye ab*  
1.SG/3.SG.M-PRÄT-IPFV-leb dann 1.SG  
*hun*  
Tochter(SG)  
*da-kas>di-n<qos*  
AG.3.SG.F-PAT.1.SG-PFV<nehm>
- (72) *ket bada: bam qan*  
Mensch(SG) 3.SG.M.sag Greisin IMP  
*da-ik>s-i<bes*  
3.SG.F-PRÄS-Ø<komm>
- (73) *en ad buŋ as*  
jetzt 1.SG 3.PL zusammen  
*di-γ-a-daq*  
1.SG-DET-PRÄS-leb
- “(54) Mein Leben. (55) Ich bin ein Mensch aus Bachta. (56) Ich lag in der Wiege, (als) mein Vater hierher an die steinige Tunguska kam. (57) Einige Jahre lebte mein Vater hier.

(58) Dann ging er flußabwärts hin (nach Bachta), ich erinnere mich ein wenig. (59) Ich ging mit den Eltern zusammen, wir (alle) gingen da flußabwärts. (60) Dann lebten wir zusammen mit den Eltern (d. h. Dann lebte ich bei meinen Eltern). (61) Mein Vater starb, (meine) Mutter starb. (62) Der Bruder meines Vaters nahm mich (auf), zog mich groß. (63) Der Bruder meines Vaters lebte (dort) einige Jahre, dann kam er hierher an die steinige Tunguska. (64) Mir war ungefähr sechzehn Jahre (d. h. Ich war etwa sechzehn Jahre alt). (65) Ich heiratete, ging zu den Leuten (meines Mannes). (66) Der Bruder meines Vaters ging wieder den Fluß abwärts, starb in Bachta. (67) Bei mir waren (d. h. ich hatte) sechs Kinder, fünf starben, nur die Jüngste blieb (am Leben). (68) Mein Mann ging irgendwo in der Armee (verloren). (69) Meine Tochter wurde groß, heiratete, ging zu den Leuten (des Mannes). (70) Ich blieb allein. (71) Allein lebte ich nicht lange, da nahm mich meine Tochter (auf). (72) (Ihr) Mann sagte: ‘Die Alte soll (zu uns) kommen!’ (73) Jetzt lebe ich bei ihnen.’ (Uljana K. Kamenskikh; aufgeschrieben von M. N. Wall, Siedlung Sulomai, 1967)

## 10. Unübliche Abkürzungen

DET	Determinativ (stammbildendes Element in Verbalformen)
R	Wurzelement des verbalen Stammrahmens
VER	Versionsaffix
NM	nicht-männlich

## 11. Zitierte Literatur

- Castren, Mathias A. (1858), *Versuch einer jenissei-ostjakischen und kottischen Sprachlehre*, hrsg. von A. Schiefner. St. Petersburg
- Donner, Kai (1955), *Ketica: Materialien aus dem Ketischen oder Jenissei-ostjakischen*. Helsinki: Suomalais-Ugrilainen Seura

Dul'zon, Andrei P. (1968), *Ketskij jazyk*. Tomsk: Izdatel'stvo Tomskogo Universiteta

Jakobson, Roman O. (1972), “Šiftery, glagol'nye kategorii i russkij glagol”. In: Uspenskij, Boris A. (Hrsg.), *Principy tipologičeskogo analiza jazykov različnogo stroja*. Moskva: Nauka, 95–113 [orig. englisch 1957]

Karger, Nikolaj K. (1934), “Ketskij jazyk”. In: *Jazyki i pismennost' narodov Severa III*. Moskau, Leningrad: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo, 223–238

Krejnović, Eruchim A. (1968), *Glagol ketskogo jazyka*. Leningrad: Nauka

Plotkin, Vul'f Ja. (1989), *Stroj anglijskogo jazyka*. Moskva: Vyššaja škola

Porotova, Tel'mina I. (1968), *Obrazovanie form čisla suščestvitel'nych ketskogo jazyka*. Dissertation, Tomsker Universität

Vall, Margarita N. & Kanakin, Igor A. (1990), *Očerk fonologii i grammatiki ketskogo jazyka*. Novosibirsk: Nauka

Van Valin, Robert D. & Foley, William A. (1980), “Role and Reference Grammar”. In: Moravcsik, Edith & Wirth, Jessica R. (Hrsg.), *Current Approaches to Syntax*. New York: Academic Press (Syntax and Semantics 13), 329–352

Verner, Genrich K. [Werner, Heinrich] (1974), “Reliktovye priznaki aktivnogo stroja v ketskom jazyke”. *Voprosy jazykoznanija* 1, 34–45

Verner, Genrich K. [Werner, Heinrich] & Živova, Galina T. (1981), “K charakteristike klassnoj sistemy v enissejskikh jazykach”. *Voprosy jazykoznanija* 5, 31–38

Werner, Heinrich (1995), *Zur Typologie der Jenissej-Sprachen*. Wiesbaden: Harrassowitz Verlag (Veröffentlichungen der Societas Uralo-Altaica, 45)

Werner, Heinrich (1997), *Die ketische Sprache*. Wiesbaden: Harrassowitz Verlag (Tunguso-Sibiriaca, 3)

Margarita Wall, Wolfsburg (Deutschland)

Igor Kanakin, Nowosibirsk (Russland)

Beispielanalyse:  
Heinrich Werner, Bonn (Deutschland)

## 129. West Greenlandic (Eskimo)

1. The language and its speakers
2. Major categories and processes
3. Morphophonology
4. Nominal inflection
5. Verbal inflection
6. Derivational morphology
7. Enclitics
8. Illustrative text
9. Uncommon abbreviations
10. References

### 1. The language and its speakers

West Greenlandic is a variety of the Inuit branch of the Eskimo languages that forms a rough continuum from North Alaska across the Canadian Arctic to the east of Greenland. Having been brought to Greenland in the rapid ‘Thule Eskimo’ expansion about a thousand years ago from northern Alaska, there is very limited mutual comprehension with adjacent Canadian dialects. The time depth to the Yupik Eskimo languages spoken in southwest Alaska and on Chukotka may be about double this, and the relationship beyond that to Aleut considerably more remote. It is spoken today by virtually all the 45,000 or so Greenlanders living on the west coast of Greenland. There are a further 3,000 who speak East Greenlandic and 750 who speak Polar Eskimo, highly divergent dialects. Of the approximately 9,000 Danes resident in Greenland very few speak Greenlandic (though it should be noted that the West Greenlanders themselves are ethnically mixed with Scandinavians and their writers are all bilingual).

The present description is of the central dialect around the capital Nuuk (Godthåb), on which the written language in use throughout the country is based (it is also the language of radio, church, and school instruction). The numbers of speakers are actually increasing and the language appears to face no threat of extinction in the near or middle future. It has a literature that includes original verse and novels going back over a century to the establishment of the illustrated journal *Atuagagdliutit* in 1861 (still earlier as regards ecclesiastical material and the first traditional legends gathered). It is, moreover, undoubtedly the earliest – and one of the best – described native North American language, with descriptions dating back to the first mis-

sionary times (e.g. Egede 1750). An important milestone in the grammatical and lexical description of the language was the grammar and dictionary of Samuel Kleinschmidt (1851 and 1871 respectively). These latter works were linguistically advanced for their time and laid down the morphophonemically transparent orthographical conventions that are still used alongside the new (more phonemic) orthography introduced in 1973, in which all new books are now being printed (and as used here).

There have been numerous more recent descriptions and practical handbooks of the language, both in Danish and English (e.g. Bergsland 1956, Rischel 1974 and the most comprehensive work to date Fortescue 1984). There is a standard Greenlandic-Danish dictionary that has been reedited and expanded several times, last by Schultz-Lorentzen, which also exists in English version (Schultz-Lorentzen 1927). There is a Danish-Greenlandic dictionary (Bugge et al. 1960) and a recently expanded Greenlandic-Danish school dictionary in the new orthography (Berthelsen et al. 1997).

### 2. Major categories and processes

West Greenlandic, like all varieties of Eskimo, is a **polysynthetic** language, typologically extreme in the number of productive bound suffixes (or ‘postbases’) that can be inserted between verbal or nominal stems and inflectional endings. There are over 400 of these in common usage, many of them corresponding to independent words in more analytic languages and having their own syntactic requirements. Derivation by suffix thus fulfils many functions more commonly expressed by external syntax elsewhere.

A particularly characteristic trait of the language is the **recursiveness** of its rich morphology, whereby a verbal stem can be nominalized (or a nominal one verbalized), then again verbalized (or nominalized) by successive suffixes up to several times within one complex word form (the minimal sentence). This can be contrasted with polysynthetic languages displaying a more rigid ‘slot-and-filler’ type morphology (see Art. 130). The language is also highly inflected, with over 300 inflectional endings forming paradigms

that indicate – for verbs – mood and subject (on transitive bases also object), person and number and – for nominal bases – case, number, and personal possessor. Especially the transitive verbal inflections display considerable ‘portmanteau’ fusion.

Greenlandic words fall into three main classes: verbs, nominals (nouns and pronouns/demonstratives), and uninflected particles. The latter category includes conjunctions, interjections, and simple adverbials, also deictic/expressive demonstrative forms like *tassa* ‘there (is)’ used in equational copular constructions or for introducing explanations (see sentence (8) in the text for an example). There are no adjectives (only appositional nominals and participial forms of descriptive verbal stems). The first two categories consist of an obligatory stem followed by an inflection (zero in the case of absolute case nouns), with the possibility of from none to at least eight or so derivational suffixes between, transparently ordered according to accumulating semantic scope ‘from left to right’. Only the ‘sentential affixes’ described in 6.2.3 must follow each other in a fixed order and do not allow further recursion. Changing the order of word-internal morphemes generally changes the meaning in predictable ways (for detailed examples see Fortescue 1980). Enclitics also occur, following any inflection.

### 3. Morphophonology

The phonemic inventory of the language is rather simple, with only three vowel morphemes, each with wide allophonic range, the most centralized/lowered ones occurring before uvular consonants (hence orthographic ‘e’ and ‘o’ for /i/ and /u/ in that context). The morphophonemic processes whereby word-internal morphemes are attached one to another are quite complex, however, and this results in considerable allomorphy for bound morphemes. Regressive consonant **assimilation**, for instance, is an important productive process, which ensures that the only consonant clusters allowed at morpheme boundaries (apart from /ts/) are geminates (plain or phrangualized if a preceding uvular is involved). The initial fricative of many suffixes following vowels alternates with the corresponding stop under assimilation with a preceding consonant stem (thus *tikikkami* ‘when he arrived’ from *tikit-* and causative inflec-

tion *-gami*). Progressive assimilation in turn reduces vowel clusters, /a/ plus /i/ or /u/ going to /aa/ (except for word-final /ai/).

Many suffixes (including all those beginning with a vowel or a double consonant) are **truncating** rather than assimilating, causing the final consonant of the base (i. e. stem plus any suffix) to drop, as in *isilerpoq* ‘he is coming in’ from *iser-* plus inchoative *ler-* plus 3rd person singular inflection *-poq*. Most /g/-initial suffixes fuse with a preceding uvular as in *arnaraa* ‘she is his mother’ from *arnaq* ‘mother’ plus *gi-* ‘have as’ plus 3rd person singular subject/3rd person singular object transitive indicative inflection *-vaa* (note the /i/-/a/ alternation and the contraction with the inflectional morpheme). A small number of suffixes causes more radical fusion with the stem in lexicalized combinations. These are the so-called ‘replacive’ suffixes such as *-lior-* ‘make’. Thus *umissiorpoq* in sentence (4) of the sample text, from *umiaq* ‘skin boat’ and *-lior-* and inflection *-poq*. The stem has an underlying sibilant between the second and third vowel which surfaces, geminated, as a result of the process.

While the majority of suffixes are either assimilating or truncating, most nominal inflections beginning with a single consonant are only selectively assimilating, namely with ‘strong’ consonant stems, whereas the final consonants of ‘weak’ stems are dropped. Many weak consonant stems further undergo gemination of the penultimate consonant when such endings are attached (fricatives often producing corresponding geminate stops), e. g. *nannumit*, ablative case of *nanoq* ‘polar bear’ and *ujaqqat* ‘rocks’, plural of *ujaraq*.

There are various other more or less productive attachment patterns that affect only limited small sets of stems or suffixes, for example **metathesis** in certain inflected nominals (thus *arlit*, plural of *aleq* ‘harpoon line’), or the alternation of /t/ and /s/ in morphemes like participial *-soq*, *-toq* after consonant but with *-soq* also following closed syllables that contain a historical /i/ (as opposed to the schwa with which it has merged). Similarly idiosyncratic is the alternation of /j/-initial suffixes like *-juma-* ‘want to’ (after a vowel stem) with /k/- and /r/-initial allomorphs as in *sinikkumavoq* ‘he wanted to sleep’ from *sinig-* and *majuarumavoq* ‘he wanted to go up’ from *majuar-*. Some suffixes contain recessive phonemes that appear in certain contexts only. The net result of such pro-

cesses can result in the obscuration of underlying morpheme shape, as for example in *sannapaa* ‘he made it for or him’ from *sana-* ‘make’, -(*u*)*t(i)-* ‘(do) for’ and transitive inflection *-paa*.

#### 4. Nominal inflection

There are three basic paradigms for nouns plus a parallel one for pronouns/demonstratives. The endings involved are given on Tab. 129.1. In general inflections and derivational suffixes are cited in the form they appear in after (most) vowel stems.

##### 4.1. Nouns

In the paradigms for **nouns**, starting with the non-possessed one, a + before an item means ‘assimilating’ and its absence ‘truncating’, while ÷ indicates ‘selectively assimilating’; recessive consonants appearing only after vowel-stems (or in the case of /u/ and /i/ in the absolute and relative only after ‘strong’ consonant stems) are in parentheses.

West Greenlandic has an **ergative morphology**, so the subject of intransitive verbs and the object of transitive ones are in the same case, namely the ‘absolutive’, as opposed to

the ‘relative’ case of transitive subjects. The latter is also used to mark the possessor in noun phrases such as *anguti-p qaja-a* ‘the man’s kayak’, where the possessum is marked for 3rd person singular possessor. There are no articles or gender distinctions in Greenlandic and the dual number, still found in northern dialects, has been lost in central West Greenlandic. The instrumental case, besides indicating means/instrument, is also used for the object of the so-called ‘half transitive’ construction described in 6.2.1 and in adverbial constructions such as *arriitsunngu-amik* ‘slowly’ in sentence (15) of the text, based on the participial form of verbal base *arriit-* ‘go slowly’ and *nnguaq* ‘small’. The allative basically indicates motion towards, the ablative motion from, the prosecutive (or viialis) motion through or via, and the equative similarity or degree (‘like’ or ‘as’).

The second paradigm for nouns on Tab. 129.1 is for absolute case possessed stems (singular or plural possessum). The possessor is indicated in the lefthand column. The category ‘4th person’ is met here for the first time: it refers to the reflexive 3rd person as in *nulia-ni asavaa* ‘he loved his (own) wife’. The subsequent paradigm is for relative case possessed endings (see sentence (16) for an example of use). The latter is also involved, in contracted shape, in combinations with cases other than the absolute and relative. Thus, for example, locative 1st person singular/2nd person singular (for both singular and plural possessum) *-nni* as in *illu-nni* ‘in my/your house(s)’, and allative 1st person plural *illut-sinnut* ‘to our house(s)’. These combined endings are also used in postpositional constructions such as *illu-p qa-a-ni* ‘on top of the house’ with positional stem *qa(a)-* ‘top’, and, in combination with the relative case of a 3rd person possessed nominal, *anguti-p qaja-ata qa-a-ni* ‘on top of the man’s kayak’.

##### 4.2. Pronouns and demonstratives

The inflection of **demonstratives** is illustrated on Tab. 129.1 with *una* ‘this/that nearby’. There are a dozen such stems in the language (used instead of 3rd person pronouns). They are differentiated according to combinations of position near/far from speaker with planar orientation above/below/inside/outside/up coast or down coast, etc. Inflection here is somewhat more idiosyncratic than for nouns. The same stems take somewhat different endings in adverbial (directional) use, as in allative *avunga*, locative *avani*, ablative

	non-possessed nouns		demonstratives	
	sing.	pl.	sing.	pl.
Abs.	–	( <i>i</i> ) <i>t</i>	<i>una</i>	<i>uku(a)</i>
Rel.	( <i>u</i> ) <i>p</i>	( <i>i</i> ) <i>t</i>	<i>uuma</i>	<i>uku(a)</i>
Instr.	÷ <i>mik</i>	÷ <i>nik</i>	<i>uuminnga</i>	<i>ukuninnga</i>
All.	÷ <i>mut</i>	÷ <i>nut</i>	<i>uumunnga</i>	<i>ukununnga</i>
Loc.	÷ <i>mi</i>	÷ <i>ni</i>	<i>uumani</i>	<i>ukunani</i>
Abl.	÷ <i>mit</i>	÷ <i>nit</i>	<i>uumannga</i>	<i>ukunanna</i>
Pros.	<i>kkut</i>	÷ <i>tigut</i>	<i>uumuuna</i>	<i>ukunuuna</i>
Equ.	÷ <i>tut</i>	÷ <i>tut</i>	<i>uumatut</i>	<i>ukunatut</i>

	absolute possessed		relative possessed	
	sing.	pl.	sing.	pl.
1.SG	<i>ga</i>	<i>kka</i>	÷ <i>ma</i>	<i>ma</i>
2.SG	( <i>i</i> ) <i>t</i>	<i>tit</i>	+( <i>r</i> ) <i>pit/vit</i>	<i>vit</i>
3.SG	<i>a</i>	<i>i</i>	<i>ata</i>	<i>isa</i>
4.SG	<i>ni/i</i>	<i>ni</i>	÷/- <i>mi</i>	÷/- <i>mi</i>
1.PL	+( <i>r</i> ) <i>put</i>	<i>vut</i>	<i>tta</i>	<i>tta</i>
2.PL	+( <i>r</i> ) <i>si</i>	<i>si</i>	<i>ssi</i>	<i>ssi</i>
3.PL	<i>at</i>	<i>ilat</i>	<i>ata</i>	<i>isa</i>
4.PL	+( <i>r</i> ) <i>tik</i>	<i>tik</i>	÷/- <i>mik</i>	÷/- <i>mik</i>

Tab. 129.1: Nominal inflections

*avannga*, and prosecutive *avuuna* from *av-* ‘that in the north/up the coast’. There are only two prefixes in the language (neither productive) and they both attach to these stems, namely deictic *aa-*, as in *aajuna* ‘there (it is)’, and anaphoric *ta-*, which in combination with *una*, for example, produces *taanna* ‘that/he/she/it aforementioned’ (its idiosyncratic relative case form is seen in sentence (10) of the illustrative text).

First and second person **pronouns** and question words *kina* ‘who’ and *suna* ‘what’ take analogous inflections, but *uanga* ‘I’ and *illit* ‘thou’ do not have distinct relative case forms. Reflexive/reciprocal stem *immi-* only occurs in the oblique cases. There is also a small group of nominal stems such as *tamar-* ‘all’ which go counter to the general ergative patterning of nominal morphology by having one form (e.g. 3rd person plural *tamaasa*) in object function and one as subject – whether of an intransitive or transitive verb (e.g. corresponding 4th person plural *tamarmik*). *ki-siat* ‘only’ in sentence (15) is of this type.

## 5. Verbal inflection

Many of the person markers seen in nominal paradigms also occur in the verbal ones on Tab. 129.2. The principal distinction here is between **superordinate** and **subordinate clause moods**. The former covers the indicative, interrogative, and imperative/optative moods, and the latter the causative, conditional, contemporative, and participial moods. All the moods except the contemporative have parallel intransitive and transitive paradigms consisting of a mood marker followed by the person/number morphemes. Intransitive endings on otherwise transitive bases give a reflexive or reciprocal sense. Greenlandic does not have subordinating particles and only few conjunctions (mostly inflected verb forms), so mood inflection is crucial for relating clause to clause. The category ‘4th person’ is, as for nominals, reflexive, but specifically refers to the subjects of the superordinate clause.

The initial /v/ of the indicative and interrogative alternates with (assimilating) /p/ after a consonant and disappears in contractions with certain preceding derivational suffixes. There is also a special negative form of the paradigm with *-nngila-* in place of *-vu-* and *-va-*.

The participial mood is like the indicative but with *-vu-/ppu-* replaced by *-su-* (alternat-

### Indicative

	detrans.	sing. trans. object		
		1.SG	2.SG	3.SG
1.SG	+vunga	–	vakkit	vara
2.SG	+vutit	varma	–	vat
3.SG	+voq	vaanga	vaatit	vaa
1.PL	+vugut	–	vatsigit	varput
2.PL	+vusi	vassinga	–	varsi
3.PL	+(p)put	vaannga	vaatsit	vaat

plural trans. object			
	1.PL	2.PL	3.PL
1.SG	–	vassi	vakka
2.SG	vatsigut	–	vatit
3.SG	vaatigut	vaasi	vai
1.PL	–	vassi	vavut
2.PL	vassigut	–	vasi
3.PL	vaatigut	vaasi	vaat

### Imperative

intrans. trans. object				
	3.SG	3.PL	1.SG	
2.SG	gitl/+na	(g)uk	kkit	nnga
1.PL	+ta	+tigu	+tigik	–
2.PL	gitsi	+siuk	+sigit	+singa
				+tigut

### Conditional

	intrans. trans. object		
	3.SG	3.PL	
1.SG	+guma	gukku	gukkit
2.SG	+guit	gukku	gukkit
3.SG	+(p)pat	+(p)pagu	+(p)pagit
4.SG	+guni	guniuk	gunigit
1.PL	+gutta	gutsigu	gutsigik
2.PL	+gussi	gussiuk	gussigik
3.PL	+(p)pata	+(p)passuk	+(p)patigik
4.PL	+gunik	gunikku	gunikkik

### Contemporative

#### subject/object

1.SG	+ (l)lunga
2.SG	+ (l)lutit
4.SG	+ (l)luni
1.PL	+ (l)luta
2.PL	+ (l)lusi
4.PL	+ (l)lutik

Tab. 129.2: Verbal inflections

ing with *-tu-* after consonants) in the intransitive and with *-gi-* (*-ki-* after bases in /k/ and /t/) instead of *-va-* in the transitive. Being subordinate, it also has transitive forms with 4th person subject or object. It has important functions in object clauses (see sentence (7)) and in temporal and (intransitive only) relative clauses.

The interrogative mood is also like the indicative except for 2nd person singular *-vit*, 3rd person singular *-va*, 2nd person plural *-visi*, 3rd person plural *-ppat* and certain 3rd person object forms (analogous rather with those of the conditional). The optative (1st and 3rd person) is as the indicative but with 1st person *-la-*, 3rd person *-li-* instead of *-vu-/va-*.

The causative and conditional (the ‘relative’ moods) largely share the same person/number markers but each has its characteristic mood marker: the former has mood marker *-ga-* instead of *-gu-* (both are assimilating and have /k/ – initial allomorphs following bases in /t/ or /k/), 3rd person forms with /mm/ rather than /pp/, and 4th person ones with /m/ rather than /n/. The object forms for other than 3rd person are analogous to those of the indicative. The causative is used to express cause or time in the past relative to a superordinate clause and the latter to express condition or time in the future. There are special forms of the causative beginning with /ng/ following suffix *-gaa-* to express contingency, as in *anigaangami* ‘every time he went out’.

The contemporative mood has the same set of endings whether intransitive (for subject) or transitive (for object), the subject of such a clause generally being coreferential with that of the superordinate clause. There are also 3rd person object forms, the same for all subjects: 3rd person singular *-(l)lugu*, 3rd person plural *-(l)lugit*. All forms are assimilating, with recessive /l/ appearing after vowel stems (but not after allomorph *-ga-* of *-gi-* ‘have as’, as in sentence (5)). There is a special negative paradigm with mood marker *-na-* instead of *-lu-*. The mood has both subordinative and coordinative functions, for example in (respectively) coreferential object or temporal clauses and in the expression of sequences of actions involving the same subject. In such **coverb constructions** it is not always possible to distinguish semantic subordination from coordination (see for example sentences (9) and (10)). In combination with causative suffix *-tit-* it produces temporal

clauses such as *sinit-sil-lugu* ‘while he slept’ (with different subject from the superordinate clause).

## 6. Derivational morphology

The following examples illustrate the kind of word-internal complexity Greenlandic words may display, the first verbal, the second nominal. Others can be found in the illustrative text. They are divided up by surface rather than underlying morpheme shape.

- (1) (a) *nannu-n-niuti-kkuminar-to-rujussu-u-vog*  
polar.bear-catch-means.for.striving.  
to-be.good.for-DETR.PART-very.  
much.so-be-3.SG.IND  
‘it is really good for catching polar  
bears with’ (of a dog)
- (b) *nuannaa-ruti-gi-nnif-figi-sinnaa-sa-a*  
be.happy-means.for-have.as-½TRANS-  
have.as.time/place.of-can-PASS.PART-  
3.SG  
‘something from which he was able to  
derive pleasure’

The derivational suffixes of West Greenlandic can be broken down into a number of subgroups, each sharing certain semantic and/or syntactic features. Some are attachable to a limited range of bases (e.g. body parts or locations) but within these natural semantic boundaries they can be said to be productive, with new peripheral combinations arising. There are numerous others which only occur in lexicalised combination with one or a handful of stems. Selected items are given on Tab. 129.3.

In Tab. 129.3, a raised + following a suffix indicates intrinsically transitive and a raised – intransitive. Suffixes beginning with an /l/ in parentheses are ‘replacive’. A useful way of looking at the difference between them is in terms of conditions on the input to the corresponding derivational rule in Functional Grammar terms. Some take for instance simple predicates, while others include whole propositions with mood and tense markers and various external satellite phrases in their scope (see Kristoffersen 1991 for a detailed discussion). It is this that is largely responsible for the notorious overlap of morphology and syntax in Eskimo languages. An example is given in 6.2.1. Within the same framework, the sentential suffixes of 6.2.3 are best treated as grammatical oper-

<i>Verbalizers</i>	
<i>nngor</i> –	‘become’
<i>u</i> –	‘be’
<i>ssaasua</i> –	‘lack’
<i>katag</i> –	‘be fed up with’
<i>gi</i> +	‘have as’
<i>qar</i> –	‘have’
<i>+si</i> –	‘get’
<i>+mukar</i> –	‘go to’
<i>nga</i> –	‘resemble’
<i>(l)ior</i> –	‘make’
<i>(l)er</i> +	‘provide with’

<i>Verbal extenders</i>	
<i>+nerar</i> +	‘say that’
<i>+(r)palug</i> –	‘look/sound like’
<i>+juma</i> –	‘want to’
<i>qqu</i> +	‘ask/tell to’
<i>+tit</i> +	‘let/cause to’
<i>+niar</i> –	‘try to’
<i>+sinnaa</i> –	‘can’
<i>+negar</i> –	‘passive’
<i>(u)t(i)</i> +	‘do with/for’

<i>Verbal modifiers</i>	
<i>+neru</i>	‘more’
<i>laar</i>	‘a little’
<i>rusaar</i>	‘slowly’
<i>rulug</i>	‘violently/hard’
<i>ler</i>	‘begin to’
<i>+sima</i>	‘perfective’
<i>+sar</i>	‘iterative/habitual’
<i>qqaar</i>	‘first’

<i>Sentential suffixes</i>	
<i>ssa</i>	‘future’
<i>+gunar</i>	‘probably/apparently’
<i>ngit</i>	‘negative’
<i>nguar</i>	‘dear or lucky one (does)’
<i>qina</i>	‘take care not to’
<i>ratar</i>	‘(and) surprisingly/finally’

<i>Nominalizers</i>	
<i>+saq/gaq</i>	‘passive participial’
<i>+soq</i>	‘active participial’
<i>+(f)fik</i>	‘place of’
<i>ssut</i>	‘reason for’

<i>Nominal extenders and modifiers</i>	
<i>kkut</i>	‘and family/companions’
<i>lik</i>	‘(one) provided with’
<i>usaq</i>	‘something resembling/model’
<i>+(r)suaq</i>	‘big’
<i>kasik</i>	‘naughty/bad/dear (ironic)’
<i>ssaq</i>	‘future’

Tab. 129.3: Selected West Greenlandic derivational suffixes

ators (as are one or two aspectual suffixes interacting with the external syntax as mentioned under 6.2.2), despite the fact that they are not obligatory and do not form clear-cut paradigms like the ‘true’ inflectional categories already treated. Semantically one can thus envisage a continuum between more and less inflectional/derivational categories here, the former appearing further away from the stem than the latter.

Structurally there is considerable recursiveness (and productivity) of the processes involved. The basic ordering within a verb form (the minimal clause) is: verbal base plus (optional) sentential affix plus inflection. A ‘verbal base’ can itself, however, be built up recursively either from a simpler verbal base – or a nominal base plus verbalizing affix – followed by a lexically ‘heavy’ verb-expanding affix, a (predicate) negation affix and one or more verb-modifying affixes, all of them optional. If there is more than one sentential affix present (before the inflection), they are in the order: tense – epistemic modality – sentential negation – subjective coloration. There are also ‘conjunctional’ affixes that appear immediately followed by particular subordinate mood inflections. As regards nominal bases, these can be built up recursively from simpler nominal bases – or from verbal base plus nominalizing affix – by a nominal-expanding affix and one or more nominal-modifying affixes. Constraints on this recursiveness are of a natural semantic and pragmatic nature, according to the meaning to be expressed, scope being cumulative ‘from left to right’ (see Fortescue 1980).

In stark contrast to its rich derivational potential the language does not allow nominal or verbal compounding at all.

### 6.1. Denominal verbal

This group of suffixes covers the following kinds of meanings: being and becoming, lacking, feeling, having, acquiring, movement to or from, acting and seeming like, doing and providing with. They all have a potential interaction with the external syntax in so far as a modifier of the nominal base to which they may attach may occur as an external argument in an appropriate case. Thus *angisuumik illoqarpooq* ‘he has a big house’, with suffix *-qar* ‘have’ and instrumental case modifier *angisoog* ‘big’ modifying stem *illu* ‘house’, and *illuuvooq angisoog* ‘it is a big house’ with *-u* ‘be’ and the postposed modifier in the absolute case. The word-order of

the latter construction is fixed, although the neutral – but flexible – word order in Greenlandic is subject-object-verb (note that modifiers always follow their head nominals in NPs). Nominal extender *-lik* ‘one provided with’, the nominal equivalent of *-qar-* here, can be seen in the same construction type in sentence (5) of the illustrative text.

Suffixes of motion towards *-mukar-* and location *-miit-* contain historical allative and locative inflections respectively and thus alternate with *-nukar-* and *-niit-* after possessed or plural bases; they may alternatively be analysed, as *-kar-* and *-it-* attach to whole inflected phrases, as in *Nuup eqqaani-ip-poq* ‘it is near Nuuk’ – compare *Nuup eqqaani* ‘near Nuuk’. Whether the bound morpheme constructions involved here should be called ‘incorporation’ (as Sadock 1985 for instance does) is a controversial matter of definition, since a Greenlandic word can never contain more than one lexical stem.

## 6.2. Deverbal verbal

### 6.2.1. Verbal extenders

The suffixes here are lexically quite substantial, covering the areas of judging and saying, wishing and waiting, causation and request, striving and intending, potentiality and valency shifters. They may bear their own intrinsic transitivity – notably the **valency shifters**, which include the so-called ‘half-transitivizing’ (or antipassive) suffixes, as in *tuttumik toqutsivoq* ‘he killed a caribou’, where the transitive stem is *toqut-* ‘kill’ and the ‘object’ – semantically indefinite – is in the instrumental case and could have been in the plural with *-nik* instead of *-mik* (compare the instrumental modifier in 6.1). The object of the corresponding transitive construction (in the absolute) would be definite. There are two passive suffixes, the dynamic and (less productive) the stative. Another important transitivity affecting suffix, *-(u)t(i)-*, promotes a non-nuclear argument – often a beneficiary, but see sentence (10) of the text – to nuclear object (i.e. one inflectionally cross-referenced on the verb).

An important group of suffixes here are the ‘double transitivizers’, which can be added to either intransitive or transitive bases, in the latter case keeping the same absolute case object as the underlying base but demoting its subject to an allative satellite by introducing its own nuclear subject. They are mostly causativizers but include the

meanings ‘think that’, ‘say that’ and ‘wait for it to happen that’. Observe the following example which illustrates how certain of these suffixes, including *nerar-* ‘say that’ (morphologized indirect speech), take a whole proposition in their scope:

- (2) *Maalia-p meeraq erniinnaq*  
*Malia-REL child soon*  
*sini-li-ssa-gunar-nerar-paa*  
*sleep-begin-FUT-probably-say.that-3.SG/*  
*3.SG.IND*  
‘Maalia said that the child would soon begin to sleep’

### 6.2.2. Verbal modifiers

This group includes suffixes of degree, manner and aspect (both phasal and quantitative). Unlike the extenders they never affect the valency of the base and do not extend their scope beyond the base to which they are attached (except for iterative/habitual *-sar-*, at least, which is best treated as an operator, being obligatory in various syntactic contexts). They may attach up to several in succession at the same level or recursively to each successive level started by a new extender or verbalizer (see the final verb form in (16), where an iterative, a degree and a habitual suffix follow each other at the same level of clausal layering and a perfective one is added after verbalizer *-u-*).

### 6.2.3. Sentential suffixes

These suffixes, while still not obligatory elements in the word, fall into ordered ‘slots’ immediately preceding the inflection in the following order: tense, (epistemic) modality, (sentential) negation, and subjective coloration. They are completely productive and modify the clause as a whole.

Negative suffix *-nngit-* can also occur earlier in a word-form as a base negator rather than a sentential one. Compare *ajunngilaq* ‘it is good’ from *ajor-* ‘be bad’ (reflecting the typical contradictory reading of sentential negation in the language) but *ajussanngilaq* ‘it will be good’ with the future suffix *-ss-* inserted between the stem and the negative suffix. So this is not simply a lexicalization, although it may act as such in some respects: *ajunngilaq* can itself be (sententially) negated to *ajunnginngilaq* ‘it is not (at all) bad’.

There are also a number of suffixes that appear just before specific sets of inflectional endings in a conjunctival sense (such as ‘after’, ‘before’, ‘while’, and ‘so that’), but these

are by and large morphemes found in other functions/meanings, e. g. aspectual (*-riar-* ‘after’ in (10) is an example: elsewhere it means ‘set about (dynamic action’).

The same historical suffix may in general appear in more than one of the semantico-syntactic sub-categories discussed above (polysemy). Thus *-sima-* appears both as a sentential suffix – either as a tense (perfect) or modality marker (indicating that the speaker was not present at the event described) – and as a verbal modifier for perfective aspect. In most contexts, where the presence of additional suffixes cannot disambiguate, there is conflation of at least two of these functions.

There are also ‘semi-lexicalized’ combinations of two or more basic suffixes (especially amongst sentential suffixes), where the meaning is predictable from the components. Such combinations may later develop idiosyncratically lexicalized meanings, such as *-nnguatsiar-* ‘it seems/probably’ from *-nnguar-* ‘the little one does’ (subjective coloration – humble of first person subjects) and *-tsiar-* ‘for a while, a bit’. Another kind of lexicalization (a kind of word-internal raising) is seen in such compound suffixes as *-qqunngit-* ‘tell not to’, where negation follows ‘tell to’, contravening the general scope principle for the ordering of successive affixes.

Note that tense is not an obligatory category and that past and present tense is not usually distinguished (though it can be – especially experiential ‘perfect’ tense). It is generally relative rather than absolute – though future tense will, all else being equal, be relative to the speech act situation.

The ‘subjective coloration’ suffixes are the most heterogenous semantically (and most ‘exotic’ as operators). They include markers of subjective valuation of the state of affairs being described, illocutionary modifiers and narrative intensifiers such as *-qi-* (*-qa-* before abbreviated indicative endings, as in sentence (13) of the text). Thus *-kasig-* (corresponding to nominal modifier *– kasik*) adds a nuance of ‘silly me/him/you’ or, ironically, ‘poor little’, often adding a humorous dimension. On the other hand *-galuar-* ‘but ...’ may indicate reservation or some contrasting implication to the predication (see sentences (6) and (9)). As a verbal modifier, earlier in the recursive structure of the complex word, it can mean ‘in vain’.

### 6.3. Deverbal and denominal nominal

Little needs to be said about the kinds of suffixes here: they are exemplified on Tab. 129.3. Note, however, that deverbal *-neq* may nominalize whole clauses, as in *ikinngitiuminut neqimik tuniussinera* ‘his giving of (the) meat to his friends’, with the 3rd person singular possessor marking attached to *tuniussineq* ‘giving’ referring to the clausal subject and *neqi* ‘meat’ as the instrumental case ‘object’. The half-transitive construction here is obligatory (as it is also in relative clauses formed on transitive verbal bases) and does not necessarily indicate indefiniteness of the object. Nominalizations with *-neq* also occur in syntagmatic construction with verbs like *ajor-* ‘be bad’ in the special sense ‘can not, does not (generally)’, as in text sentence (6).

The passive participial marker *sag/tag/gaq* (with lexicalized distribution of allomorphs) appears commonly in headless constructions such as that illustrated in (1 b). Intransitive participial *-sog* (*-toq* after consonant) happens to have the same shape as the 3rd person forms of the intransitive verbal paradigm of that name but takes ordinary nominal inflections. Headless relative clauses involving it are also common, as in *tikittoq* ‘one who has arrived’ from *tikit-* ‘arrive’.

A typical nominal extender is *-lik* and a typical nominal modifier *-(r)suag* on Tab. 129.3. The latter (which has special plural and relative singular forms *-(r)suit* and *-(r)sup*) occurs in the sample text several times between intransitive participial *sog/toq* and *u-* ‘be’ or *-ngor-* ‘become’, these combinations resulting in the equivalent of a subjective coloration (verbal deverbal) suffix.

### 7. Enclitics

**Enclitics** typically attach to the first verbal or nominal constituent in a clause or phrase. They are often conjunctional, such as *=lu* ‘and’ and *=li* ‘but’, or emphatic, such as *=mi* ‘what about, for’. Note also demonstratives like *una* used as a focus-marker in construction with a following nominal or participial mood verb, as in *Hansi=una toqukkaa* ‘it was Hans who killed it’. In conjunction with the participial (or, coreferentially, the contemporative) mood *=lu* may add a (subordinative) sense of ‘as soon as’ (see sentence (12) of the text), and *=li* can mean ‘ever since’, as in sentence (15). Quotative enclitic *=goog* (=*nngooq* with words ending in /t/ or /p/)

bears the illocutionary function of indicating that the speaker is merely mediating someone else's words, whether it be a declarative, a question or a directive. Several enclitics may occur in succession, as amply illustrated in the text.

### 8. Illustrative text

The following excerpt is taken from an oral narrative entitled "Pukkitsulik" collected in the middle of the last century (in Lyng 1978) and dating back to the time of first contact between Greenlandic Eskimos and European whalers in the seventeenth century. Only suffixes used in a fully productive and predictable manner have been separated from stems by a hyphen, thus for example *oqaluuseri*- in sentence (14), for example, is regarded as a lexicalized whole, although it can be analysed into *oqalug-* 'speak', *-useq* 'means or manner of -ing', and *-gi-* 'have as'. Enclitics are separated off by the symbol =. Unmarked absolute case on nominals is not glossed (nor singular number of case or possessum endings). Subject and object (also possessor and possessum number) in portmanteau inflections are separated by obliques.

- (3) *ukiu-u-lla-rami*  
winter-be-INTS-4.SG.CAUS
- 
- (4) *inu-il=lu=goaq tassa*  
human.being-PL=and=QUOT that.is  
*aallaar-tar-tor-sua-nngor-mata*  
go.on.hunting.trip-ITER-DETR.PART-big-  
become-3.PL.CAUS  
*Pukkitsulik umissior-poq*  
Pukkitsulik skin.boat.make-3.SG.IND
- 
- (5) *Nunaqqati-ga-lugu=goaq*  
person.from.same.village-have.as-  
(4.SG)3.SG.CONTMP=QUOT  
*piniartor-suaq nakuar-suaq*  
hunter-big strong.person-big  
*pinni-ffaarim-mik*  
beautiful-very-INSTR  
*pani-lik*  
daughter-one.provided.with
- 
- (6) *Pukkitsulli-n=ngooq*  
Pukkitsulik-REL=QUOT  
*ivanger-su-i*  
breast-big-3.SG/PL
- 
- (7) *uppas-su-i=lu*  
thigh-big-3.SG/PL=and  
*isigi-sa-raluar-lugit*  
look.at-HABIT-but ... -(4.SG)3.PL.CONTMP  
*oqar-neq*  
speak-NOM  
*aju-li-innar-sima-vog*  
can.not-begin-only-PFV/PF-3.SG.IND
- 
- (8) *natu-nngin-namiuk*  
not.know-NEG-4.SG/3.SG.CAUS  
*angus-su-ata*  
father-big-3.SG.REL  
*tunniuk-kuma-sa-nngik-kaa*  
give-want-HABIT-NEG-3.SG/3.SG.PART
- 
- (9) *piniartor-su-in=ngooq ilaatigut*  
hunter-big-PL=QUOT sometimes  
*nakuar-ta-lis-su-it*  
strong.person-pertaining.to.group-pro-  
vided.with-big-PL  
*ungasis-sor-suuar-mit*  
be.far-DETR.PART-big-ABL  
*agger-lutik*  
come-4.PL.CONTMP  
*piuma-jartor-tar-sima-galua-raat*  
want.to.get-go.and-ITER-PFV/PF-but ... -  
3.PL/3.SG.PART
- 
- (10) *oqar-nia-riar-tul=lu=goaq*  
say-try.to-after-3.PL.PART=and=QUOT  
*arnar-su-up taassuma*  
woman-big-REL that.REL  
*tui-isigut*  
shoulder-3.SG/PL.PROS  
*tigu-llugit paamut*  
take-(4.SG)3.PL.CONTMP entrance.ALL  
*miliori-ut-tar-sima-vai*  
throw-with-ITER-PFV/PF-3.SG/3.PL.IND
- 
- (11) *taamaam-man=ngooq*  
be.thus-3.SG.CAUS=QUOT  
*Pukkitsulli-p oqaasi-ssa-ni*  
Pukkitsulik-REL word-future-4.SG/PL  
*artu-li-innar-sima-ler-pai*  
be.unable.to.manage-begin-only-PFV/PF-be-  
gin-3.SG/3.PL.IND
- 
- (12) *tassa=lu=goaq=aasit*  
that.is=and=QUOT=again.as.usual

- aasa-llar-tor=lu*  
become.summer-INTS-DETR.PART=and  
*inu-it ungasis-sor-suar-mut*  
human.being-PL be.far-DETR.PART-big-  
ALL  
*nuna siner-lugu*  
land go.along.coast.of-(4.PL)3.SG.  
CONTMP  
*aallaar-tar-tor-sua-nngor-put*  
go.on.hunting.trip-ITER-DETR.PART-big-be-  
come-3.PL.IND  
*Pukkitsuli-ul=lu umia-ni*  
Pukkitsulik-REL=and skin.boat-4.SG  
*ini-llar-lugu*  
finish-INTS-(4.SG)3.SG.CONTMP
- (13) *ilaanni=gooq=aasiit*  
once=QUOT=again.as.usual  
*gernerta-qu-sua-ni*  
blue.fox.pelt-owned.thing-big-4.SG  
*ta-manna meqqor-su-isigut*  
ANA-this hair-big-3.SG/PL.PROS  
*patta-laar-ta-riannguar-lugu*  
pat-a.little-ITER-surprisingly-  
(4.SG)3.SG.CONTMP  
*oqa-qa-aq*  
say-INTS-3.SG.IND
- (14) *tassa=ana sallu-nngi-latit*  
that.is=that lie-NEG-2.SG.IND  
*umiarsu-il=lu qallunaar-su-il=lu*  
ship-PL=and white.man-big-PL=and  
*oqaluuseri-lla-rakkit*  
talk.about-INTS-2.SG/3.PL.CAUS
- (15) *arriit-su-nngua-mik*  
be.slow-DETR.PART-small-INSTR  
*oqalu-ler-poq soorlu=gooq*  
talk-begin-3.SG.IND as.if=QUOT  
*taamane-rsuar=li*  
at.that.time-big=but  
*umis-siu-le-rami*  
skin.boat-make-begin-4.SG.CAUS  
*taanna kisiat*  
that alone  
*isumaliuti-gi-sima-gaa*  
thought-have.as-PFV/PF-3.SG/3.SG.PART
- (16) *arna-ata aki-vaa*  
mother-3.SG.REL reply-3.SG/3.SG.IND  
*tassa=mi inu-it*  
that.is=for human.being-PL  
*amerlasoor-su-it qallunaar-suar-nit*  
many-big-PL white.man-big-ABL.PL

*pequiti-taar-to-rulut-tar-tor-su-u-sima-pput*  
goods-get-ITER-exceedingly-HABIT-DETR.P-  
ART-big-be-PFV/PF-3.PL.IND

“(3) It had been a long winter and summer was approaching (4) and since people were beginning to go off on hunting trips, it is said Pukkitsulik made himself a skin boat. (5) There lived in the same village as him, it is told, a powerful hunter with a most beautiful daughter. (6) Pukkitsulik would stare at her plump breasts and thighs but was unable to speak up, (7) because he knew that her father would never give her away (8) and it seems the girl did not want to get married. (9) Many fine hunters, champions among them, had come from afar to court her (10) and as soon as they tried presenting their case, it is said, the big girl would take them about the shoulders and fling them out of the entrance passage. (11) It was for this reason that Pukkitsulik had begun to be unable to express himself. (12) Then as soon as it was summer the people started to travel up and down the coast on hunting trips – and Pukkitsulik had finished his skin boat. (13) One day he began stroking the fur of his blue fox pelt and said suddenly to his mother: (14) ‘Were you telling the truth when you told about the big ships and the white men?’ (15) He spoke slowly as if he had had nothing else on his mind all the time he had been working on the skin boat. (16) His mother replied: ‘Yes, indeed many people have got lots of goods from the white men.’”

## 9. Uncommon abbreviations

CONTMP	contemporative
PROS	prosecutive
½TRANS	half-transitive

## 10. References

- Bergsland, Knut (1956), *A Grammatical Outline of the Eskimo Language of West Greenland*. Oslo: mimeographed
- Berthelsen, Christian & Jacobsen, Brigitte & Kleivan, Inge & Petersen, Robert & Rischel, Jørgen (1997), *Oqaatsit, Grønlandsk-Dansk Ordborg*. Nuuk: Atuakkiorfik
- Bugge, Aage & Fuglsang-Damgaard, Ad. & Lynga, Kristoffer & Nielsen, Frederik (1960), *Dansk-Grønlandsk Ordbog*. Copenhagen: Ministeriet for Grønland
- Egede, Paul (1750), *Dictionarium Grönlandico-Danico-Latinum*. Copenhagen: Havniae

- Fortescue, Michael (1980), "Affix Ordering in West Greenlandic Derivational Processes". *International Journal of American Linguistics* 46.4, 259–278
- Fortescue, Michael (1984), *West Greenlandic*. Beckenham, Kent: Croom Helm
- Kleinschmidt, Samuel (1851), *Grammatik der Grönlandischen Sprache*. Berlin: G. Reimer [reprinted 1968, Hildesheim: G. Olms Verlag]
- Kleinschmidt, Samuel (1871), *Den Grønlandske Ordbog Omarbejdet*. Copenhagen: L. Klein
- Kristoffersen, Lars (1991), *Verbal Derivation and Inflection in a Functional Grammar of West Greenlandic*. Ph.D. dissertation, Institute for Eskimologi, University of Copenhagen
- Lynge, Kristoffer (1978), *kalatdlit oqalugtuait oqahualavilo*. Godthab: det Grønlandske Forlag
- Rischel, Jørgen (1974), *Topics in West Greenlandic Phonology*. Copenhagen: Akademisk Forlag
- Sadock, Jerrold (1985), "Autolexical Syntax: A Proposal for the Treatment of Noun Incorporation and Other Phenomena". *Natural Language and Linguistic Theory* 3, 379–439
- Schultz-Lorentzen, Christian Wilhelm (1927), *Dictionary of the West Greenland Eskimo Language*. Copenhagen: Meddelelser om Grønland 59 [reprinted 1945: first Danish edition 1926, A. Rosenbergs Bogtrykkeri, Copenhagen]

*Michael Fortescue, Copenhagen (Denmark)*

## 130. Koyukon (Athapaskan)

1. The language and its speakers
2. Major categories and processes
3. The structure of the verbal expression
4. Nominal expressions
5. Directionals
6. Illustrative text
7. Uncommon abbreviations
8. References

### 1. The language and its speakers

Koyukon is a northern Athapaskan language spoken in the interior of Alaska by approximately 300 out of a total population of 2,300. There are three main dialects, Lower, Central, and Upper. The present description is based on the central dialect, as spoken in the villages of Huslia, Hughes, and Allakaket on the Koyukuk river and Galena and Ruby on the Yukon.

The structures and categories of Koyukon are very similar to those of other Athapaskan languages in Alaska and neighbouring Canada and further south in the USA. Within Northern Athapaskan, a group of 23 distinct languages, no clear sub-groupings can be described. One trait that distinguishes Koyukon from its neighbours is the idiosyncratic merging of the proto-Athapaskan \*c series with the lateral series. It has also lost the high-low tonal distinction typical for most non-Alaskan Athapaskan languages (though there are traces left, especially in Lower Koyukon).

It is estimated (Krauss & Golla 1981:67) that proto-Athapaskan was undifferentiated until 500 BC or later, about 1000 years after it separated from Eyak, the only remaining representative of the other branch of Athapaskan-Eyak. The conservatism and resistance to borrowing of Athapaskan languages has often been remarked upon. This is attributable to their typological idiosyncracy – in particular their elaborate morphology. Koyukon, like other Alaskan members of the family, is particularly conservative as regards the preservation of stem-final suffixation. On the other hand, it has developed – apparently in common with all northern Athapaskan languages – a considerable potential for noun incorporation not displayed by the southern languages.

The earliest written material in the language dates from the missionary era of the late 19th century, reaching a high point with the grammatical and lexical work of Jesuit Jules Jetté, which has remained partly in manuscript form only (see Krauss & Golla 1981: 74). Literacy in the language was, owing to official suppression in the schools, not allowed to develop during the first half of this century, and only in recent times has Jetté's pioneering work been carried forward – first with the introduction in 1961 of a practical orthography by Summer Institute linguist David Henry. It has been revised since then. Eliza Jones, a native speaker working at Alaska Native Language Center, Fairbanks,

has edited and translated numerous texts in the language and has recently completed a comprehensive dictionary (Jetté & Jones 2000). An English-Koyukon school dictionary (Jones 1983) and a practical grammar for the language (Thompson et al. 1983) are available. The orthography used in the present article is the same as in these works (though in the new dictionary ‘i’ and ‘a’ have been replaced by ‘e’ for schwa).

## 2. Major categories and processes

Koyukon, like all varieties of Athapaskan, displays an elaborate verbal morphology of a rare prefixing type that has been called **interrupted synthesis** (Kari 1990: 38, citing Sapir’s unpublished use of the term). This refers to the predominance of discontinuous morphemes in these languages – both in lexicalized verb ‘themes’ and in derivational strings. The principal means of morphological elaboration is prefixation, although stem-internal modification is also an important ingredient. Much of this is concerned with the fine-grained articulation of aspect. The system is notoriously full of irregularities; moreover, there is a certain degree of arbitrariness – i.e. lack of independent meaning – to many of the discontinuous elements of thematic and derivational strings. The latter may reflect the functional need to specify more concrete meanings for roots of a highly general nature by means of a phonologically limited number of prefixes.

Nouns (non-derived) are, on the other hand, very simple morphologically, Koyukon being a ‘head-marking’ language (Nichols 1986). The only other lexical categories besides verbs and nouns – including pronouns – are those of directional expressions (discussed in 5) and uninflected particles. The latter are generally of an adverbial or exclamatory nature, but also include adjectives of fixed form like attributive *kuh* ‘big’ or predicative *adzoo* ‘(is) cold’, which may or may not be followed by existential verb *laa-* ‘be’. In terms of Greenberg’s typological features, Koyukon is an SOV language with postpositions, adjectives following their heads, and possessors preceding heads.

Although the Athapaskan languages can in a broad sense be termed ‘polysynthetic’ – especially the northern ones employing noun incorporation – theirs is a very different kind of polysynthesis from that of Eskimo lan-

guages (see Art. 129). This is manifest in the contrast between the rather rigid, template-like verbal morphology of Koyukon and the extremely recursive nature of Greenlandic. Koyukon cannot verbalize nominal stems, and there can be no more than one nominalization process per word, unlike the case of Greenlandic. As regards the semantic variety of material that can be drawn into the verbal complex, however, Koyukon does not lag far behind Greenlandic. The average morpheme-per-word count is lower than in the latter since the number of derivational processes applicable to a single word-form is rather limited in Koyukon.

Another trait shared with Greenlandic is the entanglement of inflection and derivation (cf. Art. 38). The obligatory inflectional categories of Koyukon have their place between the stem and any preceding derivational material. This is at odds with the usual tendency to place derivational material closer to the stem than inflections and can be explained with reference to the nature of ‘interrupted synthesis’: derivation in Athapaskan languages is generally expressed by prefixation in combination with an effect upon either the stem itself or the ‘classifier’ (valency marker) right before it. (Compare also the discussion of ‘extended derivation’ under 3.3).

## 3. The structure of the verbal expression

**Verbal morphology** in Koyukon can be seen in terms of the building-up of layered clause structures from **basic predicates** (or ‘themes’) via successive derivational and inflectional choices towards final linearization in the manner of Functional Grammar. Such an approach anchors Koyukon ‘theme categories’ in a typology of states-of-affairs (cf. Art. 104) and, at the same time, allows a principled description of the scrambled morpheme ordering of the language in terms of the centrifugal application of expression rules (Fortescue 1992).

### 3.1. Verb themes and the organization of the lexicon

The discontinuous nature of Koyukon morphology is already seen at the level of the basic building blocks of the language, its verbal ‘themes’. These consist of a stem – or, rather, a set of stem allomorphs – plus an immediately preceding ‘classifier’ (one of four va-

lency markers obligatorily taken by a given theme). Moreover, with many themes, there are one or more discontinuous prefixes. An example of the latter possibility is theme *ni* ... *l-aan-* ‘see’ (as in sentence (8) of the sample text), where *l* is the classifier and *ni* a discontinuous thematic prefix (apparently the same historically as the gender marker of that shape in 3.6). *aan-* is the abstract root from which the relevant stem-sets are derivable (it is also found in themes meaning ‘do’).

Athapaskan languages present particular difficulties for dictionary-making owing to their overwhelmingly prefixing nature and to the character of their ‘themes’. The solution adopted by Jetté & Jones (2000) in the new Koyukon dictionary is an alphabetical listing of verbal forms by root. The core information contained under each entry is an array of stem-sets (see 3.1.2) under which specific theme strings are listed with indication of theme category membership and derivational extensions (including important nominalizations).

### 3.1.1. Roots and theme-formation

The number of verbal **roots** in the language – and even of basic **themes** containing them – is rather small, but their derivational potential is considerable. Most Koyukon themes belong to one of about eleven major theme categories according to the basic type of state-of-affairs they refer to. The relevant parameters are ± dynamic, ± control and ± telic, with a few further semantically determined subdivisions. ‘Operative’ themes thus refer to actions that normally take a certain length of time, whereas ‘successive’ ones refer to actions usually consisting of a repeated series, and ‘conversive’ ones focus on the result of an action. ‘Dimensional’ themes cover predicative adjectival meanings of all sorts.

Belonging to such a theme category is not just a semantic matter: it also entails taking one or more basic aspects, for instance the ‘durative’ in the case of ‘operative’ themes. Stative themes have fewer options here than active ones. There are also derivational strings (see 3.2) associated with individual theme categories. The categories are also cross-linked in terms of relations of a quasi-derivational nature. Thus ‘classificatory motion’ themes are paired with ‘classificatory stative’ ones (which take a basic ‘neuter’ aspect) but take derivational strings typical rather of ordinary ‘motion’ themes.

The latter two categories represent a particularly salient trait of all Athapaskan languages. The themes concerned have a very general meaning such as – stative – ‘a compact object is there’ (with inanimate subject) or – motion – ‘handle a compact object’ (with animate subject). The other ten pairs are meant for objects that are respectively rigid and stick-like; in an open container; plural (all kinds of object); enclosed or bag-like; flat and flexible; edible; burning; disorderly or scattered; mushy or wet; and granular. There are other roots with a broad sense of ‘being’, ‘doing’ or ‘happening’ which take on a more concrete sense by derivation, for example *lit-* ‘happen to singular experiencer’ as in *ts'aanislit* ‘I woke up’ with prefix *ts'aa* ‘out into the open’. Themes in specific derived forms (mainly aspectual, but incorporates and gender marking can also be involved) may become lexicalized and this results in individual defective themes or in small minor theme categories.

Each major theme category can be viewed as a prototypical action/state type associated with particular valency conditions. Some themes fit more centrally into such a schema, others are more peripheral and are more liable to be drawn into the ‘sphere of influence’ of other categories. Thus (*h*)*on-* ‘eat’ is more central to the category ‘operative’ than is *yotl-* ‘snow’, though both largely share their derivational potential. It is important to distinguish themes from roots here, since the same root may occur across theme category boundaries, sometimes in widely differing meaning. Derivational processes can often be seen to relate themes sharing roots in this way, and these can thus be said to enter into lexicalized ‘theme formation’.

### 3.1.2. Stem-sets

Every stem is associated with an array of ‘stem-sets’, root-plus-suffix allomorphs, one for each ‘tense-mode’ (see 3.4). There is one for each aspect it may appear in. Stem-set variation patterns are discussed in 3.3.1. Much of the irregularity associated with individual themes seems to result from conflict between the inherent aspectuality of more peripheral themes (especially those of the large, rather heterogenous ‘operative’ category) and the theme category to which they belong. Thus a stem variant associated with a particular aspect may intrude into a stem-set where it becomes associated with a different aspect, e.g. when a durative form of an oper-

ational theme – whose basic aspect is durative – is drawn into a momentaneous set via a derivation requiring the momentaneous aspect. Such deviations can easily become lexicalized.

There are also stem-set idiosyncracies such as irregular ablaut types that have a purely phonological origin. Suppletion is another, more limited, source of morphological irregularity. The stem *yoh-* of theme *di ... di-yoh* ‘become thus’, for example, is only found in the perfective, whereas another root altogether, *naa-*, is found in the other three tense-mode allomorphs.

### 3.2. Productive derivation

Before a ‘theme’ reaches the stage where a particular stem-set alternant is chosen and inflection may begin, one or more of a large number of derivational processes may apply to it. These fall into two main categories, ‘non-aspectual’ and ‘aspectual’ respectively. The first type generally affects the valency of the predicate (as reflected in ‘classifier’ choice) but does not determine a particular stem-set, whereas the second type does not usually (but may) affect valency and does always determine a specific aspectual stem-set choice. This is most commonly one of the variants associated with ‘momentaneous’ aspect when the theme is active or one of the ‘neuter’ variants when it is stative. The use of the expression ‘(non-)aspectual’ here should not be taken too literally: it is more a formal distinction between strings that do or do not determine a specific choice of aspect. Some ‘non-aspectual’ strings contain elements of an aspectual nature, while many of the ‘aspectual’ ones have little if anything to do with real aspect. There are over three hundred derivational strings, of which by far the greater number are ‘aspectual’. Their degree of productivity varies considerably.

#### 3.2.1. Non-aspectual derivation

‘True’ derivational processes that increase or decrease **valency**, such as causativization and passivization, are the most central here. The basic causativizing string *will*, for example, change any classifier associated with an intransitive theme to *l-* as well as introduce a pronominal object slot (and for some theme types a gender agreement slot). Valency-reduction strings *will*, on the other hand, trigger the so-called ‘D-effect’, whereby a *li* classifier is voiced to *li* and a zero-classifier changes to *di*. Another type of process in-

volves the incorporation of a nominal, as object or adverbial or, when it is a ‘force’ low in animacy, also subject (see Axelrod 1990). This is particularly common with classificatory themes. The incorporate is often a body part but may be from a wide range of inanimate noun stems (never animate and never inflected for number or possessor). Both types of process may also be involved in non-productive theme-formation (as in 3.1.1).

An example of object incorporation is found in the word for ‘porcupine’ in text sentence (11), a nominalization, immediately followed by a non-incorporative equivalent explaining the name. A purely verbal example can be seen in *neetlee'eet'onh* ‘he put his head (*tlee*) in place to rest’ with classificatory motion theme ‘o- ‘handle compact object’. Subject incorporation, with *to* ‘water’, is seen in *toyo'ol* ‘it is floating along’, literally ‘water is causing it to move’. Adjective stems may also incorporate as in *binotsil'eeditlaakk* ‘he came home soaking wet (*tsil*)’. An example of an incorporate used adverbially is seen in *yineeghaalnee'onh* ‘he bundled it up’ (*ghaal* ‘bundle’).

The remaining non-aspectual strings consist of individual prefixes such as *di(... ti)* ‘in a line’ and *no*, the so-called ‘iterative’, meaning ‘there and back, again’, both of which trigger the ‘D-effect’ though they do not semantically affect the valency of the theme. They may combine with other strings and do not determine a specific stem-set. These have been called ‘post-aspectual’ since they do not seem to enter into lexicalized theme-formation as other non-aspectual strings may. It seems preferable, however, to regard all non-aspectual derivation as prior to aspectual derivation, in so far as once a single aspectual string is applied no further derivation is possible. This is an important distinction between the two types of derivation, though in fact they do not generally interact in any ‘feeding’ relationship.

#### 3.2.2. Aspectual derivation

A sample of the many **aspectual** derivation strings in Koyukon is given on Tab. 130.1 in order to give an idea of the range of meanings and formal elements involved. They cover, for example, path, ground, phase, direction, distribution, rate, manner and aspect in the narrow sense, all elements which can be called predicate ‘satellites’ (Talmy 1985: 102). The association of a given string with a particular imperfective/perfective

(with <i>ni</i> -momentaneous:)	
<i>do</i> # <i>di</i>	'all over'
<i>P</i> -' <i>ots'a</i> ##	'to P's house'
<i>hunee</i> #	'marrying'
<i>neel-tugh</i> #	'through a portage between lakes'
<i>nilaa</i> # <i>oo-gha-ni</i>	'attaching together'
<i>kk'aa</i> # (+D-effect)	'stopping, finishing'
(with <i>gha</i> -momentaneous:)	
<i>P-a</i> # <i>di</i>	'releasing or leaving P behind'
<i>bil-di</i> #	'down over bank'
<i>diki-ggaza</i> # <i>di</i>	'in fork of tree'
<i>neekk'aats'a</i> <i>yee</i> ##	'going by an indirect route'
<i>tlee-do</i> #	'wrongly, abnormally'
(with <i>li</i> -momentaneous:)	
<i>'ada</i> # <i>oo</i>	'continuing past'
<i>do</i> # <i>oo-gha-di</i>	'sealing the end closed'
<i>neel-no-binh</i> # <i>ni</i>	'criss-crossing'
<i>P-baa</i> #	'off the end of P'
<i>P-aa-lugha</i> ## <i>no</i> #	'flipping P over'
(with persistive:)	
<i>P-i</i> # <i>k'i</i>	'go being restrained by P'
(with conclusive:)	
<i>to</i> # <i>di</i>	'treating skin or pelt'

Tab. 130.1: Koyukon aspectual derivational strings

tense-mode prefix (*li*, *gha*, *ni* or zero) for the 'default' momentaneous and neuter aspects is largely arbitrary, though the morphemes concerned once had specific meaning that may still be glimpsed, *ni* for example often being used for action terminating in some way, whereas *gha* may mean initiating a continuing action.

These strings may contain further 'conjunction' or 'disjunct' prefixes (the latter indicated by a following # – see 3.8). There may also be slots for incorporated nominals, pronominal objects, or postpositions and their own pronominal 'objects' (full nouns governed by the latter will stand immediately before the verbal complex). Examples of such incorporation are seen on Tab. 130.1 and in sentence (6) of the illustrative text, where the string is *P-aa#hut-di* 'depleting on P', which

takes zero-momentaneous aspect. The P indicates the postpositional object of incorporated postposition *aa*.

Sometimes adverbials or other external elements are included in an aspectual derivational string – they may be called 'preverbs' and must occur immediately before the verbal complex (this is what the word boundary marker ## indicates). An example is seen in (9) of the illustrative text, where the string meaning 'down the bank' is *nukk'u* ## *di-ee*, taking the zero-momentaneous or zero-neuter aspects (the conjunct prefix *di* has merged with 1st person plural marker *ts'i*).

### 3.3. Aspect choice

The point at which derivation stops and inflection begins is not always clear in descriptions of Athapaskan languages. The reason is that there is an area between 'true' derivation and 'true' inflection which partakes of the nature of both. This area of 'extended derivation' extends as far as – and includes – the assignment of a specific aspect, which is determined, as seen already, either by a specific aspectual string or by choice of a 'basic' aspect. The latter can, in Functional Grammar terms, best be expressed by abstract operators, as for true inflectional categories, since it is a matter of the obligatory selection of one of a limited number of mutually exclusive grammatical items. This contrasts with lexically expressed satellites that modify successive levels of the clause structure. Just as Koyukon can be said to have predicate-internal satellites (the highly diverse, non-obligatory aspectual derivational strings of 3.2.2) it can also be said to display predicate-internal operators, as here. However, the product of the selection here is still not a fully specified morpheme string: it is available for further optional 'superaspectual' elaboration, which in turn supplies the input to paradigmatic inflection for tense-mode and person.

#### 3.3.1. Basic aspect

There are in all 14 basic aspects, namely the momentaneous, reversative, persistive, perambulative, progressive, repetitive, and continuative (all primarily used with themes), and the neuter, transitional, durative, conclusive, semelfactive, seriative and bisective (primarily used with non-motion themes). There are a couple more stem-sets for specific theme sub-types, such as the 'onomatopoetic', a fixed set corresponding to the seriative or semelfactive of operative themes of that

	imp.	perf.	fut.	opt.
<b>momentaneous</b>				
(ni, ni)	(Ø, gha)	(Ø, Ø)	or	(li, li/Ø)
open	yh	nh	l	l(or ')
closed	L(sp)	—	S(sp)l	L(sp)
<b>neuter</b>				
(Ø, gha)	(li, gha)	(ni, gha)	or	(gha, gha)
open	nh	,	,	,
closed	—	—	—	—
<b>customary (Ø, gha)</b>				
open	yh	k	k	throughout
closed	S(sp)	(k)		
<b>root (y)o- ‘singular goes’:</b>				
neu.	yo	yo'	yo'	yo
mom.	hoyh	yo	hol	yo'
	/yo'	yo	hol	
dist.	ho'	yo	hol	hol
cont.	ho	yo	hol	hol
	/ho'	yo	hol	
per.	doyh	yo	dok	dok
cust.	hoyh	hok	hok	hok
pers.	hoh			
prog.	hot			
<b>root daatl- ‘plural go’:</b>				
neu.	daatl			
mom.	daal	daatl	dil(tl)	daal
dist.	daal	daatl	daal	daal
per.	daal	daatl	diyhtl	daal
cont.	daal	daatl	daal	daal
rev.	dil	daatl	dil	dil
dur.	daal	daatl	daal	daal
pers.	dil			
sem.	dil			
ser.	dil			
rep.	diyhtl			
cust.	diyhtl			
prog.	dil(tl)			

Tab. 130.2: Selected Koyukon stemsets

type. Other minor variants are for specific derivational strings in combination with certain theme types, such as the ‘errative’ string applied to neuter themes, but these are of very limited applicability.

On Tab. 130.2 the stem-set variations associated with two of the basic aspects plus one

‘superaspect’ are presented. The tense-mode prefixes for respectively imperfective and perfective are given in brackets in the first line (there are several possible for momentaneous and neuter, depending on derivation or theme category). The columns indicate the abstract changes to the root – the basic form of the stem – caused by choosing the tense-mode category (see 3.4) for the aspect concerned. ‘S’ means shortening, ‘L’ lengthening, and ‘sp’ spirantization of a final stop. Items in brackets only apply under certain conditions and ‘—’ indicates no change (but the apostrophe indicates a glottal stop). There is a conspicuous difference between the variation of ‘open’ (vowel- and nasal-final) roots and ‘closed’ ones. Ablaut changes to certain types of roots are not indicated. The complications here are largely due to redistributions of the handful of aspectual suffixes historically involved (see Leer 1979).

Tab. 130.2 also contains the concrete stemsets for two motion roots, an open and a closed one. There is suppletion within the first one (yo- and ho-). Some of the aspects indicated are not usually associated with motion themes: these are found only in certain derivations or in themes containing the root that actually belongs to a different (successive) theme category, for example *daatl-* in the sense ‘strike against an object’. Forms following an oblique are for combination with the conative superaspect (see 3.3.2).

### 3.3.2. Superaspect

There are four aspects which, unlike the others, can combine with a basic aspect or cumulatively amongst themselves, namely the customary, the distributive, the conative (‘try to’) and the multiple. Their behaviour in this respect can be seen from the following examples with theme *ghaa-* ‘carry on back’ and aspectual derivational string *nee* ‘to a point’ (requiring the *ni-* momentaneous aspect), where each successive superaspect overrides the stem-set of the preceding stage. Traces of each of the latter remain, however, in the form of associated prefixes.

- (1) (a) *nee-yee-neeh-ghaanh*  
‘he carried it to a point’
- (b) *nee-n-yee-l-ghaanh*  
‘he carried it to several points’
- (c) *nee-y-oo-naa-l-ghaanh*  
‘he tried to carry it to a point’
- (d) *nee-y-oo-naa-l-ghaak*  
‘he customarily tried to carry it to a point’

- (e) *nee-n-y-oo-naa-l-ghaak*  
 ‘he customarily tried to carry it to several points’

The stem is perfective in each case, but in (1 a) it is that for momentaneous aspect, in (1 b) for the distributive and in (1 c) for the conative (which happen to be the same here), and in combinations (1 d) and (1 e) it is that for the customary that dominates in both cases (the aspect with the widest scope). The conative and distributive are associated with perfective prefix *l(i)* and the customary with *gha* (all replacing the *ni* associated with the derivational string). The conative is further marked by conjunct prefixes *oo* and *niihaa* in that order, whereas the distributive has its own disjunct *ni* prefix. The *y(i)* prefix common to all the examples is the 3rd singular object marker.

Since these steps can apply after several non-aspectual derivational strings, it will be appreciated that quite complex verbal forms can be built up. There is, as might be expected, a certain amount of lexicalization involved in superspect combinations – especially those involving the conative, which is not found outside of Koyukon as a superspect and is presumably the most recent member of this category. It does not usually change closed root stems but may have its own set consisting of one or more allomorphs according to the aspect it is combined with.

#### 3.4. Tense-mode inflection

The verbal complex following choice of aspect can be called a ‘base’. All the information necessary for generating inflectional paradigms at the following stages can be read from such a string, which contains a single stem-set of allomorphs. Paradigms can thus be given for each aspect of a given theme, according to the categories of tense-mode and person-number. These are quite general patterns, irrespective of whether the theme itself is derived or basic and contains slots for object and/or gender agreement or not.

The mutually exclusive **tense-mode** choices are respectively: imperfective, perfective, optative and future, plus a fifth, the inceptive, which is a variant of the perfective (always taking a *li* prefix). It could be analysed as an aspectual derivation if it were not for the fact that it does not go with a specific aspectual stem-set but rather can apply to a wide variety of aspects just like the (other) tense-mode categories. Choice of tense-mode specifies

one stem-set allomorph in the ‘base’ plus a particular tense-mode prefix. In the case of the imperfective and perfective the latter has to be specified in the stem-set, as seen; the optative and future always take a *gha* prefix, and the future further takes the same conjunct prefix *ti* as the inceptive.

In meaning these categories cover both aspect (of a highly general ‘phasal’ sort), tense and modality. The imperfective is very often used to refer to present time and the perfective to the past, but this is probably a secondary function; the optative, on the other hand, has a basic deontic sense of ‘should’. Negation also belongs to this level of inflection, although it can be combined with any of the tense-modes. The prefix (*li*, *hi*, zero) and/or enclitic suffix (*aallaa*, optative *yu*) involved depends on the particular combination.

#### 3.5. Person-number inflection

The second stage of inflection of a base is for pronominal **subject/object assignment** – 1st, 2nd and 3rd person, singular or plural. This is a matter of prefix choice, subject markers in a slot immediately before the ‘classifier’, object markers furthest from the stem within the ‘conjunct’ complex (see 3.7). Going with the latter are also a reciprocal and a reflexive marker, which trigger the ‘D-effect’ on the classifier (they do not for that reason belong under non-aspectual derivation). There is also an indefinite marker *k'i* and an areal one *hu*, both used in object or, less commonly, subject function. The latter is also a gender agreement marker (see 3.6): a number of common prefixes in Koyukon, occupying a single morphological ‘slot’, fulfil more than one function – usually gender and lexicalised as part of discontinuous themes.

A notable anomaly in this area is the placement of 1st plural *ts'i* and 3rd plural *ha* subject markers along with the object rather than the other subject markers (this includes special 3rd plural subject plus 3rd singular/plural object combinations). The 1st plural also has another sense of ‘someone’. There are further complications with 3rd person object markers, a distinction being made between the prefix used with a 1st/2nd person subject or when the object is ‘topic’ (usually meaning higher in animacy than the subject) and that used with a 3rd person topic subject. 3rd singular subject is unmarked, as is 3rd singular object when it would cross-reference an overt nominal object. Traces of an old perfective prefix (the ‘y’-element, realized as

	imperfective	perfective
1.SG	<i>nisbaayh</i>	<i>nisbaanh</i>
2.SG	<i>neebaayh</i>	<i>neenbaanh</i>
3.SG	<i>eebaayh</i>	<i>neeabaanh</i>
1.DU	<i>ts'eebaayh</i>	<i>ts'ineebaanh</i>
2.DU	<i>nuhbaayh</i>	<i>nuhbaanh</i>
3.DU	<i>heebaayh</i>	<i>haneebaanh</i>

Tab. 130.3: Paradigms for *ni*-momentaneous of *baa-* ‘swim’

*n(i)* or *ee*) show up in certain 3rd singular subject inflections (as well as 2nd singular and 1st/3rd plural), for example in the perfective of the *ni*-momentaneous.

Sample paradigms for motion theme *baa-* ‘swim’ with zero-classifier) are shown on Tab. 130.3. The basic *ni*- momentaneous form here means ‘I am arriving swimming’, etc. Note that the plural forms are interpreted here as dual: an unrelated suppletive root is used for plural subjects. *ni* (both the aspect and the 2nd person singular marker) is variously ‘absorbed’.

### 3.6. Gender agreement

Certain bases have a slot for gender agreement with their subject or object, whether these are explicitly expressed or not. This is either part of the theme itself (notably in the case of classificatory ones) or introduced as part of some specific non-aspectual derivational string. Since the categories of covert nominal **gender** in Koyukon cut across the subject/object categories associated with classificatory themes, the combinations arising in connection with the latter are quite fine-grained and allow for considerable stretches of discourse without overt nominal arguments – especially for ‘given’ topics – since the referent can often be determined solely from gender-marking in the verbal complex. This is distinct, note, from pronominal subject and object inflection. With stative themes the gender agreement is with the subject, whereas with transitive ones it refers to the object or the form in which the object is present, e.g. within a container. One marker, *ni*, may refer specifically to the face of the subject or object (it can appear lexicalized also with themes not usually marking gender agreement).

The gender categories are each associated with a prefix: zero (the largest, default cat-

egory), *di* (which includes deep containers and their contents and anything made of wood), *ni* (which includes long, flexible, and small round objects), *hut* (area, event, or abstract), *dini* (including long, firm and compact, heavy objects), and *hudi* (weather phenomena). Only *hut* may occur as the object of a postposition (incorporated or independent). The following two examples illustrate the interaction of gender marking and classificatory themes (here *ton-* for flat or stick-like objects):

- (2) (a) *tlok neetonh*  
‘he brought a dish’ (zero gender)
- (b) *kkun' daaneetonh*  
‘he brought a stick of firewood’ (*di* gender)

### 3.7. Clausal enclitics

All productively attached morphemes in the verbal complex that follow the stem are enclitic, in a fixed order relative to each other: negative, emphatic/contrastive, relativizer/nominalizer, then interrogative. Emphatic/contrastive enclitics such as *eehu* ‘in vain’ may occur as independent particles, so enclitics – apart from the negative ones mentioned in 3.4, which must appear immediately after the stem – are usually written as separate words. Interrogative *hee*’ is written attached to the stem when no other enclitic intervenes (also to predicative adjectives of fixed form). Subordination markers following verbs like *daa* ‘if, when (in future)’, *dahoona* ‘while’ and general subordinator *ts'a*, can also be regarded as clausal enclitics. The relativizing enclitics are treated under 4.2.

### 3.8. The interdigitation of discontinuous strings

The ordering of morphemes within the verbal complex can be presented in the form of a template of slots, as on Tab. 130.4, where topmost = leftmost. This is somewhat simplified (several distinguishable slots for ‘thematic’ elements having been collapsed), but it reflects the overall order of complexity. Glosses of slots that are obligatorily filled are shown in italics. ‘Disjunct’ prefixes are separated from ‘conjunct’ ones on the basis of special morphophonemic properties of the boundary between them (as well as by the semantics). Further zones within these major areas can be discerned, for example the cluster of object pronominal slots. Some of the slots are specific to a single item or to a small

disjunct prefixes	interrogative (Q-word) obj. of postposition postposition adverbial-thematic distributive iterative incorporate
conjunct prefixes	direct object indefinite 3 pl. subject 3 pl. object 1 pl. subject area (object) thematic conative gender inceptive gender-thematic negative <i>tense-mode</i> <i>subject</i> (non-3 pl./1 pl.) <i>classifier</i> <i>stem</i>
enclitics	negative emphatic/contrastive relativizer/nominalizer subordinator interrogative

Tab. 130.4: Slots in the Koyukon verbal complex

handful of items and there is considerable mutual incompatibility between positions. A given complex verb form will have only a small number of potential slots filled. The restriction of aspectual derivation to one string per word is a principal restraint here. However, all positions can be distinguished by computational testing. The actual number of slots is somewhat fuzzy, as in all Athapaskan languages, depending in detail upon the analysis (e.g. on whether metathesis of morphemes is allowed). But the fact that such templates can be set out at all is symptomatic of the structural rigidity of Athapaskan languages as compared with the highly recursive Eskimo ones.

The way in which the fully specified but largely discontinuous strings constituting the verbal complex are linearized according to this ‘scrambled’ template can be grasped in terms of successive expression operations working outwards from the stem. Material less and less directly associated with the latter

can be seen to be attached ‘centrifugally’ according to the logic of the underlying ‘layered structure of the clause’ (Fortescue 1992: 122 ff.). Each successive choice on the way to the specification of a base, note, renders more and more of the overall template irrelevant to the derivation at hand. What the template specifically does is to take care of the interdigitation of the discontinuous prefixes whose head is either the stem itself or one of the adjacent inflectional category slots affected by a derivational string. The mutual ordering of the obligatory head items themselves is far from arbitrary and, despite overall appearances, is largely consonant with Bybee’s generalizations in this area (Bybee 1985: 196 ff.).

This can best be illustrated by following a single linearization sequence, namely for verbal complex (1 e) in 3.3.2, whose layered structure according to semantic scope can be taken to be:

(3) *PERF [CUST CON DIST [(MOM) [ghaa- (x<sub>1</sub>) (x<sub>2</sub>) nee]]]*

As it happens, stem *ghaa-* has a zero-classifier and the subject (*x<sub>1</sub>*) is 3rd person singular, so these can be ignored. The item in the input base nearest to the stem is thus the pronominal object *x<sub>2</sub>*, which will be linearized with the stem as *yi ... ghaa-* (the dots indicate disjunction). The predicate satellite *nee* is then integrated to produce *nee # yi ... (ni, ni) ghaa-(yh/nh/tl)*, where the *ni*-momentaneous aspect and associated stem-set has been determined by the satellite. Its relative position is fixed according to the template. However, the subsequent application of superspect operator *DIST* overrides *NOM* to produce *nee # ni # yi ... (Ø, li) ghaa-(-nh/kl')*. Next, the application of *CON* produces *nee # ni # yi ... ooni ... (Ø, li) ghaa-('nhl'l')*, with its own stem-set, whereafter the last superspect operator *CUST* produces *nee # ni # yi ... ooni ... (Ø, li) ghaa-(yhl/klk/k)*. Application of *tense-mode* operator *PERF* finally selects a particular stem allomorph and tense-mode prefix and the resultant string is fully specified morpheme sequence *nee-ni-yi-ooni-li-ghaak*.

The phonological rules required to produce actual surface forms – including insertion, deletion and absorption processes – are rather complicated and are dependent to a high degree on specific morphemes preceding or following. Their application seems to parallel the centrifugal linearization process illustrated above.

#### 4. Nominal expressions

Underived nouns in Koyukon are a simple matter. Nominalization of complex verbal forms can, however, produce very elaborate nominal expressions.

##### 4.1. Nouns

The only overt inflectional or derivational categories relevant to nouns are possession and, to a more limited degree, plural marking. They may also have inherent gender (see 3.6) and/or 'areal' status – for example *yah* 'house', which can also mean 'in the house' and the like, without requiring a postposition. Only words for humans and dogs take a plural marker – enclitic *kkaa* or *yoo*, the latter often in a more general, collective sense. *kkaa* is also used on proper names in the sense 'X and family/companions'. A couple of nouns have a special plural form, e.g. *solt'in* from *solt'aanh* 'woman'.

Possession is marked by a personal possessor prefix or a preceding possessor noun, which in both cases causes voicing of an initial voiceless continuant on the possessed noun and combines with a suffix *-a* on the latter (itself causing voicing of any final voiceless consonant). Body part and kinship terms are inalienably possessed, that is, they must always be marked for possessor, though this can also be indefinite prefix *k'i* as in *k'itlee* 'a head, s.o.'s head'.

Some nouns (especially body part terms) have special compounding and incorporative forms, for example *to* (or *taa*) from *too* 'water'. True compounding – as opposed to complex nominalization – is rather limited and lexicalized, as in *soltoo* 'salt water' and *hut'lil* 'sled cover' (*lil* = 'skin, cover'). Note also *tobaan(a)* 'beach' in the sample text, where the second element is usually a postposition ('at the edge of'). However, quasi-compounds consisting of possessor plus possessed nouns are common, for example *dinaa kkanaaga* 'Koyukon' (from *dinaa* plus *kkanaak*, literally 'human being speech').

Appositional adjectives of fixed form (as opposed to dimensional verbs, which may also be appositional but are always fully inflected) can be regarded as enclitic to a preceding nominal head, as in *da'aak kuh* 'big dress' (compare *koh* 'be big', the predicative, inflecting equivalent requiring gender agreement). Other nominal enclitics include *zo* 'instead' and *yaan* 'only'.

Numerals have different forms depending on whether they refer to humans or non-humans, thus *k'eelakk'ee* 'one thing/animal' but *k'eelakk'an* 'one person'. Compare also *k'eelts'in* 'in one way', *k'eeldinh* 'once, in one place' and *k'eelhu* 'in one area or direction'. Demonstratives also have special forms for humans, *goninh* 'this person' (*go* 'this'), *eyedinh* 'that person (near you)' (*ey* 'that'), and *nughuninh* 'that person (over there)' (*nugh* 'that').

##### 4.2. Nominalizations

Nominalizations can be as complex as verbs, in fact even more so since they may involve whole clauses with several independent words (and incorporates) that have been nominalized by one of several 'relativizer' enclitics. The first, (*y)ee*, is used for non-humans (the *y* is present following vowels); the *ee* may be dropped, leaving just *y* or a voiced final consonant (with or without a following *a*). For single humans (*n)inh* is used, for plural humans *na*. The first two cause voicing of a stem-final voiceless consonant. Thus *daalit'lidza* 'black bear' (*daalit'lits* 'it is black'), *taal'ona* 'fish trap' (from imperfective stem '*onh* 'compact object is there' and incorporate *taa* '(in) water'), and *nindibidza* 'flying squirrel' ('one that flutters here and there', from theme *ni*- ... *di-bits*). These are generally written as one word except when another clause enclitic intervenes as in *nanee'ots eehu na* 'the ones who came in vain' (stem '*ots*).

As examples containing more than one word observe: *t'eyagga hut'aaninh* 'Indian, native' (from *t'aa-* 'be thus' with areal prefix; *t'eyagga* 'right down there'), and *baahaa nik'idinlidige* 'pencil' (from theme *ni* # *dini* ... *li-dik* 'write' and indefinite object *k'i*; *baahaaa* = postposition plus object 'with it'). With transitive verbs note that relativizing nominalizations may refer to either the subject or the object, thus *keel yineel'aaninh* 'the one the boy is looking at' (with pronominally marked object), but *keel neel'aaninh* 'the one who is looking at the boy'.

There are also relativizers of place such as *dinh*, *hu*. See sentences (10) and (19) of the text and note also, for example, *notaaleedinh* 'Galena' ('place where strong current goes around'). Enclitic *tno* 'river' and *mina* 'lake' form compound place-names like *dolbaatno* 'Dolbi (= gosling) river', where *dolbaa(ya)* is itself a nominalization of stem *baa* 'is grey' with a derivative prefix meaning 'on top'.

Place and species names, as in all Athapaskan languages, often consist of colourful descriptions in the form of nominalizations.

## 5. Directionals

Directional expressions have their own morphological pattern that mirrors that of verbs but is of a much simpler nature. They are used much like areal nouns but generally require a prefix on the following verb. They consist of a prefix like *do* ‘nearby’, *no* ‘at intermediate distance’ and *aa* ‘far away (but in sight)’, followed by a directional stem such as *naa'(a)* ‘upstream’, *do'(u)* ‘downstream’, *yagga* ‘down’ or *an* ‘away’, plus a suffix –*ts'in* ‘from’, *t* ‘at specific place’, *ugha* ‘in general area’ or zero ‘heading towards’. Some other suffixes, notably *hu* ‘area’ and *ts'ina* ‘side’, do not require a prefix at the same time. There is some irregular lexicalization in specific directional combinations.

The system reflects riverine orientation, distance away or towards the shore and towards the middle of the river being other relevant parameters. There are similar temporal expressions consisting of stem plus suffix, e.g. with *dzaanh* ‘day’: *dzaana* ‘earlier today’, *dzaandaa* ‘later today’, and *dzaaneets* ‘midday’.

Directionals with enclitic *ts'ina* ‘on the – side of’ may function as postpositions, that is follow a governed nominal or a pronominal (possession) prefix, like all other postpositions.

## 6. Illustrative text

This is taken from the traditional story “To-baan Atsah”, ‘She’s crying on the beach’ (Attla 1983). Neuter aspect is not glossed for stative verbs, nor is 3rd singular subject or in general any zero morpheme. ‘Classifiers’ are not separated from verb stems unless derived. Discontinuous morphemes are glossed under their principal element (the first prefix for derivational strings, the stem of verb themes); the other elements of the morpheme are referenced by the same number alone (or attached to an aspect-marker or classifier required by it). Co-determination of a stem-set by a derivational string is not indicated in this way. Enclitic and compound morpheme boundaries are indicated by ‘=’. Note that the first person plural (or dual) is often used in narrative about a 3rd person protagonist

as in sentences (12) and (21). In sentence (22) there occurs another narrative convention related to this, namely the rendering of ‘he said’ by literally ‘(when) he said I said’.

Note also the 3rd person singular for 2nd person singular in sentence (26).

- (4) *Yagga gheel*  
then apparently  
*ts'uhu-t'aanh*  
and.thus-happen:MOM.PFV
- (5) *daghel yoogh tsook'aal*  
apparently long.ago old.woman  
*li-do*  
IPFV-live
- (6) *daghel go li-do*  
apparently this IPFV-live  
*ad-aa-k'-uh-dee-d-on'*  
REFL(1)-depleting.on.one(2)-INDEF(OBJ)-  
(2)-(2)-CLD(1)-eat:MOM.PFV
- (7) *at'eylo dikinh kkotsil yeen'*  
it.seems log bark:POSS only  
*ahonh*  
eat:DUR.IPFV
- (8) *huyil yoo-naan*  
and.so far-across  
*hu-n-ee-l'aanh*  
AREA-(1)-3.SG.SBJ-see:DUR.IPFV(1)
- (9) *huyil gheel yoo-naan*  
and.so apparently far-across  
*yoo-do' hu-kk'aa*  
far-downriver AREA(ABST)-wanting  
*di-nt'aa*  
thus-be:IPFV
- (10) *yoo-nts'a no-ts'in*  
far-across over.there-on.that.side  
*dikinh hoo-laanh=hu*  
tree GENDER(AREA)-be=NOM (PLACE)
- (11) *at-eylo dikahona*  
it.seems porcupine  
*dikinh ahon=a*  
wood eat:DUR.IPFV=NOM
- (12) *kk'udaa no-tlin*  
now over.there-towards.river  
*nukk'u-z-ee-'ots*  
down.bank(1)-1.DU/PL.(1)-(1)-DU.go:-  
MOM.PFV
- (13) *ts'uh to-baan ts'i-l-do*  
and water=edge 1.PL-IPFV-sit

- (14) *dahoон gheel*  
meanwhile apparently  
*saa-l-tsaah*  
1.PL.INCP(1)-(1)-cry:DUR.PFV  
*saa-l-tsaah*  
1.PL.INCP(1)-(1)-cry:DUR.PFV
- (15) *huyil hidaagh*  
and.so surprisingly  
*do-neets'a h-un*  
nearby-from.upstream AREA-to  
*bikinaal gha-baal*  
muskrat IPFV-swim:PROG
- (16) *do-tlee b-ugh*  
nearby-towards.river her-to  
*nee-yee-nee-baan*  
up.to.a.point(1)-3.SG.OBJ-PFV(1)-swim:MOM
- (17) *tsook'aal do-dee-nee*  
grandmother what-(1)-say:DUR:IPFV(1)
- (18) *do-d-ee-nee ts'a go*  
what-(1)-2.SG-say:DUR.IPFV(1) SR this  
*ni-tsah tsook'aal*  
2.SG-cry:DUR.IPFV grandmother  
*ahnee gheelhee*  
say.to:DUR.IPFV apparently
- (19) *koyaa' yoo-nts'a yoo-do*  
grandchild far-across far-downriver  
*ts'ibaa yil kk'eeyh yil neel-ta*  
spruce with birch with REC-among(1)  
*k'a-gha-naa-di-yonh=da*  
INDEF-form.solidly.into(2)-(2)-CLD(1)-grow:  
:MOM.PFV=NOM(PLACE)
- (20) *hut-kk'aa ees go*  
AREA-wanting that this  
*di-di-s-nee*  
thus-(1)-1.SG-say:DUR.IPFV(1)
- (21) *haa tsook'aal do-nee*  
haa grandmother near-behind  
*si-nee*  
1.SG-behind  
*nee-n-ee-hoyh*  
up.to.point(1)-IPFV(1)-2.SG-go:MOM  
*bee-z-nee*  
3.SG.OBJ-1.PL.(1)-say.to:DUR.IPFV(1)
- (22) *si-kaal tlee-kk'a*  
1.SG-tail out-on  
*do-l-ee-hoyh*  
up.on(1)-IPFV(1)-2.SG-go:MOM  
*yi-lnee*  
3.SG.OBJ-say.to:DUR.IPFV
- (23) *yi-n-ee-l'aanh*  
3.SG.OBJ-(1)-3.SG.SBJ-see:DUR.IPFV(1)
- (24) *huyil aa n-ugh ni-kaa' holdidl=a*  
and.so aa 2.SG-at 2.SG-tail spin=NOM  
*kk'aa-nt'aa yi-lnee*  
like-be:IPFV 3.SG.OBJ-say.to:DUR.IPFV
- (25) *at'eylo n-ugh ni-kaa' tl'aah*  
it.seems 2.SG-at 2.SG-tail sinew  
*naa-lididl=a*  
GENDER(ROPELIKE)-spin=NOM  
*kk'aa-nt'aa bee-s-nee*  
like-be:IPFV 3.SG.OBJ-1.SG-say:DUR.IPFV  
*yi-lnee=dinh*  
3.SG.OBJ-say:DUR.IPFV=NOM(TIME/PLACE)
- (26) *hut nikitl tosiyookk*  
huh nose big (?)  
*do-d-nee ahaa*  
what-(1)-say:DUR.IPFV well
- (27) *go hal bi-nikitl*  
this and.so 3.SG-nose  
*hut-dinee-koh*  
AREA-GENDER(PIPELIKE)-big:IPFV  
*gheelhee*  
apparently
- (28) *kk'uudaa no-do'*  
now over.there-downriver  
*t'aanh gha-baal*  
act.thus:DUR:IPFV PROG-swim:IPFV
- “(4) Once upon a time (5) long ago there lived an old woman. (6) There where she was living she had eaten everything there was to eat, it seems. (7) It seems she ate only log bark. (8) So she was looking across (the river) (9) and then she must have wanted to cross downriver, (10) across there on the other side where there were trees. (11) It seems she was a porcupine, an eater of wood. (12) Now she went down the bank (13) and now she was sitting on the beach. (14) Meanwhile she started to cry and cry. (15) Then to her surprise a muskrat came swimming downriver close by. (16) It swam up to her down there on the beach. (17) ‘Grandma, what are you saying? (18) Why are you crying?’, it seems he asked grandma. (19) ‘Grandchild, across and downriver, where there are spruce and birch trees all grown together. (20) That is the place I’m crying (like this) for.’ (21) ‘Haa! Grandmother, go back there behind me’, he told her. (22) ‘Climb up on my tail’, he told her. (23) She looked at him. (24) Then she said to him: ‘Ah, your tail is like rope (of twisted sinew)’. (25) Literally she was saying to him: ‘your tail is like spun sinew’. (26) ‘Well, why are you crying, old tunnel-nose?’ (27) She must have had big nostrils. (28) Now there he was swimming (away) downriver.”

## 7. Uncommon abbreviations

CLD	D classifier
CON	conative
CUST	customary
INCP	inceptive
MOM	momentaneous

## 8. References

- Attla, Catherine (1983), *Sitsiy Yugh noholnik ts'in': As My Grandfather Told it*. Fairbanks: Alaska Native Language Centre
- Axelrod, Melissa (1990), “Incorporation in Koyukon Athapaskan”. *International Journal of American Linguistics* 56.2, 179–195
- Bybee, Joan L. (1985), *Morphology*. Amsterdam, Philadelphia: Benjamins
- Fortescue, Michael (1992), “Aspect and Superaspect in Koyukon”. In: Fortescue, Michael & Harder, Peter (eds.), *Layered Structure and Reference in a Functional Perspective*. Amsterdam, Philadelphia: Benjamins, 99–141
- Jetté, Jules & Jones, Eliza (2000), *Koyukan Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center

Jones, Eliza (1983), *Junior Dictionary for Central Koyukon Athabaskan*. Anchorage: National Bililingual Materials Development Center

Kari, James (1990), *Ahtna Athabaskan Dictionary*. Fairbanks: Alaska Native Language Center

Krauss, Michael & Golla, Victor (1981), “Northern Athapaskan Languages”. In: Holm, June (ed.), *Handbook of North American Indians, Vol. VI: Subarctic*. Washington: Smithsonian Institute, 67–85

Leer, Jeff (1979), *Proto-Athabaskan Verb Stem Variation; Part One: Phonology*. Fairbanks: ANCL (Alaska Native Language Center Research Papers 1)

Nichols, Johanna (1986), “Head-marking and Dependent-marking Grammar”. *Language* 62.1, 56–119

Talmy, Leonard (1985), “Lexicalization Patterns: Semantic Structure in Lexical Forms”. In: Shopen, Timothy (ed.), *Language Typology and Syntactic Description, Vol. III*. Cambridge: Cambridge University Press, 57–150

Thompson, Chad & Axelrod, Melissa & Jones, Eliza (1983), *Han Zaadlit'ee: Scope and Sequence*. Nenana: Yukon-Koyukuk School District

*Michael Fortescue, Copenhagen (Denmark)*

## 131. Montagnais/Innu-aimun (Algonquian)

1. Introduction
2. Montagnais lexical classes
3. Semantic categories encoded by Montagnais morphology
4. Formal processes
5. Inflection
6. Word formation
7. Conclusion
8. Illustrative text
9. Uncommon abbreviations
10. References

### 1. Introduction

Montagnais forms a subgroup of the Cree-Montagnais-Naskapi language of the central branch of the Algonquian language family. There are about 10,000 speakers, living in nine communities in the south-eastern area of the province of Québec, and the central Labrador area of the province of Newfoundland, in Canada. They prefer to be known

as Innu, meaning ‘person, aboriginal person’; both the terms Innu-aimun and Montagnais are current to refer to the language. The majority speak Montagnais as a mother tongue although fewer do so in the more westerly reserves. French is the second language for Québec residents and English for residents of Labrador. Increasing bilingualism in English or French has led to loss of Innu lexical items and rapid phonological change among younger speakers. Three main dialects of Montagnais (western, central, and eastern) are distinguished by phonological and morphological criteria; for example, loss of intervocalic and final short non-rounded vowels characterizes the central dialect, as well as the introduction of tone as a grammatical marker. The first written sources of the language date from the Jesuit Relations of the early 17th century, published in Thwaites (1896). Several manuscript grammars and lexicons were compiled by missionaries in the 17th and

18th centuries; the first published book appeared in 1767 (La Brosse 1767). These indicate that while over the past three hundred years the language has undergone several important phonological and morphosyntactic changes, a 17th century Montagnais text would be comprehensible to a modern speaker of a conservative dialect. The first published description of morphology occurs in Lemoine (1901) for the western dialect; modern studies include Drapeau (1979; 1991), MacKenzie (1980), Clarke (1982), and Cyr (1996). The dialect to be described reflects most closely the central form; there is, however, considerable interdialectal and intradiialectal variation. In this article, examples are provided in the recently proposed, fairly phonemic, orthography now in use in the Québec communities (see Drapeau & Mailhot 1989; Mailhot 1997), with the addition of a circumflex to mark historically long vowels.

## 2. Montagnais lexical classes

On the basis of inflectional affixes, it is possible to distinguish a class of **nominals** (nouns and pronouns) from a class of **verbals**; the latter includes lexical items which in more familiar Indo-European languages would be categorized not only as verbs, but also as adjectives (e.g. *kâ-uâpât mîtshuâp* ‘that-it.is. white house (house that is white, i.e. white house’). Both of these word classes are distinct from a third, uninflected lexical class (termed **particles** in Algonquian scholarship), which contains lexical items that correspond to Indo-European adverbs, prepositions, or conjunctions. Like other Algonquian languages, Montagnais is polysynthetic in structure, in that a single inflected verb may function as a complete clause; verbal arguments (e.g. pronominals) need not be expressed as independent constituents (e.g. *mîneu* ‘s/he gives it to her/him’). Often, concepts that would be represented as nouns in a language such as English find verbal expression in Algonquian (e.g. ‘airplane’, ‘priest’); Montagnais lexicons typically indicate a ratio of three or four verbs to every noun.

## 3. Semantic categories encoded by Montagnais morphology

Although the morphology of Montagnais appears complex, the semantic or grammatical categories that are morphologically repre-

sented in the Algonquian language family are relatively few in number. Both nominals and verbals are marked for gender, person, and number (this last category encoding simply the opposition of singular and plural). The category of case is not systematically signalled via nominal morphology, as Algonquian nominals do not code syntactic functions; rather, such thematic roles as agent and patient are determined via, among other factors, the Algonquian person/animacy hierarchy, to be discussed below. This hierarchy is one of the most pervasive features of Algonquian morphology.

### 3.1. Gender

The term **gender** is used by Algonquianists to refer to the binary distinction of **animacy** vs. **inanimacy**. The grammaticalization of this opposition is often somewhat opaque among nominals: while living entities are typically categorized as animate, certain non-living entities – arguably, objects viewed as possessing physical or spiritual power, among them certain plants, body parts, articles of clothing, and food and household items – bear animate marking. In the verb class, Algonquian languages draw a basic distinction between intransitive (i.e. one argument) and transitive (i.e. two argument) verbs; the gender of these arguments determines four principal verb types. Intransitive verbs are marked morphologically for the gender of their single argument, resulting in an animate intransitive (AI) vs. inanimate intransitive (II) opposition (e.g. *uâpishîu* ‘s/he is white (AI); *uâpâu* ‘it is white (II)'). Transitive verbs are marked not only for an (obligatorily animate) agent, but also for the animacy of their patient or theme; this results in the opposition transitive inanimate (TI), or animate agent/inanimate theme, and transitive animate (TA), where both verbal arguments are animate (e.g. *uâpatam* ‘s/he sees it'; *uâpameu* ‘s/he sees her/him').

### 3.2. Person

Like other Algonquian languages, Montagnais morphologically encodes the basic distinction of local (i.e. 1st and 2nd) vs. non-local (3rd) **speech-act participants**. These are marked in the paradigms of the **personal pronoun** and the possessed noun, as well as in verbs. Plurals involving a first person distinguish between an inclusive (PI), which represents only local speech act participants, and an exclusive (PE), which excludes reference to

the second person (e.g. *ni-mashinaikan-inān* ‘1-book-PE (our book – ie. mine and hers/his)’; *tshi-mashinaikan-inu* ‘2-book-PI (our book – ie. mine and yours)’. An indefinite (animate) third person is also morphologically encoded in Montagnais, but is restricted to the function of subject of intransitive verbs, as well as of possessor of certain inalienable nominals (see 3.4).

Algonquian languages also contain a grammatical feature that is less usual among the world’s languages: **obviation**. While this is usually treated as a subcategory of grammatical person by Algonquianists, it has also been represented as a separate grammatical feature (e.g. Déchaine 1999). The obviative (noted 3') represents a ‘further’ third person, one that is viewed as more peripheral or less topical in any given discourse than is the normal third person, or proximate. Obviation thus provides a means of distinguishing non-coreferential third persons, both within and across sentence boundaries. At the discourse level, the choice of which third persons are to be marked as obviatives is to some degree optional, and depends on the focus of interest at any given point (cf. Goddard 1990a); none the less, there is a clear tendency for syntactic subjects to appear as proximates rather than obviatives (e.g. Dahlstrom 1991). At the clause level in Montagnais, as in other Algonquian languages, there occurs syntactically obligatory obviative marking on one of the two overt noun phrases which may be used with a transitive verb, as well as on animate nouns possessed by a third person animate (e.g. *u-nakapeshākan-a* ‘3-pants-ANIM.OBV (her/his pants)'). Double obviative marking is also syntactically obligatory in possessive constructions involving three third person animates, as in: *u-tāūt-a u-tem-inu-a* ‘3-father-ANIM.OBV 3-dog-OBV-ANIM.OBV (her/his father's dog)'.

### 3.3. The person/animacy hierarchy

Algonquian languages are characterized by the following semantic hierarchy: 2 → 1 → 3 animate → 3' animate → inanimate. Any person that is higher on the hierarchy (i.e. in the above schema, to the left) takes precedence over a person that is lower (i.e. to the right). While one effect of this hierarchy is that obviatives cannot possess proximates, its most obvious effect is manifested in the transitive verb. Here, person prefixes are not those which designate the agent argument, or the subject function, but those of the argument

that is higher on the above hierarchy; a further suffix, known as a theme sign, is required to indicate the direction or orientation of the action relative to participants (for examples, see 5).

### 3.4. Other nominal categories

In Montagnais, nouns may encode **possession** and **location**, the latter by means of a locative suffix *-it* (e.g. *mītshuāp-it* ‘in the house'). Nouns may also carry **diminutive** (e.g. *mītshuāp-iss* ‘little house'), **absentative** (e.g. *Pūn* ‘Paul', *Pūnipan* ‘the late Paul') or **pejorative** (e.g. *Pūn* ‘Paul', *Pūnish* ‘old Paul') suffixes, when these are semantically congruent. Montagnais nouns are not typically marked for **definiteness**, there being no full-fledged equivalents of definite and indefinite articles in the language; however, there is evidence that the demonstrative pronoun is increasingly assuming the function of definite article (see Axelsson & Cyr 1988).

Algonquian languages display obligatory possessor marking for a small set of nouns that otherwise cannot occur independently. Such nouns include kinship terms, as well as inalienable entities such as body parts and certain articles of clothing; with the exception of kinship terms, these dependent stems are the only nouns which co-occur with the indefinite possessor prefix *mi-*, as in *mishkāt* ‘someone's leg'.

The subsections which follow deal with categories that are regularly encoded in the Montagnais verb. These include transitivity, tense, aspect, mood, and voice.

### 3.5. Transitivity

As noted in 3.1, the distinction between **intransitive** and **transitive** is morphologically marked in the Montagnais verb. Not all verbs that are semantically transitive, however, are encoded as such; there exists a verbal subset often referred to as pseudo-transitive (VAI+O), which takes intransitive rather than the expected transitive morphology, e.g. *aiāu* ‘s/he buys (something inanimate)', *ueveshitāu* ‘s/he repairs (something)'.

Montagnais displays a number of derivational processes whereby verb valency may be increased, in that intransitive verb stems may be made transitive to express such semantic relationships as causative, benefactive, reflexive, and reciprocal (see 6). A further derivational process yields what Algonquianists term a **relational form** (cf. Wolfart 1973). Both animate intransitive and

transitive inanimate verbs undergo an obligatory increase in valency when an animate non-subject argument is introduced with a role other than direct object – for example, as possessor of an inanimate object, or subject of an embedded sentence. The resultant (intransitive) verb bears an obligatory *-w/* (=*u*) suffix (e.g. *ni-uâpat-am-u-ân u-mashinai-kan* ‘1-see.it-TI.THM-RELAT-1/2.SG 3-book (I see her/his book)’ and *ni-minuenit-am-u-ân tshe-tshítûte-t* ‘1-be.happy-TI.THM-RELAT-1/2.SG PV-leave-3.SG (I’m happy that s/he’s leaving)'). The obligatory marking for an animate argument in such instances provides further confirmation for the animate-inanimate hierarchy, since no parallel is found for inanimate arguments in such roles.

### 3.6. Tense and aspect

As is typical of Algonquian languages, Montagnais displays a fairly simple **tense** system. Tense suffixes encode the binary distinction of ‘present’ (e.g. *âkush-u* ‘sick-3.SG.PRES (s/he is sick)’) and ‘past’ (e.g. *âkushí-pan* ‘sick-3.SG.PAST (s/he was sick)'). Given the distribution of these tenses, the ‘present’ might better be termed a neutral, since depending on the context it can be interpreted with past (cf. Cyr 1990) or future reference. The Montagnais future displays no distinct suffixal coding. Rather, it is expressed by a modal preverb (see 3.7).

The grammatical **aspectual system** of Montagnais, like that of other Algonquian languages, is not extremely rich. Montagnais does not mark by means of verbal affixes a perfective/imperfective distinction. Neither does it encode grammatically such distinctions as punctual, durative, progressive or continuous. The sole grammatical aspect represented in Montagnais (Cyr 1990) appears to be the **perfect**, which is represented by the preverb *tshî*; this preverb, however, is not regularly used in all dialects of Montagnais.

### 3.7. Mood

Montagnais possesses a fairly rich system of epistemic **modal distinctions**. The speaker’s judgment as to the likelihood of occurrence of an event as yet unrealized at its point of temporal reference is conveyed, on the one hand, by a small set of verbal prefixes, including the consequential future *ka* (e.g. *ni-ka-uâpam-â-u* ‘1-FUT-see-her/him-TA.THM.1/2 >3-1/2/3.SG (I’ll be seeing her/him)'). These preverbs typically also carry a more concrete, deontic sense, as in the case of *pa* ‘would,

should, might’, which covers both likelihood and moral obligation. Epistemic modality is also regularly expressed by suffixal distinctions in Montagnais. Here, there is a fundamental opposition between what Algonquianists have termed the indicative (i.e. realis or factual) and dubitative moods, the latter used of events of which the speaker is uncertain (e.g. *nipâ-tshe* ‘sleep-3.SG.DUB (perhaps s/he’s asleep)'), or which are purely hypothetical or contrary to fact (e.g. *tshímâ tüt-ak-âkue utâ-kushít* ‘if.only do-TI.THM-3.SG.DUB.PAST yesterday (if only s/he had done it yesterday)'). In temporal or conditional subordinate clauses, an *-î* subjunctive suffix may be added to the indicative, to represent an event that is potentially realizable at a point in time beyond the temporal reference of the sentence (e.g. *nipâ-iân-î* ‘sleep-1.SG-SUBJ (if I sleep)').

The Montagnais system of mood is complicated by the existence of the verbal suffixes *-tak* and *-shapan*, which render possible the expression of epistemological modality. These **evidential** affixes (termed ‘indirect’ by Drapeau 1984; 1996; see also James et al. 2001) enable speakers to qualify the reliability of what they report: they represent events of which speakers have only second-hand knowledge, whether via indirect (often sensory) evidence, inference, or hearsay (e.g. *mishpun-îtak* ‘snow-EVID.PRES (it seems that it is snowing)’, *mishpun-îshapan* ‘snow-EVID.PAST (it appears to have snowed)'). In addition, Montagnais verbs encode a category termed ‘subjective’ by Drapeau (1986), who suggests that it is used by speakers to represent an event not as realis, but rather, as how it ‘appears’ to them, e.g. *ni-shiueni-n-ua* ‘1-hungry-1/2.SG-SUBJECTIVE (I must be hungry)’. While this form is often used to represent events that are experienced in dreams, or that are vague in memory (Cyr 1990: 64), it is by no means restricted to such contexts.

Montagnais also possesses a set of imperative mood suffixes: *nipâ-o* ‘sleep-IMP.2.SG (sleep)’, *nipâ-tâu* ‘sleep-IMP.1.PI (let’s sleep)’. These are usually treated as a separate verbal order (cf. 5) by Algonquianists.

### 3.8. Voice

Within Algonquian scholarship, there has been considerable controversy as to the existence of a genuine **passive** morphology. Although the **inverse** forms (see 5) have been occasionally viewed as passives, this view is a minority one. Like other Algonquian lan-

guages, however, Montagnais displays an ‘agentless passive’. As this form is structurally intransitive although based on a transitive stem, it involves a reduction in the valency of the verb, so that only the patient is available to serve as subject (e.g. *uápat-ákanu* ‘see.it-AI.3.INDEF (it is seen)’; *uápam-ákanu* ‘see her/him-AI.3.INDEF (s/he is seen)'). In such constructions in Montagnais, it is not possible to specify via a lexical noun phrase the thematic role of agent. Thus, there is no Montagnais passive equivalent for such English sentences as ‘The man was bitten by the dog’; rather, constructions such as these are expressed by an inverse transitive animate form.

#### 4. Formal processes

In Montagnais, as in the Algonquian language family in general, both inflection (cf. 5) and derivation (cf. 6) typically involve processes of affixation – in particular suffixation, which is, for example, used to mark the categories of gender, number, and obviation in both nominals and verbals. Prefixes are used for several functions, including the encoding of certain tense/mood/aspect distinctions (cf. 3.6 and 3.7), as well as the representation of person in possessed nominals and main-clause verbs (cf. 5).

Montagnais also makes some systematic use of morphophonemic alternation. The most obvious example is to be found in the process of initial vowel change, whereby the vowel of the initial syllable of the verb stem (or of the tense/mood/aspect prefix, should one occur) may undergo modification when, to simplify slightly, it occurs in an embedded rather than main clause (e.g. *pítuáu* ‘s/he smokes’ as opposed to *nítshissenimáu auen piátuá* ‘I know who is smoking’). While the semantic difference between changed and unchanged forms in subordinate clauses is not always fully clear, in at least one verb form a reading of iterativity is produced: a subjunctive form such as *pítuá-ián-i* ‘smoke-1.SG-SUBJ (if I smoke)’ yields *piátuá-ián-i* ‘CHGD.smoke-1.SG-SUBJ (whenever I smoke)’.

The process of **reduplication** is exploited in Montagnais, but purely from a derivational rather than an inflectional perspective. Reduplication is found in certain lexical subsets of verbs, giving an iterative meaning, as well as in numerical particles, with a distributive meaning, e.g. *mâ-míshí-u* ‘RDP-defecate-3.SG

(s/he has diarrhea)’, *nâ-neu* ‘RDP-four (four each)’.

Certain dialects of Montagnais have recently begun to exploit **suprasegmental processes**, in the form of tone alternations. These processes are compensatory, and have arisen in the central dialects as a result of deletion of inflectional suffixes, as for example in the obviative and the inanimate plural of nouns (both marked by final *-a*), in certain imperative forms (Cowan 1983; Ford 1983), and in the subjunctive.

#### 5. Inflection

The inflectional morphology of the Montagnais noun, like that of main clause verbs, is highly **agglutinative**. Affixes occur in fixed order and are fairly transparently associated with the categories of person, number, etc., as the following possessed nominal form illustrates: *ni-mashushám-inán-at* ‘1-bear.paw.snowshoe-1.PE-PL (our (exclusive) bear paw snowshoes)’. The prefix slot encodes the person of the possessor (*ni-* first person, *tshi-* second, and *u-* third). The suffix slots include the marker for plural possessor (*-inán* ‘1.PE’, *-inu/inán(u)* ‘1.PI’, *-uáu* ‘2/3.PL’), followed by the morpheme representing the plural of the possessed object. Animate and inanimate nouns are marked differently for both plural and obviative, although homophony occurs between the animate obviative and the inanimate plural markers (*-a*) (e.g. *ut-ashám-a* ‘3-snowshoe-ANIM.OBV (her/his snowshoe(s))’ and *mashinaikan-a* ‘book-INAN.PL (books)’). The slots for animate plural and obviative are mutually exclusive (e.g. *u-nipákan-a* ‘3-bedspread-ANIM.OBV (her/his bedspreads(s))’).

Within this highly agglutinative framework, there is little morphophonemic variation. Personal prefixes have epenthetic *-t-* when prefixed to all stem-initial vowels except *u/û*, e.g. *tshit-ashám* ‘2-snowshoe (your snowshoe)’, but *n-utápán* (<*ni-utápán*) ‘1-vehicle (my vehicle)’. Animate plural and locative suffixes vary according to the preceding segment; hence locative *mítshuáp-it* ‘house-LOC (in the house)’ but *nípít-t* ‘water-LOC (in the water)’.

Like other Algonquian languages, Montagnais possesses personal, demonstrative, and interrogative pronouns; there exist, however, no relative or reflexive/reciprocal pronouns, the former being expressed by a preverb and the latter, by derivational suffixes on the

verb. The personal pronouns – restricted to emphatic use, to a single word answer to a question, or to coordinate constructions – have solely animate reference. In shape they very closely resemble the nominal possessive prefixes noted above, likewise exhibiting three persons, an inclusive vs. exclusive first person plural distinction, and singular vs. plural forms. Demonstrative pronouns, found only in the third person, represent three degrees of distance (near, middle, and far) as well as absentative. They are inflected for gender, number, and obviation, as are the interrogative pronouns ‘who, what, which one’. The personal and demonstrative, as well as some interrogative sets, function as nominals e.g. *nîn* ‘me, mine’, *neme* ‘that one’, *ne tshekuán* ‘that thing’ (literally ‘that what’); the demonstratives also function as adjectivals, e.g. *ume mashinaikan* ‘this book’. Although their inflectional morphology is predominantly nominal, pronominals exhibit some interesting parallels with verbs with respect to plural suffixes; in addition, certain pronouns may occur with a verbal dubitative mood suffix, e.g. *tshekuán-itshe* ‘what-DUB (whatever one)’, as can certain temporal particles such as *anûtsh-itshe* ‘now-DUB’.

Within the Montagnais verb, there is a major inflectional distinction, based largely on the type of clause in which the verb occurs. This yields, on the one hand, an inflectional paradigm almost totally restricted to main clauses (termed by Algonquianists the **independent order**), and, on the other hand, a set of **conjunction order** forms. The latter are found in embedded clauses, as well as in three main clause environments in Montagnais: WH-questions; following certain conjunctive particles (e.g. (*e*)*kue(t)* ‘so, then’); and after the negative particle *apú* (negation not being indicated in Montagnais, as it is in some Algonquian languages, by specific verbal affixes). The occurrence of the conjunct in other types of main clause (e.g. in narratives, as in 8; cf. Cyr 1990) appears to involve discourse factors. Independent verbs are agglutinative in structure, most of their inflectional affixes, like those of nominals, being transparently associated with grammatical categories. Indeed, the inflectional affixes of independent verbs are very similar to those of nominals, in particular to the paradigm of the possessed noun, from which this system of affixes may well have developed. Conjunct suffixes are more opaque or fused, and are

hypothesized to be more archaic (Goddard 1974).

Verbal inflection encodes the person, gender, number, and obviation status of event participants, as well as, in transitive verbs, the direction of the action; through these affixes, the roles of agent and patient, and the functions of subject and object, are conveyed. The independent order transitive animate verb *tshi-ka-uâpam-â-u-at* ‘2-FUT-see.her/him-TA.THM.1/2>3-1/2/3.SG-3.ANIM.PL (you (pl.) will see them)’ illustrates the canonical shape of verbs involving a 1st or 2nd person argument (optional elements or restrictions are provided in brackets): PERSON PREFIX (independent order verbs only) (+ TENSE/MOOD MARKER) + STEM + THEME SIGN (transitive verbs only) + PERSON/NUMBER SUFFIX (+ 3RD PLURAL OBJECT MARKER). As earlier noted, the person prefix invariably marks the person that is higher on the person hierarchy, without reference to agent or patient role; the theme sign then specifies whether the direction of the action proceeds from a higher to a lower member of the hierarchy (direct) or vice versa (inverse), thereby providing the correct interpretation as to subject and object functions (e.g. *ni-uâpam-â-u* ‘1-see.her/him-DIRECT.THM.1/2>3-1/2/3.SG (I see her/him)’ and *ni-uâpam-ik<sup>“</sup>-o* ‘1-see.her/him-INVERSE.THM.3>1/2-1/2.SG (s/he sees me)'). Conjunct verbs, which do not occur with person prefixes, often use a theme sign fused with a person marker: *apú uâpam-ak* ‘NEG see.her/him-DIRECT.THM.1/2>3.1.SG (I do not see her/him)’. Features of person, number, and obviation show agreement among nouns, demonstrative pronouns, and verbs.

## 6. Word formation

This section deals with the formation of Montagnais stems, or that part of the word to which inflectional affixes are directly added. Word formation in Montagnais follows general Algonquian patterns of creating nominal and verbal stems through **derivation** and, less commonly, through **compounding**. Further details on Algonquian stem formation are to be found in Bloomfield (1946) and Goddard (1990 b).

### 6.1. Derivation

Three stem components are typically distinguished: **initials** (including the category ‘root’ established by Bloomfield 1946), **medials**, and

**finals.** All three components may consist of more than one morpheme, that is, they may be composed of combinations of initials, medials, and finals. **Primary derivation** of both nominal and verbal stems involves minimally an initial (e.g. *nit-ashám-at* ‘1-snowshoe. (INITIAL)-ANIM.PL (my snowshoes)’, to which may be suffixed a final (e.g. *píku-sham-*<sup>u</sup> ‘break.(INITIAL)-TI.FIN-TI.THM-3.SG (s/he breaks it by cutting)’), or a medial plus final. Initials may be free forms as in the noun *nípi* ‘water, lake’ or be directly followed by an inflectional suffix, as in *nípi-t* ‘water-LOC (in, on the water)’, *ními-tshe* ‘dance-3.SG.DUB (maybe s/he is dancing)’, *ními-nânu* ‘dance-3.INDEF (people (unspecified) are dancing)’. Medials occur as the optional second element of a verbal stem, and fall into two types (cf. Denny 1989) – those which are derived through the incorporation of free nouns, and those which do not occur as free nominals, but simply as verb medials, and which serve to specify or classify the material described by the verb initial. Medials derived from nouns typically refer to people, animals, body parts, edibles, and plants (e.g. *nimuáu atik*<sup>u</sup> > *ni-mû-tiku-m-â-u* ‘1-eat-MED-TA.FIN-TA.THM.1/2>3-1/2/3.SG(I eat caribou)'). Classificatory medials refer to concrete properties of objects and comprise such categories as ‘long and rigid (stick-like)’, ‘long and flexible (string-like)’, ‘sheet-like’, ‘hard/metallic’, ‘liquid’ (e.g. *tshinu-âshku-an-o* ‘long-MED(stick-like)-II.FIN-3.SG (it is a long stick)').

Finals are bound affixes which represent the last component of a stem, and may be semantically abstract or concrete. Noun finals, a closed set, typically indicate concrete properties (as do classificatory verb medials) and include such categories as ‘building, bush, person, liquid, stick-like, string-like, manufactured’. Verb finals usually differ in form according to whether the verb stem is to be processed as transitive or intransitive. Within each of these two basic categories, verb finals usually occur in corresponding pairs; for transitive verbs, these pairs indicate the animacy of the patient (e.g. *píku-shku-e-u* ‘break-TA.FIN-TA.THM.3>3'-1/2/3.SG (s/he breaks her/him, it (animate) by foot, body)’, *píku-shk-am-*<sup>u</sup> ‘break-TI.FIN-TI.THM-3.SG (s/he breaks it by foot, body)’, while for intransitive verbs they typically mark the animacy of the agent (e.g. *mikushû* ‘s/he, it (animate) is red’, *mikuâu* ‘it is red’). Intransitive finals tend to be more abstract than transitive fi-

nals; the latter, for example, often convey information about instrumentality.

Through a process of further suffixation, referred to as **secondary derivation**, the primary stems described above may change word class or, in the case of verbs, undergo changes in valency.

#### 6.1.1. Noun derivation

Noun stems may be formed from both transitive and intransitive verb stems by means of a small set of nominalizing suffixes, the most productive of which are *-kan* (designating objects and instrumentals, e.g. *utâmai-kan* ‘hit-NFIN (hammer, club)’) and *-un* (which yields abstract nouns from any semantically appropriate animate intransitive verb stem, e.g. *atusseu* ‘s/he works.’ *atusseun* ‘work’). The resulting nouns are always inanimate, and are subject to the full range of nominal inflectional processes.

Nouns may also be created through the suffixation of noun finals to free nouns: thus *mashinaikan* ‘book’ yields *mashinaikanâshk*<sup>u</sup> ‘pencil’ and *mashinaikanitshuáp* ‘office’. Certain of these finals may also be suffixed to verbs (e.g. *mítshishu* ‘s/he eats’ yielding *mítshishútshuáp* ‘restaurant’).

One extremely productive process is the use of a transitive or intransitive verb preceded by the relativizing prefix *kâ*, and followed by a conjunct order suffix (e.g. *kâuâpi-kuesht* ‘priest’ – literally ‘the one who has a white neck’). The gender of the resulting noun is that of the corresponding verb stem, that is, either animate or inanimate. Plurals are formed by means of verbal inflections, yet such forms also take certain nominal inflections, such as the locative suffix, and may have suffixed noun finals (e.g. *kâmkunueshû-tshuáp-it* ‘REL-police-NFIN(building)-LOC (in the police station)'). This is typically the means employed in Montagnais for creating neologisms, particularly for non-traditional entities, such as those involving new technologies. Borrowing from English and French is also a source of neologisms, but the extent varies greatly from one community to another; while older borrowed items have been reshaped phonologically, more recent items often retain the phonological shape of the donor language.

#### 6.1.2. Verb formation

Verbs of ‘making’ and ‘possessing’ can be derived from appropriate noun stems. The former usually consist of a stem plus the suffix

*-itsheu* (e.g. *massin-itsheu* ‘shoe-AI.FIN(make) (s/he makes shoes)'). Verbs of possession are formed by adding inflectional suffixes to a possessed third person noun: *u-massin* ‘3-shoe (her/his shoe)' yields *umassinu* ‘s/he has a shoe.'

Verbs derived from existing verbs all involve changes in valence. Intransitive verbs with animate subjects may become **causatives**, and so bring into play a second animate argument, through the addition to the intransitive verb stem of a transitive causative suffix *-ie-l-tâ-*; thus animate intransitive *pûnu* ‘s/he stops' gives rise to transitive animate *pûnieu* ‘s/he stops her/him' and transitive inanimate *pûnitâu* ‘s/he stops it'. Verbs which indicate a further goal, such as an indirect object or benefactive, are formed by suffixing the theme sign */-au-/* (= *u*) to a transitive inanimate verb, as in *atuaimu* ‘s/he points to it', *atuaimueu* ‘s/he points it out to her/him'. Transitive verbs may become intransitives by several different means. One of these is through the non-specification of the argument that serves as subject; thus the transitive stem *uâpam-*, which appears in a transitive animate verb such as *ni-uâp-am-ik<sup>u</sup>-o* ‘1-see-TA.FIN-INVERSE.THM.3>1/2-1/2.SG (s/he sees me)', corresponds to the intransitive indefinite or unspecified subject verb *ni-uâp-am-ikû-n* ‘1-see-TA.FIN-INVERSE.THM.3(INDEF)>1/2-AI.1 (someone sees me/I am seen)', in which the indefinite subject theme sign *-ikû-* is followed by the animate intransitive first person suffix *-n*.

## 6.2. Compounding

Compound stems in Algonquian are normally recognized as consisting of stem plus prefixed element; although they form a single inflectional unit, members of compounds exhibit separate word status with respect to the application of phonological rules (Goddard 1990 b). While the prefixed morpheme typically belongs to an open-ended lexical class, the above definition of compound stems would include prefixes of a more grammatical nature, such as the tense/aspect/mood prefixes outlined in 3.

Montagnais contains both nominal and verbal compounds. The noun or verb stem, which serves as the lexical head of the compound, always occurs in rightmost position; nominal compounds typically display the gender of their head noun. The prefixed ele-

ment may be a noun (e.g. *ishkuteu-ûsh* ‘boat, steamship' from *ishkuteu* ‘fire' and *ûsh* ‘canoe'); a verb (e.g. *aiamieu-mashinaikan* ‘prayer book' from *aiamieu* ‘s/he prays' and *mashinaikan* ‘book'); or a preform, this last consisting of a bound morpheme that acts as an adjectival- or adverbial-like modifier (e.g. *mishta-*, as in *mishta-mîtshishu* ‘much-s/he eats (s/he eats a lot)', *tshishe-*, as in *tshishe-utshimâu* ‘great-manager/chief (i.e. government)'). Preforms often possess a meaning of temporal or spatial directionality (e.g. *petshi-*, in *petshi-uepinam<sup>u</sup>* ‘this direction-s/he throws it (s/he throws it in this direction)'). Three and four element compounds are also to be found in Montagnais, as in *tshiâsh-innu-mashinaikan* ‘old Indian book', *matshimanitu-tshishe-utshimâu*, lit. ‘bad-spirit-great-manager (i.e. damned government)'. Typically the nominal and verbal components of all the above compounds exhibit a morphophonemic shape identical to that of their occurrence as stems; there are however some exceptions in the case of noun stems, as in *kâk<sup>u</sup>* ‘porcupine', yet *uskât-âk<sup>u</sup>* ‘porcupine leg'.

Semantically, compounds are not reliably transparent or predictable (e.g. *shîshîp* ‘duck' + *assik<sup>u</sup>* ‘container, pot' > *shîshîpassik<sup>u</sup>* ‘kettle'). Typically, however, the members of a compound display such semantic relationships as whole/part, composition, time, place, user, and goal. Morphophonemic processes involved in compounding include the addition or dropping of */w/* (= *u*) with noun stems, collapsing of sequences of vowels, and insertion/dropping of connective *i* (e.g. *nipi-u-kûkîsh* ‘water-CONN-pig (hippopotamus)', *innu-ishkueu* > *innushkueu* ‘Indian-woman', *matshi-atusseu* > *matshatusseu* ‘bad-s/he works', *mashinaikan-i-tshuâp* ‘book-CONN-building (office)'). For further information, see Drapeau (1979).

## 7. Conclusion

Like other Algonquian languages, Montagnais is highly polysynthetic. Thematic relations are largely encoded by a rich and complex verbal morphology, and are dependent on a hierarchy of person and animacy, case marking being absent from Algonquian. Within the clause, word order of major constituents is fairly free, and to some degree discourse-determined. Further details on the morphosyntax of two closely related vari-

eties, Cree and Naskapi, are to be found in Dahlstrom (1991), Wolfart (1996), Déchaine (1999), and Brittain (2001).

## 8. Illustrative text

The following text is taken from Kapesh (1979: 75), with a slightly modified orthography.

- (1) *Ne auâss nenu*  
that child that.one  
*tshetûte-t*  
CHGD.walk.along-3.CON  
*e-minukashu-t uenapissish*  
PV-be.drunk-3.SG.CON a.little.while
- (2) *pimute-u, mînuât kutak*  
walk-3.SG.INDEP again another  
*kâ-uît-eni-tâkushi-nit*  
REL-funny-MED-AI.FIN-3'.CON
- (3) *uâp-am-e-u*  
see-TA.FIN-TA.THM.3>3'-3.SG.INDEP  
*piet-âshtamite-nit.*  
CHGD.PV-come-3'.CON
- (4) *Pâssi-kan-inu*  
shoot-NFIN-OBV  
*tshikamu-tâ-nu anite*  
stick.to-AI+O.FIN-3'.INDEP there  
*uiâ-ni-t*  
body-OBV-LOC
- (5) *mâk utâmai-kan-âshku-nu*  
and hit-NFIN-NFIN-OBV  
*taku-n-am-inu.*  
hold-TI.FIN-TI.THM-3>3'.INDEP
- (6) *Ne kâ-uît-eni-tâkushi-t*  
that REL-funny-MED-AI.FIN-3.SG.CON  
*uiâp-am-â-t*  
CHGD.see-TA.FIN-TA.THM.3>3'-3.CON
- (7) *nenu-a auâss-a*  
that-ANIM.OBV child-ANIM.OBV  
*menukashu-nit,*  
CHGD.be.drunk-3'.CON  
*mishta-minu-enit-am-ü.*  
PV(big)-good-TI.FIN-TI.THM-3.SG.INDEP
- (8) *Tshessinât*  
surely  
*ni-ka-mâku-n-â-u,*  
1-FUT-grab-TA.FIN-TA.THM.1/2>3-1/2/  
3.SG.INDEP
- (9) *it-enim-e-u.*  
so-TA.FIN(think)-TA.THM.3>3'-1/2/  
3.SG.INDEP

- (10) *Pitimâ kue pâssi-tshe-t*  
right away then shoot-AI.FIN-3.SG.CON  
*anite ishpim-it*  
there above-LOC
- (11) *tshetshî-kushtâtsh-i-â-t*  
PV(so.that)-frighten-TA.FIN-TA.THM.3>  
*kuet*  
3'-3.SG.CON then
- (12) *utshi-pit-â-t*  
pull-TA.FIN-TA.THM.3>3'-3.SG.CON then  
*ut-ûtâm-u-â-t*  
RDP-hit-TA.FIN-TA.THM.3>3'-3.SG.CON
- (13) *anite u-tâshtamiku-ni-t kue*  
there 3-face-OBV-LOC then  
*it-â-t:*  
say-TA.THM.3>3'-3.CON
- (14) *kâ-mâku-n-ueshî-u-tshuâp-it*  
REL-grab-TA.FIN-AI.FIN-CONN-NFIN-LOC  
*tshi-ka-nâtshi-nipâ-n,*  
2-FUT-PV(go)-sleep-1/2.SG.INDEP  
*it-e-u.*  
say-TA.THM.3>3'-1/2/3.INDEP

'(1) The child walks along, drunk; (2) he walks for a little while, (3) then he sees another silly person coming towards him. (4) That one has a gun on his body (5) and is holding a club. (6) The silly person, when he sees (7) the child is drunk, is very happy. (8) Surely I will arrest him, (9) he thinks. (10) So right away he shoots up in the air (11) in order to frighten him, (12) then pulls him towards him and beats him (13) in the face, then says to him: (14) You are going to sleep in jail, he says.'

## 9. Uncommon abbreviations

3'	obviative agent or goal
AI	animate intransitive
CHGD	changed conjunct
CON	conjunct
FIN	final
II	inanimate intransitive
NFIN	noun final
PV	preverb
REL	relativized form
RELAT	relational form
TA	transitive animate
TI	transitive inanimate
THM	theme sign
VAI+O	pseudo-transitive

## 10. References

- Axelsson, Monica & Cyr, Danielle (1988), "Is There a Definite Article in Montagnais?". In: Rosen, Victoria (ed.), *Papers from the Tenth Scandinavian Conference of Linguistics* [Bergen, 1987]. Bergen: University of Bergen, 52–73
- Bloomfield, Leonard (1946), "Algonquian". In: Hoijer, Harry et al. (ed.), *Linguistic Structures of Native America*. New York: Viking Fund Publications in Anthropology 6, 85–129
- Brittain, Julie (2001), *The Morphosyntax of the Algonquian Conjunct Verb: A Minimalist Approach*. New York: Garland (Outstanding Dissertations in Linguistics)
- Clarke, Sandra (1982), *North-West River (Sheshatshu) Montagnais: A Grammatical Sketch*. Ottawa: National Museum of Man (Mercury Series, Canadian Ethnology Service Paper No. 80)
- Cowan, William (1983), "The Development of Suprasegmental Inflections in Montagnais". *International Journal of American Linguistics* 49.1, 64–71
- Cyr, Danielle (1990), *Approche typologique du système aspectuel montagnais, de la morphologie à la pragmatique*. Ph.D. dissertation, Université Laval [Ottawa: National Library of Canada, Microfiche TH-57474]
- Cyr, Danielle (1996), "Montagnais: An Ethnogrammatical Description". In: Maurais, Jacques (ed.), *Québec's Aboriginal Languages: History, Planning and Development*. Clevedon: Multilingual Matters, 174–203
- Dahlstrom, Amy (1991), *Plains Cree Morphosyntax*. Hamden/CT: Garland (Outstanding Dissertations in Linguistics Series 5)
- Déchaine, Rose-Marie (1999), "What Algonquian Morphology Is Really Like: Hockett Revisited". In: Bar-El, Leora & Déchaine, Rose-Marie & Reinholz, Charlotte (eds.), *MIT Occasional Papers in Linguistics* 17, 25–72
- Denny, J. Peter (1989), "The Nature of Polysynthesis in Algonquian and Eskimo". In: Gerdts, Donna B. & Michelson, Karin (eds.), *Theoretical Perspectives on Native American Languages*. Albany: State University of New York Press (SUNY Series in Linguistics), 230–258
- Drapeau, Lynn (1979), *Aspects de la morphologie du nom en montagnais*. Ph.D. dissertation, Université de Montréal [Ottawa: National Library of Canada, Microfilm QMU-2M11 451.2]
- Drapeau, Lynn (1984), "Le traitement de l'information chez les Montagnais". *Recherches amérindiennes au Québec* 14.4, 24–35
- Drapeau, Lynn (1986), "Entre le rêve et la réalité: le mode subjectif en montagnais". In: Cowan, William (ed.), *Actes du dix-septième congrès des algonquinistes* [Montréal, 1985]. Ottawa: Carleton University, 101–120
- Drapeau, Lynn (1991), *Dictionnaire montagnais-français*. Sillery, Québec: Université du Québec
- Drapeau, Lynn (1996), "Conjurors: The Use of Evidentials in Montagnais Secondhand Narratives". In: Nichols, John & Ogg, Arden (eds.), *nikotwâsik iskwâhitem, pâskihépayih! Studies in Honour of H.C. Wolfart*. Winnipeg: Algonquian and Iroquoian Linguistics, 171–194
- Drapeau, Lynn & Mailhot, José (1989), *Guide pratique d'orthographe montagnaise*. n.p.: Institut éducatif et culturel attikamek-montagnais
- Ford, Alan (1983), "Une analyse autosegmentale de quelques alternances morphologiques du cri-montagnais". In: Cowan, William (ed.), *Actes du quatorzième congrès des algonquinistes* [Québec, 1982]. Ottawa: Carleton University, 293–300
- Goddard, Ives (1974), "Remarks on the Algonquian Independent Indicative". *International Journal of American Linguistics* 40.4, 317–328
- Goddard, Ives (1990 a), "Aspects of the Structure of Fox Narratives: Proximate Shifts and the Use of Overt and Inflectional NPs". *International Journal of American Linguistics* 56.3, 317–340
- Goddard, Ives (1990 b), "Primary and Secondary Stem Derivation in Algonquian". *International Journal of American Linguistics* 56.4, 449–483
- James, Deborah & Clarke, Sandra & MacKenzie, Marguerite (2001), "The Encoding of Information Source in Algonquian: Evidentials in Cree/Montagnais/Naskapi". *International Journal of American Linguistics* 67.3, 229–263
- Kapesh, An Antane (1979), *Tante nana etutamin nitassi*. Montréal: Les éditions impossibles
- La Brosse, Jean-Baptiste (1767), *Nehiro-iriniui mas-sinaigan*. Uabistiguatsh: Broun gaie Girmor
- Lemoine, Georges (1901), *Dictionnaire français-montagnais avec un vocabulaire montagnais-anglais, une courte liste des noms géographiques, et grammaire montagnaise*. Boston: W. B. & P. Cabot
- MacKenzie, Marguerite E. (1980), *Towards a Dialectology of Cree-Montagnais-Naskapi*. Ph.D. dissertation, University of Toronto [Ottawa: National Library of Canada, Microfiche TC-53109]
- Mailhot, José (1997), *Pour une orthographe unique de la langue innue*. Sept-Iles, Québec: Institut culturel et éducatif montagnais

Thwaites, Reuben Gold (1896), *The Jesuit Relations and Allied Documents, Vol. 1-73*. Cleveland: Burrows Brothers, 1896–1901 [reprinted 1959 in 36 vols., New York: Pageant Press]

Wolfart, Hans Christoph (1973), *Plains Cree: A Grammatical Study*. Philadelphia: Transactions of the American Philosophical Society 63, Part 5

Wolfart, Hans Christoph (1996), “Sketch of Cree, an Algonquian Language”. In: Goddard, Ives (ed.), *Languages, Vol. 17: Handbook of North American Indian Languages*. Washington: Smithsonian Institution, 390–439

*Sandra Clarke & Marguerite MacKenzie,  
St. John's, Newfoundland (Canada)*

## 132. Guaraní (Tupi-Guaraní)

1. Preliminaries
2. Inflectional morphology
3. Derivational morphology
4. Word classes
5. Illustrative text
6. Uncommon abbreviations
7. References

### 1. Preliminaries

This article is a morphological description of Guaraní, a language of the Tupi-Guaraní stock, a language family originally spoken in the region known as Lowland South America (Klein 1991: 31; Payne 1990: 3). This is an extended area which covers the basins of the Amazon river and its tributaries and includes parts of the Guyanas, Brazil, Venezuela, Colombia, Ecuador, Peru, Bolivia, Argentina, and Paraguay. Paraguayan Guaraní, the variety examined here, is spoken in Paraguay, primarily in the eastern region, which houses most of the population. Closely related varieties are spoken in Northeastern Argentina, the Brazilian states of Matto Grosso and Paraná, and in the Bolivian Chaco region.

Several varieties of Guaraní were spoken by the native communities of a wide region south of the Amazon upon the arrival of the Europeans. The first written records of the language as well as the first documentation of its morphological characteristics date from the middle of the seventeenth century and come from Jesuit missionaries (Ruiz de Montoya 1639; 1640 a; 1640 b). Numerous morphological descriptions have appeared since then, most of which utilize a “morpheme-listing” approach with very little attempt at an explanatory account of the morphology as a system or as a representative instance of a language type or a linguistic area (Payne 1990: 213 f.).

In Paraguay, the language outlived most of its original speakers and is now spoken by the majority of the population, both white and mestizo. In rural areas, Guaraní is the predominant, and in many cases the only, language spoken. It is also widely used, along with Spanish, in towns and the Capital City. Only an estimated 7% of the population does not speak any Guaraní (1992 Population Census). Although Guaraní is an important symbol of identity for most Paraguayans, it has long been regarded as somewhat inferior to Spanish, which was until recently the only official instrument of written communication and instruction. Apart from some literary works (mainly poetry), and occasional documents (mostly unpublished) written in Guaraní, the language remained a primarily spoken means of communication until the Paraguayan government’s 1967 recognition of Guaraní as a “national language”. The language achieved “official language” status in 1992 and is now an obligatory subject in public schools and is increasingly used as a literary medium.

The data sources for this article include two sample texts, one of which is provided in section 5, and a second which was published elsewhere (Velázquez-Castillo 1996: 240). To the extent possible, the examples were taken from these two samples, although in some cases, when it was not possible to illustrate the issue at hand, an occasional formulated example was used. When referring to examples taken from the samples, the reference given is the glossed, translated, and numbered version provided in 5.

All Guaraní examples are transcribed following the most prevalent orthographic convention used in Paraguay. Most of the symbols used have the phonetic values generally assigned to them in standard phonetic tran-

scriptions. Note, however, the following phonetic values: <ch> alveopalatal fricative identical to the initial sound of English *ship*; <g> voiced velar fricative, roughly identical to its Spanish counterpart in words such as *amigo*; <'> glottal stop; and <y> high back unrounded vowel. Accent marks are used only to indicate a deviation from the predominant last syllable stress. Guarani has a set of six vowels, matched by a complete set of nasal counterparts, transcribed here with capital letters.

The most distinctive morpho-phonological feature of Guarani is “nasal harmony”. Nasality is contrastive in stressed syllables and tends to spread from stressed nasal vowels to other sounds of the stem and onto suffixes and prefixes, creating the following set of possible alternations: prenasalized consonants with homorganic stops, and nasals with homorganic stops and prenasals. For example, the causative *mbo-* becomes *mo-* in *mo-kA* ‘CAUS<sub>1</sub>-dry (to dry)’, the reflexive/passive *je-* becomes *ñe-* in *o-ñe-not Y* ‘3.ACT-REFL-bury (he was buried)’, and the totalitative *pa-* becomes *mba-* in *mo-kA-mba* ‘CAUS<sub>1</sub>-dry-TOT (to dry completely)’.

A set of roots often referred to as “relational” exhibit the alternation [t r h] in the first segment. For example, *tova*, *rova* or *hova* ‘face’ has the three possible forms. The *t*-form occurs when the term is used non-relationally, the *r*- and *h*- forms are used when the term is used relationally, as in *che-rova* ‘1.INACT-face (my face)’ and *hova* ‘his/her: face’. The alternation between *r*- and *h*- is determined by whether or not the possessor is 3rd person in a nominal possessive phrase, and by whether or not the subject outranks the object on the person hierarchy (1 > 2 > 3) in a predicative phrase (Velázquez-Castillo 1996: 10). Other allomorphic alternations will be noted as they become relevant in the course of the discussion.

This section concludes with a brief survey of the semantic classification of the vocabulary assumed in the description of major morphological patterns in Guarani (for a detailed classification, see Velázquez-Castillo 1996: 210–225). The classification includes the following main divisions, “objects”, “properties”, and “processes” (following Croft 1991: 63–67). Prototypical objects are signifiers whose typical designata are people, animals, plants, places, and other inanimate entities, and are characterized by a high degree of time-stability and conceptual inde-

pence. Conceptual independence is reflected in the valency of a given form; forms designating objects normally have zero valency, (0-ary). Prototypical forms designating “properties” are predicates with valency 1 (1-ary), i.e., there must be one entity to which the property is attributed. Typically, properties designate color, dimensions, age, value, physical characteristics, and human propensities. The type “process” includes actions and activities. Activities are 1-ary predicates, like properties, but they differ from properties with respect to time-stability. Unlike properties, activities typically designate dynamic events such as, translational and non-translational motion, biological processes and bodily functions. Actions are highly dynamic, 2-ary predicates which typically denote energy transfer from one participant to another. Typical members of this category include acts and events with observable physical results.

Fig. 132.1 below is intended as a rough summary of the different semantic types found in the lexical inventory of Guarani. The major semantic types are arranged along a continuum of increasing **conceptual dependence** and decreasing **time-stability** (cf. Art. 104). The left-most semantic type of the continuum has maximal conceptual independence and time-stability, while the right-most type is highly dependent and unstable in time.

The ensuing discussion of Guarani morphology describes inflectional and derivational morphology, and word classes in terms of the interplay between the inherent semantics of major lexical roots and their pragmatic function. There is an assumed natural correlation between semantic type and pragmatic function (as manifested in clausal role): object/reference, property/modification, and action/predication. Cross-linguistically, inflectional as well as derivational morphology have been observed to exhibit some degree of sensitivity to the naturalness of function/

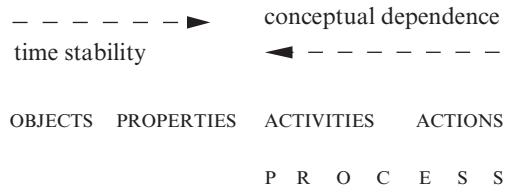


Fig. 132.1: Semantic classification of lexical inventory

meaning correlation and thus to partition the lexical inventory in a manner as to suggest the word classes that are operational in the language (Croft 1991: 62–86).

## 2. Inflectional morphology

An **inflectional morpheme** is defined as one that changes neither the basic semantic type of a lexical item nor the basic pragmatic function with which the lexical item is naturally correlated. For example, tense morphemes will be considered inflectional since they do not change, say, an action into an activity (cf. Art. 110); a causative morpheme, on the other hand, will be considered derivational since it can turn an activity into an action (cf. Art. 107). Inflectional morphology is often considered to fit the inherent semantic type of the core members of a given word class (Croft 1991: 79, 86; Bybee 1985: 13–19; Wierzbicka 1988: 485). Gender morphology, for example, which is frequently found on nouns, signals a “unique categorization” and tense morphology, usually found on verbs, fits the inherent link between their lexical content and the notion of time.

The description of Guaraní inflectional morphology begins with morphological processes that are sensitive to the pragmatic function of predication, rather than to the semantic type of lexical roots. The discussion proceeds gradually toward morphological processes that show sensitivity to the semantic type of the roots involved.

### 2.1. Sentential markers

Objects, properties, activities, and actions in predicative function can all occur with the following **sentential markers**: the negative *n-/na-* ... *i*, the interrogatives, *-pa* and *-piko*, and the reportatives, *-jeko* and *-ndaje*. Of these, only the negative attaches to the predicate itself; the others mark the modality of the whole proposition and attach to the first element of the sentence. Examples (1)–(3) illustrate the use of these morphemes.

- (1) *MoO piko upéva?*  
where INT that  
‘Where is that?’  
(Velázquez-Castillo 1996: 242 (11))
- (2) *Nde-sy piko o-ú-ma?*  
2.INACT-mother INT 3.ACT-come-PFV  
‘Did your mother return already?’

- (3) *O-ú-ma piko nde-sy?*  
3.ACT-come-PFV INT 2.INACT-mother  
‘Did your mother return already?’

The negative is a discontinuous morpheme with prefixal and suffixal components which bracket the predicate. The suffixal component precedes tense, aspect or modality morphemes. Examples (18), (19), (23), and (40) of the sample text (section 5) illustrate the use of the negative. A note on the allomorphic variation of the negative: as can be observed in the examples mentioned, the prefixal component is accompanied by the vowel /a/ in front of an inactive personal reference marker; when in front of an active personal marker, it copies the vowel of the personal marker, as in, *nde-re-japó-i* ‘NEG-2.ACT-do-NEG (you don’t do it)’. The suffixal component of the morpheme, *-i*, changes to *-ri* if the stem ends in /i/.

### 2.2. Tense/aspect/modality

It is often observed that **tense/aspect/modality inflection** is maximally relevant to actions and activities since they prototypically designate changing situations and rank low on the time-stability scale (e.g., Givón 1984: 269 f.; Croft 1991: 79, 86; Wierzbicka 1988: 486 f.). In Guaraní, predicating objects, properties, activities, and actions alike can occur with at least some of the same tense/aspect/modality markers. Example (43) contains the future marker *-ta*, on an action stem. Examples (4)–(6) show stems of various semantic types carrying this marker.

- (4) *Ko'Ero-ite nde-róga-ta.*  
tomorrow-very 2.INACT-house-FUT  
‘You’ll have a house very soon.’
- (5) *MichI-ta che-rehe ko-Kamisa.*  
small-FUT 1.INACT-at this-shirt  
‘This shirt will be small on me.’
- (6) *Mba'éicha-iko pe-hasá-ta?*  
how-INT 2.PL.ACT-pass-FUT  
‘How are you going to cross over?’  
(Velázquez-Castillo 1996: 243 (18))

The hortative prefix *ta-t-*, also called “desiderative” (Gregores & Suárez 1967: 132) occurs with stems designating properties, activities, actions, and some objects, *ta-* with objects and properties and *t-* plus personal reference prefixes with activities and actions. When combined with objects and properties, the meaning is in fact that of a desiderative: it indicates wish, but with stems indicating

activities it can be interpreted as a permissive, or a hortative, as in (7), (Velázquez-Castillo 1996: 244 (29)):

- (7) *T-a-pytá-na che ko'á-pe.*  
 HORT-1.ACT-stop-REQ I here-LOC  
 'Let me stop here.'

Other tense/aspect/modality markers that can occur with most semantic types are: *-kuri* 'recent past affirmative', *-raka'*e 'remote past reportative', *-pa* 'totalitative', *-ma* 'perfective', *-joa* 'distributive', *-se* 'desiderative', *-ramo*, *-rO* 'conditional', *-rei* 'without a purpose', *-mo'A* 'frustrative', *-ra'e*, 'mirative' (also 'recent past interrogative'), and *-ne* 'dubitative'.

Two aspectual morphemes, *-ikóni* 'frequent or continuative action irregularly distributed over time' and *-ína* 'continuative action at a given point in time' also occur with most semantic types functioning as predators. However, they cannot be characterized as affixal like the rest of the tense/aspect/modality morphemes since they exhibit root-like morphophonological behavior. Not only do they carry a primary stress but, most importantly, they can be inflected for personal reference, as illustrated in (8):

- (8) *A-ha'arO a-ína*  
 1.ACT-wait 1.ACT-PROG  
*che-irU-me.*  
 1.INACT-partner-to  
 'I am waiting for my friend.'  
 (Velázquez-Castillo 1996: 243 (17))

We have seen that tense/aspect/modality inflection occurs with a wide range of lexical semantic types. Some aspect markers, however, such as the frequentative *-mante*, the habitual *-jepi*, and the concomitant *-vo* (concurrent with another more important event), are more selective with respect to the semantic type of the stem. For example, the frequentative and the habitual cannot normally occur with objects: \**che-róga-jepi* '1.INACT-house-HABIT (I usually have a house)', or durable properties: \**che-michI-jepi* '1.INACT-small-HABIT (I'm usually small)', but can easily occur with a state-like property: *che-ró'y-jepi* '1.INACT-cold-HABIT (I'm usually cold)', and of course, it occurs naturally with activities and actions. The concomitant *-vo* can only occur with activities and actions; \**che-rógalro'y/vai-vo* '1.INACT-house/cold/ugly-CONC'.

### 2.3. Gradation

**Gradation** is often claimed to be most relevant to properties since lexical roots in this semantic category characteristically designate

opposite extremes of a property scale (cf. Art. 114). Guaraní has two inflectional morphemes which indicate gradation, the comparative *-ve* 'more', and the intensive suffixes, *-itel-iterei* 'very, genuine'. Each of these are affixed to stems designating properties, activities and actions, and even to a limited number of predication objects. Note the use of *-ve* with property, *michI-ve* 'small-more (smaller)', activity, *e-ju-vé-na* 'IMP-come-more-REQ (please come closer)', action, *a-joka-ve* '1.ACT-break-more (I broke more)'. Note also the use of *-itel-iterei* with property, *michI-iterei* 'small-INTS (very/too small)', activity, *a-karu-iterei* '1.ACT-eat-INTS (I eat too much)', and action, *a-joka-ite voi* '1.ACT-break-INTS EMPH (I really broke it)'. Frequently, when the stem designates an activity, as in the case of the stem for *eat*, the gradation applies both to the quantity of the inferable direct object and to the activity itself. In the case of actions (as in the stem for *break*), the scale applies to the degree of affectedness of the direct object, and the intensity of the process designated by the stem.

The comparative morpheme does not combine easily with objects. It would be awkward to say something like *oga-ve* 'more of a house' or *jagua-ve* 'more of a dog'. The objects that can take the comparative morpheme are those that are defined not only by their physical properties but also by a set of socioculturally defined properties, as is the case for 'woman' or 'man'. When one says, *kuña-ite-ve* 'woman-INTS-CMPR (much more of a woman)', one does not imply that the woman in question is a super woman from the biological point of view, but rather that the woman's behavior fits a set of cultural expectations associated with womanhood. Objects are not expected to be readily "gradable" because they evoke a conglomerate of properties and it would be unclear which one is being assessed (Wierzbicka 1988: 485 f.). Similarly, activities and actions are not expected to take "degree" morphology (Wierzbicka 1988: 487), but we have seen that stems in this semantic domain are consistently treated as gradable.

We have seen that the intensive and comparative morphemes always occur with a predicate, never with a lexical root in nominal function. Thus, an exact equivalent to the English, *I'm looking for a taller person* is impossible in Guaraní. One would have to say something like, *I'm looking for a man that is taller*. Even though the inherent semantics of

the lexeme does play some role in determining which forms take these morphemes, it seems that discourse function is the primary determinant.

#### 2.4. Personal reference

Like numerous other Amerindian languages, Guarani exhibits an active-stative (also called “split-S”) alignment. This alignment is most notably manifested in the existence of two sets of **cross-referencing prefixes** for the coding of active and inactive participants. Activeness can be characterized as a combination of inherent changeability of the process designated by the verb, and a subject involvement characterized by control and volitionality (Velázquez-Castillo 1991; 2002 b). The two sets of cross-referencing prefixes are shown in Table 132.1.

	inactive	active
1 singular	<i>che-</i>	<i>a-</i>
1 pl. inclusive	<i>ñane-</i>	<i>ja-</i>
1 pl. exclusive	<i>ore-</i>	<i>ro-</i>
2 singular	<i>ne-</i>	<i>re-</i>
2 plural	<i>pene-</i>	<i>pe-</i>
3 singular/plural	<i>i-/ij-/iñ-</i>	<i>o-</i>

Tab. 132.1: Personal reference markers

The **inactive** set is used to indicate the possessor in a noun phrase, as in *pende-rata* ‘2.PL. INACT-fire (your fire)’ (cf. 19); or a participant of an inactive predicate (i.e., an inactive subject or direct object). The function of predication can be performed by lexical forms of virtually every semantic type. When used in this capacity, objects and properties are marked with inactive personal prefixes, as illustrated in (32): *i-para* ‘3.INACT-ripe (is ripe)’ and in (9) below:

- (9) *Che-vare'á-ma hina.*  
     1.INACT-hungry-PFV PROG  
     ‘I’m hungry already.’  
     (Velázquez-Castillo 1996: 243 (21))

The **active** set is used with roots designating activities and actions when cross-referencing the active event participant. The use of these markers is shown in numerous other examples in the sample text of section 5, where they are glossed as “ACT”. In the case of actions, only one of the two participants gets cross-referenced on the predictor. When the two participants are of the same person, it is the agent, i.e., the active participant that

is cross-referenced, as can be observed in (21): *o-remata-uká-ta-ro* ‘3.ACT-auction-CAUS<sub>2</sub>-FUT-when (when they were going to auction (your land))’. When the participants involved are not of the same person, the participant selected for cross-referencing depends on the following person hierarchy, regardless of activeness: 1 > 2 > 3: *che-mbo-tyryry* ‘1.INACT-CAUS<sub>1</sub>-crawl (dragged me)’ (cf. 27).

The opposition active/non-active is an either/or situation only at the two extremes of the lexical inventory. For example, prototypical actions can only take the active set, while prototypical objects can only take the inactive set. In less clear-cut cases, however, forms can be conceived of either as involving change or as designating a stative relation. Contrast for example the non-active construal of, *che-monda* ‘1.INACT-steal (I’m a thief)’ and *che-karu* ‘1.INACT-eat (I’m a big eater)’ with the active construal of, *a-monda* ‘1.ACT-steal (I steal)’ and *a-karu* ‘1.ACT-eat (I eat)’.

#### 2.5. Inflection tailored to actions and activities

We have seen thus far that there are some grammatical morphemes (i.e., the active set of personal markers and some aspectual suffixes) that single out roots designating activities and actions, the prototypical semantic types of the verbal category. Three additional affixes can be counted among the morphemes that select these two semantic types: the passive/reflexive *je-* (e.g., *o-je-poi* ‘3.ACT-REFL-let:go (they hurled themselves)’, cf. 25), the imperative *e-*, as in *e-guata* ‘IMP-walk’ and *e-joká* ‘IMP-break’, and the requestative *-na*, which normally occurs with the imperative to soften a request. Example (10), (Velázquez-Castillo 1996: 243 (20)) illustrates the use of both morphemes:

- (10) *E-jupí-na che-ape-ári.*  
     IMP-climb-REQ 1.INACT-back-on  
     ‘Climb on my back (please).’

There is a small, even more selective set of affixes that can only be attached to roots designating actions. They are the reciprocal *jo-* and the prefix *poro-* ‘unspecified human object’.

Stems indicating actions are also the only ones that can take the portmanteau personal reference prefix, *ro-*, which indicates first person subject and second person object, as can

be observed in: *to-ro-mbo-hasa* HORT-1.SBJ&2.OBJ-CAUS<sub>1</sub>-pass (let me get you across)’ (Velázquez-Castillo 1996: 243 (20)).

### 2.6. Inflection specific to roots designating objects

There is a number of primarily suffixal inflectional morphemes that seem to select stems designating objects to the exclusion of all others. One finds, for example, the following set of plural and collective forms: *-kuéra* ‘PL’ (not obligatory on plural stems), *-ita* ‘a multitude of’, and *-ty* ‘collective’, primarily used with stems designating plants and trees. (21) contains an instance of the plural *-kuéra*, and (11) below (Velázquez-Castillo 1996: 245 (32)) illustrates the use of the collective *-ty*.

- (11) *o-kañ Y-ma-ne-ra'e pakuri-ty-re.*  
 3.ACT-hide-PFV-DUB-MIR pakuri-COLL-at  
 ‘He immediately disappeared among the pakuri trees.’

The suffix *-icha* ‘like, as’ can only attach to a stem designating an object. Section 5 (39, 43) contains two good illustrations of its use.

There are two diminutive suffixes (cf. Art. 99): *-mi*, which means ‘diminutive’ but has an added commiserative, or affective meaning (‘poor little ...’), and *-i*, ‘diminutive’. Both of these suffixes attach primarily to stems designating objects, but the second, *-i*, can also attach to roots of other semantic types with interesting semantic effects. When attached to roots designating objects, the diminutive usually suggests small size, or social insignificance in the case of humans: *jagua-i* ‘dog-DIM (small dog)’, *Maria-i* ‘Maria-DIM (little, insignificant Mary)’; when attached to roots designating properties, the morpheme adds an approximative meaning: *i-porA-i* ‘s/he is sort of pretty’; when attached to forms designating activities, it adds the meaning of ‘a little, not enough’: *o-karu-i* ‘3.ACT-eat-DIM (s/he eats little)’, with stems designating actions, it suggests iterations of an action, generally of insignificant effect: *oi-kutu-I pe-mesa* ‘3.ACT-prick-DIM that-table (s/he kept pricking the table)’.

The only sets of prefixes associated with roots designating objects are the demonstratives, *pe-lumi-* ‘this/those’ and *ko-lko'A* ‘this/those’, the non-active set of personal markers, and definite articles borrowed from Spanish: *la-* ‘DEF.SG’ and *lo-* ‘DEF.PL’. The possessive use of non-active personal markers is limited to object roots in nominal function.

Stems designating objects are not inflected for case but there is a set of postpositions, which are usually attached to the last element of a nominal phrase: *pe-lme-* ‘locative: in, to; human object, instrumental’ (e.g., (21), *óga-pe* ‘home-LOC (at my house)’); *-re* ‘attached to’; *-gotyo* ‘directional’; *-gui* ‘from’ ((30), *pe-ita-guy-gui* ‘that-rock-under-from (from under the rock)’); *-rupi* ‘locative, around X’.

Lastly, there are three suffixes indicating time relevance that occur primarily with stems indicating objects in nominal function: *-rA* ‘future’, *-kue* ‘former, past’ and *-re* ‘former, past’, this latter implying a dead possessor when applied to a human, e.g., *che-ména-rA* ‘1.INACT-husband-FUT (future husband)’, *che-ména-kue* ‘1.INACT-husband-PAST (my former husband)’, and *Maria ména-re* ‘María’s former husband (María is dead)’.

## 3. Derivational morphology

The definition of **derivational morphology** adopted here includes all morphological processes that shift, either the semantic type of the stem, or the function with which the stem is naturally correlated (Croft 1991: 67f.). Thus, a morpheme that has the effect of switching an activity into an action, or an object into an activity will be considered derivational. Special attention will be paid to the function changing effect of derivational morphemes; in particular, whether the morpheme in question is purely function-changing or it involves the addition of a semantic element as well. Pure function-changing morphemes will be taken as evidence of the existence of a given word-class division.

Guaraní does not have a very rich derivational morphology. The limited number of derivational morphemes that do exist are predominantly nominalizing, having the semantic effect of “reifying” inherently time-unstable concepts (cf. Art. 94). It should also be noted that most derivational morphemes are suffixal as opposed to prefixal.

### 3.1. Nominalization processes

Among the overt markers of **reification**, there is one that can be clearly characterized as purely function-changing: the nominalizer *je-*, which attaches to roots denoting action and converts them into nominals. The examples below contrast the predicative (12) and referential use (13) of *juka* ‘kill’.

- (12) *A-juká-ta peteI ovecha*  
 1.ACT-kill-FUT one sheep  
*ne-santo-ára-rehe.*  
 2.INACT-saint-day-for  
 'I'll butcher a sheep for your birthday.'
- (13) *Oi-ko peteI je-juka ange*  
 3.ACT-happen one NR-kill today  
*pyhare.*  
 night  
 'A killing took place last night.'

This nominalization process almost always implies a human agent. *je-* can also nominalize a limited number of activities, all of which are typically performed by a human agent. Action stems can also be nominalized by means of other derivational morphemes which, besides changing the basic semantic type, includes an additional semantic element: The "patientive" *tembi-ltemi-*, e.g., *tembi-u* 'PATNR-eat (food)', the "passive participle (PASS.PART)" *-py* (classified as "passive" in Gregores & Suárez 1967:127), which derives a nominal stem that can modify another nominal, e.g., *so'o joso-py* 'meat grind-PASS.PART (ground meat)'.

The agentivizer *-hal-hára* selects actions and some activities and converts them to referring nominals. When attached to action roots, the result is an agentive nominal, e.g. *monda-ha* 'steal-AGNR (thief)'; when attached to an activity root, the result is a locative nominal, e.g., *jeroky-ha* 'dance-AGNR (dancing place)'.

Another reifying morpheme is *-kue*, which selects stems designating core properties giving as a result a lexical form that designates an object, e.g. *karape-kue* 'short-NR (shortness)'. Besides this mere reifying effect, this morpheme can also add the meaning of 'part' or 'portion' to the derived form, as in: *i-porā-ngue-té-nte* '3.INACT-good-NR-INTS-only (only the good part)'. This morpheme is important in that it suggests that objects and properties do not constitute a totally undifferentiated class.

A number of roots designating human activities can be used in referential function with no overt derivational marking. This is illustrated below, where a predicative (14) and a referential use (15) of the stem *ke* 'sleep' is contrasted:

- (14) *A-ke porā ange pyhare.*  
 1.ACT-sleep good today night  
 'I slept well last night.'

- (15) *A-hecha che-ké-pe*  
 1.ACT-see 1.INACT-sleep-LOC  
*a-veve-ha.*  
 1.ACT-fly-that  
 'I saw in my dream that I was flying.'

Most stems designating inanimate activities simply resist reification. Thus, stems like *mimbi* 'shine' or *kapu* 'explode' cannot be directly used in referential function. However, onomatopoeic stems designating activities that have to do with producing sound or noise, such as *piriri* 'sound of dry leaves', *guerere* 'sound of people talking', are able to function referentially without any additional morphology. Other stems that can be readily used in nominal function are those designating state-like, physical and socioculturally defined human properties, such as *porA* 'pretty/good'.

### 3.2. Processes that derive actions

There is no purely function-changing morpheme that creates actions out of non-actions; but there are a few derivational morphemes with additional meaning that "actionalize" non-actions. The causative prefix *mbo-lmo-* 'CAUS<sub>1</sub>' attaches to roots designating objects, properties, and activities, and converts them into actions by adding a layer of **causation** to their meanings. The resulting stems have all the morphological possibilities of underived actions, including active personal markers (Velázquez-Castillo 2002 a). When added to stems designating activities and properties, the resulting meaning is 'to make somebody do X'. Examples of causativized activities can be seen in (27). When added to stems designating objects, the meaning is 'to provide somebody or something with X', as in (16):

- (16) *A-mbo-juký-ta che-rembi-'u.*  
 1.ACT-CAUS<sub>1</sub>-salt-FUT 1.INACT-PATNR-eat  
 'I will salt my food.'

Stems designating actions take a different causative, the suffix *-uka* 'CAUS<sub>2</sub>', as in (17), (Velázquez-Castillo 1996: 242 (10)):

- (17) *nde-re-je-hecha-uka-vé-i*  
 NEG-2.ACT-PASS-see-CAUS<sub>2</sub>-more-NEG  
 'you don't make yourself seen any more'

Lastly, there are two additional morphemes which derive actions: the prefix *ro-*, and the suffix *'o*. *ro-* attaches to roots designating activities to create an action "whose meaning

is that the subject performs and makes someone else perform jointly the action of the verb" (Gregores & Suárez 1967: 126), e.g., *a-ro-sapukai* '1.ACT-COMIT-yell (I yelled at somebody provoking him to yell at me too)'. The suffix *-'o*, derives action stems from stems designating objects, with the added meaning, 'to get rid of the object named by the stem', e.g., *resa-'o* 'eye-take.off (gouge out eyes)' (25). Another example is found in (27).

### 3.3. Compounding and incorporation

Like most Amazonian languages, Tupi-Guaraní languages feature a great deal of **compounding** and **incorporation**, and it has been suggested that these morphological processes constitute "the major historical source of the rich verbal morphology" found in these linguistic families (Payne 1990: 218). One example that corroborates this claim is the case of the totalitative *-pa*, mentioned in 2.2, which is clearly linked, both in form and meaning to the root *pa* 'to end, finish', which like any activity root can be used in predicative function with active personal references. e.g. *o-pa-la-tembi-'u* '3.ACT-finish DEF-PATNR-eat (the food is gone)'.

There are numerous compounds which combine roots of different sorts, but only the productive patterns will be listed here. Compound predators can be formed with stems designating properties, activities or action, plus one of the following roots: *kuaa* 'to know', *porA* 'beautiful, good', *vai* 'homely, bad', and *pota* 'want'. Examples: *a-kosina-kuaa* '1.ACT-cook-know (I know how to cook)', *re-ikuua-porA* '3.ACT-know-good (you know well)', (20), *a-karu-pota* '1.ACT-eat-want (I'm about to eat)'. In the last case, the original meaning of the verb, 'to want', has evolved to give rise to the more aspectual meaning of 'being on the point of beginning an action or activity'.

Another common type of compounding, though not fully productive is the one known as noun incorporation (Velázquez-Castillo 1996). These can be roughly categorized into two sub-types: the first subtype involves external possession and includes the incorporation of body-part terms into transitive stems to yield an active predictor, as in: *(a-)je-po-héi* '(1.ACT)-hand-REFL-wash', and body-part terms incorporated into intransitive stems designating properties and activities to form inactive predictors, as in: *(che-)py-syry* '(1.INACT)-foot-flow (to slip)' and *(che-)py'a-*

*rasy* '(1.INACT)-stomach-ill (to have a stomach-ache)'. The second type involves neither external possession nor a body-part term. Instead, a noun designating the patient type of an action is affixed to a transitive stem to form an active predictor: *(a-)mba'e-jogua* '1.ACT-thing-buy (I go shopping)'.

### 4. Word classes

If a given language is sensitive to the natural correlation between semantic type and pragmatic function of object/reference, property/modification and action/predication, one can expect the existence of function-changing derivational morphology which marks a deviation from this natural correlation. Thus, when a lexical root of the type "action" is used in a referential function, a marker is expected to indicate the unnaturalness of this correlation. Furthermore, one can expect that the inflectional morphology and function-changing derivational morphology of a given language will partition the lexical inventory in a more or less parallel manner since they would both be sensitive to the naturalness of the semantic type/pragmatic function correlation. The consistent accumulation of morphology in determinate areas of the lexical inventory would suggest then divisions pointing to the grammatical word classes that are operational in the language. With this background in mind, let us examine Fig. 132.2, which summarizes the inflectional and derivational possibilities of the different semantic types present in the lexical inventory of Guarani.

The summary of the range of morphological possibilities of this language presented in Fig. 132.2 reveals a complex division of the lexical inventory. Some morphological processes effect divisions that resemble those suggested by the traditional word classes of nouns and verbs, while other morphological processes suggest different sorts of alignments.

In Guarani, the pragmatic function of predication does not correlate naturally with any particular semantic class. Lexical roots of every semantic type can perform the function of predication with seemingly equal ease, and there is a considerable amount of inflectional morphology that co-occurs with any stem in this function, regardless of semantic class. As Fig. 132.2 shows, some morphological processes are not sensitive to the inherent

INFLECTIONAL MORPHEMES					
			reflexive, reciprocal		
			imperative, some tense/aspects/modality morph.		
inactive pers. reference		active pers. reference			
Number	gradation morph.  most tense/aspects/modality morph				
sentential markers					
OBJECTS	PROPERTIES	ACTIVITIES	ACTIONS		
CAUS <sub>1</sub>					
suffix -'o					
		agentivizer -ha			
DERIVATIONAL MORPHEMES					

Fig. 132.2: Inflectional and derivational morphemes and the lexical inventory

semantics of the lexical roots they select. For example, most of the tense/aspect/modality morphology, as well as the sentential markers, are sensitive to the function of predication rather than to the meaning of the root that happens to be performing that function. Gradation morphemes also seem to be sensitive to this function, since they combine with predicates only. But in this case, the inherent meaning of the predicate appears to be more important than in the case of tense/aspect/modality and sentential markers, since nouns generally do not allow gradation.

Fig. 132.2 also shows a considerable number of morphemes that seem to be sensitive to the lexical content of their base, some of which do suggest word classes that resemble traditionally recognized ones. There is, for example, a number of morphemes attracted to the semantic type of “actions” and some to the semantic type of “objects”, suggesting the categories of core verbs and core nouns respectively. On morphological grounds, one can propose the existence of two large word classes: nouns and verbs. The noun category includes roots designating objects which inflect for number and nominal tense, inactive personal reference, as well as some predicational tense/aspect/modality, take postpositions and the derivational suffix

-'o. The verb category includes all roots with valency 1 or greater that inflect for gradation, the hortative, personal reference, and predicational tense/aspect/modality. The category of verbs can be subdivided into a category of non-active and active verbs. The first group, also called “quality verbs” (Gregores & Suárez 1967: 137) does not have an inflectional morpheme that applies exclusively to its members, but can occur with a derivational nominalizing morpheme which suggest some level of differentiation from the nominal category, as well as from more active predators. The second group, active verbs, inflect for active personal reference and the imperative, and can occur with the agentivizing suffix -ha. The subclass of active verbs includes a further subdivision which distinguishes between transitive and intransitive active verbs. Only transitive verbs can inflect for reciprocity. On the derivational end of the morphology, transitivity seems to determine the use of the CAUS<sub>2</sub> morpheme and the nominalizers *je-*, *temi-*, and *py*.

The remainder of the meaning-sensitive morphology suggests divisions that do not line up with any of the categorial divisions discussed above. The resulting complex division of the lexical inventory seems to be due to the fact that every morphological partition

of the inventory responds to a different semantic factor. In many cases, a particular semantic factor relevant to a given morpheme affects only a portion of the lexical roots which are putative input for another morphological process. Note for example that CAUS<sub>1</sub>, which can attach to any item whose valency is less than 2, suggests a category which does not line up with any of the categories suggested by the personal reference markers, which are sensitive to degree of active involvement of event participants. It seems, then, that the multiple divisions of the lexical inventory suggested by the morphology respond to different semantic factors that do not always line up with the traditional word classes of nouns, adjectives, and verbs.

### 5. Illustrative text

The sample text presented below comes from Julio Correa's play *Sandia Yvyguy*, as it appeared in Pecci (1981, ed.: 18f.). The text is the response of an old man who was told by the young man who owns the house where the old man was living that he was free to leave if he did not like the young man's way of life. The other sample utilized for this article, but not included here, is the glossed and translated version of a folk tale (Velázquez-Castillo 1996), taken from a bilingual magazine.

- (18) *Nda-i-pó-ri á-pe*  
NEG-3.INACT-be-NEG here-LOC  
*che-mo-sE-va-’era.*  
1.INACT-CAUS<sub>1</sub>-leave-that-FUT
- (19) *Che-’Y rire ni*  
I-PRIV after DISJ  
*pende-rata-ypy-kue*  
2.PL.INACT-fire-initial-PAST  
*n-o-I-vé-i-ma-va-era-mo’á*  
NEG-3.ACT-exist-MORE-NEG-PFV-that-FUT-  
FRUST  
*ko’á-pe.*  
here-LOC
- (20) *Ha nde upé-va re-ikuua-porA*  
And you that-REL 2.ACT-know-good
- (21) *Umi Banco Agrícola-gua-kuéra*  
those bank agricultural-from-PL  
*o-remata-uka-ta-rO*  
3.ACT-auction-CAUS<sub>2</sub>-FUT-when  
*pende-hegui pende-yvy,*  
2.PL.INACT-from 2.PL.INACT-land  
*o-ho-va’e-kue nde-ru*  
3.ACT-go-REL-PAST 2.INACT-father
- (22) *che-rendá-pe óga-pe,*  
1.INACT-place-LOC home-LOC  
*o-mombe’u ché-ve.*  
3.ACT-tell 1.INACT-DAT
- (23) *Upe-ramo che-mandu’á*  
that-COND 1.INACT-remember  
*nde-aguélo rehe*  
2.INACT-grandfather about
- (24) *upé-va kuimba’e-va-’ekue, nda-ha’é-i*  
that-REL man-REL-PAST NEG-be-NEG  
*japa-kue ...*  
extra-PAST
- (25) *Rubio Ñu-rire ... upe-ramo che*  
Rubio field-after that-COND I  
*kávo ... a-pyta malerido, ha*  
corporal 1.ACT-stay wounded and  
*o-topa che-rehe ...*  
3.ACT-find 1.INACT-about
- (26) *yryvu o-veve ha*  
crow 3.ACT-fly and  
*o-je-poí-ta-ma-jepi*  
3.ACT-REFL-drop-FUT-PFV-HABIT  
*che-ári che-resa-’o*  
1.INACT-over 1.INACT-eye-take.off  
*haguA ...*  
PURP
- (27) *i-py peteI-re,*  
3.INACT-foot one-with  
*nde-aguélo ...*  
2.INACT-grandfather  
*o-je-japi-vai-pa avei ...*  
3.ACT-PASS-hit-bad-TOT also  
*o-guahE che-rendá-pe ...*  
3.ACT-arrive 1.INACT-place-LOC
- (28) *Che-mbo-tyryry,*  
1.INACT-CAUS<sub>1</sub>-drag  
*che-mo-ingé ka’aguy-pe ...*  
1.INACT-CAUS<sub>1</sub>-enter forest-LOC  
*che-raso-’o, ha guyra*  
1.INACT-worm-take.off and bird  
*rupi’á-re che-mo-ngaru ... ha*  
egg-with 1.INACT-CAUS<sub>1</sub>-eat and  
*che-mo-ingó-ve jevy ...*  
1.INACT-CAUS<sub>1</sub>-live-CMPR again
- (29) *Ava-re che-mandu’á ha*  
these-about 1.INACT-remember and  
*ha’e nde-ru amyr Yi-me:*  
I:tell 2.INACT-father late-DAT
- (30) *e-apoteka chupe-kuéra che-yvy,*  
IMP-mortgage to:them-PL 1.INACT-land  
*kóima á-pe, ha e-paga*  
have:it here-LOC and IMP-pay  
*nde-deve-ha.*  
2.INACT-OWE-AGNR

- (30) *Aga upéi o-jere-porA-ramo*  
now then 3.ACT-turn-good-COND  
*ndé-ve re-nohE-jevý-ne*  
you-DAT 2.ACT-take:out-again-DUB  
*pe-ita-guyú-gui.*  
that-rock-under-from
- (31) *upéi nde-ru o-mba'apo*  
then 2.INACT-father 3.ACT-work  
*o-ñotY heta ...*  
3.ACT-plant a:lot
- (32) *i-pará-porA-se rupi hina*  
3.INACT-ripe-good-DES around PROG  
*i-pet Y, o-u Albino*  
3.INACT-tobacco 3.ACT-come Albino  
*Jara revolución*  
Jara revolution
- (33) *ha aipo Montonero Gobiernita*  
and so:called Montonero Gobiernista  
*o-guahE o-sakea-pa*  
3.ACT-arrive 3.ACT-plunder-TOT  
*nde-rú-pe ...*  
2.INACT-father-LOC
- (34) *Rire-mí-nite o-ú-ma Banco*  
after-DIM-INTS 3.ACT-come-PFV Bank  
*Agrícola o-ipe'a*  
Agricultural 3.ACT-take:away  
*che-hegui che-róga, Francia*  
1.INACT-from 1.INACT-house Francia  
*tiempo-gua-re, ha 100 itária*  
time-from-PAST and 100 hectare  
*che-yvy ...*  
1.INACT-land
- (35) *mboariahu apI niko a-pyta ...*  
poor bare EMPH 1.ACT-stay
- (36) *ha he'i ché-ve nde-ru,*  
and 3:tell 1.INACT-DAT 2.INACT-father  
*hesay hová-re:*  
3:tear 3:face-on
- (37) *Karai Tiku ch-amigo,*  
Mr. Tiku 1.INACT-friend  
*o-u-vai-pa ñandé-ve,*  
3.ACT-come-bad-TOT 1.PI:INACT-to
- (38) *ja-ha óga-pe,*  
1.PI.ACT-go home-LOC
- (39) *ro-gue-rekó-ta*  
1.SBJ&2.OBJ-COMIT-have-FUT  
*che-ndive che-rúi-icha,*  
1.INACT-with 1.INACT-father-like
- (40) *mba'e-ve*  
thing:no-DAT  
*ne-re-ikotevE-i-chéne,*  
NEG-2.ACT-need-NEG-DUB:NEG
- (41) *rei-kuaa-porA mba'é-icha-pa*  
2.ACT-know-good thing-like-INT  
*nde-rayhu Raimunda ...*  
2.INACT-love Raimunda
- (42) *ha a-ju he-ndive nde*  
and 1.ACT-come 3.INACT-with you  
*re-nase mboyve*  
2.ACT-be:born before
- (43) *ha kóina Ága che-mo-sÉ-ta*  
and here now 1.INACT-CAUS<sub>1</sub>-leave-FUT  
*jagua-icha.*  
dog-like
- “(18) There is nobody here that will make me leave. (19) If it were not for me, not even the ashes of your fire place would exist (today). (20) And that, you know well. (21) When the men from the Agricultural Bank auctioned your land, your father came to my house to tell me (about it). (22) At that time I remembered your grandfather. (23) (Now) there was a real man, not an inconsequential one. (24) After (the battle of) Rubio Nu, when I was a corporal, I was left seriously wounded and he found me. (25) Many times, it seemed the crows would hurl themselves at me to gouge out my eyes. (26) On one foot, your grandfather, who was also badly wounded, came to me. (27) He dragged me into the forest, removed the worms from my body, and fed me bird eggs and brought me back to life. (28) I remembered all this and said to your late father: (29) ‘Mortgage my land off to them, here it is, and pay what you owe. (30) Then if things take a turn for the better for you, you might take it back from under the rock’. (31) Then your father worked and planted a great deal. (32) And just when his tobacco was about ready for the harvest, Albino Jara’s revolution came, (33) and the so-called Montonero Gobiernistas came and plundered your father’s (farm). (34) Right after that, the bank came and took away my house, which dates from the time of Francia, and my 100-hectare farm ... (35) I was left so completely poor ... (36) and your father told me, with tears on his face (37) ‘Mr. Tiku, my friend, it turned all bad for us, (38) let’s go home. (39) I’ll have you with me as if you were my father. (40) You won’t need a thing. (41) You know very well how much Raimunda loves you’. (42) And I came home with him before you were born. (43) And now you will kick me out like a dog.”
- As can be noted in this sample text, words often contain a sequence of identifiable forms

(root/s plus affixes) in a relatively loose morphological connection. For the most part, lexical and grammatical morphemes can be easily identified and segmented. It can be said then, that, in terms of traditional morphological typology, Guarani presents polysynthetic or incorporating, as well as agglutinating characteristics (see Comrie 1981: 39–49 for definitions and discussion of morphological typology).

Like most South American languages, Guarani has a predominantly suffixal morphology, but there is a respectable number of prefixes as well. As we have seen, there are prefixes for personal reference, demonstratives, causative, hortative, imperative, reflexive/reciprocal, and pre-verbal noun incorporation. Suffixes include all tense/aspect/modality morphemes, a second type of causative and most derivational morphemes.

Morphological categories and processes seem to favor the pragmatic function of predication over all others, i.e., most affixes and particles concentrate around this function, a feature that is common to South American languages in general. The predicate-internal structure is sketched in (44):

- (44) NEG-pers. ref.-*(REFL/REC)*/(obj.)  
-STEM-gradation-NEG tense/aspect/modality

There are, as we have seen, two main word classes: nouns and verbs, with the latter including two subcategories: transitive and intransitive verbs. An interesting feature of Guarani nouns is the fact that they can carry special tense marking morphemes and that they do not inflect for case or carry subject/object marking morphemes, except perhaps for a dative marker that can be observed on personal pronouns (e.g., (36) above). Grammatical relations are marked primarily on predicates via personal reference markers.

## 6. Uncommon abbreviations

CONC	concomitant
DISJ	disjunctive
FRUST	frustrative
MIR	mirative
REQ	requestative
TOT	totalitative

## 7. References

- Bybee, Joan L. (1985), *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam: Benjamins (Typological Studies in Language 9)

Comrie, Bernard (1981), *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press

Correa, Julio (1981), "Sandia Yvyvy". In: Pecci, Antonio (ed.), *El Teatro Breve del Paraguay*. Asunción (Paraguay): Cromos, 9–20

Croft, William (1991), *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press

Givón, Talmy (1984), *Syntax: A Functional-Typological Introduction, Vol. I*. Amsterdam: Benjamins

Gregores, Emma & Suárez, Jorge A. (1967), *A Description of Colloquial Guarani*. The Hague: Mouton

Klein, Harriet E. Manelis (1991), "South American Languages". In: Bright, William (ed.), *International Encyclopedia of Linguistics, Vol. IV*. New York: Oxford University Press, 31–38

Payne, Doris (1990), "Morphological Characteristics of Lowland South American Languages". In: Payne, Doris (ed.), *Amazonian Linguistics: Studies in Lowland South American Languages*. Austin: University of Texas Press, 213–241

Ruiz de Montoya, Antonio (1639), *Tesoro de la Lengua Guarani*. Madrid

Ruiz de Montoya, Antonio (1640 a), *Catecismo de la Lengua Guarani* [reprinted in 1870]. Leipzig: Oficina y Fundería de W. Drugulin

Ruiz de Montoya, Antonio (1640 b), *Arte y Bocabulario de la lengua Guarani*. Madrid: Sanchez

Velázquez-Castillo, Maura (1991), "The Semantics of Guarani Agreement Markers". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 17, 324–335

Velázquez-Castillo, Maura (1996), *The Grammar of Possession: Inalienability, Incorporation and Possessor Ascension in Guarani*. Amsterdam, Philadelphia: Benjamins (Studies in Language Companion 33)

Velázquez-Castillo, Maura (2002 a), "Guarani Causative Constructions". In: Shibatani, Masyoshi (ed.), *The Grammar of Causation and Interpersonal Manipulation*. Amsterdam, Philadelphia: Benjamins, 507–534

Velázquez-Castillo, Maura (2002 b), "Grammatical Relations in Active Systems: The Case of Guarani". *Functions of Language* 9.2, 133–167

Wierzbicka, Anna (1988), *The Semantics of Grammar*. Amsterdam, Philadelphia: Benjamins (Studies in Language Companion 18)

Maura Velázquez-Castillo,  
Fort Collins (U.S.A.)

## 133. Nahuatl (Uto-Aztecán)

1. The Nahuatl language
2. Word classes and the ‘omnipredicative’ parameter
3. Verb morphology
4. Noun morphology
5. Locative morphology
6. Derivation
7. Composition
8. Illustrative text
9. Uncommon abbreviations
10. References

### 1. The Nahuatl language

Nahuatl is the southernmost member of the Uto-Aztecán language family, which includes about 30 languages, notably Comanche and Hopi in the USA, and Tarahumara and Huichol in Mexico.

Nahuatl was the language of the Aztec empire. The period following the Spanish conquest was certainly one of genocide, but at least, in positive terms, there was interest taken by Spanish friars in the language and culture of the vanquished. The extensive literature which developed in the course of the next century represents the widest existing corpus in any one American Indian language, and brought it a prestige which explains the appellation of “Classical Nahuatl”. It also gave birth to a brilliant linguistic tradition, with some remarkable grammars such as Olmos (1547), Rincón (1595), Carochi (1645), Aldama y Guevara (1754), and a remarkable dictionary (Molina 1571).

Nahuatl still represents the major native speech community in Mexico, of between 1 and 1.5 million speakers. However, the splitting up of the Nahuatl speaking area has speeded up the process of dialectal diversification. Recently, the development of indigenous movements has encouraged contacts between Nahuatl people from various dialectal zones, and has, to some extent, fostered a collective consciousness.

It is Classical Nahuatl that will be described here. The spelling adopted is the I.P.A. with the following differences: *š* denotes an alveopalatal fricative realized [ʃ]; *c* and *č* are affricates, respectively realized [ts] and [tʃ]; *χ* is a lateral affricate realized [tl]; *y* denotes the palatal consonant [j] (not a vowel).

### 2. Word classes and the ‘omnipredicative’ parameter

#### 2.1. Word level and word classes

In Nahuatl, as in most other languages, there is both phonological and morphosyntactic evidence of the existence of a word level. The phonological criteria are: internal assimilations, initial and final exclusion of consonant clusters and of some phonemes, and penultimate stress. But, as we will see, the strongest criteria are morphosyntactic in nature.

Words fall into 5 classes: **verbs**, **nouns**, **locatives**, **demonstratives**, and **particles**. Nahuatl lacks several categories that appear as autonomous word classes in other languages. Thus, there is no clear-cut adjective class, since the usual morphosyntactic criteria fail to apply (Launey 1986: 493): it can only be said that some noun subclasses are more likely to be translated by adjectives in other languages. Nahuatl locatives correspond to most English adverbs, prepositions, or prepositional phrases (see 5.2). Particles correspond to some other adverbs and conjunctions.

#### 2.2. Predicables

The first three classes (verbs, nouns, and locatives) together form a major class of **predicables**. Since the other two classes are highly limited (2.3, 2.5), Nahuatl can be characterized as an **omnipredicative language** – an almost unrecognized yet very important grammatical type (cf. Launey 1994).

The **predicative** function is the basic one. It clearly appears when a predicate is the only word in a sentence. In the following examples, (1 a–d) are verbs, (2 a–c) are nouns, and (3 a–b) are locatives (their morphology will be discussed in 5):

- (1) (a) *Ni-čōka*.  
SBJ.1.SG-cry  
'I am crying.'
- (b) *Čōka*.  
cry  
'He/she (no gender) is crying'
- (c) *Čōka-?*.  
cry-PL  
'They are crying.'
- (d) *Ni-k-kʷa*.  
SBJ.1.SG-OBJ.3.SG-eat  
'I am eating it.'

- (2) (a) *Ni-konē-č.*  
SBJ.1.SG-child-ABSL  
'I (am) a child.'
- (b) *Ti-no-konē-w.*  
SBJ.2.SG-POSS.1.SG-child-POSS  
'You (are) my child.'
- (c) *Naka-č.*  
meat-ABSL  
'(It is) meat.'
- (3) (a) *Mōsčā.*  
tomorrow  
'(It is) tomorrow.'
- (b) *İ-čān.*  
POSS.3.SG-home  
'(It is) his home.'

### 2.3. Particles

In many cases, the predicate is preceded by one or more particles:

- (4) (a) *Ka naka-č.*  
ASRT meat-ABSL  
'It is (indeed) meat.'
- (b) *Kʷiš čōka?*  
INT cry  
'Is he crying?'
- (c) *Ok konē-č.*  
still child-ABSL  
'He is still a child.'
- (d) *San i-čān.*  
just POSS.3.SG-home  
'It is just his home.'

There are about 30 particles. They can only occur as predicate modifiers, never as the only constituent of a sentence.

### 2.4. Rhematicity, saturation, and argument positions

In sentences of the form (PARTICLE) *X* (where *X* is a predicable), *X* is always **rhematic** – i.e. it represents “new information” and can be described as **comment** or **focus** – and **saturated** – i.e. this information is provided about one or more **arguments** which have a referential value.

Arguments need not be expressed by a lexical phrase. If referentially evident, they can be left syntactically unexpressed. For instance, a speaker who says *Čōka* (1 b) probably assumes that the hearer agrees with him on who is being referred to. However, the existence of an **argument position** is always marked by a personal prefix – which has a zero-form in the 3rd person subject, see (1 b–c), etc.

Note that although (2 c) may look like the English ‘meat’ (and can likewise be an answer to a question such as ‘What is this?’ or ‘What are you eating?’), it is a full sentence, not a truncated or elliptic one. If we postulate a zero prefix in (1 b) (which can be an answer to ‘What is he doing?’), there is no reason to do otherwise in (2 c): in both cases, there is a predicate applied to a subject, which because of its referential univocity is left unexpressed. The same holds true for locative predicates: thus (3 a) can be an answer to ‘When are you leaving?’, the unexpressed subject being understood as ‘The moment I will leave ...’.

Each predicable involves a fixed number of argument positions, with a corresponding number of prefixes. Verbs can be **intransitive** (1 a–c), **transitive** (1d) or **ditransitive** (see 3.1.4), i.e. with 1, 2, or 3 argument positions (subject, 1st object, 2nd object). Nouns and locatives can be **absolute** (2 a, 2 c, 3 a) or **possessed** (2 b, 3 b), with respectively 1 or 2 argument positions (subject, possessor).

### 2.5. Demonstratives

Arguments can be expressed by lexical phrases (2.6) or by **demonstratives**. There are only two demonstratives: *in* for proximity (‘this, this one, these ones’) and *on* for distance (‘that, that one, those ones’). They can never be used as predicates, but can occur in any argument position:

- (5) \*(*ka*) *in*; \*(*ka*) *on*  
ASRT DET<sub>1</sub> ASRT DET<sub>2</sub>
- (6) (a) (*Ka*) *čōka in.*  
ASRT cry DET<sub>1</sub>  
'This (person) is crying.'
- (b) (*Ka*) *čōka-2 on.*  
ASRT cry-PL det<sub>2</sub>  
'Those (people) are crying.'
- (c) (*Ka*) *ni-k-kʷa in.*  
ASRT SBJ.1.SG-OBJ.3.SG-eat DET<sub>1</sub>  
'I am eating this.'
- (d) (*Ka*) *konē-č on.*  
ASRT child-ABSL DET<sub>2</sub>  
'That is a child.'
- (e) (*Ka*) *i-čān in.*  
ASRT POSS.3.SG-home DET<sub>1</sub>  
'This is his home.'/It is the home of this person.'

### 2.6. Argument phrases

Predicables can be used to form argument phrases. These can be **indefinite**, i.e. with no determiner (see (7 a)) or **definite**, i.e. preceded

by the determiner *in*, which in many cases can be translated by a definite article, but is actually a specific use of the demonstrative. In any cases, the sentence predicate keeps its prefix, and there is no case marking:

- (7) (a) *Ni-k-kʷa naka-č.*  
SBJ.1.SG-OBJ.3.SG-eat meat-ABSL  
'I am eating meat.'
- (b) *Ni-k-kʷa in naka-č.*  
SBJ.1.SG-OBJ.3.SG-eat DET<sub>1</sub> meat-ABSL  
'I am eating the meat.'
- (c) *Čōka (=Ø-čōka) in konē-č.*  
cry (=SBJ.3-cry) DET<sub>1</sub> child-ABSL  
'The child is crying.'

Definite phrases can precede the sentence predicate, with a slight topical effect (I will not deal here with word order in the case of a plurality of argument phrases):

- (8) (a) *In konē-č čōka.*  
DET<sub>1</sub> child-ABSL cry  
'(As for) the child (he) is crying.'
- (b) *In naka-č ni-k-kʷa.*  
DET<sub>1</sub> meat-ABSL SBJ.1.SG-OBJ.3.SG-eat  
'(As for) the meat, I am eating it.'

A definite expression *in X* can never be predicative, but only an argument. In particular, no sentence can have the form \*(PARTICLE) *in X*: Nahuatl can say 'it is meat' (*ka naka-č* 'ASRT meat-ABSL') or 'it is a child' (*ka konē-č* 'ASRT child-ABSL'), but not 'it is the meat', 'it is the child', at least not by saying only \*(*ka*) *in naka-č*, \*(*ka*) *in konē-č* (see 4.5.3).

Since there are nominal predicates, these can also have a phrasal argument, see (9 a–b). More unusual is the fact that in *in Y*, *Y* can be a verb, with a nominal or verbal sentence predicate (10 a–c):

- (9) (a) *Ka mešiʔka-č in konē-č.*  
ASRT mexican-ABSL DET<sub>1</sub> child-ABSL  
'The child is Mexican.'
- (b) *Aʔmo kʷal-li in naka-č.*  
NEG good-ABSL DET<sub>1</sub> meat-ABSL  
'The meat is not good.'
- (10) (a) *Ka konē-č in čōka.*  
ASRT child-ABSL DET<sub>1</sub> cry  
'The one who is crying is a child.'
- (b) *Ka naka-č in ni-k-kʷa.*  
ASRT meat-ABSL DET<sub>1</sub>  
SBJ.1.SG-OBJ.3.SG-eat  
'What I am eating is meat.'

- (c) *Ni-k-kaki in čōka.*  
SBJ.1.SG-OBJ.3.SG-hear DET<sub>1</sub> cry  
'I can hear the person who is crying.'

(10 a–b) must not be regarded as transformed (by emphasis or focalization) from patterns such as (7 a–b) and (7 c). In fact, each of these sentences is made up of two atomic predicates, but there is a relation of subordination between them. The main predicate is the rhematic element, just as in (1)–(3), but the argument is this time explicitly denoted by a subordinate phrase, which can be paraphrased 'the entity which can be referred to by saying that *it is N* or that *it V's*'. Only rhematic vs. non-rhematic properties, not verbal vs. nominal ones, play a part in the assignment of the syntactic functions.

## 2.7. Prefix agreement

As we can see by the object prefix *-k-* in the argument phrase *in ni-k-kʷa* 'SBJ.1.SG-OBJ.3.SG-eat' in (10 b) ('the thing such that I am eating it'), the condition for the subordination of a predicable as an argument phrase is a coindexation of two argument positions: one in the main predicate, one in the argument phrase. We can thus re-write (10 b) as:

- (11) *Ø<sub>i</sub>-naka-č in ni-k<sub>i</sub>-kʷa*

But this entails a parallel rewriting of all sentences, thus:

- (12) (a) *Ø<sub>i</sub>-čōka in Ø<sub>i</sub>-konē-č (=7 c)*
- (b) *Ø<sub>i</sub>-konē-č in Ø<sub>i</sub>-čōka (=10 a)*
- (c) *Ni-k<sub>i</sub>-kʷa in Ø<sub>i</sub>-naka-č (=7 b)*

In other words: in Nahuatl phrases like *in N*, which look so much like banal *Det-N* noun phrases, the *N* is a predicative word, fitted with a 3rd person subject marker ('the entity which is *N*'). Evidence in support for this point is the existence of argument phrases in the 1st or 2nd persons (although the referential evidence makes them unusual):

- (13) *Ni-čōka in ni-konē-č.*  
SBJ.1.SG-cry DET<sub>1</sub> SBJ.1.SG-child-ABSL  
'I, who am a child, am crying.'

## 2.8. Adverbial function

Locatives can appear in an argument phrase only if the sentence predicate is itself a locative (14). Otherwise, locatives perform an adverbial function, which is unrelated to the argument structure of the predicate, and therefore not marked by a prefix (15):

- (14) *Ø<sub>i</sub>-kʷal-kān in*  
 SBJ.3-good-DER.LOC DET<sub>1</sub>  
*Ø<sub>i</sub>-t-čān.*  
 SBJ.3-POSS.3.SG-home  
 'His home is a nice place.'
- (15) *Ni-koči in Ø-t-čān.*  
 SBJ.1.SG-sleep DET<sub>1</sub> LOC-POSS.3.SG-home  
 'I sleep in his home.'

### 3. Verb morphology

#### 3.1. Argument and directional prefixes

##### 3.1.1. General outline

**Verb prefixes** appear in a fixed order, with 6 paradigms. The first two are for definite arguments, subject and object; the third is for directional markers (3.1.5); the last three are for 'non-distinct' arguments (Langacker 1976): reflexive, unspecified human, and unspecified non-human. Only definite prefixes show a full 6-person system; reflexives vary in Classical Nahuatl, but have an invariable form *-m(o)-* in many modern dialects (see Tab. 133.1).

The /m:/n/ opposition (in *aM-*, *-kiM-*, *-oN-*) is neutralized before all consonants (16 a–b); (*i*) denotes a support vowel avoiding initial consonant clusters, or middle clusters of more than two consonants (17 a–e); (*o*) in reflexive prefixes appears before a consonant and drops before a vowel. If a verbal stem begins with /iCC-/i, the initial /i-/ drops after *-ka-* in all cases (see (23 c)), and after reflexives (with the /o/ maintained), unless the first consonant is /?/ (compare (19) and (22)).

Subject plural prefixes co-occur with plural suffixes, which vary accordingly to the verbal category (3.2). Although *t(i)-* is a general 1st person plural marker in Classical Nahuatl, some Eastern dialects have an exclusive vs. inclusive opposition: thus we find

in Isthmus Nahuatl *ni-čōka-?* 'SBJ.1.PE-cry-PL (we (= I and someone else) are crying)' vs. *ti-čōka-?* 'SBJ.1.PI-cry-PL (we (= you and I) are crying)'.

- (16) (a) *Am-olini-?*  
 SBJ.2.PL-move-PL  
 'You move.'  
*An-čōka-?*  
 SBJ.2.PL-cry-PL  
 'You are crying.'
- (b) *Ni-kim-itta.*  
 SBJ.1.SG-OBJ.3.PL-see  
 'I see them.'
- Ni-kin-kaki.*  
 SBJ.1.SG-OBJ.3.PL-hear  
 'I hear them.'
- (17) (a) *N-olini.*  
 SBJ.1.SG-move  
 'I move.'
- T-olini-?*  
 SBJ.1.PL-move-PL  
 'We move.'
- (b) *Ni-čōka.*  
 SBJ.1.SG-cry  
 'I am crying.'
- Ti-čōka-?*  
 SBJ.1.PL-cry-PL  
 'We are crying.'
- (c) *Ni-mic-kaki.*  
 SBJ.1.SG-OBJ.2.SG-hear  
 'I hear you (sg.).'  
*N-amēč-kaki.*  
 SBJ.1.SG-OBJ.2.PL-hear  
 'I hear you (pl.).'  
*Ni-k-kaki.*  
 SBJ.1.SG-OBJ.3.SG-hear  
 'I hear him.'
- (d) *K-itta.*  
 OBJ.3.SG-see  
 'He sees him.'

	subject	object	directional	reflexive	human	hon-human
singular	1 <i>n(i)-</i>	<i>-nēč-</i>		<i>-n(o)-</i>		
	2 <i>t(i)-</i>	<i>-mic-</i>		<i>-m(o)-</i>		
	3 <i>Ø</i>	<i>-k(i)-</i>	<i>-oN-</i>	<i>-m(o)-</i>	<i>-tē-</i>	<i>-ka-</i>
plural	1 <i>t(i)-</i>	<i>-tēč-</i>	<i>-wāl-</i>	<i>-t(o)-</i>		
	2 <i>aM-</i>	<i>-amēč-</i>		<i>-m(o)-</i>		
	3 <i>Ø-</i>	<i>-kiM-</i>		<i>-m(o)-</i>		

Tab. 133.1: Verb prefixes

- Ki-kaki.*  
OBJ.3.SG-hear  
'He hears him.'
- (e) *An-k-itta?*  
SBJ.2.PL-OBJ.3.SG-see-PL  
'You (pl.) see him.'
- An-ki-kaki?*  
SBJ.2.PL-OBJ.3.SG-hear-PL  
'You (pl.) hear him.'

### 3.1.2. Reflexive markers

**Reflexive markers** are far more similar to their Spanish (*se*) than to their English (*-self*) counterparts. They are used in five cases: (a) "real" reflexivity, see (18); (b) reciprocity, see (19); (c) motion verbs, see (20); (d) verbs denoting a sentiment or a physical state, see (21 a–b); (e) passive meaning (usually with an inanimate subject), see (22):

- (18) *Ni-k-pāka.*  
SBJ.1.SG-OBJ.3.SG-wash  
'I wash him/it.'
- Ni-no-pāka.*  
SBJ.1.SG-OBJ.REFL.1.SG-wash  
'I wash (myself).'
- (19) *Ti-to-tta?*  
SBJ.1.PL-OBJ.REFL.1.PL-see-PL  
'We see each other.' (cf. (17)–(18))
- (20) *Ni-k-tēka.*  
SBJ.1.SG-OBJ.3.SG-put.down  
'I put it down.'
- Ni-no-tēka.*  
SBJ.1.SG-OBJ.REFL.1.SG-lie.down  
'I lie down.'
- (21) (a) *Ti-k-āwiltia.*  
SBJ.2.SG-OBJ.3.SG-amuse  
'You amuse him.'
- Ti-m-āwiltia.*  
SBJ.2.SG-OBJ.REFL.2/3-amuse  
'You have a good time.'
- (b) *Ki-kokoa.*  
OBJ.3.SG-hurt  
'It hurts him.'
- Mo-kokoa.*  
OBJ.REFL.2/3-hurt  
'He suffers, he is ill.'
- (22) *K-i?toa on.*  
OBJ.3.SG-say DET<sub>2</sub>  
'He says that.'
- M-i?toa on.*  
OBJ.REFL.2/3-say DET<sub>2</sub>  
'That is what people say.'

### 3.1.3. Unspecified prefixes

These are used when the object is unknown or left unspecified (for further details, see Langacker 1976; Launey 1981):

- (23) (a) *Ni-mic-časo?*  
SBJ.1.SG-OBJ.2.SG-love  
'I love you.'
- Ni-tē-časo?*  
SBJ.1.SG-OBJ.INDEF.HUM-love  
'I am in love.'
- (b) *Ni-k-kʷa.*  
SBJ.1.SG-OBJ.3.SG-eat  
'I eat it.'
- Ni-ča-kʷa.*  
SBJ.1.SG-OBJ.INDEF.NHUM-eat  
'I eat.'
- (c) *Ni-tē-itta.*  
SBJ.1.SG-OBJ.INDEF.HUM-see  
'I see someone/some people.'
- Ni-ča-itta.*  
SBJ.1.SG-OBJ.INDEF.NHUM-see  
'I can see.'

Since argument positions must be marked (2.4), forms like \**ni-časo?*, \**ni-kʷa* are impossible.

### 3.1.4. Ditransitive verbs

If the three arguments belong to different paradigms of Tab. 133.1, they appear in the expected order:

- (24) (a) *Ni-tē-ča-maka.*  
SBJ.1.SG-OBJ.INDEF.HUM-OBJ.INDEF.NHUM-give  
'I give something to someone.'
- (b) *Ni-mic-ča-maka.*  
SBJ.1.SG-OBJ.2.SG-OBJ.INDEF.NHUM-give  
'I give you something.'
- (c) *Ni-k-ča-maka*  
SBJ.1.SG-OBJ.3.SG-OBJ.INDEF.NHUM-give  
*in konē-č.*  
DET<sub>1</sub> child-ABSL  
'I give something to the child.'
- (d) *Ni-k-tē-maka*  
SBJ.1.SG-OBJ.3.SG-OBJ.INDEF.HUM-give  
*in naka-č.*  
DET<sub>1</sub> meat-ABSL  
'I give the meat (to someone).'
- (e) *Ti-to-ča-maka?*  
SBJ.1.PL-OBJ.REFL.1.PL-OBJ.INDEF.NHUM-give-PL

- 'We give presents to each other.'
- (f) *Ti-k-to-maka-?*  
SBJ.1.PL-OBJ.3.SG-OBJ.REFL.1.PL-give-PL  
'We give it to each other.'
- If the two objects are definite (i.e. if both prefixes should be taken from the 2<sup>nd</sup> column of Tab. 133.1), only one can occur: two 3rd person prefixes reduce to one, and a 3rd person prefix drops if there is a 1st or 2nd person (even if there is an object phrase):
- (25) (a) *Ni-k-maka.*  
SBJ.1.SG-OBJ.3.SG-give  
'I give it to him.'  
'(I give it) would be:  
*ni-k-tē-maka*  
SBJ.1.SG-OBJ.3.SG-OBJ.INDEF.HUM-give)
- (b) *Ni-mic-maka.*  
SBJ.1.SG-OBJ.2.SG-give  
'I give it to you.'
- (26) (a) *Ni-k-maka naka-č*  
SBJ.1.SG-OBJ.3.SG-give meat-ABSL  
*in konēč-*.  
DET<sub>1</sub> child-ABSL  
'I give the child some meat.'
- (b) *Ni-mic-maka naka-č.*  
SBJ.1.SG-OBJ.2.SG-give meat-ABSL  
'I give you some meat.'

### 3.1.5. Directional markers

**Directional markers** denote a position or movement in relation to a 'center': *-oN-* marks a distance, or an increase in distance, *-wāl-* a decrease in distance. The center can be the *hic et nunc* situation (27 a), but also: an important or respected character, as opposed to a secondary one (27 b); Mexicans, as opposed to foreigners (27 c); the 'human', visible and safe world, as opposed to the supernatural, invisible or dangerous world (27d); the social or moral norm, as opposed to eccentric behavior (27e):

- (27) (a) *Cikʷini. On-cikʷini.*  
run DIST-run  
'He runs.' 'He runs away.'  
*Wāl-cikʷini.*  
PRX<sub>1</sub>-run  
'He comes running.'
- (b) *Nēč-wāl-nōca in tōnatiw.*  
OBJ.1.SG-PRX<sub>1</sub>-call DET<sub>1</sub> sun  
'The sun calls me (to him).'

- (c) *Kim-on-ilwia-?* ...  
OBJ.3.PL-DIST-say-PL  
*kin-wāl-nānkilia-?* ...  
OBJ.3.PL-PRX<sub>1</sub>-answer-PL  
'(The Aztecs) say (to the Spaniards)  
... (The Spaniards) answer them ...'
- (d) *Wāl-kīsa in tōnatiw.*  
PRX<sub>1</sub>-rise DET<sub>1</sub> sun  
'The sun rises (comes out).'  
*On-kalaki in tōnatiw.*  
DIST-set DET<sub>1</sub> sun  
'The sun sets (enters yonder).'
- (e) *K-on-iʔtoa piʔpillō-č.*  
OBJ.3.SG-DIST-say childish-ABSL  
'He says (and it is bad to do so)  
some childish (words).'

## 3.2. Verbal categories

### 3.2.1. Stem allomorphs

Verbs (unlike nouns and locatives) mark tense, mood, and aspect. Verb morphology is fairly simple, with only two irregular verbs. All regular stems end in a vowel: most in /-a/, some in /-i/, and five (plus most passives and some impersonals, 6.3.2) in /-o/ (|-ō|).

Verbal categories are marked by suffixes and/or modification of the stem, which has three possible variants: full, middle, and short. According to phonological structure, there are 4 types of regular verbs: Tab. 133.2 shows the stem allomorphs of *koči* 'sleep' (type A), *caʔci* 'shout' (type B), *čoloa* 'jump' (type C), and *kʷa* 'eat' (type D):

	A	B	C	D
full	<i>koči-</i>	<i>caʔci-</i>	<i>čoloā-</i>	<i>kʷā-</i>
middle	<i>koči-</i>	<i>caʔci-</i>	<i>čolō-</i>	<i>kʷā-</i>
short	<i>koč-</i>	<i>caʔci-</i>	<i>čolo?-</i>	<i>kʷa?-</i>

Tab. 133.2: Types of regular verbs

We can see that the full and middle stem allomorphs differ in type C only, and that in type B there are no stem allomorphs (i.e. the stem cannot be shortened).

- Type A is the most usual one. It includes stems ending in /-i/ or /-a/ preceded by a single consonant (except those listed below in type B).
- Type B includes stems ending in /-CCi/, /-CCa/; in /-o/ (in fact, |-ō|, see below); plus all stems ending in /-ka/, /-kā/, and most intransitives in /-wa/ and /-na/; plus monosyllabic stems in /-i/ (|-i|).

- Type C includes stems ending in two vowels (*/-ia/, /-oa/ i.e. |iā|, |oā|*). Note that all long vowels are shortened in final position and before */-ʔ/*.
- Type D includes monosyllabic stems in */-a/*.

### 3.2.2. Tense-aspect-mood markers

Verbal categories are morphologically constructed on one of the stem variants, in the following way:

Full stem:

- Present: no suffix; Pl. */-ʔ/*, e.g. (1 a–c)  
 Imperfect: */-ya/*; Pl. */-ʔ/* (28)  
 Potential: */-ni/*; Pl. */-ʔ/* (29)

Middle stem:

- Future: suffix *| -s-k|* (for *|k|*, which drops when final, see 3.2.3); Pl. */-eʔ/* (30)  
 Irrealis: */-skiya/*; Pl. */-ʔ/* (31)  
 Optative: no suffix; Pl. */-kān/* (32)

Short stem:

- Aorist: suffix *| -k|* (which in the singular appears only after a vowel, i.e. in type B); Pl. */-eʔ/* (33)  
 Pluperfect: */-ka/*; Pl. */-ʔ/* (34)  
 Admonitive: *| -?|* (which appears in type B only); Pl. */-tin/* (35)

(28) *koči-ya* 'sleep-IMPF',  
*caʔci-ya* 'shout-IMPF',  
*čoloā-ya* 'jump-IMPF',  
*λa-kʷā-ya* 'OBJ.INDEF.NHUM-eat-IMPF';  
*koči-ya-?* 'sleep-IMPF-PL', etc.

(29) *koči-ni* 'sleep-POT',  
*caʔci-ni* 'shout-POT',  
*čoloā-ni* 'jump-POT',  
*λa-kʷā-ni* 'OBJ.INDEF.NHUM-eat-POT';  
*koči-ni-?* 'sleep-POT-PL', etc.

(30) *koči-s* (|*koči-s-k|*) 'sleep-FUT',  
*caʔci-s* 'shout-FUT',  
*čolō-s* 'jump-FUT',  
*λa-kʷā-s* 'OBJ.INDEF.NHUM-eat-FUT';  
*koči-sk-e?* 'sleep-FUT-PL', etc.

(31) *koči-skiya* 'sleep-IRR',  
*caʔci-skiya* 'shout-IRR',  
*čolō-skiya* 'jump-IRR',  
*λa-kʷā-skiya* 'OBJ.INDEF.NHUM-eat-IRR';  
*koči-skiya-?* 'sleep-IRR-PL', etc.

(32) *mā koči* (for the particle, see below) 'OPT sleep',  
*mā caʔci* 'OPT shout',  
*mā čolo* 'OPT jump',  
*mā λa-kʷa* 'OPT OBJ.INDEF.NHUM-eat';

*mā koči-kān* 'OPT sleep-OPT.PL',  
*mā caʔci-kān* 'OPT shout-OPT.PL',  
*mā čolō-kān* 'OPT jump-OPT.PL',  
*mā λa-kʷā-kān* 'OPT OBJ.INDEF.NHUM-eat-OPT.PL'.

- (33) *koč* (|*koč-k|*) 'sleep(AOR)',  
*caʔci-k* 'shout-AOR',  
*čolo?* (|*čolo?-k|*) 'jump(AOR)',  
*λa-kʷa?* 'OBJ.INDEF.NHUM-eat(AOR)';  
*koč-k-e?* 'sleep-AOR-PL',  
*caʔci-k-e?* 'shout-AOR-PL',  
*čolo?-k-e?* 'jump-AOR-PL',  
*λa-kʷa?-k-e?* 'OBJ.INDEF.NHUM-eat-AOR-PL',  
 (34) *koč-ka* 'sleep-PLUP',  
*caʔci-ka* 'shout-PLUP',  
*čolo?-ka* 'jump-PLUP',  
*λa-kʷa?-ka* 'OBJ.INDEF.NHUM-eat-PLUP';  
*koč-ka-?* 'sleep-PLUP-PL', etc.  
 (35) *mā koč* (for the particle, see below) 'OPT sleep',  
*mā caʔci?* 'OPT shout',  
*mā čolo?* 'OPT jump',  
*mā λa-kʷa?* 'OPT OBJ.INDEF.NHUM-eat',  
*mā koč-tin* 'OPT sleep-PL', etc.

### 3.2.3. The meaning of the verbal categories

The three basic forms are the present, the future, and the aorist. The present is either a *hic et nunc* imperfective, or a 'general', customary or intemporal present.

In the future and aorist, *| -k|* is a highly polyvalent marker, which I propose to call **participial suffix**. The meaning of this suffix is roughly parallel to the English *have* (see 4.2.5; Launey 1977), so we can paraphrase:

- (36) (a) |Ø-*koči-s-k|  
          SBJ.3-sleep-FUT-PART  
          'he has to sleep' (|-s|, see 6.1.2)  
 (b) |Ø-*koč-k|  
          SBJ.3-sleep-PART  
          'he has slept'**

The future can be a 'real' temporal future, but has often a modal meaning ('be about to', 'have to') and is the usual translation of the Indo-European subjunctive and infinitive (in sentences like *ni-k-neki ti-koči-s* 'SBJ.1.SG-OBJ.3.SG-want SBJ.2.SG-sleep-FUT (I want you to sleep)', *ni-k-neki ni-koči-s* 'SBJ.1.SG-OBJ.3.SG-want SBJ.1.SG-sleep-FUT (I want to sleep)', where the second word is actually an object noun clause). The aorist, when used alone, usually corresponds to a historical past. When preceded by the particle *ō*, it has

the perfect value, which may correspond to: (a) present result of a process; (b) anteriority (in a subordinate clause); (c) present mention of a past event, e.g. *λa-kʷa?* ‘OBJ.INDEF.NHUM-eat (he ate – and this is mentioned in a chronicle)’ vs. *ō λa-kʷa?* ‘PF OBJ.INDEF.NHUM-eat (he has eaten)’, or ‘he ate – yesterday/last year/once, and I am telling you this now’.

The imperfect, the unrealis, and the pluperfect represent the transfer to the past of the three basic forms. The imperfect, like its Romance counterpart, is a past imperfective. The **unrealis** (there is no traditional term: this seems to me the closest approximation) indicates that an event, which at some past moment was thought of as probable, finally did not occur ('he nearly/almost V-ed', 'he should have V-ed – but did not'). It also appears in the main clause of conditional sentences. The pluperfect indicates that at some past moment an event had actually occurred, but that it had no durable result, being cancelled by some later event.

The usual meaning of the **optative** in the 2nd person is imperative. There is a special 2nd person subject *š(i)-* in both the singular and the plural: *ši-koči* ‘OPT.SBJ.2-sleep (sleep!)’, Pl. *ši-koči-kān* ‘OPT.SBJ.2-sleep-OPT.PL (sleep!)’. In the 1st and 3rd persons, where the particle *mā* necessarily precedes the verb, it can mean wish, indirect order, or resignation. It is also used in a subordinate conditional clause if the main clause is in the future.

The **potential** does not refer to a particular event, but (a) to a class of events, a habit, a proclivity (*koči-ni* ‘sleep-POT (he usually sleeps/he is a sleepyhead)’); (b) to a potentiality of event: in this case it has *š(i)-* (see above) in the 2nd person, and is always preceded by a particle: with *mā*, it means a contrary-to-fact wish (*mā ši-koči-ni* ‘OPT.OPT.SBJ.2-sleep-POT (If only you could sleep!)’) and with *λā* a contrary-to-fact assumption (in a conditional subordinate clause, with the unrealis in the main clause).

The **admonitive** (this term is suggested by Andrews 1975 and is more appropriate than the traditional “vetative”) is always preceded by *mā* but has the ‘usual’ 2nd person prefixes *t(i)-*, *aM-*. It is not exactly a negative imperative or optative, but rather a warning that something would be improper or dangerous (*mā ti-koč* ‘OPT SBJ.2.SG-sleep (Beware not to sleep!)’).

### 3.2.4. Purposive forms

Two sets of forms mark a movement prior to the realization of a process. This movement can be inbound (‘to come to V’) or outbound (‘to go to V’). They take the place of the ‘usual’ forms seen above and constitute a reduced system of tense-aspect-mood, with only three forms: **uncompleted**, **completed**, and **optative**. The plural is */-i?* in the uncompleted form, */-?* otherwise.

	uncompleted	completed	optative
inbound	<i>-kiw</i>	<i>-ko</i>	<i>-ki</i>
outbound	<i>-tīw</i>	<i>-to</i>	<i>-ti</i>

Tab. 133.3: Purposive suffixes

### 3.2.5. Irregular verbs

There are only two of these:

- (a) *ka?* ‘to be (somewhere)’ has two stems: |kat| (/ka?/, /kat-/) and |ye|. The first is used in the present *ka?* ‘be’, *kat-e?* ‘be-PL’, and the pluperfect (with a ‘general’ past meaning) *kat-ka* ‘be-PLUP’. The second is used to form a future, an optative, and a potential (*ye-s* ‘be-FUT’, *ye-skiya* ‘be-IRR’, *ye-ni* ‘be-POT’).
- (b) *ya+w* ‘to go’ has two stems, |ya-| and |wī|. They are combined in the present singular *ya+w*, which has *wi-?* as a plural. The optative is *ya+w* in the singular, *wī+yan* in the plural. All other forms are regularly built on the |ya-| stem, which belongs to type D. It has also a form *wī-c* ‘go-PRX<sub>2</sub> (to come)’, *wī-c-e?* ‘go-PRX<sub>2</sub>-PL’, where /-c/ is an apparently archaic synonyme of /-wāl-/ (3.1.5). There is an irregular pluperfect *wīc-a* ‘go-PLUP’. For all other forms (including the present), *wāl-la+w* (= *wāl-ya+w*) ‘PRX<sub>1</sub>-GO’ can be used.

### 3.3. Reduplication

The first syllable of the verb can be reduplicated. There are two patterns:

- (a) /CV:-/ (first vowel long or lengthened): this is an **intensive reduplication**, meaning that the process occurs in a particularly pronounced fashion or in a particularly extensive manner.
- (b) /CV?-/ (with glottal stop): this is an **expressive reduplication**, meaning either that the process occurs in a more subtle or

precise way, or that it is scattered or fragmented:

- (37) (a) *čōka* ‘cry (he cries)’  
*čō-čōka* ‘RDP<sub>1</sub>-cry (he cries with eyes out or the owl hoots)’  
*čoʔ-čōka* ‘RDP<sub>2</sub>-cry (he sobs)’
- (b) *wecka* ‘laugh (he laughs)’  
*wē-wecka* ‘RDP<sub>1</sub>-laugh (he roars with laughter)’  
*weʔ-wecka* ‘RDP<sub>2</sub>-laugh (he smiles)’
- (c) *ki-teki* ‘OBJ.3.SG-cut (he cuts it)’  
*ki-tē-teki* ‘OBJ.3.SG-RDP<sub>1</sub>-cut (he cuts a lot of it)’  
*ki-teʔ-teki* ‘OBJ.3.SG-RDP<sub>2</sub>-cut (he cuts it into small pieces)’

#### 4. Noun morphology

##### 4.1. The lack of verbal categories

Both nouns and verbs can be used as predicates or to form an argument (see 2). The main difference between nouns and verbs here is the lack of tense-aspect-mood categories in the former. To express these categories with a noun predicate, Nahuatl has the verb *ka?* (3.2.5) as a copula, e.g.:

- (38) (a) *Ni-tēkʷ-či*  
SBJ.1.SG-lord-ABSL  
‘I am a lord.’
- (b) *Ni-tēkʷ-či ni-ye-s.*  
SBJ.1.SG-lord-ABSL SBJ.1.SG-be-FUT  
‘I will be a lord.’
- (c) *Ni-tēkʷ-či ni-kat-ka.*  
SBJ.1.SG-lord-ABSL SBJ.1.SG-be-PLUP  
‘I was a lord.’

We see that in Nahuatl, dissimilar in this respect to certain more familiar languages such as Russian or Arabic, the presence of a copula does not prevent the noun predicate from marking the person (i.e. in Nahuatl, one does not say \**tēkʷ-či ni-ye-s*, \**tēkʷ-či ni-kat-ka*). This is a good indication of the fact that Nahuatl nouns lack tense-aspect, not predictability.

##### 4.2. The arguments of the noun predicate

###### 4.2.1. Absolute nouns

A one-argument structure seems to be the norm in nominal predicates, since all possessed nouns have an absolute counterpart, while many absolute nouns have no corresponding possessed form. There are three subclasses of absolute nouns: substantives, ‘short’ nouns, and participials.

**Substantives** are by far the most numerous. In the singular, they have an **absolute suffix** (for the plural, see 4.2.2). There are two such suffixes. The commonest is *-či*, which has variants *-či* after a vowel, *-či* after a consonant other than /l/, and *-li* after /l/ (39). Less common is *-in*, which is found in about 40 nouns, mostly referring to plants and animals (40).

**Short nouns** lack an absolute suffix. They are mostly ‘familiar’ nouns, names of some animal or vegetal species, or referring (by metaphor or synecdoche) to a physical or moral (mostly deprecatory) feature (41).

**Participials** end with the participial suffix (3.2.3), which always has the */-k/* allomorph after a vowel, but can appear as */-ki/* or zero after a consonant. The plural is always in */-keʔ/*, and before other suffixes (except the vocative, 4.3), the allomorph is */-kā-/* (e.g. (47 a)). With the exception of about 15 denominatives (42 a) and 4 or 5 grammatical terms (42 b), the most numerous ones are deveritative (6.1.1) or ‘possessor nouns’ (4.2.5).

- (39) *ā-či* ‘water-ABSL’  
*tōč-či* ‘rabbit-ABSL’  
*kal-li* ‘house-ABSL’
- (40) *mič-in* ‘fish-ABSL’  
*sičāl-in* ‘star-ABSL’
- (41) *čiči* ‘dog’;  
*capa* ‘dwarf’;  
*cocoka* ‘warty’  
(*cocoka-či* ‘wart-ABSL’)
- (42) (a) *istā-k* ‘salt-PART (white)’  
(*ista-či* ‘salt-ABSL’)  
(b) *iw* ‘such’ or *iw-ki* ‘such-PART’  
(*iw-keʔ* ‘such-PART-PL’)  
*iyoʔ-k-e?* ‘alone-PART-PL’

###### 4.2.2. Absolute plural morphology

Only nouns referring to animates normally have a plural form. Some inanimate nouns can be put into the plural if they are used metaphorically, e.g. *am-āwēwē-me?* ‘SBJ.2.PL-cypress-PL (you (pl.) are cypresses (i.e. protective))’. *tepē-či* ‘mountain-ABSL’ and *sičāl-in* ‘star-ABSL’ are considered as animates (*tē-tepe-?* ‘RDP<sub>1</sub>-mountain-PL (mountains)’, *sī-sičāl-tin* ‘RDP<sub>1</sub>-star-PL (stars)').

If all participials have */-k-e?* as the plural form, the situation is somewhat more chaotic for substantives and short nouns. Three suffixes can occur: */-?* after a vowel, */-tin/* after a consonant, and */-me?* after both vowels and consonants. The use of the latter, some-

what restricted in Classical Nahuatl, is far more widespread in most modern dialects. The first two can occur with a /CV:/ (3.3) reduplication, so there are 5 possible patterns:

- (43) (a) RDP + /ʔ/:  
*kō-kone-ʔ* 'RDP<sub>1</sub>-child-PL (children)'
- (b) RDP + /-tin/:  
*tō-tōč-tin* 'RDP<sub>1</sub>-rabbit-PL (rabbits)'
- (c) /-ʔ/:  
*siwa-ʔ* 'woman-PL (women)'
- (d) /-tin/:  
*okič-tin* 'man-PL (men)'
- (e) /-meʔ/:  
*pico-meʔ* 'pig-PL (pigs)'

#### 4.2.3. Possessive markers and possessor phrases

The **possessor** behaves in all respects like the second argument place of the noun predicate. It is marked (like the definite object) by a 2nd position prefix. The paradigm is:

	singular	plural
1	<i>n(o)-</i>	<i>t(o)-</i>
2	<i>m(o)-</i>	<i>am(o)-</i>
3	<i>t-</i>	<i>tM-</i>
indefinite	<i>tē-</i>	

Tab. 133.4: Possessor marking

The final /o/ appears before a consonant and drops before a vowel (in /iCC-/ stems, either the prefix /o/ or the stem /i/ may drop). A possessor noun phrase is not marked by case, but shows the same prefix coindexation as subject or object phrases:

- (44) (a) *t-kal* *in*  
(SBJ.3)<sub>i</sub>POSS.3.SG-house DET<sub>1</sub>  
*siwā-č*  
(SBJ.3)<sub>i</sub>woman-ABSL  
'(it is) the woman's house (lit. her house the woman)'
- (b) *to-kal* *in*  
POSS.1.PL-house DET<sub>1</sub>  
*ti-siwa-ʔ*  
SBJ.1.PL-woman-PL  
'(it is) our house, of us women' (cf. (13))

In spite of what is said in many grammars, there seems to be no real 'inalienably possessed' class of nouns. It is however true that kinship terms and names of body parts usu-

ally appear in the possessed form: and if no reference is made to a specific possessor, kinship terms appear with *tē-* and body parts with *t(o)-*:

- (45) (a) *tē-nān*  
POSS.INDEF-mother  
'(it is) a (someone's) mother'
- (b) *to-mā*  
POSS.1.PL-hand  
'(it is) a/the (our) hand'

Such forms are more usual than their absolute counterparts *nān-či* 'mother-ABSL', *māi-č* 'hand-ABSL', but the latter do occur in the corpus.

#### 4.2.4. Possessive suffixation

Possessed nouns always drop the absolute suffix (4.2.1) should there be one. A possessive suffix appears if there is an 'external' possessive relation (someone's thing, someone's relative ...). If the relation is 'internal' (part or abstract characteristic), there is no suffix. But this opposition is seldom visible (see below).

The possessive suffix in the singular is |-w| with four allomorphs: /-wl/, zero, /-wi/, and /-º/ (this stands for an elision of the final vowel). Their distribution is somewhat complex, but the rules seem to be:

- (a) consonant-ending nouns have zero, unless they fall under rule (e) (46 a);
- (b) vowel-ending stems, if animate, have always /-wl/ (46 b);
- (c) inanimate stems ending in short /a/ or /i/ drop this vowel (= have /-º/) unless they fall under rules (d) or (e) (46 c), but have /-w/ if they end in another vowel (46d);
- (d) if /-º/ leaves to final consonants, a final /i/ is added (or restored) (46 e);
- (e) no possessed form can have only one syllable if the absolute form has two: so /VC/ stems have /-wi/ (46 f) and /V р/ stems do not fall under rule (c) (46 g).

There are, however, a few exceptions to these rules: some stems ending in /-kal/, /-ki/ or /-či/ do not apply rule (c) (46h):

- (46) (a) *no-kal* 'POSS.1.SG-house (my house)'  
(*kal-li* 'house-ABSL');  
*no-tōč* 'POSS.1.SG-rabbit (my rabbit)'  
(*tōč-či* 'rabbit-ABSL')
- (b) *no-siwā-w* 'POSS.1.SG-woman-RELL (my wife)' (*siwā-či* 'woman-ABSL');  
*no-tīsi-w* 'POSS.1.SG-doctor-RELL (my doctor)' (*tīsi-či* 'doctor-ABSL')

- (c) *no-nak* ‘POSS.1.SG-meat (my meat)’  
(*naka-*‘meat-ABSL’);  
*no-mā* ‘POSS.1.SG-hand (my hand)’  
(*māt-*‘hand-ABSL’)
- (d) *n-ā-w* ‘POSS.1.SG-water-RELL (my water)’ (*ā-*‘water-ABSL’);  
*no-te-w* ‘POSS.1.SG-stone-RELL (my stone)’ (*te-*‘stone-ABSL’)
- (e) *no-mašči* ‘POSS.1.SG-loincloth (my loinloth)’ (*mašča-*‘loincloth-ABSL’);  
*no-čatki* ‘POSS.1.SG-burden (my burden)’ (*čatki-*‘burden-ABSL’)
- (f) *n-o?-wi* ‘POSS.1.SG-path-RELL (my path)’ (*o?-či*‘path-ABSL’)
- (g) *n-āma-w* ‘POSS.1.SG-paper-RELL (my paper)’ (*āma-*‘paper-ABSL’);  
*n-omi-w* ‘POSS.1.SG-bone-RELL (my bone)’ (*omi-*‘bone-ABSL’)
- (h) *no-meka-w* ‘POSS.1.SG-rope-RELL (my rope)’ (*meka-*‘rope-ABSL’);  
*no-teki-w* ‘POSS.1.SG-duty-RELL (my duty)’ (*teki-*‘duty-ABSL’);  
*no-šōči-w* ‘POSS.1.SG-flower-RELL (my flower)’ (*šōči-*‘flower-ABSL’)

Names of body parts only have zero and /-o/, but since none of them ends in a long vowel, this is not conclusive proof of the lack of possessive suffix. A form like *no-kši* ‘POSS.1.SG-foot (my foot)’ or *n-ikši* ‘POSS.1.SG-foot (my foot)’ (*ikši-*‘foot-ABSL (foot)’) can be due to the lack of suffix, but also to rule (d), see (46e). Since the allomorph of the participle suffix is /-kā-/ before another suffix (4.2.1), participles end in /-kā-w/ in the possessed form; but here a /-ka/ form appears if there is an ‘internal’ relation, often translatable by an abstract noun:

- (47) (a) *ī-istā-kā-w*  
POSS.3.SG-white-PART-RELL  
'his white (thing)'
- (b) *ī-istā-ka*  
POSS.3.SG-white-PART  
'his whiteness'

All plurals are in /-wān/ (possibly /-wā-n/, where /-wā-/ is an allomorph of the possessive suffix), with no reduplication. Since only animates can be put into the plural (4.2.2), this is always an ‘external’ relation:

- (48) *no-tōč-wān*  
POSS.3.SG-rabbit-RELL.PL  
'my rabbits'

*no-konē-wān*  
POSS.3.SG-child-RELL.PL  
'my children'

#### 4.2.5. Possessor nouns

The English verb *have* expresses a sort of inverse possessive relation, where the ‘possessor’ appears as a subject and the ‘possessed’ noun as an object. To express this relation, Nahuatl has ‘possessor nouns’, which can be paraphrased as ‘owner of N’ or ‘provided with N’. This is marked by either /-wa?/ or /-e?/ suffixed to the noun stem. The rule seems to be: If the possessed form of the stem is marked by /-w/ or /-wi/, or plural /-wān/ in either the singular or the plural, then the proprietive allomorph is /-wa?/ (cf. (49 a)), otherwise it is /-e?/ (cf. (49 b)); but this does not apply to ‘mother’ and ‘father’ (49 c). All these nouns are participials, so they have /-k-e?/ in the plural and /-kā-/ in derivation (54 a–b):

- (49) (a) *Siwā-wa?*  
woman-PROPR  
'He has a wife.' (cf. 46 b)  
*Tōč-wa?*  
rabbit-PROPR  
'He has a rabbit/rabbits.' (cf. 50 a)  
*Te-wa?*  
stone-PROPR  
'He has stones.' (cf. 46 d)  
*Āma-wa?*  
paper-PROPR  
'He has paper.' (cf. 46 g)
- (b) *Kal-e?*  
house-PROPR  
'He has a house.' (cf. 46 a)  
*Mašč-e?*  
loincloth-PROPR  
'He has a loinloth.' (cf. 46 e)  
*Nak-e?*  
meat-PROPR  
'He has meat.' (cf. 46 c)
- (c) *Nān-e?*,      *tat-e?*  
mother-PROPR father-PROPR  
'He has a mother (and) a father.'
- (50) (a) *Siwā-wa?-k-e?*  
woman-PROPR-PART-PL  
'They have wives.'  
*Kal-e?-k-e?*  
house-PROPR-PART-PL  
'They have houses.'

- (b) *No-kal-eʔ-kā-w*  
 POSS.1.SG-landlord-PROPR-PART-RELL  
 '(He is) my landlord.'

#### 4.3. The vocative

Nouns referring to human beings can be put into the **vocative**. These are the only oxytones in Nahuatl. In men's language, a stressed suffix /-é/ is added; in women's language, there is a stress shift onto the last syllable. The absolute suffix, if any, is maintained:

- (51) (a) *Konē-č-é!*  
 child-ABSL-VOC  
 'O child!'  
*Ko-koneʔ-é!*  
 RDP<sub>1</sub>-child:PL-VOC  
 'O children!'  
*Tēkʷ-č-é!*  
 lord-ABSL-VOC  
 'O lord!'  
*Tē-tēkʷ-tin-é!* (men's speech)  
 RDP<sub>1</sub>-lord-PL-VOC  
 'O lords!'  
 (b) *Koné-č!*  
*Kō-koné-?*  
*Tēkʷ-č!*  
*Tē-tēkʷ-tin!*  
 (id., women's speech)

#### 4.4. Abstract nouns

A suffix /-yō-/ can appear on noun stems to form substantives meaning: (a) N-qualities; (b) N's most characteristic features; (c) a set of N-s:

- (52) *siwā-yō-č*  
 woman-ABSTR.NR-ABSL  
 'femininity/vulva/the fair sex'  
*nān-yō-č*  
 mother-ABSTR.NR-ABSL  
 'motherhood/motherly love'  
*mešiʔka-yō-č*  
 mexico-ABSTR.NR-ABSL  
 'Mexican culture/product of Mexico/  
 Mexican nation'  
*istā-kā-yō-č*  
 white-PART-ABSTR.NR-ABSL  
 'whiteness'

As can be expected (4.2.4), the possessed form appears with no possessive suffix (53), but abstract nouns derived from participials have *-ka* rather than *-kā-yo* (47 b). In some cases, it is possible to oppose an 'external' and an 'internal' (with *-yo*) possession (54 a–c):

- (53) *t-siwā-yo*  
 POSS.3.SG-woman-ABSTR.NR  
 'her femininity'  
*t-nān-yo*  
 POSS.3.SG-mother-ABSTR.NR  
 'her motherly love'  
 (54) (a) *t-nak*  
 POSS.3.SG-meat  
 'his meat'  
*t-naka-yo*  
 POSS.3.SG-meat-ABSTR.NR  
 'his flesh, his body'  
 (b) *t-omi-w*  
 POSS.3.SG-bone-RELL  
 'his skeleton'  
*t-omi-yo*  
 POSS.3.SG-bone-ABSTR.NR  
 'his bone (e.g. in his plate)'  
 (c) *t-ā-w*  
 POSS.3.SG-water-RELL  
 'his water'  
*t-ā-yo*  
 POSS.3.SG-water-ABSTR.NR  
 'its juice'

There are possessor nouns in /-yo?/, which in contrast to those in /-wa?/ and /-e?/ (4.2.5) mark an internal or constitutive relation which can be paraphrased by 'made of', 'full of':

- (55) (a) *te-yo?*      *ā-yo?*  
 stone-CNST water-CNST  
 'stony'      'juicy'  
 (b) *Ti-naka-yo?-k-e?*,  
 SBJ.1.PL-meat-CNST-PART-PL  
*t-omi-yo?-k-e?*,  
 SBJ.1.PL-bone-CNST-PART-PL  
 'We are made of flesh and bones.'

#### 4.5. Grammatical nouns

##### 4.5.1. Quantifiers

These include numerals (*ōme* 'two', *ēyi* 'three', *nāwi* 'four') and non numeral (*moč* or *moči* 'all', *miyak* 'many, much') quantifiers, as well as interrogatives (*kēškič* 'how big? how much?', *kēški* 'how many?'), identifiers (*tškič* 'that big, so much', *tski* 'so many'), and others like *wēyi* 'big'. Plurals are in *-n* or *-ntin* after a vowel, *-in*, *-tin* or *-intin* after a consonant. When referring to animates, they appear in the plural:

- (56) (a) *ōme* ‘two (houses, trees, stones ... )’  
 (b) *ōme-ntin*  
 zwei-PL  
 ‘two (women, dogs ... )’

*sē* ‘one’ is apart. It has a plural *sē-me?* ‘one-PL’ meaning ‘one of them’.

#### 4.5.2. Disjunctive predicates

For this term, and the predicative status of disjunctives, see Launey (1986: 645). These include interrogatives *āk* ‘who?’, *xe?* ‘what?’, their combination with the negation *ay-āk* ‘NEG-who (nobody)’, *a?-xe(in)* ‘NEG-what (nothing)’ and the indefinites *aka?* ‘someone’ and *iča?* ‘something’.

#### 4.5.3. Identifiers

These are the so-called ‘tonic’ personal pronouns. They are formed from a predicative (*y)e?*, Pl. (*y)e?-wān, which means roughly ‘(to be) the entity involved’. In the singular a suffix *-wā-č*, in the plural *-tin* may appear with no clear change in meaning. Thus:*

- (57) *Ka ye?-(-wā-č)*.  
 ASRT (SBJ.3)IDENT(-Ø-ABSL)  
 ‘It is him/her.’ / ‘That’s it.’

This can be applied to a 1st or 2nd person, hence the full paradigm:

	singular	plural
1	<i>n-e?-(-wā-č)</i>	<i>t-e?-wān(-tin)</i>
2	<i>t-e?-(-wā-č)</i>	<i>am-e?-wān(-tin)</i>
3	<i>ye?-(-wā-č)</i>	<i>ye?-wān(-tin)</i>

Tab. 133.5: Identifiers

But the predicative structure is ambiguous: *n-e?-wā-č* ‘SBJ.1.SG-IDENT-Ø-ABSL’ can be interpreted as ‘I am that’ (with the predicate *-e?* applied to a 1st person), but can also be reinterpreted as ‘it is me’ (with a new predicate ‘be me’ applied to a 3rd person, see Launey 1986: 639–644). Of course, identifiers can be used as arguments as well as predicates (58 a–b). They are also used to avoid indefinite predicates (2.6), see (59).

- (58) (a) *Ka n-e?-wā-č in*  
 ASRT SBJ.1.SG-IDENT-Ø-ABSL DET<sub>1</sub>  
*ni-čōka.*  
 SBJ.1.SG-cry  
 ‘It is me who is crying.’  
 (b) *In n-e?-wā-č ka*  
 DET<sub>1</sub> SBJ.1.SG-IDENT-Ø-ABSL ASRT  
*ni-čōka.*  
 SBJ.1.SG-cry  
 ‘As for me, I am crying.’

- (59) *Ka ye?-wā-č in*  
 ASRT (SBJ.3)IDENT-Ø-ABSL DET<sub>1</sub>  
*konē-č*  
 child-ABSL  
 ‘It is the child (lit. the child is him).’

## 5. Locative morphology

### 5.1. Locative subclasses

Although **locatives** form a very heterogeneous class, there are two main subclasses: ‘suffixed’ locatives and ‘auto-locatives’. The only productive locative suffix is */-ko/* (after /C/ or monosyllabic stem) or */-k/* (after /V/); it appears on noun stems:

- (60) *kal-ko* ‘house-LOC (in (the) house)’  
*xe-ko* ‘fire-LOC (in (the) fire)’  
*kaši-k* ‘plate-LOC (in the/a plate)’

All the others are auto-locatives. These include:

- (a) unsegmentable ‘adverbs’ (61 a);
- (b) ‘pseudo-prepositions’, which appear possessed (see 5.2) or in some cases composed with (not suffixed to: they are stems, not affixes) noun stems (61 b);
- (c) deveritative (61 c);
- (d) denominative (61 d);
- (e) quantifiers, all derived (61 e);
- (f) disjunctive (61 f);
- (g) identifiers (61 g).

- (61) (a) *mōsčā* ‘tomorrow’  
*čahwīs* ‘to no avail’  
 (b) *teō-pan*  
 god-DER.LOC  
 ‘at the temple’  
*ā-čan*  
 water-DER.LOC  
 ‘under the water’  
*tepē-ti-kpak*  
 mountain-Ø-DER.LOC  
 ‘on top of the mountain’  
 (-ti-, a synchronically meaningless suffix, occurs with some of these units)  
 (c) *kočī-wa-yān*  
 sleep-IMPR-DER.LOC  
 ‘sleeping room’  
 (where people sleep, *kočī-wa* ‘sleep-IMPR’, see 6.3.1)  
*čā-namaka-kān*  
 OBJ.INDEF.NHUM-sell-DER.LOC  
 ‘selling place’

- (d) *miš-čān*  
cloud-DER.LOC  
'place of clouds'
- tepē-čā?*  
mountain-DER.LOC  
'mountainous country'
- (e) *ōp-pa* 'two-time (twice)'  
*ōk-čān* 'two-DER.LOC (in two places)'
- (f) *kān* 'where?'  
*īk* 'when?'  
*kēn* 'how?'
- (g) *nikān* 'here'  
*ōmpa* 'over there'  
*i?kʷāk* 'at that moment'

Some of the subclass (b) autolocatives have an absolute nominal counterpart. Thus *teō-pan* 'god-on/by' refers to a temple regarded as a place; but if it is regarded as a 'thing' (a building), the noun *teō-pan-či* 'god-on/by-ABSL' may occur. But there is no suffixed locative \**teō-pan-ko* (only *teō-pan* 'god-on/by', hence 'auto-locative').

## 5.2. Possessed locatives

Some locatives have a possessed form. In this case, there is no possessive suffix (see 4.2.4). This is true of many suffixed locatives (62) and all 'pseudo-prepositions' (61 b, 63), which do look like prepositions when a possessor phrase occurs (64 a–c):

- (62) *no-kaši-k*  
POSS.1.SG-plate-PART  
'in my plate'
- (63) *no-pampa* 'POSS.1.SG-because (because of me)'  
*ī-wān* 'POSS.3.SG-with (with him)'  
*to-pan* 'POSS.1.PL-on/by (on/by us)'  
*amo-teč* 'POSS.2.PL-about (about you)'
- (64) (a) *ī-pampa*      *in konē-č*  
POSS.3.SG-because DET<sub>1</sub> child-ABSL  
'because of the child'  
(b) *ī-wān*      *in konē-č*  
POSS.3.SG-with DET<sub>1</sub> child-ABSL  
'with the child'  
(c) *īm-pan*      *in kō-kone-?*  
POSS.3.PL-on/by DET<sub>1</sub> RDP<sub>1</sub>-child-PL  
'on/by the children'

## 6. Derivation

### 6.1. Deverbative nouns

#### 6.1.1. Active nouns

Active nouns refer to a person who is in the state or performs the action expressed by the verb stem. They are generally (but not al-

ways) translated by adjectives if they are formed on intransitive stems, by agent nouns if they are formed on transitive stems. In the latter case, they have a non-distinct (see 3.1.1) object prefix.

They are participials (see 4.2.1) and formed exactly like the aorist (see 3.2.2), except that the suffix appears as /-ki/ (instead of zero) after /C/ (except /-ʔ/). The forms are thus homonymous: in the singular in type (B) (ending in /-k/ after /V/, 3.2.1), in type (C) (ending in zero after /-ʔ/), and in all plurals (ending in /-k-eʔ/):

- (65) (a) *Koč-ki.*  
sleep-PART  
'(He is) asleep.'
- Mik-ki.*  
die-PART  
'(He is) dead.' / '(It is) a corpse.'
- (b) *Tē-yakān-ki.*  
OBJ.INDEF.HUM-guide-PART  
'(He is a) ruler.'  
(*yakāna* 'to guide')  
*Mo-pōw-ki.*  
OBJ.REFL.2/3-count-PART  
'(He is a) braggart.'  
(*mo-pōwa* 'OBJ.REFL.2/3-count (lit. he counts himself)')
- ča-namaka-k.*  
OBJ.INDEF.NHUM-sale-PART  
'(He is a) salesman (also: he sold something, from *namaka*).'  
*ča-ʔkʷilo?*  
OBJ.INDEF.NHUM-write  
'(He is a) scribe (also: he wrote, from *i?kʷiloa*).'
- (c) *Mik-k-e?*  
die-PART-PL  
'(They are) dead (also: they died).'  
*ča-namaka-k-e?*  
OBJ.INDEF.NHUM-sale-PART-PL  
'(They are) salesmen (also: they sold sth.).'  
*ča-ʔkʷilo?-k-e?*  
OBJ.INDEF.NHUM-write-PART-PL  
'(They are) scribes (also: they wrote).'

In spite of the homonymy (and a probable diachronic identity, due to the aspectual properties of the aorist, see 3.2.3), the nominal vs. verbal features are conspicuous. Thus:

- (a) the aorist, not the noun, can occur with the particle *ō* (66 a);

- (b) the transitive aorist (not the noun) can have a distinct object (66 b);
- (c) active nouns can have a possessed form (66 c);
- (d) active nouns can have a vocative (66 d);
- (e) active nouns can occur with a temporal copula (66 e):

(66) (a) *Ō tē-yakān.*

PF OBJ.INDEF.HUM-guide  
'He has ruled (people).'  
(\**ō tē-yakān-ki*)

(b) *Nēč-yakān.*

OBJ.1.SG-guide  
'He ruled me.'  
(\**nēč-yakān-ki*)

(c) *No-tē-yakān-kā-w.*

POSS.1.SG-OBJ.INDEF.HUM-guide-PART-RELL  
'(He is) my ruler.'

(d) *Tē-yakān-k-é!*

OBJ.INDEF.HUM-guide-PART-VOC  
'O ruler!'

(e) *Tē-yakān-ki*

*kat-ka.*  
OBJ.INDEF.HUM-guide-PART be-PLUP  
'He was a ruler.'  
(\**tē-yakān kat-ka*)

#### 6.1.2. Action nouns

Action nouns refer to the process itself. They are substantives (4.2.1) and usually marked by /-s-/ on intransitive verb stems, by /-li-s/ on the middle stem (see 3.2.1) of transitive stems. In the latter case, an indefinite object prefix appears: *tē*, *Āa-*, or *ne-* (not *mo-*), which is an indefinite reflexive marker (6.3.2). (In /-li-s-/; /-li-/ is probably an allomorph of /-l-/ (6.1.3, 6.3.2), see Launey 1986: 240, and for a different opinion, Dakin 1988.)

(67) (a) *koči-s-Āi*

sleep-PROC-ABSL  
'sleep'

*miki-s-Āi*

die-PROC-ABSL  
'death'

(b) *tē-yakāna-lis-Āi*

OBJ.INDEF.HUM-guide-PROC-ABSL  
'act of guiding'

*ne-pōwa-lis-Āi*

IMPR.REFL-count-PROC-ABSL  
'arrogance' ('act of counting oneself')

*Āa-2kʷilō-lis-Āi*

OBJ.INDEF.NHUM-write-PROC-ABSL  
'writing'

#### 6.1.3. Object nouns

Object nouns refer to a thing or person involved in the process or resulting from the process, and can be translated by nouns or past participles. They are substantives. Most of them are formed from transitive verb stems with the prefix *Āa-* (even if they refer to human beings) and the suffix /-l-/ on the middle stem (see 3.2.1); some stems (especially those ending in /-ka/ or /-ki/) form object names directly on the short stem (68 a–b) (for object nouns with *tē-* or *ne-*, see Andrews 1975: 233–239; Launey 1979: 1002–1004). Occasionally, object nouns appear on intransitive or impersonal verb stems: they have no prefix and generally no suffix before the absolute suffix /-Ā/ (although /-l/ may occur after /-a/, see (69)).

(68) (a) *Āa-yakāna-l-li*

OBJ.INDEF.NHUM-guide-OBJNR-ABSL  
'(person) ruled, subject'

*Āa-2kʷilō-l-li*

OBJ.INDEF.NHUM-write-OBJNR-ABSL  
'piece of writing, written'

*Āa-kʷa-l-li*

OBJ.INDEF.NHUM-eat-OBJNR-ABSL  
'food, eaten'

*Āa-čīwa-l-li*

OBJ.INDEF.NHUM-make-OBJNR-ABSL  
'made, object, creature'

also *Āa-čī-w-Āi*

'OBJ.INDEF.NHUM-make-RELL-ABSL')

(b) *Āa-namak-Āi*

OBJ.INDEF.NHUM-sale-ABSL  
'sold'

(69) *kʷīka-Ā* 'sing-ABSL (song)'

*kiyawi-Ā* 'rain-ABSL (rain)'

*yowa-l-li* 'be.dark-OBJNR-ABSL (night)'

#### 6.2. Denominative verbs

##### 6.2.1. The |t| series

These occur on noun (sometimes locative) stems, and include: (a) /-ti/, which forms intransitive verbs meaning 'to become N', 'to behave like N', sometimes 'to accomplish N' (70); (b) /-tia/, which forms either intransitive verbs meaning 'to provide someone with N' (71 a) or ditransitive verbs meaning 'to provide someone with something as an N' (71 b); (c) /Āa-/; which forms only a few transitive verbs meaning 'to consider N' (72):

(70) *Āāka-ti.*

human-DETR

'He is born (lit. becomes human).'

*Okič-ti.*

behavior-DETR

‘He behaves like a man.’

*Tē-yakān-kā-ti*

OBJ.INDEF.HUM-guide-PART-DETR

‘He performs the functions of ruler.’  
(see (71))

- (71) (a) *Ni-k-kal-tia.*

SBJ.1.SG-OBJ.3.SG-house-CAUS

‘I give him or make him a house.’

*K-īš-tia.*

OBJ.3.SG-eye-CAUS

‘He spies for him (lit. He gives him eyes).’

- (b) *Ni-mic-no-nān-tia.*

SBJ.1.SG-OBJ.2.SG-POSS.1.SG-mother-CAUS

‘I take you as my mother.’

*Ni-mic-koska-tia*

SBJ.1.SG-OBJ.2.SG-necklace-CAUS

in Šōči-č.

DET<sub>1</sub> flower-ABSL

‘I give you the flowers as a necklace.’

- (72) *Ki-časo?*-ča.

OBJ.3.SG-love-TRR

‘He loves him.’

### 6.2.2. The |w| series

Only two are commonly used: (a) /-wa/, which only appears on abstract noun stems (see 4.4) and means ‘be full of N’ (73); (b) /-wia/, which forms transitive verbs meaning ‘to apply N to’, ‘to put into N’ (74):

- (73) *Ā-yō-wa.*

water-ABSTR.NR-VR<sub>1</sub>  
‘It is full of water.’

*Aska-yō-wa.*

ant-ABSTR.NR-VR<sub>1</sub>  
‘It is full of ants.’

- (74) *K-es-wia.*

OBJ.3.SG-blood-VR<sub>2</sub>  
‘He covers it with blood.’

*Ki-ma?*pil-wia.

OBJ.3.SG-finger-VR<sub>2</sub>  
‘He points his finger at him.’  
*Ā-kimil-wia.*  
OBJ.INDEF.NHUM-bundle-VR<sub>2</sub>  
‘He carries something in a bundle.’

### 6.3. The use of derivational suffixes in the verbal voice

#### 6.3.1. Impersonal verbs

The suffix /-wa/ (see (73)) occurs on intransitive verb stems to form **impersonal verbs**, meaning ‘someone/some people/everybody

V-s’. It can bring about some alterations in the stem, such as the lengthening or transformation of the last vowel into /ō/, and/or the palatalization of the preceding consonant (75 a). With stems ending in /-a/, alternative suffixes /-o/ (/⁰/ = loss of vowel) or /-l-o/ can appear (the latter at least is probably an analogical loan from passives, see 6.3.2); the final /-o/ is certainly a post-consonant allomorph of /-wa/ (75 b):

- (75) (a) *Kočī-wa.*

sleep-VR<sub>1</sub>

‘Some people sleep.’

*Mikō-wa. (miki)*

die-VR<sub>1</sub>

‘People die.’

*Nēšō-wa. (nēsi)*

appear-VR<sub>1</sub>

‘People appear.’

- (b) *Kʷik-o. (kʷika)*

sing-VR<sub>1</sub>

‘People sing.’

*Mayāna-lo.*

hungry-VR<sub>1</sub>

‘People are hungry.’

From (73) and (75), |-wa| appears to be a sort of existential predicate (‘there is V’). We can also see that the forms of (75) are semantically equivalent to indefinite-subject verbs. But forms as \**tē-koči* never occur: (75 a–b) is used instead. Note however that this is true of human-subject verbs. In the case of a non-human subject, Āa- is used (76): this fact suggests that intransitive verbs are derived from two-place structures, one of the places being void (see Launey 1981; see also ‘unaccusativity’ vs. ‘unergativity’ in Perlmutter 1983, ed.):

- (76) *Āa-wāki.*

OBJ.INDEF.NHUM-dry

‘Things are drying up.’

#### 6.3.2. Passives

There can be no indefinite subject in transitive verbs. Instead, **passives** are used. They are usually marked by /-l-o/ on the middle stem (sometimes /-o/, and sometimes /-wa/ after /-i/), /-o/ being a post-consonant allomorph of /-wa/. In other words, the existence of the process is marked (by |-wa|, see 6.3.1), but at the same time related to an entity, so there is a reorientation (marked by /-l-/):

- (77)
- Ni-nōca-lo.*

SBJ.1.SG-call-PASS

'I am called, someone calls me.'

*N-itta-lo* or *N-itt-o.*

SBJ.1.SG-see-PASS

SBJ.1.SG-see-VR<sub>1</sub>

'I am seen, people see me.'

Ditransitive verbs (see 3.1.4) can be put into the passive. Only an object place overtly marked in the active form can be transformed into a subject; and a 'dropped' prefix does not reappear:

- (78) (a)
- Ti-ča-mak-o.*

SBJ.2.SG-OBJ.INDEF.NHUM-give-PASS

'You are given something.' (cf. 24 b)

(b) *Ti-mak-o naka-č.*

SBJ.2.SG-give-PASS meat-ABSL

'You are given meat.' (cf. 26 b) (not \**ti-k-mak-o*)

Impersonal forms can be taken from transitive verb stems with an indefinite prefix (*ne-* if reflexive):

- (79)
- Tē-nōca-lo.*

OBJ.INDEF.HUM-call-PASS

'People are called.'

*ča-kʷa-lo.*

OBJ.INDEF.NHUM-eat-PASS

'People eat.'

*Ne-pōwa-lo.*

IMPR.REFL-count-IMPR

'People are arrogant (lit. count themselves.)'

### 6.3.3. Passive and impersonal potentials

A passive potential (3.2.3) refers to a passive ability or necessity (80). An impersonal potential refers to an instrument, or more generally to a condition that brings about the process (81):

- (80)
- kʷa-lō-ni*

eat-PASS-POT

'edible'

*itta-lō-ni* or *itt-ō-ni*

see-PASS-POT

see-PASS-POT

'visible'

- (81)
- kočt-wa-ni*

sleep-IMPR-POT

'soporific (by which people sleep)'

*ča-tta-lō-ni* (or *ča-tt-ō-ni*)

OBJ.INDEF.NHUM-see-IMPR-POT

'device with which to see (e.g. magnifying glass)'

### 6.3.4. Causatives

**Causative** verbs indicate that the process is caused by an external agent. This adds one argument place to the verb. The 'new' agent appears as the subject and the 'old' subject is demoted to the status of an object. These verbs are marked by */-tia/* or */-tia/* on intransitive stems, */-l-tia/* or */-tia/* on transitive stems (with some fluctuations):

- (82) (a)
- Ni-k-kočt-tia.*

SBJ.1.SG-OBJ.3.SG-sleep-CAUS

'I make him sleep.'

(b) *Ni-mic-ča-kʷa-ltia.*

SBJ.1.SG-OBJ.2.SG-OBJ.INDEF.NHUM-

eat-CAUS

'I make you eat.'

(c) *Ni-mic-kʷa-ltia in*SBJ.1.SG-OBJ.2.SG-eat-CAUS DET<sub>1</sub>*naka-č.*

meat-ABSL

'I make you eat the meat.' (cf. (26 b))

Undoubtedly, */-tia/* is the same suffix as the denominative seen in 6.2.1. Thus (82 a) can mean 'I provide him with the process of sleeping', and there is a striking ambiguity in (82 b) ('I give him food, *ča-kʷa-l-li* 'OBJ. INDEF.NHUM-eat-OBJNR-ABSL', cf. (68 a)). Verbs of perception have both a */-l-tia/* (or */-tia/*) and a */-tia/* causatives. The former refers to an action on a person, the second to an action on an object (Launey 2002):

- (83) (a)
- Ni-k-itta-ltia in*

SBJ.1.SG-OBJ.3.SG-see-CAUS DET<sub>1</sub>*tōnatiw.*

sun

'I take him (e.g. out) to see the sun.'

(b) *Ni-k-itti-tia in*SBJ.1.SG-OBJ.3.SG-see-CAUS DET<sub>1</sub>*āma-č.*

book-ABSL

'I show him the book.'

If the causative applies to a reflexive form, the reflexive prefix is */-ne-/*; if the reflexive is 'external' to the basic form, it appears as definite:

- (84) (a)
- Ni-kin-ne-časočča-ltia.*

SBJ.1.SG-OBJ.3.PL-IMPR.REFL-love-

CAUS

'I act in such a way that they love each other.' (*mo-časočča-?* 'OBJ.REFL.2/3-love-PL')(b) *Ni-no-tē-itti-tia.*

SBJ.1.SG-OBJ.REFL.1.SG-OBJ.INDEF.

HUM-see-CAUS  
'I show myself (to someone).'

Causatives can be put into the passive:

- (85) (a) *Ni-kočī-tī-lo.*  
SBJ.1.SG-sleep-CAUS-PASS  
'Someone puts me to sleep.'
- (b) *Ni-kʷa-ltī-lo naka-č.*  
SBJ.1.SG-eat-CAUS-PASS meat-ABSL  
'Someone gives me some meat to eat.'

Intransitive inanimate-subject verbs (cf. 76) have no 'real' causative. They usually occur in pairs with a transitive counterpart, which may alter in various ways the final syllable; these pairs were called 'thematic verbs' by Whorf (1946) and extensively studied by Canger (1980):

- (86) (a) *Ni-k-wāca.*  
SBJ.1.SG-OBJ.3.SG-dry:CAUS  
'I dry it up.' (cf. *wāki* in (76))
- (b) *Ni-k-kotōna.*  
SBJ.1.SG-OBJ.3.SG-cut:CAUS  
'I cut it (e.g. thread).' (*kotōni*  
'cut:DETR')
- (c) *Ni-k-poloa.*  
SBJ.1.SG-OBJ.3.SG-destroy:CAUS  
'I destroy it.' (*poliwi* 'destroy:DETR')

### 6.3.5. Applicatives

Here again, one argument place is added. But this argument is a 'dative' one (with a benefactive or a detrimental meaning). It is marked in some verbs by */-ia/*, but in most cases by */-lia/* (probably analyzable as */-l-ia/*: once again we find */-l-/*), which can cause some minor alterations on the stem (87). Very few intransitive stems have a corresponding applicative, see however (88). If the applicative comes from a reflexive form, the reflexive prefix is indefinite (89 a); but if there is an 'external' coreference between subject and dative, it is definite (89 b). Applicatives can be put into the passive (90):

- (87) (a) *Ni-mic-čā-kōwi-lia.*  
SBJ.1.SG-OBJ.2.SG-OBJ.INDEF.NHUM-  
buy-APPL  
'I buy (*kōwa*) something for (or:  
from) you.' (also *ni-mic-čā-kōw-ia*  
'SBJ.1.SG-OBJ.2.SG-OBJ.INDEF.NHUM-  
buy-APPL')
- (b) *Ni-mic-kʷī-lia.*  
SBJ.1.SG-OBJ.2.SG-take-APPL  
'I take (*kʷi*) it from you.'

- (88) *Ni-k-čōki-lia.*  
SBJ.1.SG-OBJ.3.SG-weep-APPL  
'I bemoan it (lit. I weep for it).'
- (89) (a) *Tī-m-ināya.*  
SBJ.2.SG-OBJ.REFL.2/3-hide  
'You hide (lit. you hide yourself).'  
*Tī-k-ne-ināyi-lia.*  
SBJ.2.SG-OBJ.3.SG-IMPR.REFL-hide-  
APPL  
'You hide from him.'
- (b) *Ni-k-no-kʷī-lia.*  
SBJ.1.SG-OBJ.3.SG-OBJ.REFL.1.SG-take-  
APPL  
'I take it for myself.' (cf. (87 b))
- (90) *Tī-čā-kōwī-lī-lo.*  
SBJ.2.SG-OBJ.INDEF.NHUM-buy-APPL-PASS  
'Someone buys something for you.' (cf.  
(87 a))

### 6.3.6. Honorific verbs

They are formed by splitting the subject place (the second place being marked by a reflexive). Usually (with some exceptions) intransitive verbs use the causative as honorific, while transitive verbs use the applicative (91 a–b); this can lead to 'hyper-applicative' verbs with up to 4 prefixes (91 c). The construction cannot be used with reflexive forms: a suffix */-cinoa/* is added to the short stem (cf. 3.1.2 and (92)):

- (91) (a) *Mo-kočī-tia.*  
OBJ.REFL.2/3-sleep-CAUS  
'He (hon.) is asleep.'
- b) *Ki-mo-tti-lia.*  
OBJ.3.SG-OBJ.REFL.2/3-see-APPL  
'He (hon.) sees it.'
- (c) *Tī-k-mo-ne-ināyī-lī-lia.*  
SBJ.2.SG-OBJ.3.SG-OBJ.REFL.2/3-  
IMPR.REFL-hide-APPL-APPL  
'You (hon.) hide from him.' (cf.  
(89 a))
- (92) *Mo-kokoʔ-cinoa.*  
OBJ.REFL.2/3-ill-HON  
'He (hon.) is ill.'

## 7. Composition

### 7.1. Noun + noun

**Compound nouns** are very common in Nahuatl. Normally the second is the main noun, and the first occurs only as a stem with no absolute suffix. The semantic relation can involve one of the following: double member-

ship (an entity which is both N1 and N2), comparison, material, location, instrument, possessor, agent (Launey 1999):

- (93) *siwā-pil-li*  
woman-noble-ABSL  
'noble woman'  
*āyo-tōč-čī*  
tortoise-rabbit-ABSL  
'armadillo'  
*picō-naka-č*  
pork-meat-ABSL  
'pork meat'  
*kʷaw-kimič-in*  
wood-mouse-ABSL  
'mouse (of the) woods'  
*tē-yakān-kā-kal-li*  
OBJ.INDEF.HUM-guide-PART-house-ABSL  
'house of the leader' (cf. (65 b)), etc.

## 7.2. Noun + verb

This is traditionally called **incorporation**, and can involve either **saturation** or **modification**. In **saturating incorporation**, the noun stem refers to a generic object. Consequently, the compound verb is intransitive, while the verb stem is transitive:

- (94) *Šōči-namaka*.  
flower-sell  
'He sells flowers.'  
*ča-kʷa-l-čīwa*.  
OBJ.INDEF.NHUM-eat-OBJNR-prepare  
'He prepares food.'  
*Naka-kʷa*.  
meat-eat  
'He eats meat.'

In **modifying incorporation**, the argument pattern of the verb stem is not affected since the noun stem does not refer to an object. The semantic relation can be:

- (a) adverbial: the noun refers to an instrument, or to any kind of circumstance in which the process occurs (95 a);
- (b) predicative or comparative: the noun can be predicated, at least metaphorically, of the subject or the object (95 b);
- (c) possessive: the noun is possessed by the transitive object or by the intransitive subject (95 c);
- (d) agentive: always in the passive, the noun referring to the agent (95 d).

- (95) (a) *Ki-te-wīteki*.  
OBJ.3.SG-stone-hit  
'He hits him (with a) stone.'

*Ki-čān-kotōna*.  
OBJ.3.SG-tooth-cut  
'He tears it (with his) teeth.'

*Ki-čāl-tōka*.  
OBJ.3.SG-earth-bury  
'He buries it (in the) earth.'

- (b) *Okič-miki*.  
man-die  
'He dies (like a) man.'

*Ki-tēkʷ-čālia*.  
OBJ.3.SG-lord-install  
'He installs him (as a) lord.'

- (c) *Ni-k-keč-kotōna*.  
SBJ.1.SG-OBJ.3.SG-throat-cut  
'I cut his throat (lit. I cut him (by the) throat).'

*Yōl-miki*.  
heart-die  
'He faints (lit. he dies (by the) heart).'

- (d) *Kōwā-kʷa-lo*.  
snake-eat-PASS  
'He is bitten (lit. 'eaten') (by a) snake.'

## 7.3. Verb + verb

### 7.3.1. The /-t(i)-/ auxiliaries

About 15 intransitive and 4 transitive verbs (the latter being very uncommon) can occur as aspectual or modal auxiliaries. The morphological pattern is: Main verb (in the short stem) + /-ti-/ (/t/- before /V/) + auxiliary:

- (96) (a) *Koč-ti-ka?*. (*ka?*, see 3.2.5)

sleep-AUX-be  
'He is asleep.'

- (b) *Čōka-ti-w*. (-*w*, see 3.2.5)  
cry-AUX-go  
'He goes along crying.'

- (c) *Čōka-ti-nemi*.  
cry-AUX-move  
'He keeps crying all the time.'

- (d) *Ki-kʷi-ti-weci*.  
OBJ.3.SG-take-AUX-fall  
'He takes (*kʷi*) it suddenly.'

### 7.3.2. Other auxiliaries

Only one is common: *neki* 'to want', which can be used as transitive with an object clause, but also as an auxiliary, the main verb being in the future:

- (97)
- Ni-koči-s-neki.*

SBJ.1.SG-sleep-FUT-want:AUX  
'I want to sleep.'/'I am about to sleep.'

## 8. Illustrative text

- (98) *M-i?toa, in ok*  
(SBJ.3)OBJ.REFL.2/3-say DET<sub>1</sub> still  
*yowa-yān, in a-ya-mo*  
be.dark-DER.LOC DET<sub>1</sub> NEG-already-NEG  
*tōna, in a-ya-mo*  
(SBJ.3)be.hot DET<sub>1</sub> NEG-already-NEG  
*ka-twi, kil*  
OBJ.INDEF.NHUM-be.visible reportedly  
*mač*  
apparently  
*mo-sen-čāli?k-e?*,  
(SBJ.3)OBJ.REFL.2/3-one-set(AOR)-PART-PL  
*mo-no?nōc-k-e?*  
(SBJ.3)OBJ.REFL.2/3-RDP<sub>2</sub>-call(AOR)-PART-PL  
*in tē-teo-?* in ōmpa  
DET<sub>1</sub> RDP<sub>1</sub>-god-PL DET<sub>1</sub> there  
*Teō-tī-wa-kān,*  
god-make-IMPR-DER.LOC  
*k-i?to?k-e?*,  
(SBJ.3)OBJ.3.SG-say(AOR)-PART-PL  
*ki-mo-lwi?k-e?*:  
(SBJ.3)OBJ.3.SG-OBJ.REFL.2/3-tell(AOR)-PART-PL
- (99) *čā ū-wāl-wi-yān,*  
if OPT.SBJ.2-hither-go-OPT.PL  
*tē-teo-?-é, āk in*  
RDP<sub>1</sub>-god-PL-VOC who DET<sub>1</sub>  
*čā-māmā-s in*  
OBJ.INDEF.NHUM-carry-FUT DET<sub>1</sub>  
*tōna-s, in,*  
(SBJ.3)be.hot-FUT DET<sub>1</sub>  
*čā-twi-s?*  
OBJ.INDEF.NHUM-be.visible-FUT
- (100) *Aw niman ye ik*  
and then already hence  
*ye?wā-č onkān*  
(SBJ.3)IDENT-Ø-ABSL there  
*on-čā-?toa,*  
(SBJ.3)thither-OBJ.INDEF.NHUM-say  
*om-m-čš-keca*  
(SBJ.3)thither-OBJ.REFL.2/3-eye-raise  
*in Tēkʷsistēka-č*  
DET<sub>1</sub> Teucciztecatl-ABSL  
*k-i?to?:*  
(SBJ.3)OBJ.3.SG-say(AOR)
- (101) *Tē-teo-?-é, ka*  
RDP<sub>1</sub>-god-PL-VOC ASRT  
*n-e?wā-č ni-ye-s*  
SBJ.1.SG-IDENT-Ø-ABSL SBJ.1.SG-be-FUT

- (102)
- Ok sep-pa*

again one-time  
k-i?to?k-e?  
(SBJ.3)-OBJ.3.SG-say(AOR)-PART-PL DET<sub>1</sub>  
tē-teo-?  
RDP<sub>1</sub>-god-PL

- (103)
- Āk in ok sē?*
- 
- who DET
- <sub>1</sub>
- again one

- (104)
- Ay-āk mo-čā?paloč-ya...*
- 
- NEG-who (SBJ.3)OBJ.REFL.2/3-dare-IMPF
- 
- san moc ičāka-č*
- 
- only all man-ABSL
- 
- mo-māw-tiā-ya,*
- 
- (SBJ.3)OBJ.REFL.2/3-fear-CAUS-IMPF
- 
- cīn-čīsa-ya*
- 
- back-go.out-IMPF

"(98) It is told that when yet all was in darkness, when yet no sun had shone and no dawn had broken – it is said – the gods gathered themselves together and took counsel among themselves there at Teotihuacan. They spoke, they said among themselves: (99) 'Come hither, O gods! Who will carry the burden? Who will take upon himself to be the sun, to bring the dawn?' (100) And upon this, one of them who was there spoke: Teucciztecatl presented himself, he said: (101) 'O gods, I shall be the one.'

(102) And again the gods spoke: (103) 'And who else?' (104) No one dared ... no one else came forward. Everyone was afraid: they all drew back." (From *Florentine Codex* Book VII; free translation by C. Dibble and A. J. O. Anderson)

## 9. Uncommon abbreviations

CNST	constitutive
DER.LOC	locative derivation
IDENT	identifier
OBJNR	object nominalizer

## 10. References

- Aldama y Guevara, Joseph (1754), *Arte de la lengua mexicana*. Mexico
- Andrews, J. Richard (1975), *Introduction to Classical Nahuatl*. Austin, London: University of Texas Press
- Canger, Una (1980), *Five Studies Inspired by Nahuatl Verbs in -oa*. Copenhagen: The Linguistic Circle (Travaux du Cercle de Linguistique 19)
- Carochi, Horacio (1645), *Arte de la lengua mexicana*. Mexico [reprinted (1983), Mexico: U.N.A.M.]

- Dakin, Karen (1988), “-li in Nahuatl: the Many Faces of an Added Argument”. *Morris Swadesh in Memoriam*. Mexico: U.N.A.M.
- Langacker, Ronald W. (1976), *Non-Distinct Arguments in Uto-Aztec*. Berkeley, Los Angeles: University of California Press (Publications in Linguistics 82)
- Launey, Michel (1977), “Le pluriel transcatégoriel /-keʔ/ en nahuatl”. *Amerindia* 2, 19–45
- Launey, Michel (1979), *Introduction à la langue et à la littérature aztèques*. Paris: L’Harmattan
- Launey, Michel (1981), “Une interprétation linguistique des schémas relationnels”. *Amerindia* 6, 17–58
- Launey, Michel (1986), *Catégories et opérations dans la grammaire nahuatl*. Thèse d’Etat, Université de Paris VII
- Launey, Michel (1994), *Une grammaire omniprédictive*. Paris: CNRS-Editions
- Launey, Michel (1999), “Compound Nouns vs. Incorporation in Classical Nahuatl”, *Sprachtypologie und Universalienforschung*. 52.3/4, 347–364

- Launey, Michel (2002), “On Some Causative Doublets in Classical Nahuatl”. In: Shibatani, Masyoshi (ed.), *The Grammar of Causation and Interpersonal Communication*. Amsterdam: Benjamins (Typological Studies in Language 48), 301–318
- Molina, Alonso de (1571), *Vocabulario en Lengua Castellana y Mexicana y Mexicana y Castellana*. Mexico [reprinted (1970), Mexico: Porrua]
- Olmos, Andres de (1547), *Arte para aprender la lengua mexicana*. Mexico [reprinted in: Siméon, Rémi (1875), *Grammaire de la langue nahuatl ou mexicaine*. Paris: Imprimerie nationale]
- Perlmutter, David M. (1983, ed.), *Studies in Relational Grammar, Vol. I*. Chicago: University of Chicago Press
- Rincón, Antonio del (1595), *Arte mexicana*. Mexico [reprinted (1885), Mexico: Secretaría de Fomento]
- Whorf, Benjamin Lee (1946), “The Milpa Alta Dialect of Aztec”. In: Hoijer, Harry et al. (eds.), *Linguistic Structures of Native America*. New York: The Viking Fund (Viking Fund Publications in Anthropology 6), 367–397

*Michel Launey, Paris (France)*

## 134. Quechua (Quechua)

1. General facts on Quechua
2. Typological features
3. Word-classes
4. Personal reference
5. Pluralization
6. Tense and mood
7. Nominalization
8. Subordination
9. Other verbal affixes
10. Case and other nominal affixes
11. Independent suffixes
12. Contact phenomena
13. Illustrative texts
14. Uncommon abbreviations
15. References

### 1. General facts on Quechua

The Quechua family consists of a number of closely related linguistic varieties, alternatively referred to as ‘dialects’ or ‘languages’ according to the viewpoint of the observer. If taken as a unity, Quechua is the largest American Indian language presently spoken with an estimated total number of speakers

of 8,354,125 (Cerrón-Palomino 1987: 76). The Quechua linguistic area covers large sectors of the Bolivian, Ecuadorean, and Peruvian Andes, including some smaller regions in the Amazonian lowlands of these countries (Adelaar 2004). In addition, speakers of Quechua are found in Northwestern Argentina, in Southern Colombia and, to a very limited extent, in Northern Chile. The genetic affiliation of Quechua is uncertain. The numerous lexical, phonological, and typological similarities between Quechua and the neighbouring Aymaran languages can be attributed to intensive borrowing at the proto-language stage, followed by a state of permanent or intermittent contact which has lasted until the present day.

A modern assessment of the Quechua dialect situation was provided for in the 1960s. Broadly spoken, a subdivision of the Quechua dialects into two main groups exists. A Central Peruvian group, referred to as Quechua B (Parker 1963) or Quechua I (Tóro 1964), comprises the dialects of the de-

parts of Ancash, Huánuco (cf. Weber 1989), Junín and Pasco, and that of the Cajatambo area in the department of Lima. A second group, referred to as Quechua A or Quechua II includes the dialects of all other regions. Within the Quechua A/II group a further distinction is made between (a) a northern subgroup, including the dialects of Ecuador and Colombia, the Peruvian dialects of Amazonas, Loreto and San Martín, and the extinct dialect of the Peruvian Central coast, and (b) a southern subgroup, comprising the dialects of Southern Peru, Bolivia, and Argentina. In addition, the Quechua of Cajamarca and Lambayeque in Northern Peru and several dialects located in the high valleys of the department of Lima in Central Peru are said to constitute a third subgroup within Quechua A/II. Its members occupy intermediate positions between the two dialect branches and have been the object of conflicting classifications. One of them, the dialect of Pacaraos, spoken in the upper reaches of the Chancay valley near Lima, shares essential elements of its morphology with Quechua B/I (Adelaar 1984: 45).

The distinctive elements that separate the two main groups of Quechua dialects are essentially morphological and lexical. Considerable phonological variety is found within each of the two subdivisions, although morphological differences are important as well, in particular, within the Quechua B/I group. The general morphosyntactic structure of the Quechua dialects is remarkably similar. The Ecuadorean and Colombian varieties (cf. Cole 1982; Levinsohn 1976; Muysken 1977) constitute an exception because they have lost part of the original morphological complexity.

## 2. Typological features

Quechua is the epitome of an agglutinating language. It combines considerable morphological complexity with almost perfect regularity. Morphology occupies a central position in almost all descriptions of Quechua dialects. It is a consequence of the fact that a majority of the syntactic phenomena involve the use of one or more morphological markers. A systematic treatment of these markers in connection with their syntactic applications provides a straightforward guideline for the organization of a reference grammar. The one morphological device which is dominant in

Quechua is suffixation. There are no prefixes at all. Infixes do not occur either. However, in the Quechua B/I dialects productive affixes are sometimes inserted within petrified root-suffix combinations; e.g. Tarma Quechua *ša-ya:-mu-n* [ʃa'ja:muŋ] 'come-PROG-HIT-3.SBJ (he is coming)'. In this verb-form, the progressive aspect affix *-ya:-* precedes the suffix *-mu-* (HITHER), which is obligatorily present in the lexical base *ša-mu-* 'come' (the root *ša-* does not occur without *-mu-*). Vowel modification is a morphophonemic process which affects suffixes occurring in strings. It is found in Quechua B/I and in southern Quechua A/II. The most common type of vowel modification affects verbal suffixes with a final back vowel *-u-* [u, ɔ]. This back vowel is changed to *-a-* before a small set of suffixes, which do not constitute a phonological class but have the property of triggering vowel modification. In Quechua B/I dialects, this sort of vowel modification can operate between non-contiguous suffixes, skipping intervening suffixes which do not meet the condition for being modified. In the dialects in question, the rule of vowel modification often extends to verb roots ending in a back vowel when the final element in the root is a suffix from a historical point of view. An example is the verb root *miku-* 'eat', which at an early stage of development of the language absorbed the reflexive suffix *-ku-* (now meaningless); e.g. Tarma Quechua *mika-ya:-či-n* [mika'ja:tʃin] 'eat-PROG-CAUS-3.SBJ (he is making him eat)', where *-či-* is the triggering suffix and *-ya:-* an intervening neutral one. In the Quechua B/I dialects, which have vowel length as a distinctive element, length variation affects suffixes occurring in strings under somewhat similar conditions. Additionally, the morpheme that expresses a first person subject or possessor consists in lengthening the final vowel of the base (which is stressed in word-final position); e.g. *miku-:* [mi'ku:] 'eat-1.SBJ (I eat)'. Contrastive stress plays a role in Pacaraos Quechua. If it occurs word-finally, stress is normally a result of elision; e.g. Pacaraos Quechua *akšu-k'* [ak'ʃuk], from *akšu-kta* [ak'ʃukta] 'potato-ACC'. In Pacaraos Quechua stress is also part of the shape of the 1st person subject and possessor morpheme whenever it is word-final; e.g. *miku-'* [mi'ku] 'eat-1.SBJ (I eat)'. Stress associated with the Pacaraos Quechua 1st person morpheme is not a result of elision.

Suffixes can have different allomorphs according to the phonological context in which

they occur. The suffix *-m(i)* is a validator denoting first-hand knowledge. In most dialects, it has the shape *-m* when added to a base ending in a (short) vowel; when the base has a final consonant (or a long vowel in those dialects that distinguish length), the suffix in question appears as *-mi*; e.g. Quechua A/II Ayacucho Quechua (cf. Parker 1969) *wasi-m* ['wasim] 'house-KNWL.VAL (it is a house)'; *wasi-n-mi* [wa'simmi] 'house-POSS.3-KNWL.VAL (it is his house)'. However, when *-m(i)* is followed by one of the rare suffixes with an initial vowel, its shape is *-m-* even after a consonant; e.g. Ayacucho Quechua *-iki* in *pay-m-iki* [pa'i'miki] 'he-KNWL.VAL-OF.COURSE (it's him of course)'. With the exception of a few phonologically divergent dialects (Quechua A/II Amazonas, for instance; cf. Taylor 1975), Quechua does not allow sequences of more than two consonants, which furthermore can only occur in word-medial position (*V*: being treated as *VC*). For the same reason, nominal suffixes take a binding element *-ni-* when added to a consonantal base; e.g. Ayacucho Quechua *yawar-ni-n* [ja'warnin] 'blood-BIND-POSS.3 (his blood)'. In some dialects, historical changes in the phonological domain have created a distance between different allomorphs of a suffix; in Quechua A/II Cuzco (cf. Cusihuamán Gutiérrez 1976), the genitive case suffix is *-q* after vowels and *-pa* after consonants; e.g. *nuqa-q* [nɔqax] 'I-GEN (mine)', *qam-pa* ['qampa] 'you-GEN (yours)'. This allomorphy reflects an earlier *-pl-/pa* alternation still found in Quechua B/I Huanca (Huancayo Quechua; cf. Cerrón-Palomino 1976); e.g. *ya?pa-p* ['ja?ap] 'I-GEN (mine)', *am-pa* ['ampa] 'you-GEN (yours)'. Most other Quechua dialects have lost the short form and only use *-pa*.

As a rule, there is a one-to-one correspondence between form and meaning for each suffix. However, **portmanteau** suffixes do occur in the personal reference system of the verbs. An example from Quechua B/I is the suffix *-q* which indicates a 1st person subject acting upon a 2nd person object; e.g. Ancash Quechua *wiya-q* ['wijaq, 'wijaq] 'hear-1.SBJ&2.OBJ (I hear you)'. Also in the personal reference system, some suffixes can have a different function from their usual one when they occur in a particular combination; e.g. Ayacucho Quechua *riku-wa-n* [ri'kuwan] 'see-1.OBJ-3.SBJ (he sees me)', *riku-nčik* [ri'kuntšik] 'see-1+2.SBJ (we see him/it)', *riku-wa-nčik* [ri'ku'wantšik] 'see-3.SBJ&1+2.OBJ (he sees us)'. As the example shows, the meaning of the

suffix combination *-wa-nčik* cannot be predicted from that of the elements that make up the combination. On the other hand, it is definitely not possible to analyze *-wa-nčik* as one suffix, because the second element varies according to tense and mood (future tense *-wa-sun* [wasun], potential mood *-wa-čwan* [watʃwaŋ]). Another exception to the one-to-one relation between form and meaning is the fact that several suffixes can be used to convey one meaning. Tarma Quechua has three suffixes indicating plurality of a subject (or, occasionally, an object) in verbs; the choice is determined by the presence or absence of other suffixes in the verb form; e.g. *puri-ba:ku-n* [puri'ba:kun] 'walk-PL-3.SBJ (they walk)', *puri-rga-ya-n* [purir'gajan] 'walk-PL-PROG-3.SBJ (they are walking)', *puri-ra:ri-n* [puri'ra:rin] 'walk-PFV-PL-3.SBJ (they walked)'. The two last examples also show a different relative order of the pluralizers and the suffixes denoting progressive and perfective aspect. In these combinations, the affix order is rigid, and any deviation from it would be rejected by the speakers of Tarma Quechua.

Another morphological device frequently found in Quechua is root reduplication. It applies both to nouns and to verbs, but it remains restricted to expressive use. The reduplicated root can either be bare, or be accompanied by a suffix. In reduplicated verbs the second root is conjugated like any other verb whereas the first root remains invariable; e.g. Tarma Quechua *kaša kaša* [kaʃa'kaʃa] 'place full of thorny plants' (*kaša* 'thorny plant'); *tinka-y tinka-r* [tiŋka'i'tiŋkar] 'as with a snap of the finger' (*tinka-* 'to snap with the finger'). In the latter example, *-r* is a verbal subordinator; the element *-y* accompanies the first root in this particular type of reduplication which conveys the general meaning 'to act as if'.

Compounding is not a common process in Quechua. It is limited to certain expressions containing a substantive as their nucleus. In accordance with the general word order rules of the language, a modifier precedes its head also in compounds. Occasionally, combinations of co-ordinated substantives occur. As a rule, both constituents of a compound fully preserve their phonological characteristics (including word-stress). Therefore, the compositional character of the expressions at issue is questionable. Compound-like expressions are often place names (e.g. *aya-kuču* [aja'kutʃu] 'Ayacucho, corner of corpses').

### 3. Word-classes

Quechua morphology provides unambiguous criteria for a subdivision of the lexicon into verbs, nouns, particles, and interjections. The noun class can be further subdivided into substantives, adjectives, pronouns (demonstrative, interrogative, personal, indefinite), and numerals. The existence of a separate subclass of adverbs is doubtful because most adverbs are regularly derived from adjectives or verbs. In addition, the particles have an adverbial function. Some Quechua dialects have one or more postpositions (historically substantives), which play a role similar to that of the case suffixes except that they have the shape of a phonological word. Minor classes of interrogatives (*what? to do what? to say what?*) and deictics (*thus, to do so*) stand at right angles to the main word-class division.

The word-class division in Quechua rests upon the existence of three clearly distinct sets of suffixes. One of these sets is characteristic of the verb class, whereas a second one belongs to the noun class (and to a variable degree to each of its subclasses). Several verbal suffixes take care of the transposition of verbs to the noun-class, and there are nominal suffixes that verbalize nouns. The third set of suffixes has been known by different names, such as **independent suffixes**, **sentential suffixes**, **class-free suffixes**, or **enclitics**. It is closely linked to the organization of the sentence. Independent suffixes occur with verbs, nouns, and particles. The latter can take no other suffixes. With verbs and nouns, the verbal and nominal suffixes (if any) must be added first before independent suffixes can be attached. Interjections do not normally take any suffixes at all.

There are formal restrictions which set the verb class apart from the other classes. Verb roots always end in a vowel; they do not occur by themselves and must at the least be followed by one suffix that closes the verb form. Verbal suffixes must also end in a vowel except those that can occupy the closing position in a verb; e.g. in Ayacucho Quechua *l'amka-n* [ɿamkaŋ] 'work-3.SBJ (he works)', the suffix *-n* need not end in a vowel because it closes the verb form. All the intervening suffixes must end in a vowel, *-CV-* or *-CCV-* being the preferred shape for these suffixes; e.g. Ayacucho Quechua *l'amka-mu-n* [ɿam'kamun] 'work-THERE-3.SBJ (he works over there)', *l'amka-čka-n* [ɿam'katʃkan]

'work-PROG-3.SBJ (he is working)'. In contrast to verb roots, noun roots can end in a consonant, and they can be used by themselves without a suffix. Non-permitted consonant sequences are avoided by inserting the binding element *-ni-* (see 2 for examples). The small closed subclass of the independent pronouns is exceptional in that some of its members must be followed by a personal reference marker; e.g. Pacaraos Quechua *rapqa-yki* [rap'χaiki] 'both-poss.2 (both of you)'.

Unlike nominal and verbal suffixes, the independent suffixes never begin with a consonant cluster. In most dialects, the three validator suffixes, Ayacucho and Tarma Quechua *-m(-)/-mi* [m(i)] (first-hand knowledge), Ayacucho Quechua *-s(-)/-si* [s(i)], Tarma Quechua *-š(-)/-ši* [ʃ(i)] (hearsay), and Ayacucho Quechua *-č(-)/-ča* [tʃ(a)], Tarma Quechua *-č(-)/-či* [tʃ(i)] (conjecture) have different allomorphs according to whether there is an adjacent vowel or not. Independent suffixes can also begin with a vowel (see 2 for examples).

### 4. Personal reference

The Quechua personal reference system is a common component of both verbal and nominal morphology. It is essentially a four-term system based on the inclusion or non-inclusion of the speaker and the addressee. With substantives, personal reference markers identify a possessor; with verbs, they identify the subject and a human object, which can either be direct or indirect according to the meaning of the verb. As a rule, third person objects are not marked. The nominal possessor markers and the verbal subject markers are either identical (in Quechua B/I and Pacaraos Quechua) or partly similar (in Quechua A/II); e.g. the following paradigms of Tarma and Ayacucho Quechua *wata-* 'to tie', and Tarma Quechua *wayi*, Ayacucho Quechua *wasi* 'house': Tarma Quechua '1.SBJ' *wata-* [wa'ta:], '2.SBJ' *-nki* [wa'taŋki], '3.SBJ' *-n* ['watan], '1+2.SBJ' *-nči(k)* [wa'tantʃi(k)], Ayacucho Quechua '1.SBJ' *wata-ni* [wa'tani], '2.SBJ' *-nki* [wa'taŋki], '3.SBJ' *-n* ['watan], '1+2.SBJ' *-nčik* [wa'tantʃik]; Tarma Quechua 'POSS.1' *wayi-* [wa'ji:], 'POSS.2' *-gi* [wa'jigi], 'POSS.3' *-n* ['wajin], 'POSS.1+2' *-nči(k)* [wa'jintʃi(k)], Ayacucho Quechua 'POSS.1' *wasi-y* ['wasii], 'POSS.2' *-ki* [wa'siki], 'POSS.3' *-n* ['wasin], 'POSS.1+2' *wasi-nčik* [wa'sintʃik]. The verbal paradigm exemplified here is the unmarked (present) tense paradigm.

Personal reference suffixes may vary according to tense and mood. These variations are minor except for the future and the imperative paradigms, which have portmanteau suffixes indicating both tense/mood and person of subject. Nevertheless, the forms for '2.SBJ' future and '2.SBJ' present are identical, and the same holds for '1+2.SBJ' future and '1+2.SBJ' imperative. The future endings for Ayacucho Quechua are '1.SBJ' -*saq* [saχ], '2.SBJ' -*nki* [ŋki], '3.SBJ' -*nqa* [χa], '1+2.SBJ' -*sun* [suŋ], the imperative endings are '2.SBJ' -*y* [i], '3.SBJ' -*čun* [tʃuŋ], '1+2.SBJ' -*sun* [suŋ].

Except for marginal cases in the Quechua B/I dialects, person of object can only be indicated in combination with person of subject. The ending for '1.SBJ&2.OBJ' is a portmanteau suffix, -*q* [q, χ, x] in most of Quechua B/I, -*yki* [iki] in Quechua A/II. Apart from '1.SBJ&2.OBJ', combinations of suffixes, rather than portmanteau suffixes are used. For Ayacucho Quechua the combinations are '2.SBJ&1.OBJ' -*wa-nki* [waŋki], '3.SBJ&1.OBJ' -*wa-n* [waŋ], '3.SBJ&2.OBJ' -*sunki* [suŋki], and '3.SBJ&1+2.OBJ' -*wa-nčik* [wantʃik]. The first suffix remains unaltered, whereas the second one varies according to tense and mood; e.g. Ayacucho Quechua *yanapa-wa-y* [jana'pawai] 'help-1.OBJ-2.SBJ.IMP (help me!)'. The two suffixes need not be adjacent. They can be separated by other suffixes, e.g. Ayacucho Quechua *yanapa-wa-ptiki* [janapawap'tiki] 'help-1.OBJ-DS-2.SBJ (if you should help me)'. The suffix combinations referring to subject and object are also used with nominalized verbs. If only the subject is indicated, it is not always possible to distinguish the endings in question from the nominal possessor endings; e.g. Ayacucho Quechua *puñu-na-yki* [puñu'nałki] 'sleep-FUT.NR-2.SBJ (that you will sleep, you will have to sleep)', but also 'sleep-FUT.NR-POSS.2 (your means of sleeping, your bed)'.

The Ecuadorean-Colombian branch of Quechua A/II has a three-term rather than a four-term system of personal reference, the original 1st + 2nd person forms having taken the place of the 1st person plural. The morphological indication of the person of a possessor is no longer possible, and the possibilities of encoding person of object in the verb form are much reduced.

## 5. Pluralization

The suffixes which are used for personal reference can be accompanied by plural markers. Quechua dialects differ in their ways of indi-

cating the plurality of subject, object, and possessor. On the other hand, virtually all dialects share the pluralization of nouns and nominalized verbs by means of the suffix -*kuna*; e.g. Ayacucho Quechua *wasi-kuna* [wasi'kuna] 'house-PL (houses)', *pay-kuna* [paɪ'kuna] 'he/she-PL (they)'. Pluralization does not automatically apply to all plural entities. The emphasis is on the plurality of individual items.

The Quechua B/I dialects and Pacaraos Quechua indicate plurality of the subject by means of suffixes that precede the tense, mood, and person affixes; e.g. Pacaraos Quechua *ka-ri-rqa-yki* [karir'xałki] 'be-PL-PAST-2.SBJ (you people were)'. According to the dialect, there may be a single pluralizing suffix or several, as in Tarma Quechua (see 2). Portmanteau suffixes combining plural marking with another category occur as well; e.g. Pacaraos Quechua *ri-rka:ča-n* [rir'ka:t-ʂan] 'go-PL.PROG-3.SBJ (they are going)'. In a few contexts, the plural marker indicates plurality of object, rather than of subject; e.g. Tarma Quechua *čay-la-ta-m ni-ba:ku-x lapa-y-ta* [tʃař'latam ni'ba:kɔx la'paňta] 'that-ONLY-ACC-KNWL.VAL say-PL-1.SBJ&2.OBJ all-POSS.2-ACC (I say that to all of you)'. Plurality of possessor cannot be indicated by morphological means. The semantic difference between 1st person plus plural and 1st + 2nd person is that between an **exclusive** and an **inclusive** first person plural.

In most Quechua A/II dialects, plural of subject, object, and possessor is indicated by means of suffixes which follow the personal reference suffixes. Both the former and the latter can occur word-finally. The northern Quechua A/II dialects use -*kuna* [kuna] (Ecuadorian Highland Quechua), -*sapa* [sapa] (San Martín Quechua) (from \**sapa* 'each'), and -*žapa* [(d)ʒapa] (Cajamarca Quechua) (from \**l'apa* 'all'). The southern Quechua A/II dialects have developed an interesting system based on two pluralizers, one for second person (Ayacucho Quechua -*čik*, Cuzco Quechua -*čis*) and one for first and third person (-*ku*). The availability of these two markers allows number indication for either the subject or the object; e.g. *riku-wa-nki-ku* [rikuwaŋ'kiku] 'see-1.OBJ-2.SBJ-1.PL (you see us)', *riku-wa-nki-čik* [rikuwaŋ'kitʃik] 'see-1.OBJ-2.SBJ-2.PL (you people see me/us)'. Most dialects do not allow subject and object to be marked for plural simultaneously.

## 6. Tense and mood

All Quechua dialects distinguish several tenses in the verb system. The future tense exhibits a special conjugation (see 4) and refers to any activity after the moment of speaking. Past tense is indicated by a suffix *-rqa-* (Ancash, Ayacucho, Pacaraos, and Puno Quechua), *-rka-* (Ecuadorian Quechua) or *-ra-* (Ayacucho, Cuzco, and Tarma Quechua), followed by the personal endings of the unmarked (present) tense with slight allomorphic variation; e.g. Ayacucho Quechua *mikur(q)a-Ø* [mi'kur(χ)a] 'eat-PAST-3.SBJ (he ate)'. In Ayacucho Quechua, the allomorph for third person in the past tense is *-Ø*, not *-n*.

Many Quechua dialects distinguish a second past tense referring to an event of which the speaker or the protagonist in a narrative had not been aware of until the moment in focus (**sudden discovery tense**; cf. Adelaar 1977: 94–98). This tense is indicated by reflexes of *\*-ña(q)*, in Quechua B/I, or by the suffix *-sqa* [sχa, sqa] in Ayacucho and Cuzco Quechua; e.g. Quechua B/I San Pedro de Casas *čar-qas ka-ku-na-Ø alqu* ['tʃa:χa ka'kuna 'alχo] 'that-TOP be-CUST-SUDD-3.SBJ dog (that man had been a dog all along, so it turned out)'. Quechua B/I Ancash has a recent past tense formed with the affix *-rqu-*; e.g. *wiayarrqu-n* [wi'jarqoŋ] (Parker 1976: 109) 'hear-RECPST-3.SBJ (he has just heard)'.

Along with the synthetic tenses, compound tenses occur. As a rule, they consist of a nominalized verb form and the verb *ka-* [ka] 'to be'. Common to almost all dialects is the **habitual past**, which consists formally of 'to be' and the active nominalizer in *-q* [q, χ] (-x [x] in Ecuadorian and Tarma Quechua). As the 3rd person subject form of the verb 'to be' is normally suppressed, the pluralizing affix can be added directly to the nominalizer in Ayacucho Quechua; *rima-q-ku* [ri'maxku] 'speak-ACT.NR:HABIT-3.SBJ.PL (they used to speak)'. Some Northern and Central Peruvian dialects have additional compound tenses. These tenses have active meaning regardless of the semantic characteristics of the nominalized form underlying them. Pacaraos Quechua, for instance, has a **perfect tense** which indicates past experience or relevance to the present. It is formed on the basis of the participle in *-šqa* [ʃχa], which is normally passive when the verb base is transitive; e.g. *miku-šqa-m ka-'y* [mi'kusχam 'kai] 'eat-STAT.NR-KNWL.VAL be-1.SBJ (I have eaten it once, or: I have eaten, don't worry about feeding me)'.

Along with the indicative mood which may be considered to comprise the (mutually exclusive) tense categories treated so far, all Quechua dialects have an imperative mood (see 4) and a **potential mood** (also called **optative** or **conditional**). The latter is characterized by the presence of a suffix *-man*, which is added after the personal reference affixes of the unmarked tense with slight allomorphic variation. Many dialects have a special ending for 1st + 2nd potential, which may (or must) replace the expected suffix combination; e.g. Ayacucho Quechua *wañu-nčik-man* [wajnun'sčikman], *wañu-čwan* [wa'nutčwan] 'die-1+2.SBJ-POT (we could die)'. By adding the third person past tense form of the verb 'to be' a past potential is formed. It refers to a possibility in the past which did not materialize; e.g. Ayacucho Quechua *wañu-čwan kar(q)a-o* [wa'nutčwan] 'kar(χ)a] 'We could have died'.

## 7. Nominalization

Nominalization plays a central role both in Quechua morphology and in syntax. The nominalizing suffixes are mutually exclusive with the tense and mood suffixes and with the subordinating suffixes (see 8). They combine with personal reference markers (subject, object, and possessor) and with case affixes in the formation of different types of subordinate clauses. Most Quechua dialects have at least four nominalizers. Ayacucho Quechua, for instance, has an **infinitive** in *-y* [i] (*miku-y* ['mikuɪ] 'to eat'), an **active nominalizer** in *-q* [χ] (*miku-q* ['mikox] 'one who eats'), a **stative nominalizer** in *-sqa* [sχa] (*miku-sqa* [mi'kusχa] 'something eaten'), and a **future nominalizer** in *-na* [na] (*miku-na* [mi'kuna] 'something to be eaten, future eating'). In Ecuadorian Quechua the infinitive in *-y* has been replaced in most of its uses by *-na* (e.g. *miku-na* [mi'kuna] 'to eat'). As a rule, the infinitive refers to the activity denoted by the verb itself, the active nominalizer is co-referential with the subject of the underlying verb, and the stative and future nominalizers can refer to an object or to the fact that an event occurs. The difference between the stative and future nominalizers corresponds to that between accomplished and non-accomplished events. The future nominalizer, in particular, can also refer to place or means. Person of subject and person of object markers can co-occur with the stative and fu-

ture nominalizers. All nominalizers can take case markers.

A nominalizer *-nqa-* [ŋχ] (Tarma Quechua *-nxas-* [ŋχa]) is found in some of the Quechua B/I dialects and in Pacaraos Quechua. It is roughly equivalent to the stative nominalizer but must be combined with personal reference; e.g. Tarma Quechua *wa:ka pišta-nxa:* ['wa:ka pištan'xa:] 'cow slaughter-NR-1.SBJ (the cow I slaughtered), *wa:ka pišta-ša* [wa:ka piʃ'taʃa] 'cow slaughter-STAT.NR (the cow which was slaughtered)'.

## 8. Subordination

Along with the possibility of using nominalized verb forms for clause subordination, all Quechua dialects have a specific set of verbal subordinators. Subordinating affixes indicate that the verb in which they occur is subordinate to some other verb, and they identify the subject of a subordinate clause as the same or different in relation to the subject of the matrix clause (**switch-reference**). The **different subjects subordinator** is *-pti-* [pti] in most of Quechua B/I and in Ayacucho Quechua (*-qtis-* [χti] in Cuzco Quechua). It is obligatorily accompanied by personal reference markers including the object markers when relevant; e.g. Ayacucho Quechua *muna-su-pti-ki* [munasup'tiki] 'want-3.SBJ&2.OBJ-DS-2 (because she wants you)'. The **same subject subordinator** is *-r* [r, t] in most of Quechua B/I (-*l* [l] in Huancayo Quechua), it is *-špa* [ʃpa] or *-spa* [spa] in Quechua A/II and in Pacaraos Quechua. Ayacucho Quechua allows the redundant use of personal reference markers in combination with *-spa*. It has an additional marker for same subject and simultaneousness *-stin* [stiŋ], which is found in several Quechua B/I dialects as well.

The semantic relation between two clauses connected by a subordinator remains weakly specified. The context and the addition of independent or aspectual affixes are needed in order to distinguish between causal, concessive, conditional, and temporal clauses, all of which can contain a subordinator.

An interesting expansion of the switch-reference phenomenon is found in Ecuador. In Ecuadorean Highland Quechua (Imbabura Quechua), no person markers can be added to the different subjects affix *-(x)pi* [(x)pi], which remains distinct from its same subject counterpart *-špa* [ʃpa]. On the other hand,

purpose clauses are also marked for switch-reference, same subject being indicated by *-ngapax* [ŋgapax] (< \**nqa-paq*) and different subjects by the former 3rd person imperative suffix *-čun* [tʃun] (Cole 1982: 63 f.; Muysken 1977: 69–75).

## 9. Other verbal affixes

The verbal affixes which find their locus between the verb root on one hand, and the tense, mood, and person endings on the other, constitute the area in which Quechua morphology exhibits its greatest dialectal variety. These affixes are commonly referred to as **modal suffixes** or **(verbal) derivational suffixes**. Some affixes have the property of changing the syntactic valence of the verb base, whereas others introduce more or less important modifications to its meaning. In still other cases, the presence of a verbal affix serves a pragmatic purpose which often remains elusive. Characteristically, a single suffix may combine several of such uses according to the context in which it occurs. Idiomatic combinations of a verb root with a particular suffix are frequent. However, although some verbal affixes tend to participate in such combinations, others never do. Only the former can be considered to be truly derivational.

Combinations of affixes are frequent. The resulting sequences do not necessarily have the meaning that one would expect on the basis of the semantics of the affixes of which they consist.

Broadly speaking, the order in which the verbal affixes occur is fixed. For some parts of the affix inventory, however, the order of co-occurrence cannot be established easily. The semantic domain of an affix need not extend to the preceding base exclusively, it may also involve elements which follow the affix in question; e.g. in Tarma Quechua *maxa-čima-n* [maxa'čimaj] 'beat-CAUS-1.OBJ-3.SBJ (he makes me beat someone, he causes me to be beaten by someone)', ambiguity arises from the fact that the order *-či-ma-* is compulsory regardless of the hierarchy of the semantic elements involved in the combination. Nevertheless, affix order is relevant in the interplay of the causative suffix *-či-* with the reflexive *-ku-* and the reciprocal *-naku-*; Tarma Quechua *maxa-či-naku-n* [maxatʃi'nakun] 'they make each other beat', *maxa-naka-či-n* [maxana'katsin] 'he makes them beat each

other' (*-naku-* is modified to *-naka-* before *-či-*; cf. 2).

Verbal affixes that occur in most Quechua dialects (here given in their Ayacucho Quechua form) are *-či-* [tʃi] 'causation, permission'; *-ku-* [ku] 'reflexive, characteristic behaviour'; *-l'a-* [ka] 'just, only, politeness'; *-mu-* [mu] 'motion towards the speaker, non-motion over there'; *-naya-* [naja] (Quechua B/I *-na(:*)- [na(:)] 'physical need (used in an impersonal construction)'; *-pa-* [pa] 'repetition, action directed towards a non-characteristic object'; *-pu-* [pu] 'on behalf of a person marked as the object' (Cuzco Quechua also: 'return to former location'); *-raya-* [raja] (Quechua B/I *-ra(:*)- [ra(:)] 'continuous event or state'; *-ri-* [ri] 'inchoative, temporary change of state'; *-(y)kača-* [(i)kača] (Quechua B/I *-(y)kača(:*)- [(i)kača(:)]); *-ysi-[isi]* (Quechua B/I *-wši-* [üʃi]) 'to accompany or help with an action'. Equally widespread are the pseudo-sequences *-naku-* [naku] 'reciprocal' and *-paku-* [paku] 'taking advantage of the situation'. Many Quechua B/I dialects also have a suffix *-ka(:*)- [ka(:)] for 'potential passive or non-controlled action'. Ecuadorian Quechua and 16th century varieties of Quechua A/II have a passivizer *-ytuku-[ítuku]*. Most Quechua dialects lack a genuine passivizing element.

Aspect suffixes usually include a progressive aspect marker, Ayacucho Quechua *-čka-* [tʃka] (Cuzco Quechua *-sya-* [ʃa], Bolivian Quechua *-sa-* [sa]). In Quechua B/I the progressive marker is *-yka(:*)- [ika(:)] (Ancash and Pacaraos Quechua) or *-ya(:*)- [ja(:)] (Huancayo and Tarma Quechua). Ecuadorian Quechua has reflexes of *-ku-* [ku, xu] for progressive. Some of the Quechua B/I dialects have developed more elaborate aspect systems, involving a perfective aspect marker as well (Pacaraos Quechua *-rqu-* [r̥ʃ], Tarma Quechua *-ru-* [ru], Huancayo Quechua *-?l(u)-* [?(lu)], *-?l(u)-* [?(lu)], Eastern Ancash Quechua *-ski-* [ski]).

A set of four **directional** suffixes can be reconstructed for Quechua B/I. Only two of them, *-rku-* [rku] 'upwards' and *-rpu-* [rpu] 'downwards', have preserved their original meaning productively. The other two directional suffixes, *\*-rqu-* 'outwards' and *\*-yku-* 'inwards', occur both in Quechua B/I and in Quechua A/II, but are limited to petrified root-suffix combinations (e.g. Ayacucho Quechua *yayku-* [jaiku] 'to go in', Ancash Quechua *yarqu-* [jarqɔ] 'to go out'). As productive suffixes, they have survived with a

wide array of different meanings and uses; e.g. the reflexes of *\*-rqu-* indicate urgency (Ayacucho and Cuzco Quechua), recent past (Ancash Quechua), and perfective aspect (Huancayo, Pacaraos, and Tarma Quechua).

## 10. Case and other nominal affixes

Nominal suffixes include personal reference (cf. 4), number (cf. 5), and case markers. Other affixes are verbalizers (Ayacucho Quechua *-ča-* [tʃa] 'to make', *-ya-* [ja] 'to become') and suffixes indicating possession (Ayacucho Quechua *-yuq* [ʃɔχ] 'owner of', *-sapa* [sapa] 'rich in') and inclusion *-ntin* [ntin] 'including'. A very frequent suffix is *-l'a* [ka]. It indicates restriction ('just, only') or politeness (see 9 for its verbal counterpart). A diminutive suffix *-ča* [tʃa] is common in southern Quechua A/II.

The number of case suffixes is limited. The spatial relations conveyed by the case suffixes are 'location without a direction' (locative *-pi* [pi] in Quechua A/II, *-čaw* [tʃaú] or reflexes of *\*-čaw* in Pacaraos Quechua and Quechua B/I), 'direction towards' (allative *-man* [man]), 'direction away from' (ablative *-manta* [manta] in Quechua A/II, *-pita* [pita] in Quechua B/I, *-piq(ta)* or reflexes of *\*-piq(ta)* are found in Pacaraos Quechua [pɛχ(ta)], in Huancayo Quechua [pi?(ta), pi:(ta)]), and 'until, up to' (limitative *-kama* [kama]). The concept 'through' is indicated in different ways. Quechua B/I dialects often use the genitive case marker *-pa*; in Ayacucho Quechua and other Quechua A/II dialects the combination *-n-ta* 'POSS.3-ACC' obtains. As in other Amerindian languages, spatial relations can be specified more explicitly by means of syntactic constructions involving nouns referring to parts of the body.

Least well defined in its semantics is the accusative case marker *-ta* [ta] (historically, in Huancayo and Pacaraos Quechua *-(k)ta* [(k)ta]). It indicates a direct object or a deliberately selected fixed goal. It is also used in Quechua A/II dialects for marking adjectives that are used adverbially.

The genitive case marker is *-pa* [pa] in most dialects, whereas some (Cuzco and Huancayo Quechua) preserve reflexes of the historical alternation *-pl-pa*. Possessive constructions are formed by marking the possessor noun for genitive case and the possessed noun for person of possessor; e.g. Tarma Quechua *atux-ba nawi-n* [a'tɔxbə 'nawin]

'fox-GEN eye-POSS.3 (the fox's eyes)'. The benefactive case marker *-paq* [paq, pay] is formally akin to the genitive suffix. Although the two are kept apart in most dialects, Ecuadorian Quechua and the 16th century Lengua General use reflexes of *-paq*, along with *-pa*, for the genitive.

The last important case marker to be mentioned here is *-wan* [waŋ]. It denotes instrumental and comitative case, and it is also used as a noun co-ordinator.

Further case markers are *-rayku* [raiku] 'for the sake of' and *-pura* [pura] 'among'.

Combinations of case markers are mainly those involving genitive *-pa* and instrumental *-wan*; e.g. Ayacucho Quechua *ñuqa-pa-ta* [nɔχa'pata] 'I-GEN-ACC (affecting mine, to my place)', *ñuqa-ta-wan* [nɔχa'tawan] 'I-ACC-INSTR (and to me, to me as well)'. Case markers must follow the person and number markers on the noun. They are added to the last noun in a noun phrase, which need not be the head noun; e.g. Tarma Quechua *rigabaku-ra-o ača išgi-mu-x-ta* [zigaba:'kura 'atʃa išgi'moxta] 'see-PL-PAST-3.SBJ star fall-HIT-ACT.NR-ACC (they saw a star falling down)'.

## 11. Independent suffixes

The independent suffixes belong to the domain of morphology inasmuch they form a phonological unit with the word to which they are affixed. Their regular locus is at the end of a sentence constituent (that is, at the end of a noun phrase, at the end of a subordinate clause, or after any word which is not part of one of these entities). Among the independent suffixes are the **validators**, which indicate **data source** (see 2 and 3 for the formal characteristics of these suffixes and an inventory). The independent suffix *-ču* [tsu] marks both yes-no questions and negative sentences. In the latter case it must be accompanied by a negation particle; e.g. Tarma Quechua *mana-m muna-n-ču* ['manam mu'nantsu] 'not-KNWL.VAL want-3.SBJ-NEG (he does not want it)'. The Ancash and Huancayo Quechua dialects keep the two functions apart formally as well (Ancash Quechua has interrogative *-ku* [ku] and negative *-cu* [tsu]).

Other independent suffixes are the topic marker *-qa* [qa, χa], *-pas* [pas] or *-pis* [pis] 'also' (in Quechua B/I also *-si* [si]), *-raq* [raq, ray] 'still, first' and *-taq* [taq, tax] 'and, on the other hand'. Ayacucho Quechua and several other dialects also have *-ña* [na] 'already' (-na

[na] in Ancash Quechua). Most independent suffixes have a special meaning when added to interrogative pronouns (or verbs).

## 12. Contact phenomena

The Quechua language has a rich history of contact situations both of early and recent date. At least two mixed languages that are presently in use combine Quechua morphology with the lexicon of another language. These are the Media Lengua of the Ecuadorian Highlands (Muysken 1979) and the language of the Callahuaya herb doctors of the Charazani area in Bolivia. The lexical basis of the Media Lengua is Spanish, whereas a dialect of the extinct Puquina language provided most of the lexicon for Callahuaya; e.g. Media Lengua *inki-da azi-ngi* [in'kida a'zingi] (Muysken 1979:44), Ecuadorian Highland Quechua (Salcedo) *ima-da ura-ngi* [i'mada u'rangi] 'what-ACC do-2.SBJ (what do you do?)', Spanish *en qué* 'in what', *hacer* 'to do'; Callahuaya (Girault 1989:140) *nisi seka-ku-y-man*, Cuzco Quechua *nuqa yača-ku-y-man* ['nɔqa jača'ku'man] 'I learn-REFL-1.SBJ-POT (I could learn it)'.

The morphological parallelism which exists between Quechua and the Aymaran languages must be attributed to long-standing contacts. Most Quechua suffixes have Aymaran counterparts, which are similar in meaning and use, although in the majority of cases there is no formal resemblance. The four-person system, the sudden discovery tense and data source, which are now so characteristic of Quechua, may have been copied from Aymaran at an early moment.

Aymara verbal suffixes without a counterpart in Quechua A/II were borrowed by Quechua dialects of the Puno and Arequipa region. Within certain limits, these suffixes maintain the phonological and morphophonemic characteristics which they had in Aymara, such as the presence of ejectives and aspirated consonants and the elimination of a preceding vowel; e.g. *tiy-t<sup>h</sup>api-či-n* [tiit<sup>h</sup>a'pitʃiŋ] 'stay-TOGETHER-CAUS-3.SBJ (he allows them to stay together)'. In this example, the Aymara suffix *-t<sup>h</sup>api-* has no counterpart in other Quechua dialects. It causes the loss of the final vowel of the root *tiya-* 'sit, stay', a frequent phenomenon in Aymara but highly exotic from a Quechua point of view.

### 13. Illustrative texts

The following text fragments exemplify a Quechua A/II dialect (Ayacucho) and a Quechua B/I dialect (Tarma).

- (1) Ayacucho Quechua (from Meneses et al. 1974: 117):

<i>huk punčaw-si</i>	<i>huk warma-ča</i>
[huk pun'tʃaúsi	huk war'matʃa
one day-HEARSAY	one boy-DIM
<i>Tomas suti-yuq</i>	<i>ri-sqa<sup>(1)</sup></i>
'tomas su'tiçχ	'risχa
Tomás name-OWNER	go-NARR-3.SBJ
<i>urqu-ta</i>	<i>čita-kuna-ta</i>
ɔr'χɔta	tʃitaku'nata
mountain-ACC	lamb-PL-ACC
<i>baka-ča-n-ta</i>	<i>qati-spa-n.</i>
baka'tʃanta	xa'tispaj.
cow-DIM-POSS.3-ACC	herd-ss-3.SBJ
<i>urqu-pi</i>	
ɔr'χɔpi	
mountain-LOC	
<i>puk'l'a-ku-na-n-kama-ñ-a-taq<sup>(2,3)</sup></i>	
puk'akunaŋkama'natax	
play-REFL-FUT.NR-3.SBJ-LIM-ALR-AND	
<i>baka-qa wayqu-ta-kama<sup>(4)</sup></i>	<i>pasa-rqu-spa</i>
ba'kaχa waŋtʃota'kama	pasar'χɔspa
cow-TOP ravine-ACC-LIM	pass-URG-SS
<i>činka-rqu-sqa-Ø.</i>	<i>warma-qa</i>
-tʃinkar'χɔsχa.	war'maχa
disappear-URG-SUDD-3.SBJ	boy-TOP
<i>mana-s</i>	<i>musya-n-pas-ču.</i>
'manas	musjam'pastʃu.
not-HEARSAY	be.aware-3.SBJ-ALSO-NEG
<i>qati-mu-na-n-paq-ñ-a<sup>(5)</sup></i>	
χatimunam'paχña	
drive-HIT-FUT.NR-3.SBJ-BEN-ALR	
<i>maska-yku-pti-n-si<sup>(6)</sup></i>	<i>mana</i>
maskaikup'tniſi	'mana
look.for-INTS-DS-3.SBJ-HEARSAY	not
<i>ka-n-ču.</i>	<i>čay-l'a-raq-si<sup>(7)</sup></i>
'kantʃu.	tʃaiňka'raxsi
be-3.SBJ-NEG	that-ONLY-STILL-HEARSAY
<i>qati-n</i>	<i>baka-pa</i>
'χatiŋ	ba'kapa
follow-3.SBJ	cow-GEN
<i>yupi-n-ta</i>	<i>wiqi ñawi-ntin<sup>(8)</sup></i>
ju'pinta	'wεχe na'wintiŋ
footprint-POSS.3-ACC	eye-INCL

*waqa-ku-stin.*

waya'kustiŋ.]

cry-REFL-GER

#### Explanatory notes:

- (1) *-sqa-* indicates both narrative past and sudden discovery tense;
- (2) the combination *-na-.kama* (future nominalizer + limitative case) is interpreted as 'while, as long as, until';
- (3) *-ñ-a-taq* means 'but, on the other hand' (contrast with preceding sentence);
- (4) after another case suffix *-kama* means 'exclusively, each time' (here: 'all along, always further');
- (5) the combination *-na-.paq* (future nominalizer + benefactive case) refers to a purpose ('in order to');
- (6) the meaning of the (former) directional suffix *-yku-* (glossed 'intensive') is 'carefully, thoroughly, well';
- (7) the sole function of the demonstrative pronoun *čay* is to carry the independent suffix *-raq* 'still, first'; the expression *čay-l'a-raq* is to be translated as 'only just then, not before';
- (8) the literal meaning of *wiqi ñawi-ntin* is 'including tearful eyes'.

'One day a little boy called Tomás went to the mountains herding lambs and his little cow. But, while he was playing in the mountains, the cow had gone all the way down the ravine and disappeared. The boy was not aware of this. When he went to look for her in order to drive her home, she was not there. Only then did he follow her footprints with tears in his eyes, crying.'

- (2) Tarma Quechua (from Adelaar 1977: 362–365):

<i>pay-guna-xa</i>	<i>wača-rgu-r-xa<sup>(1)</sup></i>
[paɪgu'naxa	wačar'gurxa
he/she-PL-TOP	give.birth-SEQ-SS-TOP
<i>pampa-ta-m</i>	<i>učgu-ra-ri-i-x.</i>
pam'patam	učgu'ra:rex.
ground-ACC-KNWL.VAL	dig-PFV-PL-HABIT
<i>pampa-ta</i>	<i>učgu-rgu-r-mi</i>
pam'pata	učgu'gurmi
ground-ACC	dig-SEQ-SS-KNWL.VAL
<i>čay-man</i>	<i>xita-ya-ra-ri-i-x<sup>(2)</sup></i>
'tʃaimaŋ	xitaya'ra:rex
that-ALL	throw-INTO-PFV-PL-HABIT

<i>milwa xarača-guna-ta.</i>	<i>čay-ču</i>
milwa xaratjagu'nata.	'tʃaʃʃu
wool pelt-PL-ACC	that-LOC
<i>xučba-r-mi</i>	<i>wamra-n</i>
xɔtʃ'barmi	'wamraŋ
roll.about-ss-KNWL.VAL	child-POSS.3
<i>ga-ku-x.</i>	<i>kada bintikwatu</i>
'gakox.	'kada binti'kwatu
be-CUST-HABIT	every twenty.four
<i>ora-m</i>	<i>čuču-x.</i>
'oram	'tʃutʃox.
hours-KNWL.VAL	suckle-HABIT
<i>čay-bitā na:</i>	<i>mas-ta</i>
tʃai'bita na:	'masta
that-ABL	already more-ACC
<i>xatun-gasga-ya-tbi-n<sup>(3,4)</sup></i>	<i>na:</i>
xatungasga'jatbiŋ	na:
big-MORE-BECOME-DS-3.SBJ	already
<i>miku-y-ta</i>	
mi'kuítā	
eat-INF-ACC	
<i>yača-ka-ru-tbi-n-xa<sup>(5)</sup></i>	<i>na:</i>
jałšakarut'bıŋxa	na:
learn-NONCONT-PFV-DS-3.SBJ-TOP	already
<i>xatun ga-ya-tbi-n-xa</i>	<i>mas</i>
'xatun gajat'bıŋxa	mas
big	be-PROG-DS-3.SBJ-TOP
<i>xatun-gasga učgu-man</i>	<i>wina-ru-x.<sup>(6)</sup></i>
xatun'gasga učgu'man	wi'narox]
big-MORE	pit-ALL
	fill-PFV-HABIT

#### Explanatory notes:

- (1) the suffix *-rgu-* (originally a directional) is used in combination with the subordinator *-r* for indicating an immediate succession of events (glossed ‘sequential’);
- (2) the combination *xita-yu-* (from \**hita-* ‘to throw’ and the directional suffix \**-yku-*) means ‘to lie down’; the final vowel in *xita-ya-* has been lowered under the influence of the pluralizing suffix *-ri-*;
- (3) the suffix *-(ga)sga* (from \**-(ka)ska*) means ‘somewhat more’;
- (4) *-tbi-* is a local variant of the switch-reference marker *-pti-*;
- (5) the combination *yača-ka(:)-* (from *yača-* ‘to learn’ and *-ka(:)-* ‘non-control’) has the meaning ‘to get used to’; its verbal complement (*miku-y-ta*) is marked for infinitive + accusative case;
- (6) note the co-occurrence of perfective aspect and habitual past in *wina-ru-x.*

‘After giving birth, they would dig a pit in the ground. After digging that pit in the ground, they would lay woollen pelts in it. Their children used to be kept in that pit, while rolling about. Every twenty-four hours they would be suckled. Later on, when they were growing a little bigger ..., as soon as they had got used to eating ..., when they were big already, they would be put in a larger pit.’

#### 14. Uncommon abbreviations

ACT.NR	active nominalizer
ALR	already
BIND	binding element
CUST	customary action
HEARSAY	hearsay validator
HIT(HER)	motion towards speaker
INCL	including
INTO	inward motion
KNWL.VAL	first-hand knowledge validator
LIM	limitative case
NONCONT	non-control
OWNER	owner
SUDD	sudden discovery tense
THERE	action over there
URG	urgency

#### 15. References

- Adelaar, Willem F. H. (1977), *Quechua: Grammar, Texts, Dictionary*. Lisse: The Peter de Ridder Press
- Adelaar, Willem F. H. (1984), “Grammatical Vowel Length and the Classification of Quechua Dialects”. *International Journal of American Linguistics* 50.1, 25–47
- Adelaar, Willem F. H., with the collaboration of Pieter C. Muysken (2004), *The Languages of the Andes*. Cambridge: Cambridge University Press
- Cerrón-Palomino, Rodolfo (1976), *Gramática Quechua: Junín-Huancá*. Lima: Ministerio de Educación; Instituto de Estudios Peruanos
- Cerrón-Palomino, Rodolfo (1987), *Lingüística Quechua*. Cuzco: Centro de Estudios Rurales Andinos “Bartolomé de Las Casas” (Biblioteca de la Tradición Oral Andina 8) [2003]
- Cole, Peter (1982), *Imbabura Quechua*. Amsterdam: North Holland (Lingua Descriptive Series 5)
- Cusihuamán Gutiérrez, Antonio (1976), *Gramática Quechua: Cuzco-Collao*. Lima: Ministerio de Educación; Instituto de Estudios Peruanos
- Girault, Louis (1989), *Kallawaya: El idioma secreto de los Incas*. [La Paz]: UNICEF Panamerican

- Health Organization and World Health Organization
- Levinsohn, Stephen H. (1976), *The Inga Language*. The Hague, Paris: Mouton
- Meneses, Porfirio & Meneses, Teodoro & Rondinelli, Víctor (1974), *Huanta en la cultura peruana*. Lima: Nueva Educación
- Muysken, Pieter (1977), *Syntactic Developments in the Verb Phrase of Ecuadorian Quechua*. Lisse: The Peter de Ridder Press
- Muysken, Pieter (1979), "La mezcla de quechua y castellano: El caso de la 'media lengua' en el Ecuador". *Lexis* (Lima) 3.1, 41–56
- Parker, Gary J. (1963), "La clasificación genética de los dialectos quechuas". *Revista del Museo Nacional* (Lima) 32, 241–252
- Parker, Gary J. (1969), *Ayacucho Quechua Grammar and Dictionary*. The Hague, Paris: Mouton
- Parker, Gary J. (1976), *Gramática Quechua: Ancash-Huailas*. Lima: Ministerio de Educación; Instituto de Estudios Peruanos
- Taylor, Gerald (1975), *Le parler quechua d'Olto, Amazonas (Pérou)*. Paris: SELAF (Langues et civilisations à tradition orale 11)
- Torero, Alfredo (1964), "Los dialectos quechuas". *Anales Científicos de la Universidad Agraria* (Lima) 2.4, 446–478
- Weber, David J. (1989), *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley/CA, Los Angeles, London: Univ. of California Press (UCPL 112)
- Willem F. H. Adelaar, Leiden  
(The Netherlands)

## 135. Yagua (Peba-Yaguan)

1. Introduction
2. Morphological typology
3. Word classes
4. Morphology of nouns
5. Morphology of verbs
6. Illustrative text
7. Uncommon abbreviations
8. References

### 1. Introduction

Yagua is a VSO language of northeastern Peru spoken by some 3,000 people along the Napo, Amazon, and Yavarí rivers and their tributaries. Dialect differences are largely limited to phonetics and phonology, with very few morphological and syntactic differences. Due to extensive migrations (Chaumeil 1981), differences likely cannot be keyed to geographical areas. Yaguas will sometimes try to pass themselves off as *Mestizos* due to long-standing downgrading of the indigenous groups. Except when integrated into *Mestizo* villages, the Yagua live in villages of from one to 20 families (cf. P. Powlison 1985 for ethnographic information, which supersedes Fejos 1943). As of the early 1980's, Thomas Payne (p.c.) estimated that 75% of women and 25% of men were monolingual in Yagua. Spanish proficiency varies; we have observed a few instances of probable semilingualism stemming from the desire of Yagua-speaking

parents to have their children speak only Spanish.

Yagua is the only remaining Peba-Yaguan language. Previously there were at least two others: Yameo (Espinosa Pérez 1955) and Peba (Rivet 1911). Data on these other languages are not extensive, limited largely to word lists, some simple constructions, and short translated examples. All data on Yagua in this article are from fieldwork done by Thomas Payne and myself, and from a superb text collection compiled by Paul Powlison.

The practical Yagua orthography used here has the standard phonetic values with the following exceptions. Orthographic *j* = /h/, *v* = /w/, *c* and *qu* = /k/. There is a phonemic nasal-oral contrast for vowels. Nasal vowels are indicated by a hook below the letter, except that in order to reduce the number of diacritics which must be written, a different strategy is adopted following nasal consonants: orthographic *b* and *d* represent orally-released allophones of /m/ and /n/; vowels are phonemically nasal after *m* and *n*, and phonemically oral after *b* and *d*. Yagua has either a pitch-accent or tone system, which has never been properly analyzed. In this work I endeavor to indicate phonetic high pitch by the acute accent, and low by no marking over the vowel. Further details on phonology and

phonological processes are found in E. Powlison (1971), P. Powlison (1962), D. Payne (1990 a), T. Payne (1983), and Payne & Payne (1990). (Throughout the running text of this chapter, paradigms of Yagua morphemes are indicated in small capital letters, while individual members of a paradigm are glossed in single quotes.)

## 2. Morphological typology

Yagua is almost exclusively **suffixing**. Just one old **prefix** formative *va-* occurs in a number of abstract nouns: *vánuquh* ‘heat’, *vándú* ‘clearness’ (cf. Payne & Payne 1990: 449–451), and possibly also in the nouns *vánu* ‘male, man’, *váturá* ‘woman without children’. Set I **proclitics** indicate subjects of verbs, objects of postpositions, and possessors of nouns. Otherwise there are absolutely no prefixes.

Yagua is **polysynthetic** in the sense that the average verb in running text consists of three to four morphemes, though an attempt to establish position classes yields approximately thirteen positions in the verb. There is some verb compounding with serial-like semantics. Some suffixes most certainly derive from old roots and one could argue that there is both verb and noun compounding. Productive noun incorporation does not occur.

It is **agglutinative** in that within a word it is possible to tell where one morpheme ends and the next begins. However, there is some nasal spreading and vowels and the glide /y/ can affect vowels in contiguous morphemes. P. Powlison (1962) has also described a **metathesis** process between morpheme-final /y/ and any immediately following consonant. However, when the following consonant is coronal, one could argue that palatalization occurs rather than metathesis. Also, /y/ plus /w/ result in the bilabial fricative [β]. When it is phrase-final, /y/ is almost inaudible; it is not that it is voiceless, but there is hardly any closure to the point where a nonnative speaker may completely miss the /y/. It clearly surfaces, however, if any morpheme follows.

## 3. Word classes

Nouns and verbs are the major word classes in Yagua. There are a few adverbs and probably not more than one to three adjectives. There are a multitude of postpositions, but

most transparently derive from nouns. There is also a pronoun or clitic class.

The **pronoun/clitic** class divides into three subclasses depending on whether the form is prefical (Set I), suffixed (Set II), or free. Across all three sets the forms are segmentally identical with just two exceptions (pitch or tone may also vary among the three sets). For example, *ráy* is the free 1<sup>st</sup> person singular pronoun used in emphatic contrastive contexts. *Ray-* is the bound Set I proclitic which occurs on verbs and auxiliaries to indicate subject, on postpositions, and on nouns to indicate possessor. The **enclitic** *-ráy* indicates objects of transitive verbs. The enclitic attaches to whatever immediately precedes the lexical object. When there is no lexical object noun phrase, the same enclitic most neutrally attaches to the end of the sentence. Only for animate 3<sup>rd</sup> person singular is there segmental variation: the free and enclitic 3.SG animate (hereafter just ‘3.SG’) forms are *níil* /-níi/, while the proclitic form is *sa-*; the proclitic coreferential form is *jíiy-* and the enclitic form is *-yu*.

Semantically, the pronouns/clitics distinguish 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> animate, 3<sup>rd</sup> inanimate, and coreferential persons; **singular**, **dual** and **plural** for all **animate** persons; and **inclusive** and **exclusive** for 1<sup>st</sup> person dual and plural. The **coreferential** forms do not occur with first and second persons. The dual forms indicate semantic duality, but the 2<sup>nd</sup> and 3<sup>rd</sup> person forms are also used for singular females who have given birth (which can be specified as singular by the numeral *tíquii* ‘one’).

Diachronically the proclitics are on their way to reanalysis as inflectional agreement in that whenever a lexical noun phrase follows the head of the phrase (whether it be a verb, possessed noun, or adposition), a coreferential Set I form must be prefixed to the head. For example, all the following occur: *vánu rooríy* ‘the man’s house’, *sa-rooríy* ‘his/her house’, *sa-rooríy vánu* ‘the man’s house’, but \**rooríy vánu* (cf. T. Payne 1993; D. Payne 1986 a; and Everett 1989 for discussion of alternative orders). *Rooríy* can be translated as ‘house’, but it most centrally refers to a (traditional) home thatched with palm to the ground and now modernly to the roof of a house thatched with palm fronds.

- (1) (a) *raryooríy*  
*ray-rooríy*  
1.SG-house  
‘my house’

- (b) *vüyüyooríy*  
*vüyü-rooriy*  
1.PI-house  
'our (2 or more) house'
- (c) *nááryooríy*  
*nááy-rooriy*  
3.DE-house  
'our (2) house (but not yours)'
- (d) *núúryooríy*  
*núúy-rooriy*  
1.PI-house  
'our (more than 2) house (but not yours)'
- (2) (a) *jimeejey*  
*jiy-maay-jay*  
2.SG-sleep-CL.cloth  
'your (sg) sleeping mat'
- (b) *sáánamaajey*  
*sáána-maay-jay*  
2.DU-sleep-CL.cloth  
'your (2)', or 'one woman who has given birth's sleeping mat'
- (c) *jiryeemeejey*  
*jiryey-maay-jay*  
2.PL-sleep-CL.cloth  
'your (more than 2 people's) sleeping mat'
- (3) (a) *sa-buyqa*  
3.SG-manioc.beer  
'his/her manioc beer'
- (b) *naada-buyqa*  
3.DU-manioc.beer  
'their (2)', or 'one woman who has given birth's manioc beer'
- (c) *ribuyaq*  
*riy-buyqa*  
3.PL-manioc.beer  
'their (more than 2 people's) manioc beer'
- (d) *ra-buyqa*  
INAN-manioc.beer  
'it's manioc beer'
- (e) *sanááyqa jibuyaq*  
*sa-nááyqa jíy-buyqa*  
3.SG-stomp COREF-manioc.beer  
'She/he stomps her/his own manioc beer.'

**Nouns** cannot take the large range of verbal suffixes, but can take Set I proclitics to indicate possession. Further, noun phrases can be suffixed with postpositions as in *Tomásaroorí-mu* 'at Tom's house'. Compound nouns function as a single word phonologically

(shown by /y/ metathesis and vowel assimilations). This is most certainly the process whereby bound CLASSIFIERS have arisen (cf. 4).

**Adjectives** cannot take verbal suffixes or Set I proclitics marking possession. They can directly modify nouns but not verbs. They cannot function as a subject or object phrase syntactically, or as the object of a postposition. By these criteria the only clear adjective found in Yagua to-date is *jáámu* 'big' and perhaps *jápuu* 'small'. *Pásiy* 'small' has some nominal properties in that it can directly take postpositions. *Sámiy* 'good, well' can without further affixation modify nouns and verbs, and thus it has more adverbial properties. Other semantic concepts that would be expressed by syntactic adjectives in English are usually expressed by abstract nouns in Yagua, or else by suffixing CLASSIFIERS to the noun, which modify the noun by adding features of size, shape, function, constituting material, etc. Most postpositions derive from nouns and can be compounded, as in *sa-jutunuváá-ramurya-mu-siy* '3.SG-neck-end-LOC-ABL (from the end of his neck)'.

**Adverbs** do not take Set I clitics or verbal suffixes, but without further affixation can modify verbs. They include *sámiy* 'good, well', *munátya* 'first', *mítya* 'just', *jááriy* 'really'.

Morphologically, **verbs** are the most complex word class in Yagua, and with addition of a Set I proclitic to indicate the subject of the clause readily function as the sole element of a clause.

#### 4. Morphology of nouns

A noun root or stem can take suffixes in the following linear order: CLASSIFIER<sup>n</sup>, CONDITION SIZE QUANTITY. All except perhaps the QUANTITY suffix would be considered derivational. More than one CLASSIFIER can occur, as in:

- (4) *roorijyúdaypoóquñmíy*  
*rooriy-ju-day-poo-quñ-miy*  
house-CL.opening-CL.patch-CONDITION.rotting-SIZE.big-QUANTITY.PL  
'the several tall and rotting house doors'

Nouns with so many suffixes are infrequent, but they are completely natural.

##### 4.1. Inflectional morphology on nouns

If the Set I proclitics are viewed as inflectional, then Yagua nouns can be said to inflect for possessor. But Yagua nouns do not

have any obligatory inflections. They can optionally take QUANTITY suffixes such as *-juy* ‘dual’, *-miy* ‘plural’, *-vay* ‘plural’, *-ju* ‘heap of’, *-súma* ‘very’ which indicates augmentation as in *vánda-quíi-súma* ‘heat-SIZE.big-QUANTITY.very (light that extends a lot (e.g., from the sun))’. Dual suffixes are most frequently employed to indicate the status of a female who has given birth, rather than literal dual number. At least dual affixes can be considered inflectional, not because number must be specified (it need not be), but because they reflect lexically-specified features of the noun which are also reflected elsewhere in the clause (D. Payne 1990 a), just as masculine or feminine gender is an inherent lexical feature of nouns in Spanish. This can be seen from the behavior of CLASSIFIERS. This lexical number feature is also reflected in the clause by any clitic or pronoun element that refers to the referent.

Numerals and demonstratives must be inflected with CLASSIFIERS in order to be complete word forms. Once so inflected, they are syntactically nominal and can either occur with another noun, or can function almost pronominally as the only element in a noun phrase.

There are over 40 classifiers (D. Payne 1986 b; Powlison & Powlison 1958; cf. Art. 97). Animate CLASSIFIERS distinguish singular, dual, and plural number: *-nu*, *-nuuy*, *-vay*, respectively. Inanimate ones distinguish features such as 1, 2 and 3-dimensional shape, square vs. round, flexible vs. rigid, function, arrangement, and a number of unique items. The general CLASSIFIER *-ra* most often refers to inanimates, but can be used for animates as in *dee-ra* ‘DIM-CL (little one)’. Synchronously one could argue that CLASSIFIERS are infixated to the numerals ‘one’ through ‘four’ and numerals formed on the basis of them, and suffixed to other numerals. However, ‘two’ and ‘three’ transparently end in the dual and plural suffixes *-juy* and *-miy*.

Inflected demonstratives can also take the QUANTITY suffixes: *ru-nu-juy dapúuy-nuuy* ‘that-CL.ANIM.SG-DU hunt-CL.ANIM.DU (those two hunters)’. The demonstrative *ru* is inflected with *-nu* singular, reflecting the lexically specified number of the lexeme HUNTER. The dual CLASSIFIER *-nuuy* on *dapúuy* is not required as an inflection, but serves to derive a noun from a verb root (cf. 4.2). Similarly, *tá-nuuy-quH* ‘one-CL.ANIM.DU-one’ mixes the dual CLASSIFIER with the semantically singular

numeral ‘one’, as for example when the dual lexeme *váturuy* ‘woman who has borne children’ is being counted.

#### 4.2. Derivational morphology on nouns

Nouns can contain either an inherently nominal root (see (1), (3)), or in some restricted cases be prefixed with the formative *va-*. Or they can productively be derived from verb roots by adding a CLASSIFIER (see (2)), the infinitive *-janul-jada* (homophonous with the distant past tense suffix) as in *ramuchoonu* from *ramuchu-janu* ‘swallow-INF’, or one of several other suffixes (Payne & Payne 1990).

CLASSIFIERS also occur on stems that are already nominal in order to derive particular senses, as in *rooriy-ju-day* ‘house-CL.opening-CL.patch (door)’.

Yagua CLASSIFIERS thus cannot be specified as exclusively inflectional or derivational (D. Payne 1985 a; 1990 a). Rather, there is one set of some 40 forms that in some contexts functions inflectionally, but in other contexts functions derivationally. Inflection and derivation are thus best viewed as functions, and not as features inherent to particular morphemes (cf. Art. 38).

CONDITION and SIZE suffixes surely derive from lexical roots and have possibly developed historically from processes of incorporation, though incorporation of nouns into verbs is not synchronically viable. For example, the CONDITION affix *-poo* ‘rotting’ in (4) corresponds to the verb *poo* ‘rot’. The diminutive SIZE suffix *-dee* is found as an apparent root in the nouns *deenu* ‘boy’ and *deetu* ‘girl’ where *-nu* would appear to be the animate singular CLASSIFIER, but *-tu* does not function as a classifier (i.e., it does not inflect numerals and demonstratives). However, the form *dee* alone does not function as a noun.

### 5. Morphology of verbs

Although the verb is the most complex word class in Yagua, the ‘internal syntax’ of Yagua verbs is still not as complex as that of Mai-puran Arawakan, Panoan, or Athabaskan verbs (D. Payne 1985 c). Somewhere between 10 and 13 position classes could roughly be posited, but a strict position class approach is impossible. Further, although grammatical, it is quite uncommon for more than four affixes to occur on a verb. Morpho-semantically, a first approximation of affix order is:

- (5) ROOT-TRANSITIVITY-LOCATION-BOUNDED.MOVEMENT-COMPLETIVE-ITERATION-IMPERFECTIVITY-MODAL-TENSE

### 5.1. Postbases

A strict position class approach is impossible because of the behavior of **postbase** suffixes such as *-tániy* ‘causative’, *-rúy* ‘potential/desiderative’, the UNBOUNDED MOVEMENT suffix *-nayqa* ‘going all over’, and possibly the ‘completive’ *-muuy*. Postbases create new layers of structure, and do not constitute a single paradigm. Certain other suffixes and postbases themselves can have scope over (i) an inner structural layer established just by a root or stem, (ii) over a layer created by a postbase, or (iii) over both. Linear order in such cases generally depends on semantics. When there are differences in interpretation, the suffix to the right has wider scope:

- (6) *Saquivúchurúmyuuyanu.*

*sa-quivúy-su-rúy-muuy-janu*  
3.SG-deceive-TRR<sub>1</sub>-POT-COMPL-PAST<sub>3</sub>  
'She/he stopped wanting to deceive (someone) long ago.'

- (7) *Suumiiryútyánimyuuryúñii.*

*sa-jimiyi-rúy-tániy-muuy-rúy-níi*  
3.SG-eat-POT-CAUS-COMPL-POT-3.SG  
'She/he thinks that she/he should finish making her/him eat everything.' (lit. 'She/he wants to finish making her/him want to eat.') [Thus she/he will eat everything.]

It is thus impossible to account for affix order by fixed position classes. Rather, the following overall structure is likely (D. Payne 1990 b):

- (8) [[[STEM]-STEM.SUFFIX]-POSTBASE-POSTBASE.SUFFIX]-CLAUSAL.SCOPE.SUFFIX]

In (8), the stem suffixes and the posbase suffixes are comprised by the same paradigms, while suffixes within the verb stem (TRANSITIVITY suffixes discussed in 5.3) and those which have clausal scope, such as TENSE, are distinct.

*-Rúy* is related to the verb root *nirúy* ‘desire’. The suffix ranges in meaning from ‘desire to’, ‘want to’, ‘be able to’, ‘can’, and possibly ‘think’. Although it does not have a fixed position, it must precede IMPERFECTIVITY and preferably precedes UNBOUNDED MOVEMENT suffixes:

- (9) *Suumiiryútyítiiy.*

*sa-jimiyi-rúy-títiy*  
3.SG-eat-POT-GOING.DIRECTLY  
'She/he wants to eat going along.'

- (10) *?Suumiityítyiiryúy.*

Depending on placement of *-tániy*, scope of causation may change. The verb root plus suffixes to the left of *-tániy* are predicated of the causee. In general, *-tániy* plus suffixes to its right are predicated of the causer:

- (11) *Suumnaatyánñii Mario.*

*sa-júnaay-tániy-níi*  
3.SG-cry-CAUS-3.SG Mario  
'She/he makes Mario cry.'

- (12) *Suumnaamyuutyánñii Mario.*

*Sa-júnaay-muuy-tániy-níi*  
3.SG-cry-COMPL-CAUS-3.SG Mario  
'She/he makes Mario stop crying.'

- (13) *Suumnaatyánimyuunñii Mario.*

*sa-júnaay-tániy-muuy-níi*  
3.SG-cry-CAUS-COMPL-3.SG Mario  
'She/he stops making Mario cry.'

In (14) *-naayqa* has scope over both causing and wanting to cry and is predicated of both participants. In (15) the causee is going all over, while the causer need not be; (16) and (17) are similar.

- (14) *Suumnaaryútyáineyqaníi.*

*sa-júnaay-rúy-tániy-naayqa-níi*  
3.SG-cry-POT-CAUS-GOING.AIMLESSLY-3.SG  
'She/he makes him/her want to cry while travelling.' (Apparently both participants are travelling together.)

- (15) *Sarupiíneyqatánñii.*

*sa-rupiíy-naayqa-tániy-níi*  
3.SG-walk-GOING.AIMLESSLY-CAUS-3.SG  
'She/he makes him/her walk all over.'

- (16) *Saraqyqatáninuvñii.*

*sa-raqy-yaq-tániy-nuvñ-níi*  
3.SG-jump-DISTR-CAUS-ON.ARRIVAL.HERE-3.SG  
'She/he makes him/her dance upon arrival here.' (Whose arrival is not specified; perhaps both are arriving together.)

- (17) *Saraqyqanuvñtánñii.*

*sa-raqy-yaq-nuvñ-tániy-níi*  
3.SG-jump-DISTR-ON.ARRIVAL.HERE-CAUS-3.SG  
'She/he makes him/her come here to dance.'

The UNBOUNDED MOVEMENT suffix *-nayqa* possibly derives from *nááya* ‘to press, stomp’.

It contrasts semantically with *-tityiiy* ‘going directly’. However, *-tityiiy* is more fixed in position than *-nayqq* and thus it is less clear that *-tityiiy* should be considered a postbase.

### 5.2. Inflectional morphology

TENSE suffixes are inflectional in that either an overt TENSE suffix or zero is required on any finite verb. Zero is meaningful and thus part of the paradigm. This is not true for absence of IMPERFECTIVITY, COMPLETIVE, MOVEMENT, LOCATION, ITERATION, and suffixes of other categories. TENSE suffixes are always final (though clitics may follow TENSE).

Six distinctions in time reference are made by verb suffixes. Zero indicates non-past (present/future) and is used for the ‘historical present’ when past time is clearly set at the beginning of a text or episode. Future is differentiated from present or historical present by the irrealis auxiliary *a* (which is also used for relative future set in the past).

There are two proximate TENSES. Most neutrally *-jásiy* ‘PROX<sub>1</sub>’ indicates time earlier on the same day; *-jáy* ‘PROX<sub>2</sub>’ indicates time about one day ago. When used with the irrealis auxiliary, however, these affixes convey modal meanings in the future.

There are three past TENSES: *-siy* ‘up to a few weeks ago’, *-tiy* ‘from a few months to a year ago’, and *-jada* (dialect variant *-janu*) ‘distant’ or ‘legendary’ past time. Interpretation of the precise time indicated by each past TENSE morpheme is variable.

### 5.3. Derivational morphology on verbs

TRANSITIVITY suffixes occur within the STEM portion in (8) (D. Payne 1985 b). These change valence or the intensity of an action. Occasionally they derive verbs from nouns. With the exception of the ‘instrumental’ applicative *-ta*, they are highly restricted in terms of what verb roots they may co-occur and are idiosyncratic in meaning. The suffix *-vay* ‘intensity’ can derive a verb with a Theme role. Compare: *sa-jáy* ‘3.SG-father (his/her father)’ vs. *Rd-jáy-vey* ‘INAN-mature-INTS (It grows.)’. The suffix *-su* adds an Actor role or can derive a verb from a noun by adding both Theme and Actor: *Sajáq-vyechuníi* from *sa-jáy-vay-su-níi* ‘3.SG-mature-INTS-TRR<sub>1</sub>-3.SG (He caused him to grow.)’ The suffix *-nu* is similar to *-su*: *vátajú* ‘wound (noun)’ vs. *Sa-vátajú-nu-níi* (3.SG-wound-TRR<sub>2</sub>-3.SG) ‘She/he wounds him/her.’

The applicative *-ta* corresponds to the ‘instrumental’ (or ‘comitative’) postposition *-ta*:

- (18) *Sijchitiñíi*      *quiivá* *quiichitya*.  
*sa-jichitiy-níi*      *quiichity-ta*  
 3.SG-poke-3.SG fish knife-INST  
 ‘She/he pokes the fish with the knife.’
- (19) *Sijchitityara*      *jumurutágníi*.  
*sa-jichitiy-ta-ra*      *jumuru-táqá-níi*  
 3.SG-poke-INST-3.NAN clear-NR.INSTR-3.SG  
 ‘She/he pokes with the machete him/her.’ (i.e., ‘She/he pokes him/her with the ground-clearing instrument.’)

The suffix *-niy* also increases transitivity and probably corresponds to the postposition *-niy* ‘in’. The iterative *-tiy* is perhaps also part of this paradigm. It is lexically restricted and occurs with certain roots indicating action done with some type of motion; compare *jiitya* ‘remove (for some purpose)’ and *jiitya-tiy* ‘take all apart’. (Other ITERATIVE suffixes are productive and occur further out in the verb structure.)

The only detransitivizing suffix is the lexically restricted anti-causative *-y* (perhaps also with middle semantics). Compare *Sa-mutamaa Hilario-ra rooriy-jú* ‘3.SG-open-PF Hilario-INAN house-CL.opening (Hilario has opened the door.)’ vs. *Ra-múta-y-maa rooriy-jú* ‘INAN-open-ACAUS-PF house-CL.opening (The door has opened/has been opened.)’.

Following TRANSITIVITY suffixes come affixes that can operate on stem and postbase layers. These are highly productive and semantically predictable. They are not required and their absence does not imply a meaningful distinction within a paradigm.

As in many South American languages, **location** and **movement** are important semantic features of Yagua. LOCATION suffixes indicate that the action of the verb was done at a stationary place. These are *-ja* ‘across from’, *-imu* ‘down’ (usually downstream), and *-sa* ‘up’ (usually upstream) as in *Suyníuchusaníi* from *sa-juniúy-su-sa-níi* ‘3.SG-look-TRR<sub>1</sub>-UPWARDS-3.SG (He cared for him upriver.)’.

Two general types of movement are conveyed by verbal suffixes. A movement trajectory can be bound at its inception, its termination, or it may be punctuated, by another event. Such movement is referred to as **bounded**. Movement which is true throughout an event and which imparts imperfectivity is **unbounded**.

There are four BOUNDED MOVEMENT suffixes: *-nuvh* indicates the action is done upon arrival at the locative deictic center; *-nuvee* indicates it is done upon arrival at a location away from the deictic center. These suffixes

are related to verb roots *v̥y* or *viy* ‘arrive here’ and *veey* ‘arrive there’. The suffix *-siy* indicates the action is done upon departure from a location. This is etymologically related to the ablative postposition *-siy*. Both the postposition and the suffix may be related to the verb *siy* ‘run’ or *maasiy* ‘go out’. As a suffix, *-siy* bounds the inception of some movement relative to the action of the verb root. The suffix *-rij* indicates that an event is done while enroute to some destination. Thus, the event can bound both the termination of one part of a movement trajectory, and its resumed inception: *Sa-suuta-rij* ‘3.SG-wash-ENROUTE (She/he washed enroute.)’.

Productive ITERATIVE affixes include *-jayqa* ‘iterative’ and *-jaa* ‘iterative movement to some other location’. These do not co-occur with the proximate TENSES, though they may occur with more remote past TENSES as in *Rameeyjeyyqqisy* from *ray-maay-jayqa-siy* ‘1.SG-sleep-ITER-PAST<sub>1</sub> (A few weeks ago I slept various times.)’. Other ITERATIVE suffixes include *-saniy* ‘group action carried out by individuals successively’. Finally, *-jadapúúryii* is close to a semelfactive but also conveys the idea of ‘suddenly’. It does not (easily) occur with roots which have an inherent durative aspect such as *jimiy* ‘eat’ or *rupiy* ‘walk’ (and such collocations are unacceptable to some speakers).

The ‘distributive’ *-yaa* is problematic under a position class approach. It may occur contiguous to the root, forming well-lexicalized verb stems, as in (20). But it need not; in (21) *-yaa* occurs twice. The first instance may be lexicalized as part of how one expresses ‘tap foot’. The second instance follows the second position clitic *-numaa* ‘now’ which is not a verbal suffix per se.

- (20) *Sarááyatániñuvnii*.  
*sa-rááy-yaa-tániy-nuy-nii*  
 3.SG-jump-DISTR-CAUS-ON.ARRIVAL.HERE-  
 3.SG  
 ‘She/he makes him/her dance upon arrival here.’

- (21) *Ríjeniyqanumaayaa*.  
*ray-jíjeniy-yaa-numaa-yaa*  
 1.SG-tap.foot-DISTR-now-DISTR  
 ‘I am tapping my foot now (iteratively).’

With the exception of *-yaa*, ITERATIVITY suffixes most neutrally precede IMPERFECTIVITY, and follow bounded movement.

- (22) *Rachuutajayqanúúyada*.  
*ray-suuta-jayqa-núúy-jada*  
 1.SG-wash-ITER-IPFV-PAST<sub>3</sub>  
 ‘I used to wash all the time.’

- (23) *Samaachajqasara*.  
*sa-maay-sa-jaa-sara*  
 3.SG-sleep-UPWARDS-ITER.MOVEMENT-HABIT  
 ‘She/he always goes up (river) to sleep.’

- (24) *Naada-suuta-rij-jayqa*.  
*3.DU-wash-ENROUTE-ITER*  
 ‘They two always stop enroute to wash.’

The IMPERFECTIVITY paradigm includes *-jancha* or *-janumucha* ‘continuative’ and *-núúiy* ‘imperfective’. The gnomic habitual *-sara* (cf. (23)) cannot occur with TENSE or other IMPERFECTIVITY suffixes. Thus, its paradigmatic status is ambiguous. Here I treat it as part of the IMPERFECTIVITY paradigm. (The second position clitics *-numaa* ‘now’ and *-maa* ‘perfect’ are clearly aspectual, but are not verbal suffixes. Collocation of the IMPERFECTIVE *-núúiy* and the distant past tense *-janul-jada* is extremely likely, so much so that *-núúyanu* (from *-núúy* plus *-janu*; variant *-núúiyada*) could be a nearly-lexicalized suffix complex.

Suffixes which are not part of the IMPERFECTIVITY paradigm may have inherent aspect. In particular, ITERATIVE and UNBOUNDED MOVEMENT suffixes have imperfective meanings. COMPLETIVE and BOUNDED MOVEMENT suffixes are more perfective.

The COMPLETIVE *-muuy* is as close to a grammaticalized PERFECTIVE as can be found in Yagua. However, it puts heavy emphasis on the termination of an action and thus is not a true perfective. Semantically *-muuy* is most closely opposed to *-núúiy* ‘imperfective’. Positionally the two are distinct and *-muuy* shares some positional properties with postbases. For some speakers (but not all) *-muuy* may co-occur with and precede the HABITUAL *-sara*. For apparently all speakers it may co-occur with the CONTINUOUS *-janchal-janumucha*, either before or after it. However, it occurs after BOUNDED MOVEMENT suffixes as in (25), and preferably before UNBOUNDED MOVEMENT suffixes.

- (25) *Naada-suuta-rij-muuy*.  
*3.DU-wash-ENROUTE-COMPL*  
 ‘They two finished washing enroute.’

The COMPLETIVE also occurs before some ITERATIVITY suffixes such as *-jadapúúryii*. The latter most neutrally occur close to the verb root.

- (26) *Rachiúmyuuyadapúryijra.*  
*ray-súiy-muuy-jadapúryijj-ra*  
 1.SG-bite-COMPL-ONE.MOVEMENT-INAN  
 'I finished biting all at once (and there isn't any food left).'

In contrast to BOUNDED MOVEMENT, *-títyiiy* 'action done while going directly to some destination' and *-nayqqa* 'action done while going aimlessly all over' do not put a bound on the inception or termination of any movement trajectory. Rather, the movement is presented as simultaneous with the action of the main verb root. Alternatively, the point(s) at which the action was done relative to the movement are not of concern:

- (27) *Sa-suuta-títyiiy*  
 3.SG-wash-GOING.DIRECTLY  
*sa-rooriy-mút-ju.*  
 3.SG-house-LOC-TOWARD  
 'She/he<sub>i</sub> washed while going along directly to his/her<sub>j</sub> house.'
- (28) *Sa-suuta-nayqqa.*  
 3.SG-wash-GOING.AIMLESSLY  
 'She/he washed while going all over aimlessly.'

Unlike the BOUNDED *-rií*, UNBOUNDED MOVEMENT suffixes do not imply that the movement was necessarily interrupted while the action was performed. With certain roots such as 'wash' it is pragmatically assumed that the actor stopped to perform the action at some place along a river or stream. But with a root like 'eat' stopping is unnecessary.

*-Nayqqa* may vary in position relative to BOUNDED MOVEMENT suffixes, though differences in semantic scope are extremely subtle. In (29) there may be only one general area of locative reference, while (30) may indicate several. Lack of clarity in semantic scope is likely due to the imperfectivity inherent in *-nayqqa*: an imperfective action is somehow distributed either in time or space; the distributivity may be the dominant semantic feature, and not the spatial vs. temporal domain.

- (29) *Sa-suuta-nuvq-nayqqa.*  
 3.SG-wash-ON.ARRIVAL.HERE-GOING.AIM-  
 LESSLY  
 'She/he washes here, over here, over  
 here, whenever arriving here.'
- (30) *Sa-suuta-nayqqa-nuvq.*  
 3.SG-wash-GOING.AIMLESSLY-ON.ARRI-  
 VAL.HERE  
 'She/he came to every place to wash.'

UNBOUNDED MOVEMENT suffixes must precede IMPERFECTIVITY suffixes:

- (31) *Sasuutatítyiñúyada.*  
*sa-suuta-títyiiy-núuy-jada*  
 3.SG-wash-GOING.DIRECTLY-IPFV-PAST<sub>3</sub>  
 'She/he always goes along washing.'

Verbal suffixes which convey MOOD include *-vqqa* 'action achieved' and *-táqta* 'debitive'. Relative position and morphological properties of these suffixes need further study; *-vqqa*, at least, may precede the gnomic habitual *-sara*. It almost always occurs in negative contexts to indicate 'action not achieved':

- (32) *Dantya-múy yqqa*  
 thus-NEG 2.SG.IRR  
*játa-yqqa-vqqa-sara* *marya-day.*  
 move-DISTR-ACHIEVE-HABIT day-DAY  
 'Neither will you move (to another place) during the day.'

- (33) *Saaníduutyatáqtiíyu.*  
*sááda-jíduuty-a-táqta-iyu*  
 2.DU-prepare-DEB-REFL  
 'You two had better get yourselves ready.'

## 6. Illustrative text

Following are two short portions from an oral story *Mucatyu Munatyí* told by Manungo Díaz, who was roughly 60 years old at the time of telling. Though the story was clearly well known to the speaker, on the occasion on which this version was narrated the initial portion did not appear to be well planned. The narrator backs up and fills in bits of essential information for his audience, namely that the brother of a toucan had been swallowed by the snake. This fact is important to the entire story because later the deer (who cannot jump like the squirrel) will find himself inside the snake alongside a toucan.

This story without morpheme analysis, but with a full Spanish translation, has been printed for Yagua readers in Díaz & Peña (1982, eds.). Just a few clauses from the beginning, and from a bit later in the story, are given here. For reasons of space, for the most part the text is written representing essentially underlying forms of morphemes so that the morphology will be transparent, rather than the way Yagua is practically written

(which makes clean morpheme divisions impossible). The form here thus does not show /yl/-metathesis and operation of other rules that affect vowels. However, some derived words have essentially been lexicalized and these are not necessarily broken into underlying forms. For example, *munáátyij* ‘ancestor’ and *tááryij* ‘other’ contain roots *munáátya* and *táárya* plus an animate singular nominalizer *-i*. *Jíjta* is a second position element which, among other functions, occurs on most main events (D. Payne 1992).

- (34) *Sa-ramutyey-núuy-janu múcatyu*  
3.SG-pass-IPFV-PAST<sub>3</sub> squirrel  
*munáátyij sa-i coodiy tqaqriy*  
ancestor 3.SG-DAT snake before
- (35) *Sa-quiiñuñuúy-su-níi jánáriy*  
3.SG-deceive-TRR<sub>1</sub>-3.SG deer  
*munáátyij.*  
ancestor
- (36) *Nuvá tááryij sa-joota-janu*  
toucan brother 3.SG-begin-PAST<sub>3</sub>  
*ramuchu-janu coodiy munáátya.*  
swallow-INF snake first
- (37) *Sa-ramuchu-janu numaa-tiy-níi*  
3.SG-swallow-INF now-CONN-3.SG  
*sa-tááryij,*  
3.SG-brother  
*sa-raqy-tiiy-yqq-sa-janu*  
3.SG-jump-TRR-ITER-ITER.MOVEMENT-UP-  
WARDS-PAST<sub>3</sub>  
*jíjta sa-tááryij váriy jíy-va-siy,*  
JIITA 3.SG-brother then COREF-DAT-ABL  
*nuvá-dee-ra.*  
toucan-DIM-CL.general  
. . . (several paragraphs later) . . .
- (38) *Múcatyu joqta ramutyey-janu*  
squirrel begin pass-INF  
*sa-jajii-siy sa-siyy-siy jíjta*  
3.SG-place-ABL 3.SG-run-DEPART JIITA  
*múcatyu sa-jajii-siy coodiy*  
squirrel 3.SG-place-ABL snake  
*rúdii-íva.*  
back-DAT
- (39) *Sa-raqy-nuvee*  
3.SG-jump-ON.ARRIVAL.THERE  
*sa-jutunuvúy-ramurya-mu-siy mu*  
3.SG-neck-end-LOC-ABL LOC  
*jqq-ruday-mu níinu-tqasá-ju.*  
water-side-LOC tree-MID-TOWARD
- ‘(34) Long ago the squirrel’s ancestor passed (over) the snake. (35) (All this happened

when) he deceived the deer’s ancestor. (36) The toucan’s brother snake (had) swallowed first. (37) When he [the snake] swallowed his brother, his (remaining) brother then went jumping up away from him, (poor) little toucan. (38) The squirrel began to bound away from [the deer]. The squirrel ran upon departure from him on the snake’s back. (39) He jumped on arrival there from the end of his neck at the side of the river, to the middle of a tree.’

## 7. Uncommon abbreviations

COREF	coreferential
DEB	debitive

## 8. References

- Chaumeil, Jean-Pierre (1981), *Historia y migraciones de los Yagua de finales del siglo XVII hasta nuestros días*. Lima: Centro Amazónico de Antropología y Aplicación Práctica
- Díaz Cahuachi, Pedro & Peña Cahuachi, Hilario (1982, eds.), *Vutyuchodatera (algunos de nuestros cuentos)*. Pucallpa/Peru: Instituto Lingüístico de Verano (Colección Literaria y Cultural, Libro 5, Yagua)
- Espinosa Pérez, L. (1955), *Contribuciones lingüísticas y etnográficas sobre algunos pueblos indígenas del amazonía peruana*. Madrid: Instituto Bernardino de Sahagún
- Everett, Daniel (1989), “Clitic Doubling, Reflexives, and Word Order Alternations in Yagua”. *Language* 65, 339–372
- Fejos, Paul (1943), *Ethnography of the Yagua*. New York: Viking Fund (Viking Fund Publications in Anthropology 1)
- Payne, Doris L. (1985 a), “Inflection and Derivation: Is there a Difference?”. In: Tomlin, Russell & DeLancey, Scott (eds.), *Proceedings of the First Annual Meeting of the Pacific Linguistics Conference*. Eugene/OR: University of Oregon, 247–260
- Payne, Doris L. (1985 b), “Degrees of Inherent Transitivity in Yagua Verbs”. *International Journal of American Linguistics* 51, 19–37
- Payne, Doris L. (1985 c), *Aspects of the Grammar of Yagua: A Typological Perspective*. Ph.D. dissertation, UCLA
- Payne, Doris L. (1986 a), “Basic Constituent Order in Yagua Clauses: Implications for Word Order

- Universals". In: Derbyshire, Desmond & Pullum, Geoffrey (eds.), *Handbook of Amazonian Languages, Vol. I*. Berlin: Mouton de Gruyter, 440–465
- Payne, Doris L. (1986 b), "Noun Classification in Yagua". In: Craig, Colette (ed.), *Noun Classes and Categorization*. Amsterdam: Benjamins, 113–131
- Payne, Doris L. (1990 a), *The Pragmatics of Word Order: Typological Dimensions of Verb Initial Languages*. Berlin: Mouton de Gruyter
- Payne, Doris L. (1990 b), "Morphological Characteristics of Lowland South American Languages". In: Payne, Doris L. (ed.), *Amazonian Linguistics: Studies in Lowland South American Languages*. Austin: Univ. of Texas Press, 213–241
- Payne, Doris L. (1992), "Narrative Discontinuity vs. Continuity in Yagua". *Discourse Processes* 15, 375–394
- Payne, Doris L. & Payne, Thomas E. (1990), "Yagua". In: Derbyshire, Desmond & Pullum, Geoffrey (eds.), *Handbook of Amazonian Languages, Vol. II*. Berlin: Mouton de Gruyter, 249–474
- Payne, Thomas E. (1983), "Subject Inflection of Yagua Verbs". Ms.
- Payne, Thomas E. (1993), *The Twins Stories: Participant Coding in Yagua Narratives*. Berkeley: Univ. of California Press
- Powlison, Esther (1971), "The Suprahierarchical and Hierarchical Structures of Yagua Phonology". *Linguistics* 75, 43–73
- Powlison, Esther & Powlison, Paul (1958), "El sistema numérico del yagua (pebano)". *Tradición, Revista Peruana de Cultura* 21, 3–8
- Powlison, Paul (1962), "Palatalization Portmanteaus in Yagua (Peba-Yagan)". *Word* 18, 280–299
- Powlison, Paul (1985), *Yagua Mythology: Epic Tendencies in a New World Mythology*. Dallas: International Museum of Cultures
- Rivet, Paul (1911), "La famille linguistique Peba". *Journal de la Société des Americanistes de Paris* 8, 173–206

Doris L. Payne, Eugene (U.S.A.)

## 136. Tagalog (Austronesian)

1. Introduction
2. Morphosyntax and parts of speech
3. Formal processes
4. Voice, aspect, and mood
5. Actor involvement
6. Illustrative text
7. Uncommon abbreviations
8. References

### 1. Introduction

#### 1.1. The language and its speakers

Tagalog, a Meso-Philippine language belonging to the Western Malayo-Polynesian branch of the Austronesian family, is one of the major languages of the Republic of the Philippines (Constantino 1971: 112–118). In the 1990s, it was the native language of around 15 million speakers. Most Tagalog speakers live in the central parts of the island of Luzon, in particular in the provinces Batangas, Bulacan, Laguna, and Marinduque. But Tagalog is also widely spoken in the surrounding provinces (including the capital, Manila), and there are many settlements of

Tagalog speakers throughout the Philippine islands, especially on Mindoro and Mindanao (cf. McFarland 1983: 29, 80). Furthermore, Tagalog is widely spoken as a second language. In 1937 it was selected as the basis for the national language Filipino (formerly called Pilipino).

Since the beginning of Spanish colonisation in 1565 the Tagalog area has always been the center of political power in the Philippines, and Tagalog has thus been strongly influenced by the colonial languages, first Spanish and then American English (from 1898 to 1946). This influence, however, has been largely on the lexicon and the phonology, but not on the morphosyntax. For example, the Manila slang, called Taglish, mainly consists of English roots, but the morphology is exclusively Tagalog. Examples include *lipstick-an* 'to put lipstick on', *i-give up* 'to give sth. up', and *mag-on* 'to start dating' (cf. Cubar 1984; the affixes are explained in 4.1). As for morphosyntax, it is not unlikely that there has been some normative influence by colonial institutions and practices

because the Spanish started using Tagalog very early on as a missionary language, writing grammars and preparing catechisms in Tagalog (a *Doctrina Christiana*, dated 1593, is the oldest preserved Tagalog document). The continuing normative influence exerted by the Manila based educated classes is shown by the fact that Tagalog exhibits surprisingly little dialectal diversity (McFarland 1983: 80). Only the dialect spoken on the small island of Marinduque exhibits lexical and morphological idiosyncrasies that have caught the attention of linguists (Lopez 1970; Soberano 1980).

The most influential of the Spanish grammars is the one written by Totanes (1865) which has served as the basis for many analyses of Tagalog (e.g. Humboldt 1838; Müller 1882: 87–163; Marre 1901). The last and most comprehensive Tagalog grammar cast in a traditional framework is Blake (1925). Bloomfield's texts and grammar (1917) are the first attempt to present Tagalog in its own terms and a very fine and early piece of modern structuralist analysis. In addition to Bloomfield and Blake, further comprehensive statements of Tagalog morphology can be found in Lopez (1937), Schachter & Otanes (1972), and Wolff et al. (1991). A review of linguistic work on Tagalog is given in Constantino (1971: 118–145) and Reid (1981). The present article is based primarily on Bloomfield's data and analysis.

## 1.2. Phonology and orthography

The segmental phonology of Tagalog is relatively simple (except for problems concerning the treatment of loans for which see Matsuda French 1988: 1–17). The graphemes used in the standard orthography directly reflect phonemes and basically have their IPA values. The only exception is the use of the digraph <ng> for the velar nasal [ŋ]. The use of this digraph may be slightly confusing, since it is also used for the NP-marker [naŋ], in this case written as a separate word which in turn is not to be confused with -ng = [ŋ], the bound allomorph of the linker *na*. A genuine problem with the standard orthography is the fact that glottal stops are not written. Glottal stops regularly occur before initial vowels and intervocally, i.e., <tao> 'person' is [tá?o], <aral> 'study' is [?áral] and <magaaral> 'will study' is [mag?á?áral]. As the last example shows, base-initial glottals

are not dropped when prefixes are added. Although the phonemic status of these initial and intervocalic glottals is far from clear much recent writing on infixation in Tagalog assumes without any discussion that they are non-phonemic (e.g. Crowhurst 1998: 590 who misrepresents her older sources with regard to this point). Word-finally, the problem is further confounded by the fact that here final vowels (optionally followed by [h]) contrast with glottal stops. Thus, <baga> 'glowing charcoal' is [bágah], but <baga> 'lungs' is [bágá?]. Deviating from standard orthography a final glottal stop will be marked in this article by '>'.

Stress is also unmarked in standard orthography. Its analysis is somewhat controversial. Some authors (e.g. Schachter & Otanes 1972: 15–18; Wolff et al. 1991: 12) consider vowel length the primary phenomenon, while others consider vowel length an epiphenomenon of stress (cf. Bloomfield 1917: 141f.; Matsuda French 1988: 63f.). In this article, the latter view has been adopted, but since stress assignment in Tagalog is not fully understood, all remarks pertaining to this phenomenon must be regarded with caution. Stress clearly is phonemic in Tagalog, compare *búkas* 'tomorrow' and *bukás* 'open', and plays an important rôle in affixation (see 3.4). Bases with stress on the penultimate syllable are called barytone bases, those with stress on the final syllable oxytone bases. Primary stress on the penultimate syllable will remain unmarked (thus *bukas* for 'tomorrow'), elsewhere it will be marked by the acute accent. The grave accent marks secondary stress.

Phonological rules of relevance to the present exposition are:

- /d/ often becomes /r/ intervocally;
- phrase-final vowels are followed by a weak glottal fricative (cf. [bagah] above);
- /u/ and /i/ in phrase-final syllables are lowered to [o] and [e], respectively;
- word-final glottal stops regularly disappear before clitics and the linker (cf. Bloomfield 1917: 136; Wolff et al. 1991: 12).

## 2. Morphosyntax and parts of speech

Tagalog distinguishes between two parts of speech (Bloomfield 1917: 146): **full words** and **function words** (or particles). Some function

words mark morphosyntactic slots, then usually being proclitics. Others are second position clitics (these are not discussed here; see Schachter & Otanes 1972: 187–193, 433–435 and Kroeger 1993: 118–123, 152–154). Full words may be extensively affixed and occur in the limited set of morphosyntactic slots provided for by the function words. There are five morphosyntactic slots for full words in a Tagalog clause, four of which are illustrated by the following example:

- (1) *i-ni-abót ng*  
 CV-RLS(UGR)-within.reach GEN  
*màng-ga-gamot sa sundalo ang itlób*,  
 AV-RDP<sub>2</sub>-medicine LOC soldier SPEC egg  
*at ang pare at siyá ay nág-hintáy*  
 and SPEC priest and 3.SG PM RLS.AV-wait  
*ng sà-sabih-in ng sundalo*.  
 GEN RDP<sub>1</sub>-say-PV GEN soldier  
 ‘The physician handed the egg to the soldier, and the priest and he waited for what the soldier would say.’

Except for the first word, all full words in this clause are preceded by a function word. The clause initial position is the predicate position, which is unmarked unless it is preceded by a topical constituent. In this case, the predicate is marked by the predicate marker *ay* ‘PM’, as in the second part of example (1). The other markers have the following functions:

*sa* is a general locative preposition marking oblique arguments and adjuncts. It is the final constituent of all of the more specific prepositions in Tagalog such as *hanggang sa* ‘until’ or *para sa* ‘for’.

*ang* is often called a topic (or a subject) marker in the literature. The notion of subject in Tagalog is highly controversial (cf. Schachter 1976; 1995; Drossard 1984: 73–78; Foley & van Valin 1984: 134–148; DeWolf 1988: 144–150; Kroeger 1993; Naylor 1995). The fact that *ang* by itself marks neither topics nor subjects is evident from clauses containing two *ang*-phrases:

- (2) *ang mga buhók lamang ang*  
 SPEC PL hair only SPEC  
*p<in>ù-putol ng patálím*  
 <RLS(UGR)>RDP<sub>1</sub>-cut GEN blade  
 ‘only the hair was cut by the blade’

In this example, the first *ang*-phrase functions as the predicate, the second as the predication base or subject (subject<sub>2</sub> in the terminology of Matthews 1981: 104–113). Subjects in this sense have to be configurationally

defined in Tagalog as the *ang*-phrase which occurs either after the predicate or before the predicate marker *ay*. But this does not yet clarify the function of *ang*.

Basically, the function of *ang* is similar to that of an article (the standard analysis in traditional Tagalog grammars, cf. Blake 1925: 205 f.). It is, however, not a definite article, but includes all kinds of referentially specific expressions (definite, specific-indefinite, generic; for details, see Adams & Mandler-Ramer 1988 and Himmelmann 1991: 8–16), and is therefore glossed as SPEC here.

*ng* [naŋ] ‘GEN’ marks genitive attributes. In the literature it is common to differentiate between *ng* marking non-topic agents, direct objects, instruments, manner, etc., but there is little empirical support for such distinctions (cf. Naylor 1980: 37–42).

The fifth morphosyntactic slot, not illustrated by examples (1) and (2), is constituted by the linker (or ligature) *na* (-*ng* after vowels, /n/ or glottal stop). This particle links the elements of a modifying construction such as *ulól na unggó* ‘foolish CONN monkey’, but also occurs in compounds such as *puno-ng-saging* ‘tree-CONN-banana’ and in complement clauses (see Gonzales 1971). The order in modifying constructions is not fixed in Tagalog, thus *unggó-*ng* ulól* is equally possible for ‘foolish monkey’. The difference between the genitive marker *ng* and the linker pertains to referentiality as shown by the following ‘minimal pair’:

- (3) (a) *bata-*ng* dalaga*  
 child-CONN young.woman  
 ‘girl’  
 (b) *bata’ *ng* dalaga*  
 child GEN young.woman  
 ‘child of the young woman’

Full words are not formally differentiated with respect to the five morphosyntactic positions just illustrated. Any full word, regardless of its affixation, may occur in any position (provided its meaning is appropriate). That is, with regard to their syntactic distribution the large class of full words cannot be further divided into classes such as nouns, verbs, and adjectives (cf. Lemaréchal 1982; 1989; Himmelmann 1991; to appear; Gil 1993; Shkarban 1995; Naylor 1995; Art. 72). For example, the words *sásabihin* in (1) and *piníputol* in (2) which are inflected for aspect and mood (see 4.2) and which are often called verbs, occur – without any further

derivation – in a *ng-* and an *ang-*phrase, respectively. Similarly, so-called nouns may occur underived in predicate position, e.g.

- (4) *Isdá' ang pág-kain niyá.*  
 fish SPEC GER-eat 3.SG.POSS  
 'His meal was fish.'

A brief look at the list of affixes in Schachter & Otanes for what they call 'nouns' (1972: 97–106), 'adjectives' (1972: 198 f.; 216–229), and 'verbs' (1972: 344–355) immediately reveals that a subclassification of full word bases on the basis of morphological evidence is also not a straightforward enterprise since basically the same set of affixes is involved in all these formations. Nevertheless, bases appear to differ in their morphological potential. A detailed study of the different morphological classes, however, still remains to be done (see Wolff 1993 and Himmelmann to appear for some preliminary suggestions).

There is no doubt that both unaffixed bases and affixed words differ in their semantics in that some denote actions, others things, yet others states, etc. A base such as *bili* clearly designates the action of buying, *bahay* the entity 'house', *bago* the quality 'new'. The distinction between these different kinds of concepts is sometimes grammatically marked by different stress patterns (cf. 3.4). It is not clear, however, whether these conceptual classes are in any way directly relevant to stating the (segmental) morphosyntactic regularities of Tagalog. Therefore, given the absence of clear-cut formal evidence for such categories, the terms for lexical categories such as noun, verb, adjective, etc. are avoided in this article and the terms action, entity, etc. are used whenever it is convenient to make reference to conceptual classes.

### 3. Formal processes

#### 3.1. Affixation

The number of affixes in Tagalog is fairly small, but each of the major three affix categories is represented: there are two suffixes (*-in* and *-an*, see 4.1), two infixes (*-in-* and *-um-*, see 4.1 and 4.2), and about a dozen prefixes. Most of these affixes can be combined with each other so that a large number of complex formations results (Bloomfield 1917: 317–319 lists some 200 formations). This number is further increased by the fact

that affixes may co-occur with other formal processes such as reduplication (see 3.3) and stress shifting (see 3.4).

**Prefixes** exhibit some word-like properties in that (with few exceptions) they (a) do not display any fusional characteristics (not even resyllabification takes place, which, however, is not surprising given the fact that all bases start with a consonant (cf. 1.2)), (b) mostly carry their own stress, and (c) may occasionally be used in isolation (cf. Bloomfield 1917: 213; Matsuda French 1988: 89 f.; Rubino 1998). Exceptions are, on the one hand, the prefixes *i-* (4.1) and *ka-* (5.1), the first of which is never stressed and both of which are sometimes fused with the base, e.g. *ka-ibigan* 'ass-like-LV (friend)' is usually pronounced [kaybigan] rather than [ka?ibigan] (Bloomfield 1917: 139 f.).

The other exception is the prefix *pang-* (and related *mang-* and *nang-*), where the prefix-final nasal regularly assimilates to, and sometimes substitutes for, the initial consonant of the base. This prefix-type is very common in western Austronesian languages and is often analysed as containing an archisegmental nasal (and then is represented as /paN-/). The assimilation and substitution regularities may be summarised as follows (for the semantics of this derivation, see 5.2):

- N is regularly deleted before base-initial nasals: *maN* + *manhid* → *mamanhid* 'get numb'.
- N never substitutes for a glottal fricative or for glides: *maN* + *hiyá'* → *mànghiyá'* 'humiliate', *maN* + *walís* → *màngwalís* 'hit with a broom', *maN* + *yari* → *màngyari* 'happen'.
- Voiceless obstruents (apart from the glottal fricative) are regularly substituted for: *maN* + *pili'* → *màmili'* 'choose (several things)', *maN* + *takot* → *mànakot* 'frighten several people', *maN* + *kabayo* → *màngabayó* 'ride on horseback', *maN* + *sakít* → *mànakít* 'cause pain'. Glottal stop is sometimes not substituted for; the factors involved are unclear. In fact, two derivations are occasionally possible from the same base. Compare [*mànjaso*] 'hunt with dogs' with [*mâj?aso*] 'ridicule people' (base *aso* 'dog').
- Voiced obstruents and the lateral are generally not substituted for: *maN* + *daya'* → *màndaya'* 'cheat people' (but *maN* + *dikit* → *mànikít* 'adhere'), *maN* + *gamót* → *mànggamót* 'practice medicine (profes-

sionally)’ *maN + laró*’ → *mànlaró*’ ‘amuse others’. For /b/ both options are common and, in a few cases, possible for the same base. Thus from *bakyá*’ ‘wooden shoe’ it is possible to derive either *màmbakyá*’ ‘hit with a wooden shoe’ or *màmakyá*’ ‘wear wooden shoes’.

In the derivation of so-called instrumental adjectives substitution is said to be optional, e.g. both *pansuklái* and *panuklái* mean ‘for use in combing’ (Schachter & Otanes 1972: 218–221; 319 f.).

**Suffixes** exhibit regular fusional tendencies in that the stem-final consonant (including the optional [h] after vowel-final bases) is regularly resyllabified with the suffix and the base-final vowel is often syncopated, e.g. *lakás + an* → *lak.sán* ‘strengthen (x)’, *bili + in* → *bil.hín* ‘buy (x)’ (cf. Matsuda French 1988: 12 f.). There are further sporadic irregularities such as *tawan-an* ‘laugh at (x)’ instead of the expected \**tawah-an*, and *tamn-án* ‘plant in (x)’ instead of the expected ?*tanim-án* or \**tanm-án* (cf. Bloomfield 1917: 214; Schachter & Otanes 1972: 375–380).

**Infixes** are always inserted after the first consonant of the base, e.g. *b<um>ili*’ '<AV> buy' from *bill*. All other formative processes precede infixation, infixes thus also occur in prefixes, e.g. *p<in>ág-tapat-án*’ '<RLS(UGR)>GER-fronting-LV (confided to (x))', and reduplicated syllables (cf. 3.3). In formations involving the infix -in- and stem-initial glides or liquids, the stem-initial consonant and the nasal of the infix are regularly metathesized, thus \**l-in-uto*’ but *ni-luto*’ ‘RLS(UGR)-cook’. With stem-initial /w/ metathesis is optional, for example, both *ni-wisik-án* and *w<in>isik-án* for ‘RLS(UGR)-sprinkle-LV/<RLS(UGR)>sprinkle-LV (sprinkle on (x))’. When co-occurring with the prefix *i-* (i.e. [?i]), -in- is infixated into the following syllable, e.g. *i-b<in>igáy*’ ‘CV-<RLS(UGR)> give’ from *bigáy*. In stems beginning with a liquid, a glide or a glottal, metathesis again regularly occurs, compare *i-ni-hatíd*’ ‘CV-RLS(UGR)-convey’ (\**ihihatid*), *i-ni-abót* [?i-ni-?abót]’ ‘CV-RLS (UGR)-within.reach’ (\**i-in-abot*).

### 3.2. Consonant alternation

A number of prefixes display regular alternation of the initial consonant (cf. de Guzman 1978: chapter 3.3): the /p/-initial form is the basic form (used as gerund (4.3) or imperative), /m/ marks non-realism and /n/ realism (cf.

4.2). Examples are *pag-/mag-/nag-*, *paN-/maN-/naN-*, *paki-/maki-/naki-*. The alternation between the two nasal forms also occurs with the stative or potentiative prefix (*ma-/na-*, cf. 5.1), but here no /pl/-initial basic form exists. Furthermore, a small number of /p/-initial bases (which probably contain a fossilised prefix) exhibits this alternation, e.g. *pakiníg makiníg nakiníg* ‘listen’, *panoód manoód na-noód* ‘watch’. This alternation probably developed from infixed formations by clipping the first (unstressed) syllable. Thus, *mag-* probably derives from \**pumag-* and *nag-* from \**pinag-*. Note, however, that *pinag-* is still a productive formative, i.e. the RLS (UGR) form of *pag*-derived stems (cf. 5.2).

### 3.3. Reduplication

There are three kinds of **reduplication** processes in Tagalog. In two of these processes the first consonant and vowel of the base are copied (CV-reduplication). The two processes differ with regard to the fact that in one the reduplicated syllable is assigned stress (RDP<sub>1</sub>), and in the other it is not (RDP<sub>2</sub>): *mang-gàgamót* ‘will practice medicine’ vs. *màng-ga-gamot* (or *mànggagamót*) ‘one who makes cures, physician’. If a base starts with a consonant cluster, only the first consonant is copied, cf. *trabaho* ‘work’ → *mag-tà-trabaho* ‘will work’.

The third process consists in copying the first two syllables of the base (RDP<sub>3</sub>). In the case of disyllabic bases, the complete base is copied, e.g. *lakad* ‘walk’ → *mag-lakàd-lakád* (RDP<sub>3</sub> with stress shift) ‘do a little walking’, *mura* ‘cheap’ → *mura-mura* ‘rather cheap’. In the case of bases containing more than two syllables, the second syllable is copied only up to its peak, which is then assigned secondary stress, viz. *tahimik* ‘peaceful’ → *tahi-tahimik* ‘rather peaceful’, *baluktód* ‘crooked’ → *balù-baluktód* ‘variously bent’. RDP<sub>2</sub> and RDP<sub>3</sub> generally apply to bases only, while RDP<sub>1</sub> freely applies to prefixes as well (see below).

Each type of reduplication may occur only once in a derivation but different types may be combined with each other, as will be seen shortly.

Tagalog reduplication phenomena have figured prominently in the discussion of the status of reduplication (cf. Art. 57) as well as the related issues concerning the boundaries between morphology and phonology and the internal structure of the morphological component (cf. Marantz 1982: 438 f., 451 f., 473–479; Carrier-Duncan 1984; Matsuda French

1988: 19–61). The argument is concerned with the interrelation of (regular) affixation and reduplication. The Tagalog evidence, however, seems inconclusive in this regard. The following is a (partial) list of the relevant phenomena:

- Formations involving nasal substitution (cf. 3.1) suggest that allomorphy rules precede reduplication, e.g. *maN + RDP<sub>1</sub> + putol* → *mamūmutol* ‘will cut (a lot)’.
- Infixation, on the other hand, seems to follow reduplication (cf. Matsuda French 1988: 38), i.e., *pumūputol* ‘cut’ seems to be derived from *pūputol*, otherwise one would expect \**pūpumutol*.
- In formations involving polysyllabic prefixes or a combination of prefixes, it is usually the second syllable of the prefix that is reduplicated (by RDP<sub>1</sub>, receiving word level stress, cf. Matsuda French 1988: 44–52), viz. *mag + pa + putol* → *magpāpaputol* ‘will cause to be cut’, *maka + putol* → *makākaputol* ‘will be able to cut’, or *maka + pag + pa + putol* → *makākapāgaputol* ‘will be able to cause to be cut’. A general exception to this regularity are formations involving the prefix *i-*, which is never reduplicated and does not count in determining the second syllable, cf. *ma + i + pag + luto*’ → *maipāpagluto*’ ‘will be able to cook for (x)’ and *i + paki + pa + putol* → *ipakikipaputol* ‘will be asked to be caused to be cut’. In several instances, however, either the second syllable of the prefix or the first syllable of the base may be reduplicated, an example being *makapūputol* ‘will be able to cut’ which alternates with the form *makākaputol* just mentioned (for more examples, cf. chart in Schachter & Otanes 1972: 369). This raises the problem of where to place RDP<sub>1</sub>-reduplication in the derivation of these complex formations.
- The relation of suffixation and RDP<sub>3</sub> is also problematic. In some cases suffixation (and the related morphonological processes) clearly apply before RDP<sub>3</sub>: *tingín + an* → *tingnán* ‘look at (x)’ → *ting-nán-tingnán* ‘look at (x) a little’. In other cases, it is the other way around: *sakit* → *sakit-sakit* → *màgsakit-sakitan* ‘pretend to be sick’. If suffixation preceded RDP<sub>3</sub>, \**magsakisakitan* would be expected (trisyllabic base), cf. *mag + salitá’ + an* → *màgsalí-salitaan* ‘talk a little to each other’.
- RDP<sub>1</sub> and RDP<sub>3</sub> may co-occur. In this case either RDP<sub>3</sub> precedes RDP<sub>1</sub>, e.g. *magsà-sakit-sakitan* ‘will pretend to be sick’ and *magsàsalí-salitaan* ‘will talk a little to each other’, or they may apply simultaneously (at different locations), viz. *mag + pa + ka + ingat* → *magpápakaingatingat* ‘will be extremely careful’. Carrier-Duncan (1984: 269) claims that there are also cases where RDP<sub>2</sub> precedes RDP<sub>3</sub>: *mag + kaN + punit* → *magkámpupunit* ‘will tear spontaneously’ → *magkámpupù-pupunit* ‘will (intensively) tear spontaneously’.
- RDP<sub>1</sub> and RDP<sub>2</sub> may co-occur as well, in which case RDP<sub>2</sub> precedes RDP<sub>1</sub>. Thus from *takbó* ‘run’ *mág-ta-takbó* ‘run wild’ is derived by prefixing *mag-* and RDP<sub>2</sub>. From the latter the imperfective aspect *mag-tà-ta-takbó* is derived by RDP<sub>1</sub>.

### 3.4. Stress shift

Both primary and secondary **stress assignment** in Tagalog may be connected with a variation in meaning and this assignment is at least partially independent of segmental processes (cf., for example, the difference between RDP<sub>1</sub> and RDP<sub>2</sub> mentioned in 3.3). The data, however, are not clear, since most sources (apart from Bloomfield 1917 and Wolff et al. 1991) do not mark stress consistently (cf. 1.2). Two examples are given to illustrate the possibilities and complexities involved.

There is a substantial number of bases which differ only with respect to stress and which are clearly semantically related (unlike the pair *bukas* ‘tomorrow’/*bukás* ‘open’ mentioned in 1.2). Examples include *abot* ‘overtake’ vs. *abót* ‘within reach, reach for, pass’, *alam* ‘knowledge’ vs. *alám* ‘known’, *buhay* ‘live, life’ vs. *buháy* ‘alive’, *bunot* ‘pull out’ vs. *bunót* ‘pull out a lot/repeatedly’, *isip* ‘think, thought’ vs. *isíp* ‘think hard/with deliberation’, *lakad* ‘walk’ vs. *lakád* ‘on foot, barefooted’, *tulog* ‘sleep’ vs. *tulóg* ‘asleep’ (cf. Bloomfield 1917: 215 f.). Wolfenden (1961: 12) characterises the meaning shifts involved as (a) accomplishment (resultative) or (b) intensification. While resultative pairs are widely attested, it is unclear whether the much more sporadic pairs not belonging to this type can all be subsumed under ‘intensification’ as the handful examples just given should make clear.

The complex interaction of stress assignment and affixation is illustrated by the suffix

*-an* (cf. Bloomfield 1917: 250–262). If this suffix marks locative voice (cf. 4.1), primary stress usually shifts one syllable to the right (i.e. to the ultimate or penultimate syllable of the derived word). Examples are *táwag* ‘call’ → *tawágán* ‘call (x)’, *gupít* ‘cut (with scissors)’ → *gupítán* ‘cut the hair of (x)’, *bílì* ‘buy’ → *bilhán* ‘buy from (x)’, *sáma* ‘go along, accompany’ → *samáhan* ‘accompany (x)’. If *-an* derives expressions denoting either a collective action or the place where something (an entity or an action) is located, stress in oxytone bases remains on the same syllable as in the underived word: *iyák* ‘cry, weep’ → *iyákan* ‘a crying of many’, ‘buy’ → *bilhán* ‘place where to buy, market’, *aklát* ‘book’ → *aklátan* ‘library’, *litsón* ‘a roast pig’ → *litsúnan* ‘place to roast pigs or a barbecue (= roast-pig party)’. For barytone bases there are two possibilities: either stress shifts one syllable to the right (to the penultimate syllable of the derived word), which is often accompanied by secondary stress on the first syllable of the derived word, e.g. *sámáhan* ‘a going together of many, company’, or it is shifted to the ultimate syllable of the derived word (i.e., the suffix is stressed), e.g. *básá* ‘read’ → *basahán* ‘a reading-room, library’ (vs. *basáhan* ‘read sth. to (x)’), *lábán* ‘contrary, fight’ → *labanán* ‘a fighting of many, battle, war’ (vs. *labánan* ‘fight/oppose (x)’). Stress shift to the penultimate syllable of the derived word is used (with very few exceptions) for collective action expressions, while stress shift to the suffix is more commonly (though by no means exclusively) used for ‘place where’ derivations. Note that these are only regularities; differences in meaning are not always accompanied by formal differences, e.g. *dílo* ‘end’ → *dulíhan* ‘terminate (x)’ (locative voice), but also ‘end part, back yard’ (‘place where’; cf. Bloomfield (1917: 261) who lists a number of words which formally appear to be locative voice derivations, but the meaning of which does not fit this categorisation).

#### 4. Voice, aspect, and mood

Tagalog and the other Philippine languages are most famous for a phenomenon variously called voice, orientation, case marking on the verb (cf. Blake 1906; Ramos 1974), or ‘focus’, a term introduced in the late 1950s to underscore the exceptional nature of the phenomenon (cf. Llamzon 1973: 168). The last term is widely used to refer to the pragmatic

phenomenon of highlighting new or contrastive information. ‘Focus’-affixes in Philippine languages do not have such a highlighting function. The participant ‘focussed’ on by these affixes is usually given information and often remains unexpressed. Therefore, this term is avoided here. Instead, **voice** is chosen because it is the least misleading term (see 4.1).

The literature on voice in Tagalog is fairly extensive (see – in addition to the general reference works mentioned in 1.1 – Müller 1882: 136–142; Marre 1901: 574–582; Blake 1906; Scheerer 1924; Capell 1964; Llamzon 1973; 1976: 89; Wolfenden 1961: 14–16; Ramos 1971: 21–23, 56–69; 1974: 19–40; Foley 1976: 105–113; McFarland 1976: 16–24; de Guzman 1978: chapter 3; Drossard 1984: 34–51; Himmelmann 1987: 92–125; DeWolf 1988; Shibatani 1988). Since voice marking is formally tied to aspectual and modal distinctions, these three categories will be treated together in one section.

##### 4.1. Voice

Tagalog predicate expressions usually display a voice affix that indicates the semantic rôle of one of the participants involved in the state of affairs denoted by the predicate. There are four such affixes, as illustrated by the following examples:

- (5) *t<um>angó’ ang unggó’*  
     <AV>nod SPEC monkey  
     ‘the monkey nodded in assent’
- (6) *dikdik-in siyá sa lusóng*  
     crush-PV 3.SG LOC mortar  
     ‘(that) he (i.e., the turtle) be crushed in a mortar’
- (7) *hulug-an mo akó!*  
     drop-LV 2.SG.POSS 1.SG  
     ‘drop me (some)!’
- (8) *kung i-tà-tanim niyá ang*  
     if CV-RDP<sub>1</sub>-plant 3.SG.POSS SPEC  
     *kaniyá-ng ka-parté*  
     DAT.3.SG-CONN ASS-part  
     ‘if he would plant his part (for him)’

As briefly shown in 2, the NP-markers *ang* and *ng* in Tagalog do not signal semantic rôles. Rather, the voice affixes indicate the semantic rôle of the participant which appears in the *ang*-phrase (*siyá* in (6) and *akó* in (7) are *ang*-forms of the pronoun). Thus, the infix *-um-* in (5) indicates that it is the monkey who does the nodding, and in (6) the suffix

*-in* indicates that the turtle is going to be the undergoer of the crushing (rather than the actor), etc.

Before discussing some of the more remarkable features of this voice marking system in more detail, it should be noted that the **actor voice** marking infix *-um-* does not only occur in clauses with a subject which is in full control of an action. It is also used for subjects which are involved in a process, as in *p<um>ulá* ‘become red’ or *l<um>ù-lutang* ‘be floating’. Furthermore, it occurs in subject-less expressions for natural events such as *um-ulán* ‘rain’ or *l<um>indól* ‘earthquake’.

In addition to *-um-*, the prefix *mag-* (realis *nag-*) also marks actor voice, cf. *nag-là-laró* ‘*silá* ‘RLS.AV-RDP<sub>1</sub>-play 3.PL (they are playing)’. Following de Guzman (1978: chapter 3), this prefix is analysed here as involving the prefix *pag-* used in gerund formation (see 4.3), actor voice (and mood) being signalled by consonant alternation (cf. 3.2). The difference between the two actor voice affixes is further commented upon in 5.2.

One of the remarkable features of voice marking in Tagalog, which sets it apart from voice marking systems in many other languages, is the fact that both actor as well as **undergoer voices** involve overt morphological marking, while in languages such as English only undergoer orientation (passive) is explicitly marked. In other words, actor voice and undergoer voice are equally marked in Tagalog (at least in morphological terms).

Another peculiarity is the fact that there is not only one affix for undergoer orientation. Instead, three different ways in which the undergoer may be involved in a given state of affairs are distinguished:

- *-in* (**patient voice**) indicates a directly affected undergoer, such as the turtle in (6), the hair in (2), or *ító* in *inum-in mo ító* ‘drink-PV 2.SG.POSS PROX (drink this)’.
- *-an* (**locative voice**) is used for recipients (see (7)), addressees, beneficiaries, and the location where an action takes place, e.g.:

(9)	<i>ni-lakar-an</i>	<i>ko</i>	<i>ang</i>
	RLS(UGR)-walk-LV	1.SG.POSS	SPEC
	<i>ma-bató-ng</i>	<i>kalye</i>	
	STAT-stone-CONN	street	

‘I walked on a stony road’

More generally, it is used for indirect undergoers, i.e. undergoers which are not

directly affected by the action denoted by the predicate, as in *inum-án mo ító* ‘drink-LV 2.SG.POSS PROX (drink from/some of this)’ or *buks-án mo ang pintó* ‘open-LV 2.SG.POSS SPEC door (open the door)’.

- *i-* (**conveyance voice**) indicates an undergoer that is moved (a displaced theme), such as the egg in (1) or one half of the banana tree in (8). It is also used for the instrumental rôle, instruments thus being conceived of as moving undergoers:

(10)	<i>Ang iták ay i-p&lt;in&gt;utol</i>		
	SPEC	bolo	PM CV-<RLS(UGR)>cut
	<i>ko</i>	<i>ng</i>	<i>saging</i> .
	1.SG.POSS	GEN	banana

‘I cut bananas with the bolo.’

Furthermore, *i-* may also indicate the beneficiary of an action with a few bases (e.g. *i-bilí* ‘buy for (x)’), a usage not easily accounted for by any of the analyses proposed for this prefix (cf. Himmelmann 1987: 103–22, 139 f.).

A third cross-linguistically remarkable feature of the voice affixes is that they may be applied to all kinds of bases without any further derivation. That is, the above affixes cannot only be attached to bases denoting actions but also to ones denoting things (e.g. *bató* ‘stone’ → *batuh-in* ‘throw stones at (x)’), masses (e.g. *tubig* ‘water’ → *tubig-an* ‘add water to (x)’), states (e.g. *bago* ‘new’ → *baguh-in* ‘change (x)’ or *i-bago* ‘move (x) to another position’), or animate beings (e.g. *langgám* ‘ant’ → *langgam-in* ‘be infested with ants’). Of course, the derivational possibilities depend on the semantic compatibility of base and affix, and thus are more restricted with regard to, for example, expressions for human beings than for action expressions.

It is common to treat voice with respect to action expressions as inflection, and voice with regard to non-action expressions as derivation. There is, however, no formal evidence to support this distinction. On the contrary, the analysis of voice as inflection leads to extremely complex systems of ‘verb’ classes in Tagalog. That there is little clear-cut evidence for such classifications is shown by the fact that the proposed classifications differ extremely. Blake (1925), for example, proposes 17 classes, de Guzman (1978) about 80 (cf. Himmelmann 1987: 69, 129–145). The main empirical observation here is that there is no simple classification for action expressions

with regard to their voice marking. In particular, there are no productive inflectional paradigms for voice, as suggested by the commonly used ‘paradigmatic’ examples in the literature. Instead, derivations from all kinds bases are only partially predictable on the basis of their semantics and exhibit a large number of idiosyncrasies, which again suggests derivation rather than inflection. See McFarland (1976) and Ramos & Bautista (1986) for instructive surveys of those derivations which are actually attested for a given action base (see also Art. 38).

With regard to the three features just mentioned, Tagalog voice marking has much in common with nominalising morphology in other languages. Like much of the morphology used for deriving nouns (or, in many languages, participles) from verbs, the voice affixes change the orientation of a given base in such a way that it may be used to refer to one of the participants involved in the state of affairs denoted by the base (cf. Lehmann 1984: 151 f., who introduces the term **orientation** for analysing nominalisation strategies). In this view, *-um-* is an actor orienting infix which derives from a base such as *tangó* ‘nod, nodding in assent’ a word *tumangó* which could be glossed as ‘one who nods, nodder’. This expression no longer directly denotes the action of nodding, but rather the participant who nods. That is, in the Tagalog clause (5) *tumangó ang unggó* both *tumangó* and *uggó* refer to the same entity. Imitating the equational structure of this clause it could be rendered as ‘nodd-er in assent (was) the monkey’. Similarly, example (6) is ‘he (be) crush-ee in the mortar’, example (7) ‘I (be) the place of your dropping/your droppery’, and example (8) ‘if his plant-ee (would be) his part’ (cf. DeWolf 1988: 157 f.). Note, however, that Tagalog voice affixes are not nominalising in a morphosyntactic sense, since they do not change the syntactic category of the base (both base and derived word are full words which can be used in any of the five major morphosyntactic slots discussed in 2). That is, the similarity of Tagalog voice affixes and noun or participle-forming morphology is primarily a semantic one.

With regard to this semantic similarity, it should be noted that the voice marking formatives do not only occur in expressions which denote actions, processes or states (or, more precisely, a participant involved therein). Most of the voice marking affixes also occur in formations which clearly denote

entities. The major possibilities pertaining to locative *-an* are mentioned in 3.4. Actor voice marking *mag-* (plus RDP<sub>2</sub>) occurs in formations denoting professionals, thus from *nakaw* ‘steal’ *mág-na-nakaw* ‘thief’ may be derived. This form differs only with respect to stress from the action denoting formation *mag-ná-nakaw* ‘will steal’ (cf. 3.3 and Bloomfield 1917: 242 f.; Schachter & Otanes 1972: 103). Furthermore, *mag-* is used together with kinship terms to indicate two (with RDP<sub>2</sub> several) persons between which the relation designated by the base holds: *mág-iná* ‘mother and child’ (< *iná* ‘mother’), *mág-pípinsan* ‘several cousins’ (cf. Bloomfield 1917: 242; Schachter & Otanes 1972: 102). The suffix *-in* may denote entities undergoing the action denoted by the base, e.g. *aral* ‘study’ → *aralin* ‘lesson’. Again, it is stress that (often) differentiates action and thing, compare *aralin* with *aralin* ‘study (x)’. Other examples are *kumpuni* ‘repair’ → *kùmpunihin* ‘things to repair’ vs. *kumpunihin* ‘repair (x)’; *kain* ‘eat’ → *kanin* ‘boiled rice’ or ‘eat (x)’ (no difference whatsoever, compare also *kakanín* ‘sweets’); *inóm* ‘drink’ → *inumín* ‘drinking water’ or ‘drink (x)’, cf. also *inumin* ‘beverage’ (cf. Bloomfield 1917: 247; Schachter & Otanes 1972: 99 f.). There are no derivations with *-um-* or *i-* which denote entities in a similar way.

#### 4.2. Aspect and mood

Voice marked forms occur in two aspects (perfective and imperfective) and two moods (realis and non-realis). **Imperfective** aspect is indicated by RDP<sub>1</sub> (**perfective** aspect is unmarked), **realis** mood is indicated by the infix *-in-* or consonant alternation (/m/ → /n/, cf. 3.2), **non-realis** again being unmarked. These formations are illustrated in Tab. 136.1 with the paradigms for the base *bili* ‘buy’, one of the few bases which occurs with all voice affixes. A number of different analyses and terminologies have been proposed for these forms (see Werlen in Bader et al. 1994: 95–100 and Kroeger 1993: 15–18 for overviews). The terminology used here reflects the formal make-up of the paradigms but it is not quite felicitous in all instances with regard to the function of the forms.

Aspect-mood formation is highly regular both formally and semantically and thus clearly is an instance of inflection. It exists for every voice-marked form and is also found for other affix combinations, e.g. *maki-*, *makiki-*, *naki-*, *nakiki-* (see 5.2). With

	AV ('buy')	AV ('sell')	PV	LV	CV
NON.RLS/PFV	<i>b&lt;um&gt;ilí</i>	<i>màg-bilí</i>	<i>bilh-ín</i>	<i>bilh-án</i>	<i>i-bilí</i>
NON.RLS/IPFV	<i>bilí</i>	<i>mag-bilí</i>	<i>bìlh-ín</i>	<i>bìlh-án</i>	<i>i-bilí</i>
RLS/PFV	<i>b&lt;um&gt;ilí</i>	<i>nàg-bilí</i>	<i>b-in-ilí</i>	<i>b-in-ilh-án</i>	<i>i-b-in-ilí</i>
RLS/IPFV	<i>b&lt;um&gt;ibilí</i>	<i>nag-bilí</i>	<i>b-in-ibilí</i>	<i>b-in-ibilh-án</i>	<i>i-b-in-ibilí</i>

Tab. 136.1: Aspect-mood paradigms for *bilí* 'buy'

a few exceptions not dealt with here, the contexts of use for each form are the same regardless of the meaning of the base. Thus, NON.RLS/PFV (also called *basic form*; cf. Schachter & Otanes 1972: 66 f.) is used in hypothetical and complement clauses and in commands (cf. (6) and (7)), NON.RLS/IPFV (Schachter & Otanes' *contemplated aspect*) is used for future events (cf. (8)), RLS/PFV (Schachter & Otanes' *perfective aspect*) and RLS/IPFV (Schachter & Otanes' *imperfective aspect*) for past and present events, respectively (see illustrative text in 6).

In addition, there is a recent perfective formation (with prefix *ka-* + RDP<sub>1</sub>) which involves no voice marking and does not allow for subjects (*ang*-phrases). Hence, all participant expressions are either genitive or locative marked:

- (11) *Ka-là-laró ko pa*  
RECENT.PFV-RDP<sub>1</sub>-play 1.SG.POSS still  
*lamang sa bata*.  
only LOC child  
'I have just finished playing with the child.'

As may be immediately observable there are some asymmetries in the paradigms in Tab. 136.1. Strictly speaking, the realis patient voice forms do not have a marker for voice and, similarly, there is also no marker for actor voice in the NON.RLS/IPFV form of the *um*-paradigm. These asymmetries are probably relevant for unravelling the diachronic development of the voice marking system. Their relevance for the synchronic analysis of the system is still in need of further exploration (see Himmelmann 1987: 157–171; Blake 1988: 79 f. for some discussion). In this regard it may be noted that although most action expressions in Tagalog are voice and hence also aspect-mood marked, it is possible to use bases denoting actions without further affixation. In such uses action bases may be semantically undergoer oriented (often with a resultative connotation). For example, in (11) *an-*

*táy* could be replaced with the patient voice form *inantáy*.

- (12) *Antáy ko ang sagót mo.*  
wait 1.SG.POSS SPEC answer 2.SG.POSS  
'I wait for/expect your answer.'

But unaffixed bases may also occur in imperatives with actor orientation, e.g. *hintáy ka* 'you wait' (which could also be rendered with *mag-hintáy ka*), and for denoting a state of affairs without orientation:

- (13) *Iyák ang sagót niyá sa akin.*  
cry SPEC answer 3.SG.POSS LOC DAT.1.SG  
'His answer to me was crying/to cry.'

In this last usage, unaffixed bases are similar to gerunds, to which we now turn.

#### 4.3. Gerunds

For bases denoting a state of affairs it is possible to derive a form which is neither voice marked nor aspectually and modally inflected. This is done by prefixing *pag-* to the base according to the following correspondence rules which hold between actor voice and gerundial forms (cf. Schachter & Otanes 1972: 160):

ACTOR VOICE	GERUND
<i>-um-</i>	<i>pag-</i>
<i>mag-</i>	<i>pag-RDP<sub>2</sub></i>
<i>mang-</i>	<i>paN-RDP<sub>2</sub></i>

Tab. 136.2: Gerund formation

**Gerunds** are further derivable from stative expressions (cf. 5.1, prefix *ma-* is substituted by *pagka-*) and joint action expressions (cf. 5.2, prefix *maki-* replaced by *pakiki-*). In general, gerunds may not be used in predicate position, since they are not oriented towards one of the participants of the state of affairs denoted. Only in a clause such as *pág-lu-luto ng pagkain ang trabaho niyá* 'GER-RDP<sub>2</sub>-cook GEN food SPEC work 3.SG.POSS (his job is

cooking food)' may a gerund be used predicatively (*pagkain* – which may also mean ‘eating’ – is a semantically specialised gerund from *kain* ‘eat’). Gerunds are most commonly used in noun phrases, e.g. *nág-umpisáng pág-si-sigáw* ‘RLS.AV-begin GEN GER-RDP<sub>2</sub>-shout ((the turtle) began shouting/to shout)’ and *nàng-galing sa pág-su-sugál niyá* ‘RLS.AV-come.from LOC GER-RDP<sub>2</sub>-gamble 3.SG.POSS ((this) is due to his gambling)’, or in subordinate clauses (usually with a temporal meaning) *pág-balik ni Gabby sa Pilipinas* ... ‘GER-return GEN.PN Gabby LOC Philippines (when Gabby returned to the Philippines ...)’. As shown by the preceding examples, all participants involved in the state of affairs denoted by a gerund have to be expressed in genitive or locative phrases.

A special perfective form of the gerund indicates that the event took place before that of the main clause. It involves the prefix *ka-* (which may be optionally reduplicated) following the general gerund formatives mentioned above (cf. Schachter & Otanes 1972: 161). Compare *pág-punta* ‘going’ with *pág-(ka)-ka-punta* ‘having gone’. Gerunds may become the basis for further derivations involving voice, aspect, and mood (see 5.2).

Formations with *paN-* without reduplication regularly denote instruments used in the state of affairs denoted by the base, e.g. *pámutol* ‘cutting instrument’ (< *putol* ‘cut’), *pànghampás* ‘a whip’ (< *hampás* ‘whip’), *pàngkapé* ‘means for buying coffee’, etc. (cf. Bloomfield 1917: 224 f.).

## 5. Actor involvement

Although much less widely discussed, Tagalog morphology concerning the way an actor is involved in a given state of affairs is even more elaborate than the voice morphology. The basic split here is that between dynamic and stative or potentive formations (for statives see also Drossard 1984: 64–72). The **dynamic** forms are morphologically unmarked

and have been dealt with in 4. That is, an action expression marked for voice (and aspect and mood) generally implies a volitional actor who is in full control of the action (the major exception being some expressions for processes marked with *-um-* briefly mentioned in 4.1).

### 5.1. Stative and potutive

States of affairs which do not involve a controlling actor are expressed by a related but different set of formations. Two different scenarios have to be distinguished here. On the one hand, the state of affairs may be such that it excludes the involvement of an actor for principled conceptual reasons. This is typically the case for **stative** expressions such as ‘be hungry’, ‘be angry’, ‘be adrift’, and the like. On the other hand, the state of affairs may be such that in principle it allows for controlling actors but in the specific instance at hand the conceptually possible controlling actor is not in full control of the event. This is the case when someone happens to do something without having the intention to do it. Formations expressing this second possibility are called **potutive**.

A completely regular correspondence exists between dynamic and potutive formations. That is, for each dynamic form there is a corresponding potutive form. The major potutive formatives are *maka-* for actor voice and *ma-* for the undergoer voices. For details compare Tab. 136.1 with Tab. 136.3.

The typical use of potutive forms is for involuntary actions:

- (14) *Biglá niyá-ng*  
 sudden 3.SG.POSS-CONN  
*nà-bigkás iyón:*  
 RLS.POT.PV-enunciation DIST  
 ‘(Terrified) she suddenly exclaimed  
 this: ...’

This includes actions done accidentally, i.e. the actor may be in control of the action but did not really intend its outcome:

	AV(‘buy’)	AV(‘sell’)	PV	LV	CV
NON.RLS/PFV	<i>maka-bilí</i>	<i>maka-pagbili</i>	<i>ma-bilí</i>	<i>ma-bilh-án</i>	<i>ma-i-bilí</i>
NON.RLS/IPFV	<i>maka-bibilí</i>	<i>maka-pagbibili</i>	<i>ma-bibili</i>	<i>ma-bibilh-án</i>	<i>ma-i-bibilí</i>
RLS/PFV	<i>naka-bilí</i>	<i>naka-pagbili</i>	<i>na-bilí</i>	<i>na-bilh-án</i>	<i>na-i-bilí</i>
RLS/IPFV	<i>naka-bibilí</i>	<i>naka-pagbibili</i>	<i>na-bibili</i>	<i>na-bibilh-án</i>	<i>na-i-bibilí</i>

Tab. 136.3: Potutive aspect/mood paradigms for *bilí* ‘purchase, sale’

	STAT	STAT.LV	STAT.CV	STAT.AV
NON.RLS/PFV	<i>ma-galit</i>	<i>ka-galit-an</i>	<i>i-ka-galit</i>	<i>maka-galit</i>
NON.RLS/IPFV	<i>ma-gàgalit</i>	<i>ka-gàgalit-an</i>	<i>i-ka-gàgalit</i>	<i>maka-gàgalit</i>
RLS/PFV	<i>na-galit</i>	<i>kina-galit-an</i>	<i>i-kina-galit</i>	<i>naka-galit</i>
RLS/IPFV	<i>na-gàgalit</i>	<i>kina-gàgalit-an</i>	<i>i-kina-gàgalit</i>	<i>naka-gàgalit</i>

Tab. 136.4: Voice and aspect/mood paradigms for stative bases (base *galit* ‘anger’)

- (15) *Na-i-luto ko na.*  
 RLS.POT-CV-cooked 1.SG.POSS now  
 ‘I happen to have cooked it already  
 (by mistake).’

It also includes perceptions over which the actor (= experiencer) has no control as in:

- (16) *doón ay nà-kita nilá*  
 DIST.LOC PM RLS.POT.PV-see 3.PL.POSS  
*ang isá-ng ma-lakí-ng higante*  
 SPEC one-CONN STAT-size-CONN giant  
 ‘there they saw a great giant ...’

In a second, somewhat different use potentiative forms express the ability of an actor to perform the action in question:

- (17) *kung inyóng*  
*kung inyó-ng*  
 if 2.PL.DAT-CONN  
*mapagtisán iyán*  
*ma-pag-tiis-an iyán*  
 POT-GER-suffer-LV that  
 ‘if you are able to endure this ...’

- (18) *at hindí makabaríl sa kanyá.*  
 at hindí maka-baríl sa kanyá  
 and NEG POT.AV-gun LOC 3.SG.DAT  
 ‘(The man got bitten by the ants) and  
 wasn’t able to shoot at him.’

Turning now to stative expressions, these also come in four different voices, two of which are formally identical to the potentiative voice forms. Because of the formal similarities, the same labels have been chosen here for glossing these stative voices. However, their syntax and semantics differ quite clearly from the potentiative formations so that the labels are not really indicative of their functions.

In the basic voice form for statives, which is simply called *stative* here, the subject is a theme, i.e. an entity which is in, or currently is undergoing, a given state. The forms are identical to the patient voice forms of the potentiative paradigm, i.e. base plus prefix *ma*. The forms of the stative actor voice are morphologically identical to the potentiative actor

voice forms, i.e. base plus *maka*. The stative locative and conveyance voices are marked by the prefix *ka*- to which the basic voice affixes *-an* and *i-*, respectively, are added. See Tab. 136.4 for an overview of the forms.

The only really productive formation is the basic stative formation. Almost every Tagalog content word base can be prefixed with *ma*- and then expresses a state:

- (19) *na-galit siyá*  
 RLS.STAT-anger 3.SG  
 ‘she was/got angry’

With bases such as *basag* ‘crack, break’, which allow both a state and an action reading, the form *nabasag* is ambiguous: It can mean ‘be in a broken state’ (stative) or ‘happen to break/able to break’ (potentiative patient voice). In context, these readings are generally distinguished by the presence of an overt actor expression in the potentiative use (*nabasag niyá* ‘s/he happened to break it/was able to break it’).

The stative locative voice is common with bases expressing emotions. The subject expresses the person or thing at which the emotion is directed:

- (20) *kinagalitan siyá ng*  
*in-ka-galit-an siyá ng*  
 RLS(UGR)-STAT-anger-LV 3.SG GEN  
*Nanay nanay*  
 mother mother  
 ‘mother was angry with him/her’

Frequently, stative locative voice derivations take on some more specialised meanings. Thus, *kagalitan* also means ‘to reprove, to scold, to rebuke’. In addition, stative locative voice derivations are possible with a (relatively small) number of stative expressions which do not pertain to emotions. They then denote the place at which a given state occurs (e.g. *ka-matay-an* ‘place where someone died’, *ka-hulug-an* ‘place where someone falls’).

The stative conveyance voice is also most common with bases expressing emotions. Stative conveyance voice formations always have the connotation of causation, that is, the subject specifies the reason for the emotion:

- (21) *ikinagalít niyá akó*  
*i-in-ka-galit niyá akó*  
 CV-RLS(UGR)-STAT-anger 3.SG.POSS 1.SG  
 'she got angry at me (I was the reason for her being angry)'

The stative conveyance voice is found with a somewhat broader range of bases than the stative locative voice. These include *ikabasá* 'get wet on account of', *ikabagsák* 'fall on account of', *ikatawa* 'laugh on account of', *ikaiyák* 'cry on account of', etc.

The stative actor voice is very similar in meaning to the stative conveyance voice since it also specifies the cause for a given state. But the two formations differ in their grammar and productivity. In the stative conveyance voice construction (as in the stative locative voice construction), the theme argument (i.e. the one who experiences an emotion in the case of emotions) is grammatically coded as a genitive argument. In the stative actor voice construction it is a locative argument:

- (22) *lahát ng kanyáng sabihin*  
*lahát ng kanyá-ng sabi-in*  
 all GEN 3.SG.DAT-CONN statement-PV  
*ay nakagàgalit sa*  
*ay naka-RDP<sub>1</sub>-galit sa*  
 PM RLS.STAT.AV-RDP<sub>1</sub>-anger LOC  
*akin*  
*akin*  
 1.SG.DAT  
 'everything he says irritates me'

The subject expression in the stative actor voice construction usually refers to an inanimate cause (some state of affairs or a thing). With regard to productivity, the stative actor voice forms are the least common of all stative formations and whenever they occur they often take on somewhat specialised meanings (thus *makagalit* is 'irritate, antagonise, give offence' rather than a plain 'make angry'). Furthermore, the stative actor voice derivations are often conventionalised in one of the four aspect/mood forms, for example, *naka-àawa* 'arousing pity, pitiable' (<*awa*' 'mercy, compassion'), *nakàka-lító* (or *naka-lilitó*

'confusing' (< *litó* 'confused, at a loss'), or *nakàka-gandá* (or *naka-gàgandá*) 'beautifying' (< *gandá* 'beauty').

## 5.2. Further modes of actor involvement

The following prefixes mark further modes of actor involvement:

- *pa-* is a **causative** prefix and compatible with all voice affixes (see McFarland 1984 for ample exemplification and discussion). Actor voice is used when the causer is the subject, patient voice when the causee is the subject, e.g. *p<in>a-talim niyá itó* '<RLS(UGR)>CAUS-sharp 3.SG.POSS PROX (he made this sharp)' and *pa-patul-in mo si Huán ng kugon* 'CAUS-cut-PV 2.SG.POSS PN Juan GEN species.of. grass (have Juan cut the cugon-weeds)'. Patient voice in non-causative constructions becomes conveyance voice in causative ones. Compare the following two examples:

- (23) *Nànakaw-in ba natin*  
*RDP<sub>1</sub>-nakaw-in ba natin*  
*RDP<sub>1</sub>-stealing-PV INT 1.PI.POSS*  
*ang bangkáy ni Andrea?*  
*ang bangkáy ni Andrea*  
 SPEC corpse GEN.PN Andrea  
 'Will we steal Andrea's corpse?'

- (24) *Ipanànakaw ba ulí'*  
*i-pa-RDP<sub>1</sub>-nakaw ba ulí'*  
*CV-CAUS-RDP<sub>1</sub>-stealing INT again*  
*sa atín ang bangkáy*  
*sa atín ang bangkáy*  
 LOC 1.PI.DAT SPEC corpse  
*ni Andrea?*  
*ni Andrea*  
 GEN.PN Andrea  
 'Is (he) asking us to steal Andrea's corpse again?'

- *paki-* indicates that the actor joins an ongoing action (**sociative**). It may also be used to make a polite request or to indicate that something is done as a favour. Next to actor voice *maki-*, all undergoer voices are possible, for example:

- (25) *i-p<in>àki-hulog ni*  
*CV-<RLS(UGR)>SOCIAT-fall GEN.PN*  
*Pedro ang aking sulat*  
 Pedro SPEC DAT.1.SG.CONN letter  
 'Pedro mailed my letter (along with his)'

- *si-* only co-occurs with *mag-* and indicates **plurality** of (individual) actors, e.g. *nàg-si-ilag silá sa bayan* ‘RLS.AV-PL-flee 3.PL LOC town (they all fled from the town’).

Further differences in the conceptualisation of an action are indicated by the prefixes *pag-* and *paN-*. As illustrated in 4.3, these prefixes are used to derive gerunds. Such gerunds, which do not have an inherent orientation, are also compatible with voice affixes. Thus, contrasting sets such as the following occur: *p-um-utol* ‘cut’, *mág-putol* ‘cut several things’, *màmutol* ‘cut selectively or in quantity’. Similar contrasts involving undergoer voices are rare (an example is *tapakan* ‘step on’ vs. *pág-tapakan* ‘step on repeatedly’). The use of voice marked *paN*-derivatives is not very common and generally indicates intensive, distributive or repeated action, e.g. *bumilí* ‘buy’ vs. *nàmili* ‘shop’, *humampás* ‘hit with a whip’ vs. *màngampás* ‘whip people, go whipping’.

The major contrast is the one between *-um-* and *mag-* (cf. Blake 1925: 248 f.; Lopez 1937: 46–49; Pittman 1966; Schachter & Otanes 1972: 292 f.; Drossard 1984: 87–92; Himmelmann 1987: 185–188). Often *mag*- indicates the greater frequency or intensity of an action, cf. *bumasa* ‘read’ vs. *magbasa* ‘to read a lot/study’. A similar formation, i.e. the prefixing of *mag-* plus RDP<sub>2</sub>, is possible in principle for any base to indicate intensive or repeated action (cf. Schachter & Otanes (1972: 337 f.); Bloomfield (1917: 237–239) specifies stress shifts which may also occur). *mag-* and *-um-* may even co-occur to indicate a high degree of intensity, cf. *mág-um-aral* ‘study diligently’ or *mag-s-um-igáw* ‘shout (long and very loud)’. In other instances, the contrast seems to pertain to transitivity: *t<um>ayó’ kami* ‘<AV>stand.upright 1.PE (we stood up)’ vs. *nag-tayó’ kami ng bahay* ‘RLS.AV-stand.upright 1.PE GEN house (we erected a house)’. Note that in the corresponding undergoer voices no *pag-* appears: *i-t<in>ayó’ niyá ang bahay* ‘CV-<RLS(UGR)> stand.upright 3.SG.POSS SPEC house (he erected the house)’. In this type of example, the base denotes some kind of position or motion, and the *um*-form denotes an actor who moves himself, the *mag*-form an actor who moves something. A similar contrast exists with respect to bases denoting qualities,

e.g. *um-init* ‘become/get hot’ vs. *mag-init* ‘make hot, heat’. Much quoted, but unique is the contrast between *bumilí* ‘buy’ and *mág-bilí* ‘sell’.

Apart from such contrasting sets there are also several bases which only allow voice marking for *pag*-derived stems. For example, from *bawal* ‘prohibited’ neither \**bumawal* nor \**ibawal* may be derived, but only *magbawal* and *ipagbawal*. These bases do not display a common semantic or phonological feature. Other examples are *bilín* ‘order, instruction’, *kanuló* ‘betrayal’, *lingkód* ‘servant’, etc. (cf. McFarland 1976 Appendix II; Himmelmann 1987: 151). Another group of bases – which again does not exhibit a common denominator – allows actor orientation only with *mag-*, while undergoer orientation is possible without prior derivation. For example, from *luto* ‘cook’ \**lumuto* may not be derived but only *magluto*, while the undergoer voice form is simply *iluto*’ or *lulu’in*. Other bases belonging to this group are *dasál* ‘prayer’, *hugas* ‘wash’, *punas* ‘wipe off’, *libíng* ‘burial’, *bayad* ‘payment’, *kahoy* ‘wood’, *hubád* ‘naked’, etc. (cf. Himmelmann 1987: 179 f.).

## 6. Illustrative text

The standard Tagalog orthography is used with the modifications noted in 1.2.

*Ang ulól na unggó’ at ang*  
SPEC foolish CONN monkey and SPEC  
*ma-runong na pagóng. Minsan ang*  
STAT-knowledge CONN turtle once SPEC  
*pagóng habang na-li-ligo’ sa*  
turtle while RLS.AV-RDP<sub>1</sub>-swim LOC  
*ilog, ay nakà-kita siyá ng*  
river PM RLS.AV.POT:RDP<sub>1</sub>-see 3.SG GEN  
*isa-ng puno-ng-saging na*  
one-CONN tree-CONN-banana CONN  
*l<um>ù-lutang at t<in>à-tangáy*  
<AV>RDP<sub>1</sub>-float and <RLS(UGR)>RDP<sub>1</sub>-carry.off  
*ng agos. H<in>ila niyá sa*  
GEN flow <RLS(UGR)>pull 3.SG.POSS LOC  
*pasigan, dàtawát hindi niyá*  
river but NEG 3.SG.POSS  
*ma-dalá sa lupa’. Dahil dito*  
POT.PV-carry LOC earth cause LOC.PROX  
*t<in>awag niyá ang ka-ibig-an*  
<RLS(UGR)>call 3.SG.POSS SPEC ASS-like-LV

niyá-*ng*        *unggó'* at *i-ni-alay*  
 3.SG.POSS-CONN monkey and CV-RLS(UGR)-offer  
 niyá        *ang ka-putol ng*  
 3.SG.POSS SPEC ASS-cut GEN  
*puno-*ng*-*saging**        *kung i-tà-taním*  
 tree-CONN-banana if CV-RDP<sub>1</sub>-plant  
 niyá        *ang kaniyá-*ng**        *ka-parté*.  
 3.SG.POSS SPEC DAT.3.SG-CONN ASS-part  
*T<um>*angó'* *ang unggó'* at*  
*<AV>nod*        SPEC monkey and  
*h<in>*até'**        *nilá*        *sa gitná' mulá'*  
 <RLS(UGR)>divide 3.PL.POSS LOC middle begin  
*sa mag-kàbilá-*ng**        *dulo ang puno ng*  
 LOC AV-other.side-CONN end SPEC tree GEN  
*saging. In-angkín*        *ng unggó'*  
 banana RLS(UGR)-appropriate GEN monkey  
*ang ka-putol na*        *may mga dahon*,  
 SPEC ASS-cut CONN EXIST PL leaf  
*dahil sa panukala'* niyá        *na iyón*  
 cause LOC plan 3.SG.POSS CONN DIST  
*ay tù-tubo'*        *na ma-butí*        *kaysa*  
 PM RDP<sub>1</sub>-grow CONN STAT-good than  
*ka-putol na*        *wala-*ng**        *dahon.*  
 ASS-cut CONN NEG.EXIST-CONN leaf  
*Nang maka-raán*        *ang ilang*        *araw*,  
 when AV.STAT-way SPEC some:CONN day  
*ang puno ng unggó'*        *ay namatáy*,  
 SPEC tree GEN monkey PM RLS.STAT:dead  
*yamang ang sa pagóng ay t<um>ubo'*  
 whereas SPEC LOC turtle PM <AV>grow  
*hanggang sa mag-bunga. Ang mga saging*  
 until LOC AV-fruit SPEC PL banana  
*ay na-hinóg,*        *dàtапwát hindí*  
 PM RLS.STAT-ripe but NEG  
*ma-akyát*        *ng pagóng. Dahil dito*  
 POT.PV-climb GEN turtle cause LOC.PROX  
*t<in>awag*        niyá        *ang kaniyá-*ng**  
 <RLS(UGR)>call 3.SG.POSS SPEC DAT.3.SG-CONN  
*ka-ibig-ang*        *unggó'* at  
 ASS-like-LV.CONN monkey and  
*i-ni-alay*        niyá        *ang ila-*ng**  
 CV-RLS(UGR)-offer 3.SG.POSS SPEC few-CONN  
*bunga ng saging*        *kung à-akyat-ín*  
 fruit GEN banana if RDP<sub>1</sub>-climb-PV  
 niyá        *ang puno'. Ang unggó'* *ay*  
 3.SG.POSS SPEC tree SPEC monkey PM  
*um-akyát at*        *k<um>ain ng makákaya*.  
 AV-climb and <AV>eat GEN utmost  
*S<in>abi*        *ng pagóng:* "Hulug-an"  
 <RLS(UGR)>say GEN turtle drop-LV

mo        *akó.*" *Dàtапuwát*  
 2.SG.POSS 1.SG but  
*i-s<in>agót*        *ng unggó':*  
 CV-<RLS(UGR)>answer GEN monkey  
*"Balát man*        *at ma-linamnám ay*  
 skin though and STAT-delicious PM  
*hindí kitá*        *hù-huhug-an.*" *Ang*  
 NEG 1.SG.POSS:2.SG RDP<sub>1</sub>-drop-LV SPEC  
*pagóng ay na-galit*        *at nag-sabug*  
 turtle PM RLS.STAT-angry and RLS.AV-scatter  
*siyá ng tiník sa paligíd*        *ng*  
 3.SG GEN spine LOC surroundings GEN  
*puno'. Nang l<um>uksó ang unggó'*  
 tree when <AV>jump SPEC monkey  
*na-tiník*        *siyá. P<in>àg-bintang-án*  
 RLS.STAT-spine 3.SG <RLS(UGR)>GER-suspect-LV  
*niyá ang pagóng at kaniyá-*ng**  
 3.SG SPEC turtle and DAT.3.SG-CONN  
*h<in>anap*        *upang pa-rusah-an*  
 <RLS(UGR)>look.for so.that CAUS-suffer-LV  
 niyá.        *Nà-huli*        niyá        *ang*  
 3.SG.POSS RLS.POT.PV-catch 3.SG.POSS SPEC  
*pagóng sa kabilá'*        *ng isa-*ng**        *tuód.*  
 turtle LOC other.side GEN one-CONN stump  
*S<in>abi*        niyá        *sa pagóng:*  
 <RLS(UGR)>say 3.SG.POSS LOC turtle  
*"kitá ay aking*        *pa-rù-rusah-an.*  
 1.DI PM DAT.1.SG.CONN CAUS-RDP<sub>1</sub>-suffer-LV  
*Mamili*        *ka sa dalawá. Dikdik-ín*  
 AV:choose 2.SG LOC two crush-PV  
*kitá*        *sa lusóng o lunur-in*  
 1.SG.POSS:2.SG LOC mortar or drown-PV  
*kitá*        *sa ilog?"* *Ang ma-runong*  
 1.SG.POSS:2.SG LOC river SPEC STAT-knowledge  
*na pagóng ay nàg-umpisá* *ng*  
 CONN turtle PM RLS.AV-begin GEN  
*pàg-si-sigáw*        *at h<in>ilóng*  
 GER-RDP<sub>2</sub>-shout and <RLS(UGR)>request  
 niyá        *sa unggó'* *na, kung*  
 3.SG.POSS LOC monkey CONN if  
*ma-à-are',*        *ay dikdik-ín siyá sa*  
 STAT-RDP<sub>1</sub>-possible PM crush-PV 3.SG LOC  
*lusóng. Dàtапwát i-s<in>agót*        *ng*  
 mortar but CV-<RLS(UGR)>answer GEN  
*unggó': "I-bì-bigáy ko*        *sa iyó*  
 monkey CV-RDP<sub>1</sub>-give 1.SG.POSS LOC DAT.2.SG  
*ang pa-rusa*        *na hindí mo*        *gustó."*  
 SPEC CAUS-suffer CONN NEG 2.SG.POSS liking  
*At i-ni-hagis*        niyá        *sa ilog*  
 and CV-RLS(UGR)-throw 3.SG.POSS LOC river

*ang pagóng. Nang d<um>apo' ang pagóng*  
SPEC turtle when <AV>alight SPEC turtle

*sa tubig ay nág-si-sigáw siyá at*  
LOC water PM RLS.AV-RDP<sub>2</sub>-shout 3.SG and  
*s<in>abi niyá sa unggó':*  
<RLS(UGR)>say 3.SG.POSS LOC monkey  
“*Salamat, ka-ibig-an. Itó ang*  
thank ASS-like-LV PROX SPEC  
*aking tirah-an!”*  
DAT.1.SG.CONN dwell-LOC

Bloomfield's (1917) translation:

“Once upon a time, when the turtle was swimming in the river, he saw a banana-tree adrift and being carried along by the current. He dragged it to the beach, but was not able to carry it up to the solid ground. Therefore he called his friend, the monkey, and offered him a half of the banana-tree, if he would plant his part for him. The monkey agreed, and they divided the banana-tree at the middle, half-way from either end. The monkey took the half which had leaves, because he thought it would grow better than the half which had none.

When a few days had passed, the monkey's tree died, while that of the turtle grew until it bore fruit. The bananas grew ripe, but the turtle could not climb for them. Therefore he called his friend, the monkey, and offered him some of the fruits of the banana, if he would climb the tree. The monkey climbed up and ate for all he was worth.

Said the turtle: ‘Throw me some.’

But the monkey answered: ‘Though sweet the skins, I'd throw you none.’

The turtle got angry and scattered spines round the foot of the tree. When the monkey jumped down, he landed on the spines. He suspected the turtle and looked for him, in order to punish him. He found the turtle behind a stump.

Said he to the turtle: ‘I am going to punish you. Choose between the two: shall I Bray you in a mortar or drown you in the river?’

The clever turtle began to shout and begged the monkey, if it were possible, to Bray him in a mortar.

But the monkey answered: ‘I shall give you the punishment you don't want.’

And he threw the turtle into the river.

When the turtle arrived in the water, he set up a shout and said to the monkey: ‘Thank you, friend! This is my home.’” (Bloomfield 1917: 16)

## 7. Uncommon abbreviations

AV	actor voice
CV	conveyance voice
LV	locative voice
PM	predicate marker
PN	proper noun
PV	patient voice
SOCIAT	sociative
SPEC	specifier

## 8. References

- Adams, Karen L. & Manaster-Ramer, Alexis (1988), “Some Questions of Topic/Focus Choice in Tagalog”. *Oceanic Linguistics* XXVII, 79–101
- Bader, Thomas & Werlen, Iwan & Wymann, Adrian (1994), *Towards a Typology of Modality*. Bern: Institut für Sprachwissenschaft (Arbeitspapier 32)
- Blake, Barry J. (1988), “Tagalog and the Manila-Mt Isa Axis”. In: *La Trobe Working Papers in Linguistics* 1. Melbourne: La Trobe University, 77–90
- Blake, Frank R. (1906), “Expression of Case by the Verb in Tagalog”. *Journal of the American Oriental Society* 27, 183–189
- Blake, Frank R. (1925), *A Grammar of the Tagalog Language*. New Haven: American Oriental Society (American Oriental Series 1)
- Bloomfield, Leonard (1917), *Tagalog Texts with Grammatical Analysis*. Urbana: The Univ. of Illinois (University of Illinois Studies in Language and Literature 3.2–4)
- Capell, Arthur (1964), “Verbal Systems in Philippine Languages”. *Philippine Journal of Science* 93, 231–249
- Carrier-Duncan, Jill (1984), “Some Problems with Prosodic Accounts of Reduplication”. In: Aronoff, Mark & Oehrle, Richard T. (eds.), *Language, Sound, Structure. Studies in Phonology Presented to Morris Halle by His Teacher and Students*. Cambridge/MA, London: MIT Press, 260–286
- Constantino, Ernesto (1971), “Tagalog and other Major Languages of the Philippines”. In: Sebeok, Thomas A. (ed.), *Current Trends in Linguistics, Vol. VIII.1*. The Hague, Paris: Mouton, 112–154
- Crowhurst, Megan J. (1998), “Um Inflection and Prefixation in Toba Batak”. *Language* 74, 590–604

- Cubar, Ernesto H. (1984), "From Borrowed Noun to Verb: A Study of Functional Shift in Philippine Languages". In: Paz, Consuelo J. & del Corro, Anicia (eds.), *Studies on Philippine Minor Languages*. Quezon City: University of the Philippines, 119–133
- DeWolf, Charles M. (1988), "Voice in Austronesian Languages of Philippine Type: Passive, Ergative, or Neither?". In: Shibatani (ed.), 143–193
- Drossard, Werner (1984), *Das Tagalog als Repräsentant des aktivischen Sprachbaus*. Tübingen: Gunter Narr
- Foley, William A. (1976), *Comparative Syntax in Austronesian*. Ph.D. dissertation, University of California, Berkeley [UMI 77–4453]
- Foley, William A. & van Valin, Robert (1984), *Functional Syntax and Universal Grammar*. Cambridge etc.: Cambridge Univ. Press
- Gil, David (1993), "Tagalog Semantics", *Berkeley Linguistic Society* 19, 390–403
- Gonzales, Andrew B. (1971), "The Tagalog Nasal Ligature: A Reinterpretation". *Philippine Journal of Linguistics* 2.1, 28–43
- Guzman, Vide P. de (1978), *Syntactic Derivation of Tagalog Verbs*, Honolulu: Univ. Press of Hawaii (Oceanic Linguistics Special Publication 16)
- Himmelmann, Nikolaus P. (1987), *Morphosyntax und Morphologie: Die Ausrichtungsaffixe im Tagalog*. München: Fink (Studien zur Theoretischen Linguistik 8)
- Himmelmann, Nikolaus P. (1991), *The Philippine Challenge to Universal Grammar*. Köln: Institut für Sprachwissenschaft (Arbeitspapier 15)
- Himmelmann, Nikolaus P. (to appear), "Lexical Categories and Voice in Tagalog". In: Austin, Peter & Musgrave, Simon (eds.), *Voice and Grammatical Functions in Austronesian*. Stanford: CSLI
- Humboldt, Wilhelm von (1838), *Über die Kawi Sprache, Vol. II*, Berlin: Dümmler
- Kroeger, Paul R. (1993), *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI
- Lehmann, Christian (1984), *Der Relativsatz*. Tübingen: Gunter Narr
- Lemaréchal, Alain (1982), "Sémantisme des parties du discours et sémantisme des relations". *Bulletin de la Société de Linguistique de Paris* 77.1, 1–39
- Lemaréchal, Alain (1989), *Les parties du discours: Sémantique et syntaxe*. Paris: P.U.F.
- Llamzon, Teodoro A. (1973), "The Four Focus Transformations of Tagalog". In: Gonzales, Andrew B. (ed.), *Parangal kay Cecilio Lopez*. Quezon City: Linguistic Society of the Philippines, 168–183
- Llamzon, Teodoro A. (1976), *Modern Tagalog: A Functional-Structural Description*. The Hague, Paris: Mouton (Janua Linguarum, Series Practica 122)
- Lopez, Cecilio (1937), "Preliminary Study of the Affixes in Tagalog". In: Constantino, Ernesto (1977, ed.), *Selected Writings of Cecilio Lopez in Philippine Linguistics*. Diliman: Univ. of the Philippines Press, 28–104
- Lopez, Cecilio (1970), "On the Boak Tagalog of the Island of Marinduque". *The Archive* (University of the Philippines) 1.2, 1–53
- Marantz, Alec (1982), "Re Reduplication". *Linguistic Inquiry* 13, 435–482
- Marre, Aristide (1901), "Grammaire tagalog, composé sur un nouveau plan". *Bijdragen tot de Taal-, Land- en Volkenkunden* 53, 547–92
- Matsuda French, Koleen (1988), *Insights into Tagalog: Reduplication, Infixation, and Stress from Non-linear Phonology*. Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington (Publications in Linguistics 84)
- Matthews, Peter H. (1981), *Syntax*. Cambridge: Cambridge University Press
- McFarland, Curtis D. (1976), *A Provisional Classification of Tagalog Verbs*. Tokio: Institute for the Study of Languages & Cultures of Asia & Africa (Monograph Series 8)
- McFarland, Curtis D. (1983), *A Linguistic Atlas of the Philippines*. Manila: Linguistic Society of the Philippines
- McFarland, Curtis D. (1984), "Tagalog Causative Verbs". *Philippine Journal of Linguistics* 15, 19–47
- Müller, Friedrich (1882), *Grundriss der Sprachwissenschaft, II, Band II. Abt., Die Sprachen der Schlichthaarigen Rassen*. Wien: Alfred Hölder
- Naylor, Paz B. (1980), "Linking, Relation-Marking, and Tagalog Syntax". In: Naylor, Paz B. (ed.), *Austronesian Syntax. Papers from the Second Eastern Conference on Austronesian Languages [Ann Arbor 1976]*. Ann Arbor: Center for South and Southeast Asian Studies (Michigan Papers on South and Southeast Asia 15), 33–49
- Naylor, Paz B. (1995), "Subject, Topic, and Tagalog Syntax". In: Benett, David & Bynon, Theo-

- dora & Hewitt, George B. (eds.), *Subject, Voice and Ergativity*. London: School of Oriental and African Studies, 161–201
- Pittman, Richard (1966), “Tagalog -um- and -mag-: An Interim Report”. In: *Papers in Philippine Linguistics* 1, Canberra: Pacific Linguistics (Series A.8), 9–20
- Ramos, Teresita V. (1971), *Tagalog Structures*. Honolulu: Univ. Press of Hawaii
- Ramos, Teresita V. (1974), *The Case System of Tagalog Verbs*. Canberra: Pacific Linguistics (Series B.27)
- Ramos, Teresita V. & Bautista, Maria L. S. (1986), *Handbook of Tagalog Verbs: Inflections, Modes, Aspects*. Honolulu: Univ. of Hawaii Press
- Reid, Lawrence A. (1981), “Philippine Linguistics: The State of the Art: 1970–1980”. In: Hart, Donn W. (ed.), *Philippine Studies: Political Science, Economics, and Linguistics*. Northern Illinois University, Center for Southeast Asian Studies Occasional Paper 8, 223–230
- Rubino, Carl R. G. (1998), “The Morphological Realization and Production of a Nonprototypical Morpheme: The Tagalog Derivational Clitic”. *Linguistics* 36, 1147–1166
- Schachter, Paul (1976), “The Subject in Philippine Languages, Topic, Actor, Actor-Topic or None of the Above”. In: Li, Charles N. (ed.), *Subject and Topic*. New York etc.: Academic Press, 491–518
- Schachter, Paul (1995), *The Subject in Tagalog: Still None of the Above*. Los Angeles: UCLA/Department of Linguistics (UCLA Occasional Papers in Linguistics 15)
- Schachter, Paul & Otanes, Fe T. (1972), *Tagalog Reference Grammar*. Berkeley etc.: Univ. of California Press
- Scheerer, Otto (1924), “On the Essential Difference Between the Verbs of the European and the Philippine Languages”. *Philippine Journal of Education* 7, 1–10
- Shibatani, Masayoshi (1988), “Voice in Philippine Languages”. In: Shibatani (ed.), 85–142
- Shibatani, Masayoshi (1988, ed.), *Passive and Voice*. Amsterdam, Philadelphia: Benjamins (Typological Studies in Language 16)
- Shkarban, Lisa I. (1995), *Grammatičeskij stroj tagal'skovo jazyka*, Moskva: Nauka
- Soberano, Rosa (1980), *The dialects of Marinduque Tagalog*. Canberra: Pacific Linguistics (Series B.69)
- Totanes, Sebastian de (1865), *Arte de la lengua tagala y manual tagalog*. Binondo: Miguel Sanchez y Ca [1745]
- Wolfenden, Elmer (1961), *A Re-statement of Tagalog Grammar*. Manila: SIL and Institute of National Language
- Wolff, John U. (1993), “Why Roots Add the Affixes with which they Occur”. In: Reesink, Ger (ed.), *Topics in Descriptive Austronesian Linguistics*. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, 217–244
- Wolff, John U. with Centeno, M. T. C. & Rau, D. V. (1991), *Pilipino through Self-Instruction*, Vol. 1–4. Ithaca: Cornell Southeast Asia Program

Nikolaus P. Himmelmann, Bochum  
(Germany)

## 137. Diyari (Pama-Nyungan)

1. Introduction
2. Morphological type
3. Parts of speech
4. Nominal morphology
5. Pronoun morphology
6. Verb morphology
7. Predicate determiner morphology
8. Particles
9. Post-inflectional morphology
10. Illustrative text
11. Uncommon abbreviations
12. References

### 1. Introduction

The Diyari language was traditionally spoken in the far north of South Australia, to the east of Lake Eyre along the lower reaches of Cooper Creek – a mostly dry watercourse that runs from western Queensland into Lake Eyre. This is one of the driest and hottest regions of Australia with an average annual rainfall of about 100 mm (or 4 inches) and summer temperatures regularly reaching 45

degrees. However, the whole of the area was exploited by the Diyari and is covered with placenames and mythological sites.

The neighbours of the Diyari were the Thirrari and Arabana (west), Kuyani and Adnyamathanha (south), Yandruwandha and Yawarrawarrga (east), and the Ngamini and Wangkangurru (north). The Diyari language is closely genetically related to Thirrari and Ngamini, and more distantly to Yandruwandha and Yawarrawarrga. Austin (1990) shows that these languages form a sub-group with Pitta-Pitta (western Queensland) and Wangkumarra (western New South Wales). The more distant genetic connections of this subgroup have yet to be worked out in detail (see Bowern 1998; 2001); however, it is clear that they all belong to the Pama-Nyungan family which occupies the southern seven-eighths of the Australian continent (Dixon 1980; Blake 1988; Bowern & Koch 2004).

When first contacted by white settlers in the 1860's the Diyari probably numbered over 1000; today there is no-one alive who can speak the language fluently. Mr. Ben Murray (traditional name *parlku-nguyuthangkayiwarna*) was born in 1891 and was the very last speaker of the Thirrari language. His life history, including a series of stories told by him in Diyari and Thirrari, appears in Austin et al. 1988 (cf. also Murray & Austin 1981; 1986).

Unlike many Australian Aboriginal languages, Diyari has a long history of use as a written language. In 1869 German Lutheran pastors established a mission among the Diyari at Lake Killalpaninna on Cooper Creek; the mission and associated sheep station flourished until its closure by the South Australian government in 1914. The missionaries studied the Diyari language and used it in their work and their daily lives, including preaching in Diyari and teaching it in the mission school. They prepared primers, school books, dictionaries and grammars of Diyari, and translated a large number of Christian works into the language, including hymns and the Old and New Testaments.

One of the most remarkable missionaries was the Rev. J.G. Reuther who not only translated the New Testament into Diyari, but also compiled a 14 volume manuscript on Diyari language, culture, mythology and history that includes a grammatical description (employing traditional Latin-based categories), and a massive four volume dictionary (Reuther 1981).

Recent research on Diyari began in the 1970's when tape recordings and notes were made by Luise Hercus, David Trefry and Peter Austin. Austin (1978) is a reference grammar of Diyari (revised and published as Austin 1981). Austin is preparing a dictionary and collection of stories told by the last speakers, and a study of Diyari literacy based on letters written by native speakers.

## 2. Morphological type

Like all Pama-Nyungan languages, Diyari is entirely suffixing with a highly agglutinative word structure. Morpheme shape is generally invariant or else conditioned by phonological (and rarely morphological) factors. Word order is relatively free, though less so than in many other Australian languages. Diyari is also consistently **dependent marking** (Nichols 1986) with grammatical functions coded on dependents only. Approximately half the Pama-Nyungan languages have bound pronominal agreement suffixes attached to the finite verb, to a sentence-level auxiliary element, or else to the first constituent of a main clause, and hence show **double marking**. Bound pronominals are lacking from Diyari and its neighbours (see bound pronoun distribution maps in Blake 1977, Dixon 1980).

Case affixes are attached to the last element of a noun phrase only, unless the constituents are separated by other material. Diyari is thus a **phrase marking** language (Blake 1987: 78; Dench & Evans 1988), not **word marking** like many other Pama-Nyungan languages.

Diyari shows **Suffixaufnahme** (Austin 1995; Plank 1990), i.e. double marking of case in possessive constructions, where the genitive case-marked element may take the case marker assigned to the phrase. This occurs when the possessor follows the possessee.

Words in Diyari must begin with one and only one consonant and must end in a vowel. Intervocalic consonant clusters are limited (Austin 1981: 26) and no vowel sequences are permitted. Bound morphemes may begin with consonant clusters, but must end in a vowel. The transcription adopted here follows Australianist conventions: *th, nh, lh* are lamino-dental stop, nasal and lateral respectively; *j, ny, ly* are lamino-palatal; *rd, rt, rn, rl* are apico-domal (retroflex). There are three r-sounds: a post-alveolar continuant *r*, a tap *rr*, and a trill *rrh*.

Morphophonemics is relatively limited and is discussed in the relevant sections below (cf. 4.2.2). Reduplication (involving copying of the first CV(C)CV of the stem) occurs with both nominals and verbs, coding ‘diminutive’ with nominals and ‘iterative’ with verbs (cf. illustrative text, (8)).

### 3. Parts of speech

Seven word classes may be set up for Diyari:

- (a) **nominals** – words which inflect for case according to their relationship to the predicate or to other nominals in the clause. This category covers entity concepts (‘nouns’), property concepts (‘adjectives’), certain temporal and spatial locationals, and personal names. Adjectives and nouns have the same inflectional behaviour, but can be distinguished as subclasses of nominals on syntactic grounds (Austin 1981: 33 f.).
- (b) **pronouns** – inflect for case and are specified for person (1st, 2nd, 3rd) and number (singular, dual, plural). In the first person non-singular there is an inclusive-exclusive contrast; there is also a set of special relationship pronouns with a defective paradigm.
- (c) **verbs** – obligatorily inflect for tense/mood or dependent clause categories. There are main and auxiliary verbs, the latter occurring after main verbs and marking tense/aspect.
- (d) **predicate determiners** – closed class with two members that are inflected for deixis and adverbially modify predicates.
- (e) **particles** – modal elements that cannot take inflectional or word-forming affixes.
- (f) **co-ordinators** – phrasal and clausal conjunctions and disjunctions that allow no affixation.
- (g) **interjections** – comprise a whole utterance by themselves and are uninflected (see Austin 1981: 37).

Diyari has a set of post-inflectional affixes that can attach to words of any class (except co-ordinators and interjections). These code various pragmatic (information status) categories (see (9)).

### 4. Nominal morphology

Nominal words minimally consist of a root which may be inflected for case (4.2). One or more derivational suffixes may be attached

to roots to form stems; some of these derive nominals (4.1.1), while others cause a change in category (4.1.2). On morpho-syntactic grounds nominals can be divided into: adjectives, common nouns, proper nouns (personal and place names), and location nominals. Case is realised differently according to nominal sub-class.

#### 4.1. Word formation

There are eleven derivational suffixes that may attach to nominal roots to form stems. Eight of these derive nominal stems (4.1.1); the other three derive verbs (4.1.2).

##### 4.1.1. Non-category changing

Affixes deriving nominals are:

- (a) *-kanyji* ‘EXCESS’ derives a stem meaning ‘animate being excessively concerned with or possessed of’ qualities denoted by the root, e.g. *yapa-kanyji* ‘fear-EXCESS (fearful one)’, *parla-kanyji* ‘sexual.desire-EXCESS (sex maniac)’.
- (b) *-yija* ‘HABIT’ produces ‘animate being habitually associated with’ the root nominal, e.g. *kupula-yija* ‘alcohol-HABIT (drunkard)’, *nhaka-yija* ‘there-HABIT (one from there)’.
- (c) *-lha* ‘CHAR’ added to a locational noun or a common noun referring to a geographical feature it forms ‘entity which inhabits, grows or is found in’ a place, e.g. *pantu-lha* ‘salt.lake-CHAR (salt lake dweller)’, *marda-lha* ‘hill-CHAR (hill dweller)’. With temporal location nominals it indicates characteristic time association, e.g. *warru-lha* ‘long.ago-CHAR (old, ancient)’.
- (d) *-wurlu* ‘DU’ derives a dual stem from a common noun or adjective root, e.g. *kanku-wurlu* ‘boy-DU (two boys)’, *kapi-wurlu* ‘egg-DU (two eggs, testicles)’. Number is not an obligatorily marked category in Diyari; unmarked common nouns may be interpreted as singular, dual or plural according to context. Dual and plural derived stems inflect like pronouns (see 5).
- (e) *-wara* ‘PL’ derives a plural stem from a common noun or adjective root, e.g. *kanku-wara* ‘boy-PL (boys)’, *pirta-wara* ‘tree-PL (trees)’.
- (f) *-nithu* ‘PROPR’ meaning ‘with’ or ‘having’ is suffixed to common noun or adjective roots to produce stems that can serve as

- adnominal modifiers, e.g. *nganka-nthu* ‘beard-PROPR (bearded)’, *nhuwa-nthu* ‘spouse-PROPR (married)’ (Austin 1981: 42, 140–145). Most Australian languages have a complementary suffix, the PRIVATE indicating ‘lacking, without’, Diyari has an adjective *pani* ‘none’ to express this concept, e.g. *ngapa pani* ‘water without (waterless)’ (cf. Austin 1981: 43 f.).
- (g) *-mara* ‘KINPROPR’ can be added to common nouns or adjectives with two effects: as a synonymous alternative to *-nthu* ‘PROPR’, or suffixed to a kinship term to derive a collective noun stem where *A-mara* means ‘group of people one of whom is called A by the others’, e.g. *nhuwa-mara* ‘spouse-KINPROPR (man and wife/wives)’, *kaku-mara* ‘elder.sister-KINPROPR (group of sisters)’. Australian languages typically have one or more affixes that attach to kinship terms to derive such collective nouns, often called **dyadic** or **triadic kin terms** (Heath et al. 1982), and sometimes with different functions in address and reference (Dench 1982).
  - (h) *-nja* ‘NUMBER’ added to numeral adjectives to mean ‘just so many in number’, e.g. *parrkulu-nja* ‘three-NUMBER (just three)’.

Nominal stems may be composed of two of these affixes with (a) or (b) or (c) followed by (d) or (e) or (f) or (g) (affix (h) has never been observed in combination), e.g. *kupula-yijawara* ‘alcohol-HABIT-PL (drunkards)’, *yapakanyji-mara* ‘fear-EXCESS-KINPROPR (with a fearful one)’. Other combinations and longer affix sequences are avoided by speakers.

#### 4.1.2. Category changing

Three affixes may be added to nominal roots to derive verb roots. Two derive intransitive roots and one transitive verbs:

- (a) *-rri* ~ *-ri* ‘INCH’ derives inchoative intransitive verbs meaning ‘become’ (Austin 1981: 164–167). The allomorphs are phonologically conditioned: *-ri* is added to roots containing an intervocalic apical consonant or cluster, *-rri* is used elsewhere, e.g. *wardu-ri* ‘short-INCH (become short)’, *pandrra-ri* ‘cooked-INCH (become cooked, cook)’, *ngumu-rri* ‘good-INCH (become good)’, *thungka-rri* ‘rotten-INCH (become rotten)’ (cf. illustrative text, (6), (7), (19)).

- (b) *-nga* ‘PRODUCT’ derives an intransitive verb meaning ‘to produce ...’ when added to nouns referring to sounds, e.g. *karta-nga* ‘crack-PRODUCT (to crackle, snap)’, *kandrru-nga* ‘snore-PRODUCT (to snore)’. There is also the form *thupu-nga* ‘smoke-PRODUCT (to smoke, give off smoke)’.
- (c) *-nganka* ‘CAUS<sub>3</sub>’ derives a transitive verb meaning ‘to cause to be ...’, e.g. *mukanganka* ‘sleep-CAUS<sub>3</sub> (to put to sleep)’, *kirrhi-nganka* ‘clever-CAUS<sub>3</sub> (to teach)’ (cf. illustrative text, (2), (8)). This affix is homophonous with the main verb *nganka* ‘to make, cause’.

#### 4.2. Inflection

Nominals and pronouns inflect for case with only the last of a sequence of nominals in a noun phrase bearing the case marker, the others being unmarked. Case inflection, as in most Pama-Nyungan languages, is of the ‘split-ergative’ type (cf. Art. 102; Silverstein 1976; Dixon 1979); the distinction between intransitive subject (**nominative**), transitive subject (**ergative**), and transitive object (**accusative**) is not overtly marked for all nominal types. There is formal syncretism conditioned by animacy of the nominal or pronoun referent (cf. Blake 1987: 20–23). In Diyari there are three paradigms:

- (a) singular common nouns and male personal names syncretise nominative and accusative (giving an ergative-absolutive pattern);
- (b) non-singular first and second person pronouns (see 5) syncretise ergative and nominative (giving a nominative-accusative pattern);
- (c) all others have three separate forms for the three cases, i.e. female personal names, non-singular common nouns (derived by the affixes *-wurlu* and *-wara*, cf. 4.1.1), singular first and second person pronouns, and all third person pronouns (see 5).

There are four other case categories: ablative, allative, locative and dative. All nominals have a separate ablative case form but the others are subject to syncretism: for singular common nouns allative and dative fall together, all other non-locational nominals have a single form for locative and allative. Case allomorphy is described in the following sections.

#### 4.2.1. Location nominals

**Location nominals** have the unmarked stem as their **locative** case form. They include temporal locationals, e.g. *warru* ‘long ago’, *thangkuparna* ‘tomorrow’, and spatial locationals. The latter consist of the cardinal directional terms (*thinankarra* ‘north’, *kunankarri* ‘south’, *yantakarra* ‘west’, *thirrhiwa* ‘east’) and the roots *nhaka* ‘there’ and *nhingki* ‘here’. The root *nhingki* obligatorily selects a deictic marker of relative distance from the speaker (-*rda* ‘VICIN’, -*ya* ‘NEAR’, -*parrha* ‘THERE’, -*wa* ‘DISTANT’) before being suffixed for case. The case affixes are allative -*nhi* for *nhaka* and *nhingki*, -*ya* for other locationals, and ablative -*ndrru*, e.g. *thangkuparna* ‘tomorrow(LOC)’, *nhaka* ‘there(LOC)’, *warru-ndrru* ‘long.ago-ABL’, *thinankarra-ya* ‘north-ALL’, *nhingki-rda-nhi* ‘here-VICIN-ALL’.

#### 4.2.2. Personal names, common nouns, and adjectives

Seven cases are distinguished for personal names, common nouns and adjectives, with some formal syncretism (4.2). There are two morphophonological rules that affect inflected forms (Austin 1981: 28 f.):

- (a) in tri-syllables ending in *i* or *u* the final vowel is neutralised to *a* when a suffix is added;
- (b) final *u* of -*wurlu* ‘DU’ becomes *a* when followed by a suffix.

The cases are:

- (a) **Nominative** codes intransitive subject function. Allomorphs are -*ni* added to female personal names, -*nha* added to male personal names; all others use the bare stem form.
- (b) **Ergative** codes transitive subject and instrumental functions. There are three allomorphs:
  - *ndrru* added to female personal names;
  - *li* added to male personal names, non-singular common nouns, singular common nouns ending in *a*, and singular common nouns of three syllables ending in *u* or *i* (optionally for tri-syllables ending in *i*);
  - *yali* added to singular common nouns of two, four or five syllables ending in *u* or *i* (and optionally for tri-syllables ending in *i*).

Words exemplifying these allomorphs are given in Tab. 137.1 (see also illustrative

text). Note that most Pama-Nyungan languages have ergative allomorphs -*ngku* and/or -*lu* (often the former is used with disyllables and the latter with longer stems; see Dixon 1980, Blake 1987). Diyari and its neighbours show no evidence of these.

- (c) **Accusative** codes transitive object function. Singular common nouns and male personal names syncretise this case with the nominative, all others add -*nha*. Accusative is commonly -*nha* or -*nya* throughout Australia (Dixon 1980).
- (d) **Locative** codes functions including location in time or space, accompaniment, object of comparison and complement of certain intransitive verbs (Austin 1981: 123–130). Allomorphs are:
  - *nhangu* added to female personal names;
  - *ngu* added to male personal names and non-singular common nouns;
  - *nhi* added to singular common nouns.

Most Pama-Nyungan languages show locative allomorphs identical to ergative except in having final *a* instead of *u* (e.g. -*ngka* and/or -*la*). Diyari and its neighbours lack this parallelism.

- (e) **Allative** codes location in time or space towards which motion or an action is directed. It is formally the same as LOCATIVE for female and male personal names and non-singular common nouns, and is -*ya* for singular common nouns.
- (f) **Dative** codes purpose, benefactive and genitive (alienable possession) functions. The realisation is the same as allative for singular common nouns, -*nhangka* for female personal names, and -*rni* for male personal names and non-singular common nouns. Most Pama-Nyungan languages have a dative case form -*ku*; this is lacking in Diyari and its relatives, though found in Arabana and Wangkangurru to the west.
- (g) **Ablative** codes source location from which motion or an action proceeds and indirect cause. The suffixes are: -*ngundrru* added to male and female personal names and to non-singular common nouns, -*ndrru* added to singular common nouns.

Tab. 137.1 sets out examples of these allomorphs.

stem	ergative	nominative	accusative	locative	allative	ablative	dative
person-DU	<i>karnawu-rlali</i>	<i>karnawurlu</i>	<i>karnawurla-nha</i>	<i>karnawu-rlangu</i>	<i>karnawu-rlangu</i>	<i>karnawu-rlangundru</i>	<i>karnawu-rlarni</i>
person-PL	<i>karnawara-li</i>	<i>karnawara</i>	<i>karnawara-nha</i>	<i>karnawa-rangu</i>	<i>karnawa-rangu</i>	<i>karnawara-ngundru</i>	<i>karnawa-rarni</i>
woman's name	<i>jirimiri-rrindru</i>	<i>jirimirrini</i>	<i>jirimirrinha</i>	<i>jirimirri-nhangu</i>	<i>jirimirri-nhangu</i>	<i>jirimirri-ngundru</i>	<i>jirimirri-nhangka</i>
man's name	<i>wangami-rrili</i>	<i>wangami-rrinha</i>	<i>wangami-rrinha</i>	<i>wangami-rringu</i>	<i>wangami-rringu</i>	<i>wangamirri-ngundru</i>	<i>wangamirri-rrirni</i>
woman	<i>wilhali</i>	<i>wilha</i>	<i>wilha</i>	<i>wilhanhi</i>	<i>wilhaya</i>	<i>wilhandru</i>	<i>wilhaya</i>
sister's husband	<i>kardiyali</i>	<i>kardi</i>	<i>kardi</i>	<i>kardinh</i>	<i>kardiya</i>	<i>kardindru</i>	<i>kardiya</i>
boy	<i>kankuyali</i>	<i>kanku</i>	<i>kanku</i>	<i>kankunhi</i>	<i>kankuya</i>	<i>kankundru</i>	<i>kankuya</i>
mother's mother	<i>kanhiniyali</i>	<i>kanhini</i>	<i>kanhini</i>	<i>kanhinanhi</i>	<i>kanhinya</i>	<i>kanhinandru</i>	<i>kanhinaya</i>
Murnpeowie (place)				<i>mampinhi</i>	<i>mapmiya</i>	<i>mampindruru</i>	
long ago				<i>warru</i>	<i>warrya</i>	<i>warrundru</i>	
there				<i>nhaka</i>	<i>nhakanhi</i>	<i>nhakandru</i>	

Tab. 137.1: Case paradigms

## 5. Pronoun morphology

Pronouns inflect for case and distinguish three persons (1st, 2nd, 3rd) and three numbers (singular, dual, plural). There are two types of pronouns: regular personal pronouns and special relationship pronouns.

Special relationship pronouns are a marginal set with a defective paradigm – they occur only in the dual absolute (intransitive subject and transitive object functions) and are clearly borrowed from Wangkangurru. They code same or different (matrilineal) moiety membership. Special pronouns reflecting kinship are found in many Australian languages (Heath et al. 1982, eds.).

For regular personal pronouns there is an inclusive-exclusive contrast in the first person non-singular (present in about half the languages of Australia, cf. Dixon 1980), and a feminine/non-feminine natural sex based gender contrast in the third person singular. Third person pronouns have demonstrative functions: they co-occur with nouns in noun

phrases (cf. illustrative text, (12), (19)) and may take the deictic suffixes attached to spatial location nominals (4.2.1). Tab. 137.3 sets out the pronoun paradigm.

Dative pronouns are used to indicate alienable possession in Diyari. Body-part possession is coded by apposition of the possessor and the possesum, both being marked for case (cf. two instances in illustrative text, (11)). Dative pronouns may precede the head (cf. illustrative text, (3), (11), (13)) or follow it, in which circumstance they bear the case marker of the noun phrase as a whole, the head remaining unmarked (a type of Suffixaufnahme; Austin 1981: 62). An example is illustrative text, (9) (note the shift of final *i* to *a*, as with tri-syllabic nominals).

## 6. Verb morphology

Verbs in Diyari inflect for tense, mood or dependent clause categories. They sub-classify into **main verbs** and **auxiliary verbs**; the latter are optional, code tense/aspect and do not undergo word formation processes. Main verbs are strictly sub-categorised into intransitive (taking a nominative subject), transitive (taking ergative and accusative) and di-transitive (taking ergative and two accusative). Most Australian languages have several morphologically conditioned verb conjugations (Dixon 1980: 382 ff.); in Diyari and its neighbours all verbs inflect in a single paradigm.

person	same moiety	different moiety
first	<i>ngalantha</i>	<i>ngalilakiya</i>
second	<i>ampalantha</i>	<i>ampalakiya</i>
third	<i>pulalantha</i>	<i>pulalakiya</i>

Tab. 137.2: Special relationship pronouns

person	ergative	nominative	accusative	dative	locative	ablative
1:SG	<i>ngathu</i>	<i>nganhi</i>	<i>nganha</i>	<i>ngakarni</i>	<i>ngakangu</i>	<i>ngakangundrru</i>
2:SG	<i>yundrru</i>	<i>yini</i>	<i>yinantha</i>	<i>yingkarni</i>	<i>yingkangu</i>	<i>yingkangundrru</i>
3:SG:F	<i>nhandrru</i>	<i>nhani</i>	<i>nhanha</i>	<i>nhangkarni</i>	<i>nhangkangu</i>	<i>nhangkangundrru</i>
3:SG:NF	<i>nhulu</i>	<i>nhawu</i>	<i>nhinha</i>	<i>nhungkarni</i>	<i>nhungkangu</i>	<i>nhungkangundrru</i>
1:DL:INCL	<i>ngaldrra</i>	<i>ngaldrra</i>	<i>ngaldranha</i>	<i>ngaldrrarni</i>	<i>ngaldrrangu</i>	<i>ngaldrrangundrru</i>
1:DL:EXCL	<i>ngali</i>	<i>ngali</i>	<i>ngalinha</i>	<i>ngalirni</i>	<i>ngalingu</i>	<i>ngalingundrru</i>
2:DL	<i>yula</i>	<i>yula</i>	<i>yulanha</i>	<i>yularni</i>	<i>yulangu</i>	<i>yulangundrru</i>
3:DL	<i>pulali</i>	<i>pula</i>	<i>pulanha</i>	<i>pularni</i>	<i>pulangu</i>	<i>pulangundrru</i>
1:PL:INCL	<i>ngayana</i>	<i>ngayana</i>	<i>ngayananha</i>	<i>ngayanarni</i>	<i>ngayanangu</i>	<i>ngayanangundrru</i>
1:PL:EXCL	<i>ngayani</i>	<i>ngayani</i>	<i>ngayaninha</i>	<i>ngayanirni</i>	<i>ngayaningu</i>	<i>ngayaningundrru</i>
2:PL	<i>yurra</i>	<i>yurra</i>	<i>yurranha</i>	<i>yurrarni</i>	<i>yurrangu</i>	<i>yurrangundrru</i>
3:PL	<i>thanali</i>	<i>thana</i>	<i>thananha</i>	<i>thanarni</i>	<i>thanangu</i>	<i>thanangundrru</i>

Tab. 137.3: Pronouns

### 6.1. Word formation

Verb stems may be derived by the addition of suffixes to roots. Some suffixes change category, deriving nouns (6.1.2), while others derive verbs (6.1.1). Among the latter, some affixation changes transitivity, some does not.

#### 6.1.1. Non-category changing

There are eight affixes which may be added to verb roots to derive verb stems; five of these affect **transitivity** and three do not. Main verbs can be sub-classified into ten morpho-syntactic groups according to their co-occurrence with these five affixes. The classification is set out in Tab. 137.4.

### Intransitive

class	<i>-lka-</i>	<i>-ipa</i>	<i>-ma-</i>	example
1A	+	—	—	<i>wapa-</i> ‘to go’
1B	+	+	—	<i>tharrka-</i> ‘to stand’
1C	—	+	—	<i>purnka-</i> ‘to grow’
1D	—	—	—	<i>yatha-</i> ‘to speak’
1E	—	—	+	<i>pali-</i> ‘to die’

### Transitive

class	<i>-tharrhi-</i>	example
2A	reflexive	<i>nandrra-</i> ‘to hit’
2B	anti-passive1	<i>karlka-</i> ‘to await’
2C	anti-passive2	<i>thayi-</i> ‘to eat’
2D	passive	<i>thinthia-</i> ‘to lose’
2E	none	<i>yurlka-</i> ‘to swallow’

Tab. 137.4: Verb classes

The affixes and their functions are as follows. Firstly, those which affect transitivity:

- (a) *-lka* ‘APPL’ added to class 1 A or 1 B intransitive roots derives an applied transitive stem indicating accompaniment, e.g. *wapa-lka-* ‘go-APPL (to go with, take)’, *tharrka-lka-* ‘stand-APPL (to stand with)’.
- (b) *-ipa* ‘CAUS<sub>1</sub>’ added to 1 B or 1 C roots derives a causative transitive stem (with replacement of the root-final vowel with *i*), e.g. *tharrkipa-* ‘stand:CAUS<sub>1</sub> (to stand (it) up)’, *thurraripa-* ‘lie:CAUS<sub>1</sub> (to lay (it) down)’. This affix also has a benefactive function (see (f)).
- (c) *-ma* ‘CAUS<sub>2</sub>’ added to 1 E roots derives a causative transitive stem, e.g. *pali-ma-* ‘die-CAUS<sub>2</sub> (to kill)’.
- (d) *-tharrhi* has a number of functions according to the transitive verb sub-class:
  - with 2 A or di-transitive roots it derives an intransitive reflexive stem, e.g. *nandrra-tharrhi-* ‘hit-REFL (to hit oneself)’;
  - with a 2 B root it derives an intransitive anti-passive activity stem, e.g. *karlka-tharrhi-* ‘await-APASS (to wait for)’;
  - with a 2 C root it derives an intransitive anti-passive unspecified object activity stem, e.g. *thapa-tharrhi-* ‘drink-APASS (to have a drink)’;
  - with a 2 D root it derives an intransitive passive process stem, e.g. *thintha-tharrhi-* ‘lose-PASS (to get lost)’.

This affix also has an aspectual function where there is no change in transitivity. It can be added to a reduplicated root of

- any class to derive a durative stem (iterative or continuous according to the root semantics), e.g. *ngara-ngara-tharrhi-* ‘RDP-hear-DUR (to listen)’, *yatha-yatha-tharrhi-* ‘RDP-speak-DUR (to gossip)’.
- (e) *-mali* ‘REC’ added to a transitive or ditransitive root derives an intransitive stem with reciprocal meaning, e.g. *nandrра-mali-* ‘hit-REC (to hit one another)’.

Affixes not affecting transitivity are:

- (f) *-yirrpа* ‘BEN’ attaches to a transitive or ditransitive root producing a stem indicating altruistic action done for the benefit or someone other than the subject, e.g. *pardaka-yirrpа-* ‘carry-BEN (to carry for someone)’. The transitiviser *-ipa-* is a synonymous alternative, e.g. *pardakipa-* ‘carry:BEN (to carry for someone)’.
- (g) *-ingа* ‘PROL’ may be added to roots of any transitivity to derive a stem indicating relative motion of the subject with respect to some other referent (this affix changes root-final vowel to *i*), e.g. *mаnинга-* ‘take-PROL (to take going past)’, *thar-ingа-* ‘go.up-PROL (to go up past)’ (cf. illustrative text, (20)).
- (h) *-nthi* ‘CONSEC’ added to roots of any class derives a stem meaning action done as a consequence of some other action, e.g. *wapa-nthi-* ‘go-CONSEC (to go afterwards)’.

Complex words composed of roots plus two or three derivational suffixes have been attested. The general order is transitiviser plus benefactive plus de-transitiviser, e.g. *yakkalkipa-mali-* ‘ask:BEN-REC (to interrogate one another)’, *wapa-wapa-lka-tharrhi-* ‘RDP-go-APPL-DUR (to be taking)’, *parni-mа-yirrpа-mali-* ‘smell-CAUS<sub>2</sub>-BEN-REC (to smell one another)’.

### 6.1.2. Category changing

Noun stems may be derived from verbs by:

- (a) suffixing *-ni* to the verb stem (which may be preceded by its object if transitive) to derive an instrumental nominal, e.g. *pawa daka-ni* ‘seed grinder’, *warlpardaka-tharrhi-ni* ‘cover-REFL-NR (cover, blanket)’;
- (b) suffixing *-yija* ‘HABIT’ (cf. 4.1.1) to the participial inflected stem (6.2.1) to derive an agentive nominal, e.g. *yindrrа-rna-yija* ‘cry-PART-HABIT (crier)’.

## 6.2. Inflection

Main verb inflections divide into those added to **main clause verbs** (finite, marking tense/mood, and non-finite (6.2.1)), and those added to **dependent clause verbs**. The latter code interclausal reference and linkage categories (6.2.2).

### 6.2.1. Main clause

There are four finite tense/mood inflections and two non-finite inflections:

- (a) *-yi* ‘PRES’ marks non-past tense and imperfective aspect, including present tense, immediate future, and generic statement. It is common in narrative texts for actions in sequence in the main story line (see illustrative text).
- (b) *-ya* ‘PAST’ marks past tense and perfective aspect, situations completed at the moment of speaking.
- (c) *-ya* ‘IMP’ marks imperative mood for verbs ending in *i* and *u*; verbs ending in *a* use the bare stem for imperative (see illustrative text, (6), (18)). Following ‘IMP’ a subject number marker may be optionally suffixed: *-lu* ‘DU’ or *-ni* ‘PL’ (see illustrative text, (6)). Optionally an emphatic *-mayi* may be added (cf. illustrative text examples).
- (d) *-yathimayi* ‘OPT’ marks a situation the speaker feels ought to hold.

Non-finite inflections are added when the main verb is followed by an auxiliary (6.3).

- (e) *-rna* ‘PART’ is the participial inflection used when most past auxiliaries follow.
- (f) *-lha* ‘FUT’ is the future inflection when the future (and one exceptional past) auxiliary follows.

### 6.2.2. Dependent clause

Verbs in dependent clauses take non-finite inflections that are portmanteau morphs coding type of clause linkage (including relative tense) and cross-clausal referential (non-) identity of subjects (switch-reference). There are four clause types; three distinguish same-subject and different-subject reference. The categories are (see also Austin 1981: 187–229):

- (a) **implicated clauses** – these code relative future tense and are typically purposive. The inflections are *-lha* ‘IMPLSS’ for same-subject (see illustrative text, (8), (11), (15), (18)) and *-rnanthu* ‘IMPLDS’ for dif-

- ferent-subject (see illustrative text, (6), (9), (10), (17)). Note that 'IMPLSS' is homophonous with 'FUT' and 'IMPLDS' consists of 'PART' plus 'PROPR'.
- (b) **relative clauses** – these code relative past tense and imperfective aspect. They have both adnominal and adverbial semantic functions. The inflections are *-rna* 'RELSS' for same-subject (cf. illustrative text, (8)), and *-rnanhi* 'RELDs' for different subject (cf. illustrative text, (19)). Note that 'RELSS' is homophonous with 'PART' and 'RELDs' consists of 'PART' plus 'LOC'.
- (c) **sequential clauses** – these code relative past tense and perfective aspect. They have both adnominal and adverbial semantic functions. The inflections are *-rnandrru* 'SEQSS' for same-subject, and *-ni(ngurra)* 'SEQDS' for different-subject. Note that 'SEQSS' consists of 'PART' plus 'ABL' and 'SEQDS' is homophonous with 'NR', optionally followed by *ngurra* which is 'ABL' in neighbouring languages.
- (d) **lest clauses** – express an undesirable state of affairs and are typically used in admonitives. There is no switch-reference and just one suffix *-yathi* (cf. optative in 6.2.1).

### 6.3. Auxiliary verbs

Main verbs may be optionally followed by an auxiliary verb, marking tense (and in one case, aspect). There are five past tense auxiliaries coding relative temporal distance from the present, and one future. Habitual aspect is also marked by auxiliary. Each auxiliary is homophonous with a main verb, and places requirements on the inflection it takes and the non-finite inflection assigned to the main verb which precedes it. These are set out in Tab. 137.5.

auxiliary	homophonous root	function	non-finite inflection	auxiliary inflection
<i>wanthi-</i>	to search	distant past	PART	PRES
<i>wapa-</i>	to go	habitual	PART	PRES
<i>wapa-</i>	to go	intermediate past	PART	PAST
<i>parrha-</i>	to lie	recent past	PART	PAST
<i>wirrh-i-</i>	to enter	yesterday past	FUT	PRES
<i>warra-</i>	to throw	immediate past	PART	PRES
<i>ngana-</i>	to be	future	FUT	PRES

Tab. 137.5: Auxiliary verbs

### 7. Predicate determiner morphology

There are two adverbial predicate determiners: *yaru-* 'like that' (typically with cataphoric reference) and *yani-* 'like this' (typically with immediate context anaphoric reference). These two take the deictic suffixes added to third person pronouns and spatial location nominals (see 4.2.1). e.g. *yani-parrha* 'like.this-THERE (like this)', *yaru-ya* 'like.that-NEAR (like that)' (cf. illustrative text, (12)).

### 8. Particles

Particles are uninflected words with modal functions, having scope over a phrase or whole clause (Austin 1981: 169–177). They include: *wata* 'NEG', *pulu* 'CANNOT', *kanyji* 'CAN', *kara* 'PERHAPS', *maja* 'ALREADY', *ngampu* 'ALMOST', *windri* 'ONLY', *pinti* 'RUMOURED', and *ngarla* 'BUT'.

### 9. Post-inflectional morphology

Post-inflectional suffixes may be attached to words of any class (except co-ordinators and interjections) following their inflection. They code pragmatic functions reflecting informational status of discourse elements (Austin 1981: 177–186), and include: *-rlu* 'STILL', *-yarri* 'RESEMBLE', *-tha* 'OLD.INFORMATION', *-lha* 'NEW.INFORMATION', *-ldrра* 'CONTRAST', *-ku* 'SENSORY.EVIDENCE', *-matha* 'IDENTIFIED', and *-yina* 'TAG' (see illustrative text, (20) for *-tha*).

### 10. Illustrative text

This is an extract from the only mythological text recorded in Diyari. The full text is in Austin (1981: 236–251); this extract covers lines 28–47.

- (1) *Kardi-yali wama*  
sister's.husband-ERG snake(ACC)  
*dukara-yi*  
take.out-PRES
- (2) *Ya malthi-nganka-rna*  
and cool-CAUS<sub>3</sub>-PART
- (3) *Nhungkarni nhuwa yatha-yi*  
3:SG:NF:DAT spouse(NOM) say-PRES
- (4) *Ngali kanku-yali mawa-li*  
2:DE:NOM boy-ERG hunger-ERG  
*ngana-yi*  
be-PRES
- (5) *Thari-yali kalapa-yi*  
youth-ERG answer-PRES
- (6) *Walya karlka-lu-mayi nganthi*  
a.while wait:IMP-DU-EMPH meat(NOM)  
*malthi-rri-rnanthu*  
cool-INCH-IMPLDS
- (7) *Pula kutha-rri-yi*  
3:DU:NOM anger-INCH-PRES
- (8) *Thari-yali nganthi*  
youth-ERG meat(ACC)  
*wardu-wardu-nganka-rna thari-yali*  
RDP-short-CAUS<sub>3</sub>-RELSS youth-ERG  
*wani-yi thayi-lha*  
begin-PRES eat-IMPLSS
- (9) *Ngarda-nhi yingki-yi nhuwa*  
then-LOC give-PRES spouse  
*nhungkarna-li thayi-rnanthu*  
3:SG:NF:DAT-ERG eat-IMPLDS
- (10) *Kaku-yali kanku yingki-yi*  
elder.sister-ERG boy(ACC) give-PRES  
*nganthi nhinha thayi-rnanthu*  
meat(ACC) 3:SG:NF:ACC eat-IMPLDS
- (11) *Ngarda-nhi nhungkarni*  
then-LOC 3:SG:NF:DAT  
*kardi mara*  
sister's.husband(NOM) hand(NOM)  
*wirrhi-yi dukara-lha*  
enter-PRES take.out-IMPLSS  
*nhungkangundrru marna-ndrru*  
3:SG:NF:ABL mouth-ABL
- (12) *Yaru-ka nhawu kanku*  
like.that-TOKEN 3:SG:NF:NOM boy(NOM)  
*mawa-li ngana-yi*  
hunger-ERG be-PRES
- (13) *Kaku-yali ngarda-nhi kurrukurru*  
elder.sister-ERG then-LOC secret  
*yingki-yi nhinha nhangkarni*  
give-PRES 3:SG:NF:ACC 3:SG:F:DAT  
*ngathata kanku*  
younger.sibling boy(ACC)
- (14) *Kardi-yali marla*  
sister's.husband-ERG more(ACC)  
*dukara-yi*  
take.out-PRES
- (15) *Ngarda-nhi thana yirrji-yi*  
then-LOC 3:PL:NOM get.up-PRES  
*wapa-lha*  
go-IMPLSS
- (16) *Ngarda-nhi thana paya kuparrhu*  
then-LOC 3:PL:NOM bird young(ACC)  
*nhayi-yi*  
see-PRES
- (17) *Kardi yatha-yi*  
sister's.husband(NOM) say-PRES  
*kanku-nhi patharra-nhi kari-rnanthu*  
boy-LOC box.LOC climb-IMPLDS
- (18) *Mirri kari-ya-mayi warla-ya*  
above climb-IMP-EMPH nest-ALL  
*kapi mani-lha*  
egg(ACC) get-IMPLSS
- (19) *Ngarda nhawu kanku*  
then 3:SG:NF:NOM boy(NOM)  
*kathi-yi pirta ngarla*  
climb-PRES tree(NOM) but  
*mirri-ri-rna thara-rnanhi*  
above-INCH-PART go.up-RELDS
- (20) *Nhawu mirri-tha thara-inga-rna*  
3:SG:NF:NOM above-OI go.up-PROL-PART  
*tharrhi-yi*  
cont-PRES
- '(1) The brother-in-law took out some carpet snake (from the fire). (2) And cooled (it). (3) His wife said: (4) We two are hungry, the boy and I. (5) The young man answered: (6) You two wait a while for the meat to cool. (7) The two of them became angry. (8) Breaking up the meat the young man began to eat (it). (9) Then (he) gave (some meat) for his wife to eat. (10) The elder sister gave the boy the meat to eat. (11) Then the brother-in-law's hand went in and took (it) out of his mouth. (12) Thus the boy was hungry. (13) Then the elder sister secretly gave (some meat) to him, her younger brother. (14) The brother-in-law took more out. (15) Then they got up to go. (16) Then they saw some young birds (in a nest in a tree). (17) The brother-in-law told the boy to climb the box tree. (18) Climb up above to the nest to get the eggs. (19) Then the boy climbed but the tree went up and up (at the same time). (20) He went up and up at the top (of the tree).'

## 11. Uncommon abbreviations

CHAR	characteristic
EXCESS	excessive
IMPLSS	implicated clause: same-subject
IMPLDS	implicated clause: different-subject
KINPROPR	kinship proprietive
NF	non feminine
OI	old information
PROL	prolative
RELSS	relative clause: same-subject
RELDs	relative clause: different-subject
SEQSS	sequential clause: same-subject
SEQDS	sequential clause: different-subject
VICIN	vicinity

## 12. References

- Austin, Peter (1978), "A Grammar of Diyari, North-east South Australia". Ph.D. thesis, Australian National University
- Austin, Peter (1981), *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press
- Austin, Peter (1990), "Classification of Lake Eyre Languages". *La Trobe University Working Papers in Linguistics* 3, 171–201
- Austin, Peter (1995), "Double Case Marking in Kanyara and Mantharta Languages". In: Plank, Frans (ed.), *Agreement by Suffixaufnahme*. Oxford University Press, 363–379
- Austin, Peter (in prep), Diyari Ethnographic and Narrative Texts. ms
- Austin, Peter & Hercus, Luise & Jones, Philip (1988), "Ben Murray (Parlku-Nguyu-Thangkayi-warna)". *Aboriginal History* 12, 115–188
- Blake, Barry J. (1977), *Case Marking in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies
- Blake, Barry J. (1987), *Australian Aboriginal Grammar*. London: Croom Helm
- Blake, Barry J. (1988), "Redefining Pama-Nyungan: Towards the Pre-history of Australian Languages". *Aboriginal Linguistics* 1, 1–90
- Bowern, Claire (1998), *The Case of Proto-karmic*. B.A. honours thesis, Australian National University
- Bowern, Claire (2001), "Karnic Classification Revisited". In: Simpson, Jane & Nash, David & Laughren, Mary & Austin, Peter & Alpher, Barry (eds.), *Forty Years On: Ken Hale and Australian Languages*. Canberra: Pacific Linguistics, 245–261
- Bowern, Claire & Koch, Harold (2004), *Australian Languages: Classification and the Comparative Method*. Amsterdam: Benjamins
- Dench, Alan C. (1982), "Kin Terms and Pronouns of the Panyjima Language of Northwest Western Australia". *Anthropological Forum* 5, 105–120
- Dench, Alan & Evans, Nicholas (1988), "Multiple Case-marking in Australian Languages". *Australian Journal of Linguistics* 8, 1–47
- Dixon, R. M. W. (1979), "Ergativity". *Language* 55, 59–138
- Dixon, R. M. W. (1980), *The Languages of Australia*. Cambridge: Cambridge University Press
- Heath, Jeffrey & Merlan, Francesca & Rumsey, Alan (1982, eds.), *The Languages of Kinship in Aboriginal Australia*. Sydney: Oceania Publications
- Murray, Ben & Austin, Peter (1981), "Aborigines and Afghans: Diyari Texts". *Aboriginal History* 5, 71–79
- Murray, Ben & Austin, Peter (1986), "Afghan Story" and "Paradise Crossing". In: Hercus, Luise A. & Sutton, Peter (eds.), *This is what Happened*. Canberra: Australian Institute of Aboriginal Studies, 129–132, 293–296
- Nichols, Johanna (1986), "Head-marking and Dependent-marking Grammar". *Language* 62, 56–119
- Plank, Frans (1990), "Suffix Copying as a Mirror-image Phenomenon". *Linguistics* 28, 1039–1045
- Reuther, J. G. (1981), *The Diari*. Canberra: Australian Institute of Aboriginal Studies
- Silverstein, Michael (1976), "Hierarchy of Features and Ergativity". In: Dixon, R. M. W. (ed.), *Grammatical Categories in Australian Languages*. Canberra: Australian Institute of Aboriginal Studies, 112–171

Peter Austin, London (Great Britain)

## 138. Wambon (Awyu)

1. Introduction
2. Verbs
3. Nouns
4. Pronouns
5. Adjectives
6. Demonstratives
7. Numerals
8. Adverbs
9. Nominal case suffixes
10. Connectives
11. Concluding remarks
12. Illustrative text
13. Uncommon abbreviations
14. References

### 1. Introduction

There are about 1,000 languages spoken on New Guinea and surrounding islands; around 250 of these belong to the Austronesian family. The other 750 languages, called Non-Austronesian or Papuan languages, cannot be said to derive from one ancestral language and are organized in more than sixty language families (cf. Wurm 1982; Foley 1986).

Wambon is a Papuan language of the Awyu family of Irian Jaya, the easternmost province of Indonesia on the western half of New Guinea. It is spoken by about 4,000 people in the Upper-Digul area. The southern and western neighbours of Wambon are the Awyu languages Mandobo (Drabbe 1959), Kombai (De Vries 1993 a) and Tsakwambo; the eastern and northern neighbours are the Ok family languages Sait and Muyu.

Wambon is an oral language. About 20% of the speakers can write but they do this only in Indonesian, the national language, taught in school. Local, pidginised varieties of Indonesian, sometimes called Bazaar Malay, are used for intertribal contacts and for contacts with petty traders, government and mission personnel.

Wambon has two main dialects, Digul-Wambon and Yonggom-Wambon (Drabbe 1959). Here, we summarise the morphology of the major word classes of the Digul dialect, starting with the verb, the core of Wambon grammar. For more background information, see De Vries & Wiersma (1992).

### 2. Verbs

#### 2.1. Clause-chaining

Wambon exhibits the New Guinea type of **clause-chaining**. (This type of clause-chaining occurs especially in the grouping of 500 Pap-

uan languages called the Trans-New Guinea phylum, proposed as a genetic grouping in the work of McElhanon et al. 1975. For New Guinea clause-chaining, see Longacre 1972 and Foley 1986.) Longacre (1972: 2) describes New Guinea clause-chains as structures with an engine at the end, and a bunch of cars hooked on, preceding it. This comparison points to the essential division between the fully inflected distinctive verb forms of the last clause, the final verbs, and the less inflected dependent verb forms which occur medially in the chain.

Dependent (medial) verb forms in New Guinea clause-chains typically express **switch-reference** and **temporality**. Temporality markings indicate interclausal relations of sequence and simultaneity of events. Switch-reference markings indicate whether the following clause has the same subject or a different subject from that of the preceding clause in the chain.

The basic opposition in the verb morphology of Wambon is not between final and medial forms but between independent and dependent forms: chain-final clauses must have independent verb forms but independent forms are also frequently used chain-medially. The chain-medial use of independent verbs is switch-reference neutral. Consider the following Wambon clause-chain. (Wambon has a complex morphophonemics; see De Vries and Wiersma (1992: 8–10). Some of the processes occurring in the examples are: intervocalic voicing and fricativisation of /p/ (>/v/) and /k/(>/kh/), intervocalic voicing and lateralisation of /t/ (>/l/), and intervocalic insertion of /n/ as a transitional nasal, glossed as -Ø-):

(1) *Ap nda-khe-n-o*

<i>house</i>	<i>come-NON.1.SG.PRES-Ø-CONN<sub>1</sub></i>
<i>kit-mo-mbel-o</i>	<i>nekho-salip</i>
<i>finished-VR-SS.SEQ-CONN<sub>1</sub></i>	<i>his-wife</i>
<i>lokha-l-o</i>	<i>kap-kap</i>
<i>speak-NON.1.SG.NFUT.DS-CONN<sub>1</sub></i>	<i>man-man</i>
<i>lokha-l-o</i>	<i>lokha-l-o</i>
<i>speak-NON.1.SG.NFUT.DS-CONN<sub>1</sub></i>	<i>lap-ka-lembō</i>
<i>do-go-NON.1.PL.PAST</i>	

'Having come home he informed his wife and she informed the people and they left.'

In this clause-chain there are five verbs. Two of these are independent forms, the chain-fi-

nal *lapkalembo* and the chain-medial switch-reference neutral form *ndakheno*. The other three verbs are dependent forms which can only be used in chain-medial clauses and are inflected for switch-reference and temporality.

Chained clauses are not embedded clauses. The combination of dependency and non-embeddedness is an essential characteristic of clause-chaining (see De Vries 1993 b for clause-combining in Awyu languages). Chaining linkage contrasts with subordinate linkage in Wambon; subordinate clauses take independent verb forms and nominal inflections such as case suffixes and the topic-marking demonstrative *-eve* (see 6); they cannot take the connective *-o* which links chained clauses. Chained clauses cannot take nominal case inflections and the marker *-eve*.

## 2.2. Inflection

As for the inflection of dependent forms, there are two kinds of dependent forms in Wambon: different subject forms and same subject forms.

Same subject forms come in two types, both non-finite, temporality specific forms and temporality neutral forms.

Temporality specific forms consist of the verb stem plus an invariable temporality ending, marking either simultaneity (*-no*) or sequence (*-mbet*) relations between consecutive events. Temporality neutral same subject forms consist of just the verb stem or the verb stem plus the same subject marker *-lo*.

Different subject forms also come in two types. Depending on the tense and mood of the chain-final independent verb, future or non-future different subject forms are used. When the tense of the final verb is future or the mood intentional, the future different subject forms are used. In all other cases the non-future different subject forms are used. Since Wambon has three tenses in independent verbs (past, present, future), the non-future versus future distinction in different subject forms is a tense-neutralisation, making these forms less inflected than the independent forms.

The future/different subject forms consist of future stem plus person-number plus connective *-o*. (Wambon has tense-related suppletive verb stems). The non-future/different subject forms consist of past stem plus *-t* 'DS' plus person-number plus *-o* 'CONN<sub>1</sub>'. In the example text (see section 12) *nembelo* in chain (20) is an example of a same subject sequen-

tial form and *tamjalo* of a different subject form.

Independent verb forms are inflected with suffixes (with the exception of the imperative which is expressed by the prefix *na-*) for tense (present, past, future), number (singular and plural), person (first person versus non-first person), mood (imperative, intentional, hortative), status (counterfactual, conditional, uncertain future) and negation. The tense distinctions of Wambon are defined in terms of both the moment and the day of speaking. Present tense is used for events taking place at utterance-time or before that time on the utterance-day. In past narratives, the present tense may be used for chain-medial independent forms, in a tense-neutral fashion (see e.g. the text in 12). Past tense is used for events taking place before the utterance-day (= yesterday and the time before yesterday). Future tense is used for events after utterance-time. Completive and durative aspects are expressed by periphrastic constructions employing medial verbs or aspectual adverbs. Habitual-iterative aspect is expressed by derived (reduplicated) verbs.

Wambon independent and dependent forms take the same set of person and number endings:

(2)	SG	1:	<i>-ep</i>
		NON.1:	<i>Ø</i>
	PL	1:	<i>-ev-a</i>
		NON.1:	<i>Ø</i> (Present and FUT/DS)
			<i>-e</i> (Past and NFUT/DS)
			<i>-na</i> (Future)

The plural marker in the first person plural is *-a* and the /p/ of the first person marker is fricativised and voiced between vowels.

## 2.3. Word formation with verbalising suffixes

Adjectives can be productively turned into verbs with either *-mo* or *-ke* as verbalisers. For example:

- (3) adjective: *kok* 'ill'
- ke* verb: *koke* 'to be/become ill'
- mo* verb: *kokmo* 'to be/become ill'

There do not seem to be differences in meaning between the *-mo* derived verbs and *-ke* derived verbs of adjectival origin. Once turned into verbs, dependent same subject forms of the verbalised adjectives are used as manner adverbials. For example in (4) the verb *matetmo* is a same subject form of the

verb *matetmo* ‘to be/become good’. This verb is derived from the adjective *matet* ‘good’:

- (4) *Jakhov-e matet-mo ka-lempo?*  
they-CONN<sub>2</sub> good-vr.ss go-NON.1.PL.PAST  
'Did they travel well?'

The verb *kitmo* ‘to be finished’, a *-mo* formation based on the adjective *kit* ‘finished; ready; sufficient’, is frequently used as a periphrastic verb expressing completive aspect. See e.g. the use of this verb in (22) of the example text in 12.

Indonesian verbs are productively integrated into the Wambon lexicon with *-mo* and *-ke*. For example, the Indonesian verb *belajar* ‘to learn’ receives the form *mbelajatmo* ‘to learn’.

Although there are many noun-based *-mo* and *-ke* verbs, there is no longer a productive *-mol-ke* verbalisation of nouns. For example, the verb *takhimo* ‘to buy’ is derived from the noun *takhet* ‘kauri-shell; money’.

From verbs, habitual-iterative verbs are productively derived by reduplicating the first syllable of the verb and adding *-mo* or *-ke*. For example, from *ndat* ‘to listen’ the verb *ndatndatmo* ‘to listen repeatedly/habitually’ is derived.

### 3. Nouns

The morphology of nouns is very simple: nouns may take possessive prefixes, there is a plural formation with kinship nouns and there is derivation of nouns from verbs with the nominaliser *-mop*.

Kinship nouns have plural forms which consist of the stem plus the plural suffix *-ngguye*, for example *mbap* ‘father’ and *mbap-ngguye* ‘father-PL’. Some kinship nouns have irregular plural forms, for example *nambut* ‘my-brother.in.law’ and *na-mbul-alile* ‘my-brother.in.law-PL’.

Occasionally, reduplication is used for plural: *kap-kap* ‘man-man (people)’ (cf. (22)).

As for word formation, nouns can be (productively) derived by adding the nominalizing suffix *-mop* to the future-stem of verbs. For example, from *en-* ‘to eat’, future/past stem *ande-* the noun *ande-mop* ‘eat-NR (food)’ is derived.

### 4. Pronouns

Wambon has a basic set of personal pronominal forms, the set (5), on which emphatic, reflexive and bound possessive forms are based:

- (5) SG 1: *nuk* ‘I’  
2: *nggup* ‘you’  
3: *nekhep* ‘he/she/it’  
PL 1: *nokhop* ‘we’  
2: *nggokhop* ‘you’  
3: *jakhop* ‘they’

The emphatic forms are formed by adding *-otke* to the basic forms of (5). (In this process there occur some morphophonemic changes which we shall not discuss here):

- (6) SG 1: *novotke*  
2: *nggovotke*  
3: *nekhovotke*  
PL 1: *nokhovotke*  
2: *nggokhovotke*  
3: *jakhovotke*

Reflexive pronouns are formed by adding the reflexive suffix *-ta* to the basic forms of (5).

Possessive pronominal forms are bound forms, also based on the forms of (5), except for the first person singular form *na-* ‘my’ and the second person singular form *ngga-* ‘your’ which are special forms:

- (7) SG 1: *na-* ‘my’  
2: *ngga-* ‘your’  
3: *nekho-* ‘his/her/its’  
PL 1: *nokho-* ‘our’  
2: *nggokho-* ‘your’  
3: *jakho-* ‘their’

### 5. Adjectives

Adjectives are words which take the optional intensifier *-mbalin*. For example, *kaimo* ‘good’ and *kaimo-mbalin* ‘good-very (very good)’. Adjectives have plural forms that are rather infrequent. They are formed by reduplication of the first syllable of the adjective, for example *kowanggai* ‘big’ and *ko-kowanggai* ‘PL-big (big)’. There are no morphological superlative and comparative, these notions being expressed by periphrasis.

### 6. Demonstratives

The demonstrative *nombo* ‘this (proximity of speaker)’ contrasts with *ep* ‘that (not in proximity of speaker)’.

When used attributively as in (8), the demonstrative must take the connective *-o*, when used independently they must take *-e*, as in (9). Examples:

- (8) *ev-o lan*  
that-CONN<sub>1</sub> woman  
'that woman'

- (9) *ev-e lan*  
 that-CONN<sub>2</sub> woman  
 ‘that is a woman’

The demonstratives with *-o* may also be used as definite articles. The demonstrative forms with *-e* are also used as spatio-temporal adverbials meaning ‘here/now’ and ‘there/then’ and as markers of textual relations (topic/setting). Especially *ev-e* ‘that-CONN<sub>2</sub>’ functions as a topic marker in which case it cliticises to the setting clause or topic phrase. Consider the following example:

- (10) *Kikhuv-e*  
*Digul-CONN<sub>2</sub>*  
*nde-tkhekkel-ev-e*  
 rise-NON.1.SG.COND-that-CONN<sub>2</sub>  
*ev-e Manggelum*  
 then-CONN<sub>2</sub> Manggelum  
*ko-nok-si-v-a.*  
 go-NEG-NEG.INTENT-1-PL  
 ‘If the Digul rises, then we do not want to go to Manggelum.’

## 7. Numerals

The Wambon number system is a **body-part tally system** of a type common in New Guinea. The nouns denoting body-parts have the numbers as their secondary meanings. Thus e.g. *javet* has two meanings: ‘upper arm’ and ‘nine’. In the numeral use of the words, conventional gestures must accompany the use of the number words: the fingers are bent and the parts of the arm and head are touched with the outstretched middle finger and/or index finger (cf. Tab. 138.1):

1	<i>sanop</i>	little finger
2	<i>sanopkunip</i>	ring finger
3	<i>takhem</i>	middle finger
4	<i>hitulop</i>	index finger
5	<i>ambalop</i>	thumb
6	<i>kumuk</i>	wrist
7	<i>mben</i>	lower arm
8	<i>muyop</i>	elbow
9	<i>javet</i>	upper arm
10	<i>malin</i>	shoulder
11	<i>ngokmit</i>	neck
12	<i>silutop</i>	ear
13	<i>kelop</i>	eye
14	<i>kalit</i>	nose

Tab. 138.1: Wambon numeral corresponding body-part

The nose is the turning point, after which counting goes down again via the right-hand side of the body, by prefixing *em-* ‘the other side’ to the body-part/number words. Thus, 15: *em-kelop* ‘eye of the other side’; 16: *em-silutop* ‘ear of the other side’.

This continues down again until the little finger of the right side of the body is reached, the highest number of the system being 27: *emsanop*.

The body-part/numbers are integrated in noun phrases as numeral modifiers by adding the comitative suffix *-kup* ‘with; also’. For example:

- (11) *ap hitulop-kup*  
 house (index.finger)four-with  
 ‘four houses’

The number words *sanop* ‘little finger; one’, *sanopkunip* ‘ring finger; two’ cannot be used as numeral modifiers in noun phrases. Instead, the only two “real” numerals are used:

- (12) (a) *ap ndominuk*  
 house one  
 ‘one house’  
 (b) *ap ilumo*  
 house two  
 ‘two houses’

## 8. Adverbs

Adverbs are a very marginal word category in Wambon since *-mo* and *-ke* derived verbs express adverbial meanings of manner (see 2.3) and demonstrative forms have spatio-temporal adverbial meanings. There is a small closed class of invariable adverbs including *sonop* ‘again’, *ndo* ‘just’, *osak* ‘again; also’ and *wasi* ‘quickly; already’. These adverbs occur clause-initially but also pre-verbally. Consider the following example:

- (13) *Hitulop-e osi ka-levambo.*  
 Thursday-CONN<sub>2</sub> again go-1.PL.PAST  
 ‘On Thursday we went again.’

*Hitulop* ‘indexfinger/four’ is also used to denote the fourth day of the week. The week starts with *sanop* ‘little finger/Monday’, etc.

## 9. Nominal case suffixes

Wambon has three nominal case suffixes: *-ka*, *-sikhi* and *-kot*. The element *-ka* is suffixed to inanimate circumstantial adjuncts such as time, place and instrument:

- (14) *Kikhup-ka okima-levambo.*  
 Digul-in take.bath-1.PL.PAST  
 ‘We took a bath in the Digul.’

In (15) *-ka* (*-ngga* after morpheme-final vowels) marks an adverbial time clause:

- (15) *Nukh-e ande-lepo-ngga ev-o*  
 I-CONN<sub>2</sub> eat-1.SG.PAST-when that-CONN<sub>1</sub>  
*kav-e nde-tmbo.*  
 man-CONN<sub>2</sub> come-NON.1.SG.PAST  
 ‘When I ate, that man came.’

The suffix *-ka* forms complex locational relators with the nouns *wamip* ‘inside’, *hitop* ‘bottom’, *kuk* ‘front’, *linggit* ‘back’, *palip* ‘top’. For example:

- (16) *ap-wamip-ka*  
 house-inside-in  
 ‘in the house’

The nominal case suffix *-si(khi)* is a nominal case marker with cause, reason and purpose phrases but it marks also human “goals” such as addressed (with speech act verbs), recipient and beneficiary. Consider the following example:

- (17) *Keno-sikhi nda-khe?*  
 what-for come-NON.1.SG.PRES  
 ‘Why does he come?’

The suffix *-kot* marks cause and reason:

- (18) *Nukh-e jambolok-mo-knde-p-kot*  
 I-CONN<sub>2</sub> ill-VR-PRES-1.SG-because  
*nekhev-e nda-no-khe.*  
 he-CONN<sub>2</sub> come-NEG-NON.1.SG.PRES  
 ‘Because I am ill, he does not come.’

## 10. Connectives

An interesting aspect of Wambon morphology is the set of very frequent connectives *-a*, *-o*, and *-e* which create syntactic cohesion in various syntactic domains. In the framework of this article we cannot discuss those functions and we just give a number of examples of some of the functions. Consider (19):

- (19) *Nuk oy-a*  
 I pig-CONN<sub>3</sub>  
*temke-n-o*  
 shoot.NON.1.SG.PRES-Ø-CONN<sub>1</sub>  
*kav-e khetak-nok-nde-p.*  
 man-CONN<sub>2</sub> see-NEG-PRES-1.SG  
 ‘I do not see the man who shot the pig.’

In (19) the relative clause (*oya temke*) is linked to its syntactic head noun *kave* ‘man’

by the connective *-o*. The first nominal phrase of every relative clause receives the relative clause delineator *-a*. The connective *-e* also occurs in (19), linking the nominal phrase with *kave* as headnoun to the verb, the head of the clause. The connective *-e* may replace (but not co-occur with) the nominal case suffixes (cf. 9).

The connective *-o* is used within noun phrases to connect modifiers to the headnoun. But *-o* is also used to link chained clauses (see (20)–(22)) and to coordinate nouns.

## 11. Concluding remarks

Wambon represents the morphological type which is sometimes called the New Guinea type of clause-chaining. In this type the verbal morphology is complex, with dependent and independent forms. The dependent forms express switch-reference and temporality and are used as the non-embedded “cars” hooked to the “engine”, the chain-final verb. The morphological differentiation between independent forms and dependent verb forms is small in Wambon, compared with many other clause-chaining languages of New Guinea.

The nominal morphology is simple; nouns are not inflected, apart from a marginal plural formation with kinship nouns. Adjectives have no morphological comparative or superlative forms. The demonstrative system is a simple two-way deictic system (+/- proximate to speaker) but demonstrative forms have important secondary functions, marking definiteness, topic phrases and adverbial setting clauses. The numerals of Wambon exemplify the body-part tally system which is widespread in New Guinea.

## 12. Illustrative text

- (20) *Savanop ko-khe-n-o*  
 Savanop go-NON.1.SG.PRES-Ø-CONN<sub>1</sub>  
*kavat hetokho-p ne-mbel-o*  
 trap see-1.SG.INTENT say-ss.SEQ-CONN<sub>1</sub>  
*kono hetakhe-n-e*  
 and see.NON.1.SG.PRES-Ø-CONN<sub>2</sub>  
*lava-tmbo-nggelo odo uto*  
 take-NON.1.SG.PAST-and next go.in.ss  
*tamja-l-o odo*  
 shoot-DS.NFUT.NON.1.SG-CONN<sub>1</sub> and  
*oy-e ndano in-ande-tmbo.*  
 pig-CONN<sub>2</sub> next hit-eat-NON.1.SG.PAST

- (21) *Jama-l-o*  
thus.do-NON.1.SG.NFUT.DS-CONN<sub>1</sub>  
*et-mbel-o ap*  
leave-SS.SEQ-CONN<sub>1</sub> house  
*nde-tmbo*  
come-NON.1.SG.PAST
- (22) *Ap nda-khe-n-o*  
house come-NON.1.SG.PRES-Ø-CONN<sub>1</sub>  
*kit-mo-mbel-o nekho-salip*  
finished-VR-SS.SEQ-CONN<sub>1</sub> his-wife  
*lokha-l-o*  
speak-NON.1.SG.NFUT.DS-CONN<sub>1</sub>  
*kap-kap*  
man-man  
*lokha-l-o*  
speak-NON.1.SG.NFUT.DS-CONN<sub>1</sub>  
*lap-ka-lembo*  
do-go-NON.1.PL.PAST

(20) ‘Savanop went to see his trap and he saw that he had trapped (a pig) and he approached and shot it but the pig attacked him. (21) Thus it (the pig) did and he (Savanop) left and went home. (22) Having come home he informed his wife and she informed the people and they left.’

This short text consists of three clause-chains, connected by means of **generic verb linkage**, linking (20) and (21), and **tail-head linkage**, linking (21) and (22) (see De Vries 1993 a: 69–75 for these common Papuan linkage types). Tail-head linkage is a recapitulative cohesion device by which the chain final clause is repeated as the chain initial clause of the next clause-chain.

The use of the verb ‘to say’ in (20) is part of the **quotative** construction which Wambon, like many other Papuan languages, employs to express purpose and intention in the form of “inner speech” (cf. De Vries 1993 a: 91–104).

The connective *-o* is used to link the chained clauses in (20)–(22). There is one clause (in (20)) having *-e* as connective clitic (*hetakhene*) signalling the nominal status of the perception clause, a strategy occurring in several Awyu languages with perception verbs which can be rendered literally as ‘(what) he saw (was that) he had trapped (a pig)’.

The verb *inandetmbo* in (20) is an example of a serial root construction in which two verb roots are juxtaposed to form one complex predicate which is the head of one clause. This **serialisation** of verb roots is a productive morpho-syntactic process in Wambon as in many other Papuan languages.

The text illustrates the central role of verbal morphology in Wambon in both event

and participant cohesion. The participants in the text, after initial introduction by nouns, are not tracked by pronouns or nouns but by agreements and switch-reference. Not surprisingly, the verb has functions in Wambon which are carried out by other word classes in other languages, like adverbs and conjunctions. Conjunctions like *kono* and *ndano* in (20) at closer inspection turn out to be verb forms: they are same subject simultaneity forms of the motion verbs *nda* ‘to come’ and *ko* ‘to go’. Such forms are used in Awyu languages as discourse relational verbs expressing speaker continuation.

### 13. Uncommon abbreviations

INTENT	intentional
NON.1	second/third person

### 14. References

- Drabbe, Peter (1959), *Kaeti en Wambon: Twee Awju-dialecten*. The Hague: Nijhoff
- Foley, William A. (1986), *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press
- Longacre, Robert E. (1972), *Hierarchy and Universality of Discourse Constituents in New Guinea Languages: Discussion*. Washington: Georgetown University Press
- McElhanon, K[arl] & Voorhoeve, C[ornelius] L. & Wurm, S[tephen] A. (1975), “The Trans-New Guinea Phylum in General”. In: Wurm, Stephen A. (ed.), *New Guinea area Languages and Language Study, Vol. I*. Canberra: A.N.U. Press (Pacific Linguistics C-38), 299–322
- Vries, Lourens J. de (1993 a), *Forms and Functions in Kombai, an Awyu Language of Irian Jaya*. Canberra: A.N.U. Press (Pacific Linguistics. Series B, 108)
- Vries, Lourens J. de (1993 b), “Clause Combining in Oral Trans-New Guinea Languages”. *TEXT* 13, 481–502
- Vries, Lourens J. de & Wiersma, Robinia (1992), *The Morphology of Wambon of the Irian Jaya Upper-Digul Area*. Leiden: KITLV
- Wurm, Stephen A. (1982), *Papuan Languages of Oceania*. Tübingen: Gunter Narr

Lourens de Vries, Amsterdam  
(The Netherlands)

## 139. Turkana (Niloctic)

1. The verb
2. The noun and other categories
3. Category shift
4. The morphology-phonology interface
5. The morphology-syntax interface
6. Illustrative text
7. Uncommon abbreviations
8. References

Turkana ([ŋáttürk(w)ánà]), is spoken in northwestern Kenya by approximately 270,000 people. It belongs to a cluster of closely related dialects spoken by distinct ethnicities in bordering areas of Ethiopia, Kenya, Sudan, and Uganda which are generally referred to as the Teso-Turkana group. Together with the Lotuko-Maa group and the Bari group, Teso-Turkana forms the Eastern Niloctic branch of Niloctic, one of the major subgroupings, in terms of number of languages, within the Nilo-Saharan stock. Although Turkana is used in written form in primary education in Kenya, there is minimal literacy and only a fairly recent tradition of writing. The earliest studies of the language date back to the 1920s; more recent studies include those of Heine (1980), Best (1983), Dimmendaal (1983), and Barrett (1988; 1991).

Section 1, “the verb”, addresses verb structure and a number of more general morphological properties of Turkana. Section 2, “the noun and other categories”, provides an introduction to structural features of nouns and other word-classes. Section 3, “category shift”, introduces the reader to the relational properties between categories. Finally, sections 4 and 5, “the morphology-phonology interface” and “the morphology-syntax interface”, deal with the interaction of morphology with phonology and syntax. The text in section 6 presents a brief sample of transliterated, connected Turkana speech.

### 1. The verb

Many investigators have remarked on the complexity and richness of the verb systems of Niloctic languages. For example, in their description of the Southern Niloctic language Nandi Creider & Creider (1989: 75) observe that “the complexity of the verbal word in Nandi is, for the non-Nandi staggering”.

Similar observations can be made for Turkana. There are at least three reasons for this complexity. Verb roots may be preceded by markers for negation, person (subject and/or object), and causative (in that order), and be followed by derivational suffixes, as well as aspect, number and voice markers, as shown below. Also, a number of distinct categories are involved in determining the shape of a single formative, i.e. there is a considerable degree of morphological allomorphy. In addition, there is morphophonemic alternation, due to vowel harmony, fusion of vowels, and tone rules.

Verb roots in Turkana fall into two **morphological classes**, here called class 1 and class 2. On the basis of morphophonemic alternations triggered by various prefixes added to verb roots, it can be deduced that class 2 verbs have an initial high front vowel *i* or *ɪ* (depending on vowel harmony) whereas class 1 verbs do not: e.g. class 1 *-nam* ‘eat (solids)’, class 2 *-imoj* ‘eat (fluids)’. (In the examples below, low tones are marked with ` , whereas high tones are left unmarked.) Such a verb root (or stem, i.e. a root plus derivational extension) cannot constitute a word by itself in Turkana, except as the imperative form of a verb like ‘get, fetch’ which is *ɔm* ‘get (it)!’. This situation shows iconicity in that the semantically least marked forms (presumably because of their high frequency) are also the morphologically least marked ones. Regular verbs are marked in the imperative by the prefix *ta-/tɔ-/to-* (depending on vowel harmony) for class 1, and *k-* for class 2 verbs: *ta-nam* ‘eat (solids)!', *k-imoj* ‘eat (fluids)!'. For other moods either an infinitive marker occurs or a person marker is required as an inflectional morpheme to complete a word:

- (1) (a) *aki-nam*  
INF-eat<sub>1</sub>,  
‘to eat (solids)’
- (b) *à-nam* [àŋám]  
1-eat<sub>1</sub>,:PAST  
‘I have eaten (solids)’
- (c) *ak-imoj*  
INF-eat<sub>2</sub>,  
‘to eat (fluids)’
- (d) *à-imoj* [èmój]  
1-eat<sub>2</sub>,:PAST  
‘I have eaten (fluids)’

Fusion as a morphological technique, as with the indicative mood form ‘I have eaten (fluids)’ above, complicates the segmentation of surface words into exhaustive strings of morphemes in Turkana, but this of course is a more general property of fusional (or inflecting) languages.

**Derived verbs** (stems) can be formed either by reduplicating the verb root or by expanding the root with a derivational suffix in order to express categories like habitual, ventive, itive, dative, inchoative, affective, or factitive. Each of these sememes is realized by a morpheme that has phonologically and morphologically conditioned allomorphs. For example, the form of the habitual marker is variously *-een-*, *-aan-*, *-oon-* or *-een(i)*. Roots or stems can also be preceded by a causative marker whose various allomorphs include *-ita-* and *-is-* among others. The following infinitive verb forms illustrate the causative as well as some of the derivational suffixes that can be added to the class 1 verb *-nam* and the class 2 verb *-imoj* introduced above:

(2) Causative	<i>ak-ita-nam</i> INF-CAUS-eat <sub>1</sub> , ‘to feed (solids)’
	<i>ak-is-imoj</i> INF-CAUS-eat <sub>2</sub> ‘to feed (fluids)’
Ventive	<i>a-nam-ōn(i)</i> INF-eat <sub>1</sub> -VEN ‘to wear out’
Itive	<i>a-nam-ār(i)</i> INF-eat <sub>1</sub> -IT ‘to cheat; to wear out’
Intensive	<i>ak-imoj-o-moj</i> INF-eat <sub>2</sub> -Ø-RDP ‘to rinse the mouth’
Habitual	<i>ak-imuj-een(i)</i> INF-eat-HABIT ‘to eat <sub>2</sub> regularly’
Dative	<i>ak-imoj-akin(i)</i> INF-eat <sub>2</sub> -DAT ‘to start eating’

Semantic labels such as “ventive” or “itive” refer to proto-typical meanings of movement towards or away from a deictic center, but metaphorical extensions are frequently observed, as in the examples above. Also, the abovementioned suffixation and reduplication processes are used productively (i.e. they can be applied to new forms, such as borrowings from Swahili), but their actual occurrence is sometimes precluded on semantic grounds. The ventive or itive markers, for ex-

ample, can be used in combination with the intensive (i.e. reduplicated) form of the verb *-imoj*, but not with the corresponding root form of the same verb: *ak-imojomoj-ōn(i)* (‘flush out the mouth’, or *ak-imojomuj-aär(i)* ‘1. to rinse; 2. to eat after a long time of fasting’, but not *\*ak-imoj-ōn(i)*).

The Turkana verb is conjugated using inflectional markers for the first, second, and third **person**; these precede the root or stem. **Number** suffixes are usually only found in the second and third person plural, thereby distinguishing these from the corresponding singular form. Number-marking inflections follow derivational suffixes as well as aspect markers. The form of the person and number markers depends on the phonological environment, as well as on the paradigm in which a verb occurs. For example:

- (3) (a) *i-nam-ar-it'*  
2-eat<sub>1</sub>-IT-ASPECT  
‘you (sg.) are cheating’
- (b) *i-nam-ar-it-o'*  
2-eat<sub>1</sub>-IT-ASPECT-PL  
‘you (pl.) are cheating’
- (c) *i-nam-a'*  
2-eat<sub>1</sub>-IT  
‘you (sg.) have been cheating’
- (d) *i-nam-ā-si'*  
2-eat<sub>1</sub>-IT-PL  
‘you (pl.) have been cheating’

The Turkana verbal system has the additional dimension of **object marking**. A first or second person object is marked by the prefix *k-*, which precedes the subject-marking prefix; if there is a first or second person object, but the subject is a third person, the subject prefix is absent, and the corresponding marker for first or second person object occurs in the position where the third person marker would occur otherwise. This kind of system, which is more common in Nilo-Saharan, occasionally results in formal neutralisations as well as in neutralisations of inflectional features, i.e. in syncretism:

- (4) (a) *k-à-mm-à*  
PAST-I-love-VOICE  
1. ‘I love you (sg./pl.)’
- (b) *k-à-mm-à*  
PAST-me-love-VOICE  
2. ‘(s)he loves me’

The verb forms above also illustrate the relational category of **middle voice**, which contrasts with the active and the passive (or im-

personal active) voices of Turkana. Middle and passive voice are expressed by a series of suffixes following a root or stem. Aspect markers follow the root or stem, but they precede voice and number markers. The marker for passive voice in Turkana is found in combination with transitive as well as intransitive verbs (see also Dimmendaal 1991):

- (5) (a) *ɛ-nam-i-o'*  
           3-eat<sub>1</sub>-ASPECT-VOICE  
           '(s)he/they is/are being eaten'  
 (b) *ɛ-ləs-i-o'*  
           3-go-ASPECT-VOICE  
           'one is going; people are going'

The **Aktionsart** inherent to verbs also plays an important role with aspect marking (cf. Art. 109), as shown by the following examples with the progressive aspect marker *-it*:

- (6) *à-ŋɔp-it'*  
       3:PAST-disappear-ASPECT  
       '(s)he disappeared (and reappeared)'  
 (7) *à-twàn-it'*  
       3:PAST-die-ASPECT  
       '(s)he almost died'

As opposed to **aspect** marking (and other inflectional and derivational processes described above), the expression of tense is not linked to a sequential slot in a series of segmental morphemes; the binary distinction between [+past] and [-past] tense in Turkana is expressed by tonal inflection of a verb form as a whole; in addition, the third person subject prefix varies between *a-* for [+past] and *ɛ-/e-* for [-past]. Compare:

- (8) (a) *ɛ-lòs-i'*  
       3-go-ASPECT  
       '(s)he is going'  
 (b) *à-lòs-i'*  
       3:PAST-go-ASPECT  
       '(s)he went'

With **mood** distinctions in Turkana (e.g. indicative versus subsecutive, which is used in narrating a story) it becomes even more evident that not all morphosyntactic or semantic oppositions are necessarily identifiable with a single morphological slot or position class; instead, the word as a whole may operate as the signaling unit. The second person subject prefix, for example, is *i-/i-* in the indicative and *ta-/tɔ-/to-* in the subsecutive. But the formal expression of tense or aspect also varies across moods: thus, imperfective past

indicative, *i-lòs-i* 'you went' contrasts with narrative *to-lot* 'then you went'. The indicative mood – as opposed to all other moods – constitutes the core of the paradigmatic structure; it shows the widest range of morphological distinctions expressed on the verb. The subsecutive, by contrast, lacks tense marking, and only first, but not second person objects are expressed on the verb. The indicative has also formed the basis for certain analogical realignments in the history of Teso-Turkana (Dimmendaal 1991).

## 2. The noun and other categories

As with verb roots, most nominal roots in Turkana cannot be used alone as independent words. For example, with very few exceptions, lexical nouns are inflected for **gender** by a prefix (although gender marking is absent from verbs and pronouns). A different set of gender markers is used in the singular and the plural, thus indicating gender and number at the same time. In addition, a noun stem is inflected with a plural suffix when the singular is the basic form, or with a singulative suffix when the plural is basic; alternatively, there is sometimes replacement (Dimmendaal 1983: 208–258).

singular	plural
<i>a-mok-àt</i>	<i>ŋa-møk'</i>
<i>e-com'</i>	<i>ŋi-com-in'</i>
<i>e-wùko'</i>	<i>ŋi-wuko-i'</i>
<i>i-moru'</i>	<i>ŋi-mòr(u)</i>
<i>a-tera-n(i)</i>	<i>ŋa-tera-k'</i>

Tab. 139.1: Number marking

Overt **singulative** marking – used by taking plural forms as basic (as with 'sandal, shoe' above) and inflecting the singular – is common in Turkana when a noun refers to entities which naturally occur in pairs or large numbers. This kind of number inflection is widespread in the Nilo-Saharan family (Dimmendaal 2000). Not all nouns in Turkana are overtly marked for number by one of the above suffixes. Some nouns are either inherently singular or plural (as shown by their gender prefix) and so never alternate; such nouns are typically, although not exclusively, mass nouns or collective nouns, e.g. *a-sinyon(i)* 'sand', *ŋa-kipl* 'water'. There is also a group of nouns showing alternates for number that are not marked by segmental

morphemes. In the form for ‘pebble’ above, the voicing of the final vowel in the singular is the result of tonal inflection. Word-final vowels are devoiced before a pause, unless they are followed by a floating low or high tone (Dimmendaal & Breedveld 1986). Thus, in the case of ‘pebble’ the voicing of the final vowel indicates the presence of a word-final floating tone. The floating tone and hence voicing of the final vowel is correlated with a particular tonal configuration which is one of the exponents of number inflection in Turkana.

As the examples above illustrate, the gender prefixes show syncretism between masculine and neuter in the plural. Kinship terms and names of various animals take different sets of gender prefixes: *na-* (fem.), *lò-/lo-* (masc.), preceded by *ta-* in the plural. Gender prefixes also play a role in derivation; compare *e-wìko* ‘lung’ with *lò-wìko* ‘pneumonia’, or *e-com* ‘baboon’ with *i-com* ‘young/small baboon’. Synchronously these gender prefixes do not bear any relation to definiteness or referentiality in Turkana. But their historical link with demonstratives is obvious. Compare, for example, the feminine marker *na-* with the [+near] feminine demonstrative form *na*, or the masculine marker *lò-* with the [+near] masculine demonstrative form *lo*.

There are relatively few **adjectives** in Turkana. Their number marking resembles that of nouns, since it makes use of singular and plural suffixes, as well as replacement:

singular	plural	
<i>-kitèt-et</i>	<i>-kitèt(i)</i>	‘new’
<i>-lèm(u)</i>	<i>-lemu-yèk</i>	‘hornless’
<i>-yoro-t'</i>	<i>-yoro-k'</i>	‘old (of inanimate things)’

Tab. 139.2: Adjectives

Turkana clearly avoids monomoraic words (see also 4). Thus, minor categories like adjectives, numerals, adverbs of time, place or manner and prepositions all consist of at least two moras (this often coinciding with two syllables):

- (9) numeral: *a-pey'* ‘one’  
               adverb: *inaa'* ‘over there’  
                     *mòyi'* ‘tomorrow’

There are a few exceptions to this prosodic principle, however:

(a) There are monomoraic prepositions like *a-*, *ka-*, and *ni-* which always cliticize onto the nouns they govern, however. Prepositions that do not cliticize onto the noun they modify are historically or synchronically derived from related nouns, and consequently consist of at least two moras.

(b) Ideophonic adverbs, expressing the perception or emotional reaction to size, speed, colour, etc. may be monomoraic, forming a separate prosodic constituent of their own. An example:

- (10) *iryonò                    cuc'*  
          3.be.black   IDEO  
          ‘it is black as coal’

(c) Particles expressing attitudes of the speaker towards propositions in an utterance. These attitude markers, which play an important role in discourse, are particularly frequent with illocutionary acts such as questions and imperative constructions (Dimmendaal 1996; see also the illustrative text in 6).

### 3. Category shift

Nouns in Turkana (as in most languages) may be said to typically express referents, whereas the other major word class in the language (that of verbs) typically expresses predication or events having a temporal dimension (cf. Art. 94; Art. 104). Thus, all conjugated verbs can distinguish between [+past] and [−past] tense. A comparison between stative verbs and nominal modifiers such as true adjectives (as a part of speech) in Turkana suggests that the latter typically describe inherent, permanent, and often irreversible states or property concepts that are not contingent on time. For example, ‘(being) small’ is categorized as a stative verb, whereas ‘(being) bald, without horns’ is categorized as an adjective in Turkana.

In order to modify the syntactic behaviour and semantic characteristics of such parts of speech, category shift (from noun to verb, verb to noun etc.) can be applied. Thus, nouns as well as adjectives can be changed into verbs in order to express a stative or dynamic event, via the addition of a habitual (-*aan*-) or dative (-*kin*-) verb extension:

- (11) *ɛ-mɔɪt                    ‘enemy’*  
          *ɛ-mɔɪt-aan-à            ‘it/(s)he is hostile’*

- lèm(u) ‘hornless’  
 è-lèmw-aan-à ‘it is hornless’

There are few cases in Turkana where the direction of the derivational relation cannot be determined, since both of two related forms require a morphological marker, as in *e-mur-on(i)* ‘diviner’ vs. *a-mur-òrè* ‘to divine’.

There are various processes whereby verbs can be changed into nouns in Turkana. Such nouns express various semantic roles, e.g. agent, instrument, location, or an abstract noun, as the following examples with the verb root *-nam* illustrate:

- (12) (a) *e-ka-nam-an(i)*  
          M.SG-AGNR-eat<sub>1</sub>-AGNR  
          ‘glutton; s. o. who gets dowry of cattle’
- (b) *a-nam-èt'*  
       F.SG-eat<sub>1</sub>-INSTNR  
       ‘instrument for eating; generation’
- (c) *e-nam-rr'*  
       M.SG-eat<sub>1</sub>-SG  
       ‘salary; wage’
- (d) *ŋi-nam-èn*  
       M.PL-eat<sub>1</sub>-PL  
       ‘food; eatables’

Nominalized verbs may “inherit” the verbal causative marker, derivational suffixes, or the aspect marker *-rl/-it*, but tense, voice, person marking or mood distinctions do not occur in nominalized verbs. An example:

- (13) *ŋa-k-is-icwom-aan-akin-èt'*  
       F.PL-Ø-CAUS-spear-HABIT-DAT-INTS  
       ‘mortuary ceremony (lit. cause to spear often for somebody [in order to wipe away tears of bereavement])’

(The prefix *k-* above is part of the nominalization but does not seem to have a separate function synchronically.)

Verbs are changed into **nominal modifiers**, in order to describe a property attributed to a noun (Dimmendaal 1983: 318–332). Some of these attributive (“adjectival”) verb forms, as well as adjectives can be used adverbially by prefixing *ni-/ni-*:

- (14) (a) *-lɔn-a-n(i)*  
          far-ADJR-SG  
          ‘distant (attribute)’
- (b) *ni-alɔnan(i)*  
       ADVR-distant  
       ‘distantly’

- (c) *-lèm(u)*  
       hornless  
       ‘hornless (adjective)’
- (d) *ni-lèm(u)*  
       ADVR-hornless  
       ‘blindly, without proper thinking’

A number of **prepositions** in the language can be shown to be derived from, or related to, nouns historically; these prepositions sometimes lack the nominal gender prefix found in the corresponding noun; this derivational relationship is not productive.

- (15) *kwap'* ‘under’  
       *a-kwap'* ‘land’  
       *nà-àbor'* ‘behind’  
       *a-àbor'* ‘buttocks’

There does not seem to be a productive derivational process in the language whereby adverbs are changed into verbs, or adverbs into nouns, except for a few ideophonic adverbs (expressing a sensory experience), a number of which can be changed into class 2 verbs: *butubut'* ‘sound of gurgling liquid’, and *akibutubut'* ‘to empty a bottle containing a liquid’.

Finally, there are recursive elements which suggest a derivational relation historically but which lack any function or meaning synchronically in the language. For example, the initial high front vowel of class 2 verbs (cf. 2), or intrusive elements in nouns such as *-si-/si-* are “empty morphs” synchronically, but probably were separate morphemes historically:

- (16) (a) *e-si-dbɔŋɔrɔr(i)*  
          M.SG-Ø-elbow  
          ‘elbow’
- (b) *e-si-girigir(i)*  
       M.SG-Ø-crest  
       ‘crest of baboon’

**Compounding** of roots or stems is not used as a productive type of word formation in Turkana; it is restricted to the formation of names of geographical locations, animals, or persons. The corresponding meaning of such nouns is often non-compositional (whereas vowel harmony does not apply between juxtaposed roots):

- (17) (a) *nà-morù-toŋa'*  
          ‘Namorutunga (name of a village)’
- (b) *lò-kòri-ŋaj'*  
       ‘Lokorinyang (personal name of someone owning an ox with a yellow, patched fur)’

As may be expected, absence of compounding as a productive word formation process corresponds to a large set of lexical forms, whose meaning is rendered by compounds in, for example, English. This can be illustrated with the following (stative) verbs, all related to the bovine idiom:

- (18) (a) *-lɪŋà*  
‘black headed and white bodied’
- (b) *-ŋòlè*  
‘white headed and black bodied’
- (c) *-moray'*  
‘fat-necked or fat-headed’

#### 4. The morphology-phonology interface

For several bound morphemes above, at least two allomorphs exist which differ in the quality of their vowels. This is due to a **vowel harmony** system which is based on the position of the tongue root – a common feature among African languages. In general, all vowels within a word belong to one of two sets of vowels: a set of [–Advanced Tongue Root] vowels, containing *i*, *ɛ*, *a*, *ɔ*, *ʊ*, and a set of [+Advanced Tongue Root] vowels, comprising *i*, *e*, *o*, *u*. As a corollary, vowels in bound morphemes alternate between *i* and *i*, *ɛ* and *e*, etc. For example, the neuter gender prefix is *i-* before a root with [–ATR] vowels, and *i-* before an [+ATR] root. The feminine gender prefix is always *a-*, however, and does not alternate with other vowels in prefixes or roots. In suffixes *a* alternates with *ɔ* or *o*.

A number of suffixes only occur with [+ATR] vowels; for example, the passive marker has a number of morphologically conditioned allomorphs, e.g. *-i* or *-o*, but they all contain [+ATR] vowels. These suffixes cause the preceding [–ATR] vowels to shift to [+ATR]:

- (19) *è-gyel-i-o'*  
3:PAST-buy-ASPECT-VOICE  
(s)he/it is bought; they are bought’

Because of this category shift caused by [+ATR] suffix vowels, it can be concluded that alternating affixes either contain [–ATR] vowels underlyingly which shift to their corresponding [+ATR] counterpart in the environment of an [+ATR] vowel, or that alternating affixes are unspecified for the feature [ATR] (Noske 1996; 2000).

Apart from vowel harmony, the **syllable** is the primary factor in conditioning allomorphy in Turkana. The infinitive prefix of class 1 verbs has a short allomorph *a-* and a longer form, *aki-/laki-*. The distribution of these allomorphs is governed by the number of moras in the stem. If the stem is monomoraic, the allomorph *aki-/laki-* occurs; if the verb is bimoraic the short variant *a-* is found.

- (20) (a) *aki-nam'*  
INF-eat<sub>1</sub> ‘to eat’
- (b) *a-ŋaam'*  
INF-yawn ‘to yawn’

If a derivational suffix is attached to a monomoraic verb stem, the short variant of the infinitive prefix occurs, as in *a-nam-är(I)* ‘to acquire, absorb, win’. The distribution of these allomorphs suggests that Turkana places a restriction on the minimal size of a verb stem or root: Verb stems have to be at least bimoraic. If a root is monomoraic *-ki-/ki-* is inserted to satisfy the bimoraic stem template (Noske 1991).

The **reduplicative system** is another area of morphology in which the segmental shape of a morpheme is adapted to a template. Turkana has two verb forms which are constructed by reduplicating the root: The intensive and the resultative. The intensive is formed by complete reduplication of the root:

- (21) (a) *a-pet-e-pet'*  
INF-kick-Ø-RDP  
‘to kick repeatedly’
- (b) *a-ɛn-ɛ-ɛn'*  
INF-tie-Ø-RDP  
‘to fasten’

An **epenthetic vowel** occurs between the root and the reduplicative suffix which assimilates completely to the surrounding vowels. In these examples vowel epenthesis seems to be motivated by the syllable structure constraints of Turkana. Without the epenthetic vowel the reduplicated string (\**a-petpet*) could not be syllabified. (Turkana bars obstruents in the syllable coda word-internally). However, even if the verb ends in an open syllable, an epenthetic vowel occurs between the root and the reduplicative suffix:

- (22) (a) *a-da-i-da'*  
INF-thresh-Ø-RDP ‘crumple’
- (b) *ak-ilō-i-lo'*  
INF-defeat-Ø-RDP ‘shake’

In these examples vowel epenthesis does not result from syllabification, since reduplication of an open syllable gives rise to a well-formed syllabic sequence (e.g. *a-dada'*). In addition, in the case of *ak-ilō-i-lō*, the epenthetic vowel fails to assimilate to the surrounding vowels. If assimilation took place the epenthetic vowel would be syllabified with the preceding open syllable. This shows that the epenthetic vowel is not motivated by the syllable structure constraints of Turkana, but that it is an integral part of the intensive verb form. The **intensive** consists of a bimoraic foot template which is filled by copying the root. If the root material does not suffice to fill the template, a high front vowel is inserted (Noske 1991).

The occurrence of a nominal number marker depends primarily on the syllable structure of the base form. In addition, the semantics of a noun need to be taken into consideration. An entirely different set of number markers occurs with abstract nouns than with concrete nouns. Only concrete nouns will be considered here.

The nominal number suffix *-inl-in* marks plurals. It is suffixed to nouns consisting of a monomoraic closed syllable:

- |                        |                   |
|------------------------|-------------------|
| (23) (a) <i>e-com'</i> | <i>ŋi-com-in'</i> |
| M:SG-baboon            | M:PL-baboon-PL    |
| (b) <i>e-pem'</i>      | <i>ŋi-pem-in'</i> |
| M:SG-platform          | M:PL-platform-PL  |

This suffix is also found with stems that consist of a bimoraic, closed syllable and bisyllabic stems that end in a closed syllable. Nouns with this syllable structure can also be inflected with other number markers. The occurrence of this suffix is therefore not entirely predictable. However, no noun stems consisting of a light, closed syllable have been found which form the plural with a marker other than *-inl-in*.

The suffixes *-itl-it*, *-et*, *-atl-ɔtl-ot*, and *-utl-ut* always mark the singulative.

- |                      |                     |               |
|----------------------|---------------------|---------------|
| (24) <i>ɛ-cɔk-it</i> | <i>ŋi-cɔk'</i>      | 'fruit stone' |
| <i>ɛ-kasùkow-ut'</i> | <i>ŋi-kasùkowu'</i> | 'old man'     |
| <i>a-bòŋ-et</i>      | <i>ŋa-bòŋe'</i>     | 'thousand'    |
| <i>a-war-èt</i>      | <i>ŋa-wàr</i>       | 'cow dung'    |

The most common singulative marker amongst these is *-itl-it* (which is also widespread in other Niloctic and Nilo-Saharan languages). Other suffixal forms probably came about as distinct singulative markers as a result of paradigmatic displacement. As the ex-

ample for 'old man' and 'thousand' show, the suffix vowel has assimilated to the root-final vowel. However, in many instances the base-final vowel that conditioned the vowel assimilation was lost in the uninflected form, as in the case of 'cow dung'. Consequently, *-et* must be considered as a separate singulative marker synchronically.

Turkana also has the suffixes *-a/-o/-ɔ*, and *-i/-i*, which mark both the singulative and the plural, depending on what the base form of the noun is. These suffixes do not have a constant meaning and their function can only be defined as signalling the morphologically marked category of number as such. Suffixation by way of these suffixes is in fact the most common way of marking number in Turkana. The suffix *-a/-o/-ɔ* is used if the base form is at least bimoraic and ends in a closed syllable. It is also found with nouns ending in an open syllable with a high front vowel in the nucleus. Note that in all examples the suffix has a high tone.

singular	plural	
<i>a-ŋasep'</i>	<i>ŋa-ŋàsèp-a'</i>	'placenta'
<i>e-putìr</i>	<i>ŋi-pùtìr-o'</i>	'warthog'
<i>a-lagam-a'</i>	<i>ŋa-lagam'</i>	'kind of necklace'
<i>a-kopit-o'</i>	<i>ŋa-kòpit</i>	'rope'
<i>ɛ-kɔr(i)</i>	<i>ŋi-kɔrì-a'</i>	'ratel'
<i>e-ris(i)</i>	<i>ŋi-rìsy-ò</i>	'cheetah'

Tab. 139.3: Suffixes *-a/-o*

The suffix *-i/-i* occurs if the noun stem is at least bimoraic and ends in an open syllable with any vowel other than a high, front vowel in the nucleus.

singular	plural	
<i>e-suro'</i>	<i>ŋi-suru-ì</i>	'dik-dik'
<i>a-gùlù</i>	<i>ŋa-gulu-i'</i>	'water jar'
<i>e-pòno-i</i>	<i>ŋi-pòn(o)</i>	'lip'
<i>a-làma-i</i>	<i>ŋa-làm(a)</i>	'wild cherry'

Tab. 139.4: Suffixes *-i/-i*

The two number suffixes are in complementary distribution with *-a* occurring only after a consonant and a high front vowel and *-i/-i* occurring elsewhere. They can therefore be treated as allomorphs of the same morpheme, even though they cannot be related to each other by means of phonological rules. This is a rather rare case, then, of phonologi-

cal factors governing the distribution of morphs which are themselves phonologically quite unlike (see also Carstairs 1988: 70 and Dimmendaal 1987 for a discussion). Since Turkana does not have any diphthongs, suffixation of this number marker always adds a syllable to the existing noun. The allo-morph *-a* is chosen when the noun ends in a high front vowel to create an additional syllable and not a long (or complex) vowel.

In general, Turkana nouns receive a low tone on the second mora of the stem if they are marked for number. However, tonal marking varies depending on the syllable structure of the noun and the choice of the number suffix.

### 5. The morphology-syntax interface

Turkana is a verb-initial language, i.e. the syntactic subject and object tend to follow the main verb or auxiliary verb, as in (27) of the illustrative text in 6. Alternative orders, e.g. where either the subject or the object precedes the verb, are used to express topic continuity (as in (28)) as well as other pragmatic roles. These alternative word orders are therefore highly context-sensitive. Whereas pronominal marking for subjects (and objects in the case of first and second person) on verbs is obligatory in conjugated verbs, independent subject and object pronouns again may be used for pragmatic reasons (as in (30)). From a syntactic point of view, then, there is a unidirectional relation between the (obligatory) pronominal subject or object prefix on a verb and an independent subject or object pronoun if it occurs, since the latter is superfluous from a syntactic point of view. It follows that subject and object markers on a conjugated verb in Turkana should not be called agreement markers, but, rather, cross-reference markers.

Nominal modifiers such as adjectives and numerals in Turkana take gender-sensitive **agreement** prefixes, whereas the form of demonstratives co-varies with the gender or number of the nouns they modify (cf. the examples in (36)). Indexing the different nominal modifiers through prefixes, thereby binding the modifier to a noun, constitutes a true case of agreement marking, since the marking is assigned as a reflection of the properties of the head noun.

- (25) (a) *aite' na-pus(i) na'*  
cow F-grey this:F  
'this grey cow'

- (b) *emoj' lo-pus(i) lo'*  
bull M-grey this:M  
'this grey bull'

Syntactic **subjects** (both noun phrases and independent pronouns) take **nominative** case marking when they follow the (auxiliary or main) verb. When they precede the latter, they take **absolutive** case, as do objects (or nouns or pronouns in isolation). Nominative case is formed by changing the tonal pattern of nouns and pronouns in isolation, by way of tonal displacement triggered by a low tone case marker (Dimmendaal 1983: 261–264). However, because of lexical irregularities in the nominative tone patterns of certain nominal paradigms, nominative case marking cannot or can no longer be treated as an entirely syntactic phenomenon *pur sang*.

absolutive	nominative	
<i>ŋi-tɔŋa'</i>	<i>ŋi-tòŋà</i>	'people' (cf. 27)
<i>ɛ-kasùkow-ut'</i>	<i>ɛ-kasùkow-ùt</i>	'old man'
<i>ŋa-kiṭɔ'</i>	<i>ŋa-kiṭò</i>	'firewood'
<i>kèci'</i>	<i>kèc(i)</i>	'they' (cf. 30)
<i>ŋwɔ̃n'</i>	<i>ŋwɔ̃n(i)</i>	'we (inclusive)' (cf. 32)

Tab. 139.5: Absolutive and nominative

Similarly to the nominative case, the **genitive** and **instrumental** case in Turkana are formed through tonal modification of the noun (and modifiers, if they occur).

	singular	plural
absolutive	<i>a-kiṭɔ-r'</i>	<i>ŋa-kiṭɔ</i>
instrumental	<i>a-kiṭɔ-r</i>	<i>ŋa-kiṭɔ</i>
genitive	<i>à-kiṭɔ-r</i>	<i>ŋa-kiṭɔ</i>

Tab. 139.6: Instrumental and genitive

The **locative** and **ablative** are marked by a different set of gender prefixes, together with a change in the tonal pattern of the noun. In addition, the ablative case form is preceded by a proclitic *à* for nouns, pronouns, and their modifiers, or *ka* for adverbs and demonstratives.

	singular	plural
absolutive	<i>a-kiṭɔ-r'</i>	<i>ŋa-kiṭɔ</i>
locative	<i>nà-kiṭɔ-r'</i>	<i>na-kiṭɔ</i>
ablative	<i>à na-kiṭɔ-r</i>	<i>à na-kiṭɔ</i>

Tab. 139.7: Locative and ablative

As these examples illustrate, the lexical property of gender and number, and the syntactic (relational) property of case are identified cumulatively in Turkana by way of prefixes as well as tonal inflection. Note also the typologically rare property of expressing case by means of prefixes, rather than suffixes. As argued by Nichols (1986), head-marking languages favour verb-initial structures. Turkana is no exception to this cross-linguistic tendency, but it is somewhat exceptional in that it combines head marking on the verb with dependent marking (by way of core and peripheral case marking). As argued in Dimmendaal (to appear), this property is shared with a number of other languages in the area, which form a transitional zone in this respect between strongly head-marking languages towards the south and southwest, and dependent-marking languages towards the north and north-east.

A number of nouns in Turkana have an irregular locative case form; they lack a gender prefix and display a low high tone pattern: *a-kàr(ε)* ‘well (absolutive)’, *kàrε* ‘at/ to the well’ (not \**nà-kar(ε)*).

With regard to **clause linkage** it is noted that Turkana does not use the connective ‘and’ to link up clauses or sentences; instead, the narrative mood is used with verbs to express subsecutive action (as in (30)). This feature is widespread among African languages. An additional morpho-syntactic “africanism” found in Turkana is the form and use of the complementizer ‘that’, which is based on the verb ‘to say’ (as in many African languages). Its infinitive form (*atamar(i)*) is used with complements in indirect speech, whereas the conjugated form occurs with direct speech.

A rather intricate aspect of the interaction between morpho-syntax (or categorization) and semantics is illustrated in the following examples:

- (26) (a) *à-rài` e-buw` ε-ka-naman(i)*  
3:PAST-be M.SG-hyena eater  
*ŋi-barèn à ŋi-twɔŋà*  
cattle of people  
‘a/the hyena ate the cattle of people’
- (b) *a-lómón-et (ŋési) ni-kalapatàn(i)*  
‘after it had come out’

In (26a), Turkana uses an agent noun *ekaman(i)* (derived from a verb) rather than a conjugated verb. Since eating cattle is a habit or more permanent property of the species hyena, no tense marking seems to be required

in such a proposition. The use of the corresponding noun (which lacks the verbal property of tense marking) fits in with the universal proto-typical discourse role of nouns, which describe referents or more permanent entities, while verbs tend to describe events (or temporary properties) cross-linguistically. In (26b), a derivational noun instead of a conjugated verb is used in an adverbial clause. In Turkana, perfective or completive aspect in an adverbial clause is expressed obligatorily through an instrumental or locative noun (which is marked by the suffix *-et*). That the expression of “completion of an event” does not require tense marking (as a proto-typical property of verbs) again seems to be related to the proto-typical features of nouns, which contrast with verbs in this respect.

Finally, a rather interesting aspect of the association between form and meaning (which is relevant to Turkana morphology and its interaction with syntax) is illustrated by the following examples: *ak-igaràr(i)* ‘spread (fingers or wings)’, *ak-igarar-a-òn* ‘come with hands or wings spread out’. In the second example, two semantic notions, namely motion and manner, are conflated in a single verb form, rather than being expressed in a syntactically independent way, for example, by means of a verb of motion plus a separate manner phrase. This aspect of lexical composition and lexicalization patterns, has been barely studied for either Turkana or any other African language.

## 6. Illustrative text

We illustrate some of the phenomena that were discussed in 5 by means of the following text.

- (27) *kòlɔŋ` à aŋoròt à-yàkà-si`*  
long ago 3.PAST-be-PL  
*ŋi-tòŋà*  
M.PL-people:NOM  
*à-nàr-it-ae`*  
3.PAST-call-ASPECT-PAST  
*ŋi-tùrkana`*  
M.PL-Turkana:ABS
- (28) *ŋi-twɔŋa` ŋulu`*  
M.PL-people:ABS those(ABS)  
*à-lóm-it-o` nà-mɔn(i).*  
3.PAST-live-ASPECT-PL LOC.SG-forest
- (29) *à-rài ni-a-rimùn-e-ta*  
3.PAST-be M.PL-3.PAST-gather-ASPECT-PL  
*a-kìmøj ŋési` ak-irrikàrè*  
F.SG-food:ABS it:ABS INF-hunt  
*ŋi-tyaŋ(i).*  
M.PL-animal:ABS

- (30) *à-lòs-èèn-e-te*                    *kec(i)*  
       3.PAST-go-HABIT-ASPECT-PL they:NOM  
       *nà-mòn(i)*                    *ta-ar-a'*  
       LOC.SG-forest 3.NARR-kill-PL  
       *ŋr-tyay(i)*                    *kà-maa'*.  
       M.PL-animal:ABS ABL-there
- (31) *à-nàm-è-te'*                    *kèc(i)*  
       3.PAST-eat<sub>1</sub>-ASPECT-PL they:NOM  
       *a-kirij'*                    *è-jon'*                    *kà-inyò*  
       F.SG-meat 3-be.raw because  
       *pe-a-yèn-è-te'*  
       NEG-3.PAST-know-ASPECT-PL  
       *a-kim(i)*.  
       F.SG-fire:ABS
- (32) *sodì ε-ce-kàsùkowut'*  
       then M:SG-other-old.man:ABS  
       *to-lim(ù)*                    *ta-ma'*                    *no' mono'*  
       3.NARR-say 3.NARR-say why ATT  
       *ki-nam-i-a'*                    *ŋwòn(i)*  
       we-eat<sub>1</sub>-ASPECT-SUBJ 1.PL:NOM  
       *a-kirij'*                    *è-jon'*  
       F.SG-meat:ABS 3-be.raw
- (33) *à-pòt(ò)*                    *luçè*  
       3.PAST-come other:NOM  
       *to-kyëna-ta'*                    *è-bàasì*  
       3.NARR-laugh-PL 3-say.PL  
       *i-bay-aan-à*                    *iyòŋ*.  
       2-stupid-HABIT-VOICE you:NOM
- (34) *naat' ni kɔlɔŋ' i-te-ò*  
       where that long.ago you-see-VOICE  
       *iyòŋ*                    *ŋi-tɔŋa'*  
       you.SG:NOM M.PL-people:ABS  
       *e-pèy-è-te'*                    *a-kirij'*  
       3-roast-ASPECT-PL F.SG-meat:ABS
- (35) *sodì à-bo' ε-kàsùkowùt*                    *ŋolo'*  
       then 3-come M.SG-old.man:NOM that  
       *k-rlilŋ.*  
       3.NARR-be.quiet
- (36) *àni ε-naŋ-ì*                    *nà-tapàrac(ù)*  
       when 3-reach-ASPECT LOC-morning  
       *à-bo'*                    *ta-ŋara'*  
       3.PAST-come 3.NARR-call  
       *ŋi-tɔŋa'*                    *lu-ce'* *daaŋ(i)*  
       M.PL-people:ABS other all  
       *to-limw-ok(i)*                    *ta-ma'*  
       3.NARR-tell-DAT 3.NARR-say  
       *k-ŋjɔl-ikì-s(i)*                    *à-lem-un-it'*  
       IMP-look-DAT.PL I-take-VEN-ASPECT  
       *ayɔŋ*                    *ŋa-kìts'*                    *ŋa-arey'*.  
       I:NOM F.PL-stick:ABS LOC-two
- (37) *sodì à-bo' ñès(i) tɔ-ŋara'*  
       then 3.PAST-come he:NOM 3.NARR-call  
       *ε-ce-kàsùkowut'*  
       M.SG-other-old.man:ABS
- (38) *to-limw-ok(i) ta-ma'*                    *bwa'*,  
       3.NARR-tell-DAT 3.NARR-say come  
       *èpir-u-tu'*                    *a-kim(i)*  
       we.NARR:light-VEN-PL F.SG-fire:ABS  
       à                    *na-kitɔ*                    *ka' nu'*.  
       PREPOSITION F.PL-stick ABL those
- (39) *à-pòt' à-bo' a-kim(i)*  
       then 3.PAST-come F.SG-fire:NOM  
       *ki-ipìr-akì-si'*                    *ŋa-kìts'*  
       3.NARR-ignite-DAT-PL F.PL-stick:ABS  
       *ŋona'*.  
       those
- (40) *à-pòt' ŋi-tòŋà daaŋ(i)*  
       3.PAST-come M.PL-people: NOM all  
       *ta-lakar-o-s(i)*  
       3.NARR-happy-VOICE-PL
- “(27) Long ago, there were people called the Turkana. (28) Those people lived in the forest. (29) Their way of gathering food was by hunting animals. (30) They used to go into the forest and kill animals there. (31) They ate the meat raw, because they did not know fire. (32) Then, a certain old man said: ‘Why for goodness sake are we eating the meat raw?’ (33) The others started laughing and said: ‘You are a fool!’ (34) ‘Where did you ever see people roasting meat?’ (35) Then that old man became silent. (36) The next morning, he called all the other people, and said: ‘Look, I brought two sticks.’ (37) Then he called for another old man and said: ‘Come, and let us get fire from these sticks’. (38) They set the sticks on fire. (39) Then the fire emerged from those sticks. (40) Then, all the people turned out happy.”

## 7. Uncommon abbreviations

ATT	attitude marker
IDEO	ideophone
IT	itive

## 8. References

- Barrett, Anthony (1988), *English-Turkana Dictionary*. Nairobi: Macmillan
- Barrett, Anthony (1991), *Turkana-English Dictionary*. Nairobi: Macmillan
- Best, Günter (1983), *Culture and Language of the Turkana of NW Kenya*. Heidelberg: Carl Winter Universitätsverlag

- Carstairs, Andrew (1988), "Some Implications of Phonologically Conditioned Suppletion". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology 1988*. Dordrecht: Foris Publications, 67–94
- Creider, Chet & Creider, Jane Tabsubei (1989), *A Grammar of Nandi*. Hamburg: Helmut Buske Verlag (Nilo-Saharan: Linguistic Analyses and Documentation 4)
- Dimmendaal, Gerrit J. (1983), *The Turkana Language*. Dordrecht: Foris (Publications in African Languages and Linguistics 2)
- Dimmendaal, Gerrit J. (1987), "Drift and Selective Mechanisms in Morphological Changes: The Eastern Nilotic Case". In: Ramat, Anna Gialcolone et al. (eds.), *Papers from the 7th International Conference on Historical Linguistics, Pavia 1985*. Amsterdam: Benjamins (Current Issues in Linguistic Theory 48), 193–210
- Dimmendaal, Gerrit J. (1991), "The Geometry of Verb Paradigms in Teso-Turkana". In: Plank, Frans (ed.), *Paradigms: The Economy of Inflection*. Berlin: Mouton de Gruyter (Empirical Approaches to Language Typology 9), 275–306
- Dimmendaal, Gerrit J. (1996), "Attitude Markers and Conversational Implicatures in Turkana Speech Acts". *Studies in Language* 20.2, 249–274
- Dimmendaal, Gerrit J. (2000), "Number Marking and Noun Categorization in Nilo-Saharan Languages". *Anthropological Linguistics* 42.2, 214–261
- Dimmendaal, Gerrit J. (to appear), "Head Marking, Dependent Marking and Constituent Order in the Nilotic Area". In: Voeltz, Erhardt (ed.), *Typological Studies on African Languages*. Amsterdam: John Benjamins
- Dimmendaal, Gerrit J. & Breedveld, Anneke (1986), "Tonal Influence on Vocalic Quality". In: Bogers, Koen & Mous, Maarten & van der Hulst, Harry (eds.), *The Phonological Representation of Suprasegmentals*. Dordrecht: Foris, 1–33
- Heine, Bernd (1980), *The Non-Bantu Languages of Kenya*. Berlin: Dietrich Reimer (Language and Dialect Atlas of Kenya 2)
- Nichols, Johanna (1986), "Head-marking and Dependent-marking Grammar". *Language* 62, 56–119
- Noske, Manuela (1991), "Metrical Structure and Reduplication in Turkana". In: Bender, M. Lionel (ed.), *Proceedings of the Fourth Nilo-Saharan Conference, Bayreuth, August 30–September 2, 1989*. Hamburg: Helmut Buske Verlag (Nilo-Saharan: Linguistic Analyses and Documentation 7), 245–262
- Noske, Manuela (1996), "[ATR] Harmony in Turkana". *Studies in African Linguistics* 25.1, 61–99
- Noske, Manuela (2000), "[ATR] Harmony in Turkana: The Case of Faith Suffix >> Faith Root". *Natural Language and Linguistic Theory* 18, 771–812

*Gerrit Dimmendaal, Cologne (Germany)  
Manuela Noske, Seattle (U.S.A.)*

## 140. Twi (Kwa)

1. Sprachgebiet, Forschungsstand und soziolinguistische Angaben
2. Typologischer Charakter
3. Silben- und Morphemstruktur
4. Wortarten
5. Wortbildung
6. Illustrativer Text
7. Unübliche Abkürzungen
8. Zitierte Literatur
  
1. Sprachgebiet, Forschungsstand und soziolinguistische Angaben

Unter dem Namen Twi werden verschiedene Dialekte, v. a. das Akuapem und Asante, zusammengefaßt. Es bildet zusammen mit dem Fante das Akan, die in Ghana am weitesten

verbreitete Sprache. Das Akan wird im südlichen bis mittleren Ghana von ca. 45 Prozent der Bevölkerung des Landes als Muttersprache gesprochen; 70 Prozent sprechen es als Zweitsprache. Es hat eine Sprecherzahl von ca. 9 Millionen. Bei einem v. a. in älterer Literatur weiter gefaßten Akan-Begriff sind eng verwandte Völker und deren Sprachen der östlichen und zentralen Côte d'Ivoire integriert. In genetischer Hinsicht ist es dem Potou-Tano-Zweig des New Kwa der Niger-Congo-Sprachfamilie zugehörig (vgl. Stewart 1989).

Die morphologische Skizze behandelt das Akuapem-Twi, das Mitte des 19. Jahrhunderts durch Missionare schriftlich fixiert und

wissenschaftlich bearbeitet worden ist (s. Riis 1853). Für das Asante-Twi und das Fante war ebenfalls eine eigene schriftliche Form entstanden. Jahrelange Bemühungen des Akan Orthography Committee mündeten 1978 in die Schaffung einer einheitlichen Akan-Orthographie. Die Erforschung des Akan ist mit dem Namen des Missionars J. G. Christaller verbunden, dessen Hauptwerke (Grammatik 1875 und Wörterbuch 1881) zu den Pionierarbeiten der afrikanischen Linguistik im 19. Jahrhundert zählen.

## 2. Typologischer Charakter

Das Twi ist in typologischer Hinsicht als eine isolierend-agglutinierende Sprache zu charakterisieren. Silbe und Morphem sind vielfach identisch, dann einsilbig (vgl. 3); es gibt jedoch formbildende Affixe. Besonderes Merkmal der Agglutination auf nominaler Ebene sind Präfixe als Reste eines nominalen Klassensystems (vgl. Art. 98 und 4.1). Das Twi ist eine Tonsprache mit zwei distinktiven Tönen, die einerseits lexikalische, andererseits grammatische Funktion haben. Tonträger sind Vokale und Nasale. Daneben ist als Intonationsmerkmal das Downdrift-Phänomen, d. h. die stufenweise Absenkung der Hoch- und Tieftöne, wenn diesen ein oder mehrere Hoch- bzw. Tieftöne vorangehen, charakteristisch (vgl. Stewart 1965; Dolphyne 1988). Downdrift erstreckt sich über den gesamten Satz.

## 3. Silben- und Morphemstruktur

Die charakteristischen Silbenstrukturen des Akuapem-Twi, die meistens gleichzeitig auch Morphemstrukturen repräsentieren, sind /CV/, /V/ und /C/ (vgl. Dolphyne 1988). Vokale fungieren v. a. als grammatische Morpheme, wie nominale Präfixe und Tempus/Aspekt/Modus-Marker (PF -*a*, PRÄT -*è*, -*i*). Silbische Konsonanten werden durch Nasale sowie finales -*w* repräsentiert. Nasale fungieren entweder als Grammeme (Nominalpräfixe, Tempus/Aspekt/Modus-Marker, wie {OPTATIV}, oder sind finaler Bestandteil eines /CVC/-Morphems (*dàñ* ‘umdrehen’, (*ɔ-*)*kwáñ* ‘Weg’). /CV/-Strukturen haben grammatische Funktion (Pronomina, Tempus/Aspekt/Modus-Marker, Nominalpräfixe) oder repräsentieren lexikalische Morpheme (*kɔ* ‘gehen’, (*ɔ-*)*bá* ‘Kind’). Zu den häufigsten

mehrsilbigen Morphemtypen, die vorwiegend Lexeme darstellen, gehören die folgenden:

- (a) /CVC/: als CVN *nom* ‘trinken’; als CVw *tow* ‘werfen’
- (b) /CVV/: *boá* ‘helfen’, *bué* ‘öffnen’
- (c) /CVCV/: *bisá* ‘fragen’, *katá* ‘bedecken’, *srú* [*srá*] ‘besuchen’
- (d) /CVCVC/: *hyerén* ‘scheinen’, *hurúw* ‘springen’

Die Kombination von grammatischen und lexikalischen Morphemen zu einem Wort folgt dem Prinzip der morphologischen Agglutination.

Als eine für alle Akan-Dialekte typische Eigenschaft des Wortes ist die Vokalharmonie herauszustellen, die eine ausgeprägte allomorphe Variation bedingt. Die Vokalharmonie wirkt sowohl auf der nominalen Ebene (Präfix + Stammmorphem) als auch auf der Ebene des ein Wort repräsentierenden verbalen Ausdrucks (Pronomen + Tempus/Aspekt/Modus + Verb). Alle Vokale sind, durch unterschiedliche Zungenwurzelstellung bedingt, in zwei sich in ihrer Qualität unterscheidende Reihen eingeteilt:

- [i, e, æ, o, u] [+gespannt] [+ATR]
- [I, ε, a, ɔ, ɒ] [+ungespannt] [-ATR]

Lediglich [æ] repräsentiert kein Phonem. Die regressiv wirkende Vokalharmonie geht vom lexikalischen Morphem aus, dessen Vokalqualität die der vorangehenden Formative beeinflußt, so daß jedes Wort Vokale nur einer Reihe enthält; vgl. die Verbformen *ɔ-béba* [-ATR] (‘er/sie wird kommen’), *o-bé-hu* [+ATR] (‘er/sie wird sehen’) sowie die Nomina (*e-*)*dá* [-ATR] ‘Tag’ und (*e-*)*hú* [+ATR] ‘Furcht’. Über Wortgrenzen hinaus wirkt die Vokalharmonie in der Regel über eine Silbe.

## 4. Wortarten

### 4.1. Substantiv

Die Morphologie des Substantivs weist auf ein durch Präfixe geprägtes Nominalklassen-System hin, das jedoch in seiner morphologischen Ausprägung, der jeweiligen semantischen Klasseninhalten sowie der vor allem auf syntaktischer Ebene relevanten Kongruenz reduziert ist. Die Präfixe dienen zur Numerusbildung ({SINGULAR}/{PLURAL}) sowie zur nominalen Derivation (s. 5.1.1). Im wesentlichen entsprechen bei der SG/PL-Distinktion zwei {PLURAL}-Präfixen fünf singularische.

- {SINGULAR}: *ɔ-lo-*, *ɛ-le-*, *a-/æ-*, *n-/m-/lŋ-*, *o-*
- {PLURAL}: *a-/æ-*, *n-/m-/lŋ-*

Die allomorphen Vokalvarianten sind der Vokalharmonie geschuldet, die nasalen der Assimilation an den konsonantischen Anlaut des lexikalischen Morphems. Die Präfix-Entsprechungen ergeben folgende Genera (nach Danso 1983):

{SING}		{PLUR}		
<i>ɔ-lo-</i>	<i>a-/æ-</i>	<i>ɔ-héne</i>	<i>a-héne</i>	'Häuptling'
		<i>ɔ-túo</i>	<i>a-túo</i>	'Gewehr'
		<i>n-/m-</i>	<i>ɔ-bá</i>	'Kind'
<i>ɛ-le-</i>	<i>a-/æ-</i>	<i>ɛ-káw</i>	<i>a-káw</i>	'Schuld'
		<i>n-/m-</i>	<i>ɛ-dá</i>	'Tag'
		<i>ɛ-le-</i>	<i>ɛ-só</i>	'Zahn'
<i>a-/æ</i>	<i>a-/æ-</i>	<i>a-ní</i>	<i>a-ní</i>	'Auge'
		<i>n-/m-</i>	<i>a-bofrá</i>	'Kind'
<i>o-</i>	<i>a-/æ-</i>	<i>gyatá</i>	<i>a-gyatá</i>	'Löwe'
		<i>n-/m-</i>	<i>kurow</i>	'Dorf'
<i>n-/m-</i>	<i>n-/m-</i>	<i>n-hómá</i>	<i>n-hómá</i>	'Buch'

Tab. 140.1: Präfix-Entsprechungen

Den Präfixen ist nur noch selten ein bestimmter semantischer Inhalt zuzuordnen. Bei Personen tritt vorwiegend die Opposition {SINGULAR} *ɔ-lo-*, {PLURAL} *a-/æ-* oder *n-/m-* auf. Kollektiva und Flüssigkeiten nehmen meistens ein nasales Präfix an, vgl. *n-sú* 'Wasser', *n-tá* 'Speichel'. Präfixe fallen, mit Ausnahme von *a-/æ-* und *n-/m-*, im Syntagma unter regelmäßigen Bedingungen aus. Das betrifft v. a. die sog. assoziativen Konstruktionen: *mé ná* 'POSS.1.SG-Mutter (meine Mutter)' (*ɛ-ná* Mutter).

Die in begrenztem Umfang auftretenden Suffixe dienen, in Kombination mit jeweils spezifischen Präfixen, der Derivation (vgl. 5.1.1).

Da das Twi als nicht-flektierende Sprache keine Kasus kennt, werden diese durch syntaktische Mittel gekennzeichnet. Relevant für morphologische Prozesse (vgl. auch 5.2) ist die Wiedergabe des {GENITIV} durch die Aufeinanderfolge zweier Substantive in einem Possessor-Possessum-Verhältnis (s. Art. 103) ohne Verwendung eines segmentalen Morphems, jedoch einhergehend mit tonalen Veränderungen.

Zur Klasse der Substantive gehören auch die v. a. als Postposition fungierenden relationalen Nomina, die das Possessum einer solchen Substantiv-Verbindung darstellen, z. B.:

*ɛ-hó* 'Außenseite', *e-mú* 'Innenseite', *ɛ-só* 'Oberseite'. Ebenso repräsentieren Numeralia Substantive, erkennbar an ihrem nominalen Präfix, das jedoch entsprechend den in der Sprache geltenden Regeln ausfällt: *a-nán* 'vier'; *e-dú* (Plural *a-dú*) 'zehn': *a-du-a-sa* 'dreißig'; *ɔ-há* (Plural *a-há*) 'Hundert': *a-há-n-krón* 'neun Hundert'. Numeralia folgen in einer Nominalphrase den Adjektiven: *m-mofrá n-kétewá dú* 'zehn kleine Kinder'.

#### 4.2. Pronomina

Die pronominalen Formen in Subjekt- und Objektfunktion bzw. in Funktion eines Possessors stimmen in segmentaler Hinsicht im wesentlichen überein (Ausnahmen: '3.SG', '1.PL', '3.PL'). Ihre konkrete Funktion leitet sich von ihrer syntaktischen Position ab: präverbal als Subjekt, postverbal als Objekt, pränominal als Possessor. Bei den als Subjekt fungierenden Anaphorika ist die ursprüngliche Nominalklassen-Unterscheidung auf eine Opposition zwischen {BELEBT/UNBELEBT} reduziert: *ɔ-bà* '3.SG{BELEBT}-komm(HABIT) (er/sie kommt)', *ɛ-dú* '3.SG '{UNBELEBT}-ankomm(HABIT) (es kommt an)'. Subjektpronomina unterliegen der Vokalharmonie (vgl. 3). Ihre tonale Charakteristik ist vom Verbparadigma abhängig, z. B. (im Akuapem-Twi) Hochton im {HABITUALIS} (*ɔ-kò* 'er geht') und {OPTATIV} (*ɔ-n-kò* 'er möge gehen'), Tiefton im {PRÄTERITUM} (*ɔ-kò-ɛ* 'er ging'). Possessivpronomina sind, ebenso wie Nomina in assoziativen Konstruktionen, Tonveränderungen unterworfen.

#### 4.3. Adjektive

Bei Adjektiven ist zwischen ihrer attributiven und prädikativen Gebrauchsweise zu unterscheiden. Bei attributiver Funktion ergibt sich die für alle Modifikatoren eines Substantivs (mit Ausnahme des {GENITIVS}) charakteristische postnominale Stellung. Auch hier ergeben sich Tonveränderungen: Tilgung des Hochtons auf dem Nomen, hochtonige Realisierung der 1. Silbe des Adjektivs: *ɔ-nipa boné* 'sg-Mensch schlecht (schlechter Mensch)'; *ɔ-dan késé* 'sg-Haus groß (großes Haus)'. Prädikativer Gebrauch ist durch die Verbindung mit der Kopula *yε* gekennzeichnet: *ɔ-dán yε késé* 'sg-Haus KOP groß (das Haus ist groß)'. Einsilbige Adjektive sind in der Regel auf prädikativen Gebrauch beschränkt (mit Ausnahme von *pa* 'gut'); als Attribut werden sie redupliziert: *ɔ-béa yε fέ* 'sg-Frau KOP hübsch (die Frau ist hübsch)', aber: *ɔ-bea fέfέ* 'sg-Frau hübsch (die hübsche

Frau’); *a-duań ye dé* ‘SG-Essen KOP süß (das Essen schmeckt gut)’; *a-duan dédé* ‘SG-Essen süß (schmackhaftes Essen)’. Eine dreifache Wiederholung des Stammes – wobei der Vokal der 2. Silbe gelängt ist – drückt Intensität aus bzw. ergibt adverbiale Formen, vgl. folgende Formen: *dedéede* ‘(sehr) süß’, *féfíféfí* ‘(sehr) hübsch’, *pápaapa* ‘(sehr) gut’. Im Plural existieren drei Varianten. Während in der Regel die Adjektive den pluralischen Substantiven unverändert folgen, vgl. *m-mofra bóné* ‘PL-Kind schlecht (schlechte Kinder)’, nehmen einige entweder Präfix *a-* oder {NASAL}- an: *a-dán á-késé* ‘PL-Haus PL-groß (große Häuser)’; *a-dán n-kétewá* ‘PL-Haus PL-klein (kleine Häuser)’.

In der Funktion prädikativer Adjektive treten stative Verben oder auch dynamische Verben, letztere in Verbindung mit dem Perfekt-Marker *-a-*, auf, wodurch ein Zustand als Resultat einer Handlung zum Ausdruck gebracht wird.

- (a) Stative Verben: *ware* ‘lang/weit sein’: *ɔ-kwán nó ware* ‘SG-Weg DEF weit sein (der Weg ist weit)’.
- (b) Dynamische Verben: *bré* ‘müde werden’: *wɔ-a-bré* ‘3.PL-PF-müde werden (sie sind müde)’.

Auch Experiencer-Konstruktionen drücken prädiktative Adjektive aus. Als Experiencer fungiert ein Nomen oder Pronomen, das den Possessor eines Körperteil-Nomens (= Possessum) darstellt; das Verb ist mit dem perfektivischen *-a-* verbunden, vgl. *né wèrə á-hòw* ‘3.SG.POSS Herz PF-trockn (er ist ärgerlich)’.

Die Komparation von Adjektiven wird nicht durch morphologische Mittel ausgedrückt, sondern mit Hilfe einer seriellen Verbkonstruktion (vgl. 4.5.6), in der *sen* ‘übertreffen’ als zweites Verb steht: *o-n-yé n-sén wo* ‘3.SG-NEG-gut.sein NEG-übertreff 2.SG (er/sie ist nicht besser als du)’.

#### 4.4. Determinatoren

Zum Ausdruck der Determination bzw. Referentialität nominaler Ausdrücke ist im Twi zwischen vier Morphemen zu unterscheiden: Nullmorphem, *yí*, *bí* und *nó*. Ihnen sind folgende Funktionen zuzuweisen:

- (a) Nullmorphem: generisch; *a-gyinamoá* ‘Katze/Katzen’
- (b) *bí*: indefinit; *a-bofrá bí* ‘ein Kind’
- (c) *yí*: deiktisch; es bezeichnet die räumliche und zeitliche Nähe; *ɔ-dán yí* ‘dieses Haus’, *sáá béré yí* ‘zu dieser Zeit’

#### (d) *nó*: definit; *a-bofrá nó* ‘das Kind’

Die betreffenden Determinatoren stehen am Ende einer auch durch andere Modifikatoren qualifizierten Nominalphrase: *ɔ-bea fífé yí* ‘diese schöne Frau’; *yen a-dán á-késé dú nó* ‘unsere zehn großen Häuser’

#### 4.5. Verb

Am Verb selbst vollziehen sich keine morphologischen Prozesse. Nur in ganz geringem Maße sind Spuren derivativer Bildungen festzustellen (vgl. 5.1.2). Der verbale Ausdruck, im Twi ein orthographisches Wort repräsentierend, besteht aus Subjektpronomen, Formativ und Verb. Die Formativen, die mit einer Ausnahme dem Verb vorangehen, bringen die grammatischen Kategorien Tempus, Aspekt und Modus zum Ausdruck und markieren die Polarität. In vielen Fällen ist eine Interaktion lexikalischer Bedeutungen von Verben mit Tempus/Aspekt/Modus-Markern zu konstatieren, was ihre gegenseitige Inkompatibilität bedingt.

Neben den segmentalen Formativen bilden suprasegmentale Merkmale einen konstitutiven Bestandteil jeder Verbform. Bei Fehlen eines segmentalen Morphems können sie ausschließliches Unterscheidungsmerkmal von Verbformen sein, vgl. {STATIV} und {HABITUALIS}.

##### 4.5.1. Tempus- und Aspekt-Marker

###### (a) Habitualis:

Silbe	Ton			
Verb	Pron. Verb			
1-silbig	H	T	ɔ-kɔ	‘er geht’
2-silbig	T	T-H	ɔ-sómá	‘er schickt’
	H	T-T	ɔ-nwènè	‘er webt’

Tab. 140.2: Bildung des Habitualis

Diese, ohne segmentales Formativ gebildete, Form drückt in Verbindung mit dynamischen Verben einen {HABITUALIS} sowie auch generische Sachverhalte aus. Sie trägt primär aspektuellen, imperfektiven, Charakter: *ó-dí pááno* ‘3.SG-ess(HABIT) Brot (er/sie ißt Brot)’; *a-gyinamoá kyérè n-kurá* ‘sg-Katze fang(HABIT) PL-Maus (die Katze fängt Mäuse)’.

(b) **Stativ:** Der Stativ ist ebenfalls eine imperfektive Verbform, die sich formal vom {HABITUALIS} nur durch seine generelle Tiefotonigkeit unterscheidet. In ihm stehen ausschließlich stative Verben: *o-gyina hó* ‘3.SG-

steh{STAT} dort (er/sie steht dort); *ɔ-wɔ há* ‘3.SG-sein(LOKAT) hier (er/ sie ist hier)’.

(c) **Präteritum:** Die die Vergangenheit bezeichnende Verbform hat primär temporalen Charakter und ist die einzige, die durch ein Suffix markiert ist. Es existieren zwei Morphemvarianten, deren Form syntaktisch bedingt ist; sie hängt vom Vorhandensein einer dem Verb folgenden Ergänzung ab:

- bei Verb ohne Ergänzung: *-èl-i:* *me-kó-è* ‘1.SG-geh-PRÄT (ich ging)’.
- bei Verb mit Ergänzung: Längung des auslautenden Verbvokals: *me-kó-ɔ hó* ‘1.SG-geh-PRÄT dort (ich ging dorthin)’.

(d) **Futur:** Zukünftige Sachverhalte werden durch ein hochtoniges *-bé-/bé-* gekennzeichnet: *ye- bέ-dua n-nuá* ‘1.PL-FUT-pflanz PL-Baum(AKK)(wir werden Bäume pflanzen)’. Daneben kann *-bé-* modale Bedeutung haben und eine Vermutung über das erwartete Geschehen ausdrücken: *ebià ye-bé-ba* ‘vielleicht 1.PL-FUT-komm (vielleicht kommen wir)’.

(e) **Progressiv:** Die tonale Charakteristik des imperfektiven {PROGRESSIV}-Markers *-re-* hängt von der Silbenstruktur des Verbs ab: Hochton bei einsilbigen Verben (*ɔ-ré-yé a-dwíma* ‘3.SG-PROG-mach SG-Arbeit (er/sie ist beim Arbeiten)’, Tiefton bei zweisilbigen Verben (*ye-re-didí* ‘1.PL-PROG-ess RDP (wir sind beim Essen)’. Die temporale Locierung progressiver Sachverhalte ergibt sich aus dem Kontext; ein Progressiv der Vergangenheit kann jedoch auch durch die dem Subjekt vorangehende hochtonige Vergangenheitspartikel *ná* markiert werden: *ná ye-re-didí* ‘wir waren beim Essen’. *-re-* blockiert die Vokalharmonie.

(f) **Perfekt:** Ein Perfekt wird durch Funktionsmorphem *-a-* ausgedrückt, dessen Ton je nach Verbindung mit ein- oder zweisilbigen Verben variiert: *ye-á-hù* ‘1.PL-PF-seh (wir haben gesehen)’, *ye-à-bísá* ‘1.PL-PF-frag (wir haben gefragt)’.

(g) **Venitiv/Allativ:** Gehen ein in der Regel tieftöniges *bè* oder *kò* einem dynamischen Verb voran, wird eine Bewegung zum Sprecher hin (*bè*) bzw. vom Sprecher weg (*kò*) als Voraussetzung für die durchzuführende Handlung ausgedrückt. Diese Venitiv- bzw. Allativformen können mit Tempus/Aspekt/Modus-Markern verbunden sein, wobei eine

Kombination mit dem futurischen *bé* ausgeschlossen ist. Beide *be*-Marker (*bé* = Futur und *bè* = Venitiv) sind aus *ba* ‘kommen’ grammatisiert: *o-be-kyiá-à yen* ‘3.SG-VEN-begrüß-PRÄT 1.PL (er/sie kam uns begrüßen)’; *w-a-kó-hwé sini* ‘3.SG-PF-ALL-anseh Film (er ist einen Film ansehen gegangen)’. Die entsprechenden Konstruktionen stellen Verkürzungen aus seriellen Verbkonstruktionen dar (vgl. 4.5.6), in denen *ba* und *kò* als Vollverben auftreten; *ɔ-n-kó n-kó-hwéhwe* ‘3.SG-NEG-geh{HABIT} NEG-ALL-such{HABIT} (er/sie geht nicht (um zu) suchen)’. Das vor dem zweiten Verb der seriellen Konstruktion stehende Venitiv- bzw. AllativmorpheM verdeckt die aus der Sicht des Sprechers für die Durchführung der Handlung notwendige Bewegung des Kommens oder des Gehens.

(h) **Nahe Zukunft:** Die Kombination von *-re-* {PROGRESSIV} + *be* {VENITIV} ergibt eine grammatisierte spezielle Verbform, ein nahes Futur: *ye-re-be-tú kwán* ‘1.PL-PROG-ven-herauszieh Weg (wir werden gleich verreisen)’.

(i) **Konsekutiv:** Eine ebenfalls mit *-a-* gebildete Verbform (s. Perfekt in (f)) hat syntaktische Funktion, indem sie einen Konsekutiv bezeichnet: *ɔ-da a-wiá na w-a-yé ε-há a-nadwó* ‘3.SG-lieg(STAT) SG-Mittag und 3.SG-KONSEK-mach SG-Jagd SG-Nacht (er/sie schläft mittags und jagt abends)’. Konsekutives *-a-* folgt obligatorisch dem {PROGRESSIV}, {FUTUR} und {NAHEN FUTUR}. Eine finale Interpretation erscheint häufig der Semantik der Konstruktion angemessen: *brá na ye-a-hwé no* ‘komm-(IMP) und 1.PL-KONSEK-anseh 3.SG (komm, damit wir ihn betrachten)’.

#### 4.5.2. Modi

Das Twi verfügt über drei Modi. Neben den zum {INDIKATIV} (vgl. 4.5.1) gehörenden Formen sind der {IMPERATIV} und {OPTATIV} zu unterscheiden.

(a) **Imperativ:** Er tritt nur im {SINGULAR} auf und wird durch den Verbstamm repräsentiert; einsilbige Verben sind hochtonig, bei folgender Ergänzung nehmen sie Tiefton an: *kó!* ‘geh!'; *kò hó!* ‘geh dorthin!’

(b) **Optativ:** Ein stets hochtoniger, dem Verb vorangehender homogener {NASAL} bildet den {OPTATIV}; außer mit der 2.SG tritt er in allen Personen auf: *ɔ-ń-kyérε yen a-dán* ‘3.SG-OPT-zeig 1.PL PL-Haus (er/sie möge/soll uns Häuser zeigen)’. Mit Hilfe der {OPTATIV}-

Form der 2.PL wird der pluralische {IMPERATIV} ausgedrückt: *mó-m-fá* ‘2.PL-OPT-nehm (nehmt!)’.

#### 4.5.3. Negation verbaler Formen

Formales Kennzeichen jeder negativen Verbform ist ein tieftoniger homorganer {NASAL}. Das Twi verfügt über mehr affirmative als negative Formen, da eine der Negativ-Formen, die neben einer spezifischen Tonstruktur lediglich den negierenden {NASAL} als segmentales morphologisches Kennzeichen besitzt, semantisch mehreren affirmativen entspricht. Für den {HABITUALIS}, {STATIV}, {PROGRESSIV} sowie den {KONSEKUTIV} in einer seriellen Verbkonstruktion nach einem {FUTUR}, {NAHEN FUTUR} und {PROGRESSIV} ist in der Negation folgende tonale Struktur charakteristisch: {PRONOMEN} Tief (außer 2. Personen), {NASAL} Tief, 1-silb. Verb: Hoch, 2-silb. Verb: Tief-Hoch, vgl: {NEG/HABITUALIS} *ɔ-n-kó* ‘3.SG-NEG-geh (er/sie geht nicht)’; {NEG/STATIV} *ɔ-n-yé ɔ-héne* ‘3.SG-NEG-KOP (er/sie ist kein Hauptling)’; {NEG/KONSEK} in einer seriellen Verbkonstruktion *ɔ-ré-ñ-tumí n-kó* ‘3.SG-PROG-NEG-könn NEG-geh (er/sie wird nicht gehen können)’.

Einige affirmative Morpheme unterliegen bei ihrer Kombination mit dem negierenden {NASAL} einem Funktionswandel. So bezeichnet das präteritale Suffix *-el-i* bei seiner Verbindung mit dem {NASAL} ein negiertes {PERFEKT}: *m-á-hú* ‘1.SG-PF-seh (ich habe gesehen) > *mi-n- hú-i* ‘1.SG-NEG-seh-PRÄT (ich habe nicht gesehen)’. Negierender {NASAL} + perfektivisches *-a-* ergibt das negierte Präteritum: *mi-hú-i* ‘1.SG-seh-PRÄT (ich sah) > *m-a-n-hú* ‘1.SG-PF-NEG-seh (ich sah nicht)’. Das {FUTUR} wird im Akuapem-Twi mit Hilfe des {PROGRESSIV} *-re-* negiert: *ɔ-bé-fa* ‘3.SG-FUT-nehm (er/sie wird nehmen) > *ɔ-ré-m-fá* ‘3.SG-PROG-NEG-nehm (er/sie wird nicht nehmen)’.

#### 4.5.4. Genera verbi

Das Twi kennt kein grammatisiertes {PASSIV}. Als funktionales Äquivalent dient die 3.PL: *wɔ-kyér-è ɔ-wɔ* ‘3.PL-fang-PRÄT SG-Schlange (eine Schlange wurde gefangen)’. Ein Zustandspassiv kommt zum Ausdruck, wenn ein Mittelverb in seiner intransitiven Gebrauchsweise mit {PERFEKT} *-a-* verbunden ist. In diesem Fall ist das Nicht-{AGENS} Subjekt: *n-tamá nó á-suáne* ‘PL-Kleid DEF PF-reiß (die Kleider sind zerrissen)’. Vgl. dagegen: *w-á-suáne n-tamá* ‘SG-PF-zerreiß PL-Kleid (er/sie hat die Kleider zerrissen)’.

#### 4.5.5. Syntaktische Kategorie der Valenz

Bei der für das Twi gültigen grundlegenden Wortfolge SV(O) geht das Subjekt dem Verb voran. Die Position des Objekts bleibt bei intransitiven einwertigen Verben unbesetzt: *Kofi sɔré-è* ‘K. aufsteh-PRÄT (Kofi stand auf)’. Für dreiwertige Verben, deren Objekte die semantischen Rollen {RECIPIENS} und {PATIENS} tragen, wie *ma* ‘geben’, *kyéré* ‘zeigen’, *fem* ‘leihen’, gilt die Aufeinanderfolge {RECIPIENS}-{PATIENS}, wobei die Verbindung dreiwertiger Verben mit zwei Objekten Beschränkungen unterworfen ist. Nur, wenn das {PATIENS} indefiniten Charakter trägt, kann es als direktes Objekt dem indirekten {RECIPIENS} folgen: *yé-mà no n-hóma bi* ‘1.PL-geb{HABIT} 3.SG(IO) SG-Buch INDEF(wir geben ihm ein Buch)’, aber *\*yé-ma no nhoma no* ‘1.PL-geb {HABIT} 3.SG SG-Buch DEF(wir geben ihm das Buch)’. Bei definitem {PATIENS} ist eine serielle Verbkonstruktion (vgl. 4.5.6) obligatorisch, in der das direkte Objekt dem defekтивen Verb *de* ‘nehmen’ folgt und das indirekte Objekt dem handlungstragenden Verb angeschlossen ist: *yé-de n-hóma no ma no* ‘1.PL-nehm SG-Buch DEF geb(HABIT) 3.SG (wir geben ihm das Buch)’.

Ein wesentliches Charakteristikum zweien und dreiwertiger Verben im Twi ist die Tatsache, daß sie mit einer Ergänzung eine semantische Einheit, einen neuen verbalen Begriff, bilden können, z. B.: *fa so* ‘nehmen Oberseite’ = ‘besiegen’: *ɔ-fá-à wɔn so* ‘3.SG-nehm-PRÄT POSS 3.PL Oberseite (er/sie besiegte sie)’. Bei dreiwertigen Verben gibt die dem Verb unmittelbar folgende Ergänzung an, auf welche Person oder Sache sich das Geschehen bezieht. *tow... tuo* ‘werfen ... Gewehr = schießen’; *bu...fɔ* ‘beugen ... Schuld = verurteilen’: *wo-bú-ù no fɔ* ‘3.PL-beug-PRÄT 3.SG Schuld (sie verurteilten ihn)’.

#### 4.5.6. Serielle Verbkonstruktionen

Die Aneinanderreihung zweier oder mehrerer Verben ist eine in den westafrikanischen Kwasprachen charakteristische Erscheinung, vgl. auch Osam (1994). In formaler Hinsicht sind solche Konstruktionen dadurch gekennzeichnet, daß im wesentlichen die Verben mit denselben Formativen (Tempus/Aspekt/Modus-Markern, Polarität) verbunden sind und das zugrundeliegende, nicht ausgedrückte Subjekt des 2. Verbs mit dem Subjekt oder Objekt des 1. Verbs identisch ist. In funktional-semantischer Hinsicht sind die seriellen Verbkonstruktionen von formal-identischen

koordinierenden, die Aufeinanderfolge von Handlungen ausdrückenden Konstruktionen abzugrenzen.

(a) Zwei Verben nehmen eine spezifische, z. T. idiomatische Bedeutung an und bilden einen neuen Lexikoneintrag: *gye... di* ‘nehmen essen > glauben’; *sua... hu* ‘lernen sehen > beherrschen’; *mi-suá-à kásá hú-i* ‘1.SG-lern-PRÄT Sprache seh-PRÄT (ich beherrschte die Sprache)’.

(b) Die Kombination zweier Verben ist funktional-grammatisch determiniert, indem eine spezifische grammatische Funktion zum Ausdruck kommt, wie {INSTRUMENTALIS}, {BENEFAKTIV}. Dabei gibt es Beschränkungen, die die Fähigkeit der Verben betreffen, als 1. oder 2. Verb aufzutreten, z. B.:

	1. Verb	2. Verb
INSTR	<i>de</i> ‘nehmen’	Aktionsverb
BEN	Aktionsverb	<i>ma</i> ‘geben’

Tab. 140.3: Serielle Verbkonstruktionen

Z. B.: *o-wu ma no* ‘3.SG-sterb(HABIT) geb(HABIT) 3.SG (er/sie stirbt für ihn)’; *me-de sekán twá-à nám* ‘1.SG-nehm Messer schneid-PRÄT Fleisch (ich schnitt mit dem Messer Fleisch)’.

(c) Ein Verb der Konstruktion wirkt modifizierend und drückt häufig Adverbien aus: *nya* ‘erhalten’ > ‘schon’, *san* ‘zurückkehren’ > ‘wieder’. Solche Verben gehen entweder dem handlungstragenden Verb voran oder folgen ihm: *ye-á-nyá á-kò hó* ‘1.PL-PF-erhalt PF-geh dort (wir sind schon dorthin gegangen)’.

#### 4.5.7. Grammatikalisierung von Verben

Verben, die in seriellen Verbkonstruktionen auftreten, sind sehr häufig Grammatikalisierungsprozessen unterworfen. Dabei ist, aus insbesondere funktionaler, aber auch formaler Sicht, ein eindeutiger Übergang zu anderen Wortarten hin, vor allem zu Präpositionen, Konjunktionen und Adverbien, zu beobachten. Sie werden weitgehend flektiert, d. h. mit Tempus/Aspekt/Modus-Markern verbunden. Der Grammatikalisierungsprozeß ist im einzelnen jedoch unterschiedlich stark ausgeprägt, z. B.:

(a) Serielle Verbkonstruktionen zum Ausdruck eines {BENEFAKTIV}: *wu* ‘sterben’ ... *ma* ‘geben’ > ‘sterben für’. Die Bedeutung von *ma* ‘geben’ ist völlig zurückgetreten; es über-

nimmt hier die Funktion eines {BENEFAKTIV}. *ma* wird jedoch wie ein Verb flektiert.

(b) Verben zur lokalen, temporalen oder quantitativen Fixierung des ausgedrückten Geschehens; sie dienen zur Wiedergabe von Präpositionen und Konjunktionen, z. B.: *fi* ‘herkommen von’ > ‘von, aus, seit’; *ko-si* oder *be-si* ‘{ALL} bzw. {VEN}-steh(bleib)’ > ‘bis’: *w-á-kàn n-hóma á-fi ne m-fiasé a-be-sí n'a-wièi* ‘3.SG-PF-les SG-Buch PF-herkomm 3.SG.POSS SG-Anfang PF-VEN-steh 3.SG.POSS SG-Ende (er/sie hat das Buch von Anfang bis Ende gelesen)’. Daneben existieren folgende Formen:

- (1) (a) *wakan nhoma fi mfiase abesi awiei*  
          (b) *wakan nhoma fi mfiase besi awiei*

Das Nebeneinanderbestehen unterschiedlicher morphologischer Ausprägungen beim Verb macht den vor sich gehenden Grammatikalisierungsprozeß deutlich.

#### 4.6. Adverb

Adverbien stehen entweder in Satzanfangs- oder in Satzendstellung. Unter semantischem Aspekt sind lokale (*há* ‘hier’, *hɔ* ‘dort’), temporale (*ɔkyena* ‘morgen’), kausale (*énti* ‘deshalb’) und modale Adverbien (*dennéennen* ‘schwer’) zu unterscheiden. In morphologischer Hinsicht gibt es neben reinen Adverbien solche, die andere Wortarten (*ntém* ‘Schnelligkeit’) oder Zusammensetzungen (*enó nti* ‘deshalb’) repräsentieren.

#### 4.7. Modalwort

Modalwörter sind im Unterschied zu den modalen Adverbien satzwertig; sie modifizieren den gesamten Satz: *ebiá* ‘vielleicht’, *nokwárem* ‘tatsächlich’.

#### 4.8. Konjunktion

Entsprechend ihren syntaktischen Eigenschaften, d. h. ihrem Fügungswert, sind koordinierende (*na* ‘und’, *mímóm* ‘sondern’) und subordinierende (*sé* ‘daß’, *efisé* ‘weil’, *esiánè* ‘se... nti’ ‘weil’, *kosí sé* ‘bis (daß)’, *sénèa* ‘wie’ Konjunktionen zu unterscheiden. Die subordinierenden sind meistens auf Grammatikalisierungen von Verben zurückzuführen, s. *e-fi-i sé mi-duá-è nò* ‘3.SG-herkomm-PRÄT daß 1.SG-pflanz-PRÄT DEF (seit ich pflanzte)’; *ɔ-yé-è a-dwúma ko-sí-i sé w-a-bré* ‘3.SG-mach-PRÄT SG-Arbeit ALL-steh-PRÄT daß 3.SG-PF-müde werd (er/sie arbeitete, bis er/sie müde war)’.

#### 4.9. Partikel

Es sind verschiedene Partikeln zu unterscheiden: *nà* = Fokusmarkierer, Fokuspartikeln wie *mmom* ‘vielmehr, eher’, Fragepartikeln (*so* = Satz einleitend, *anáa* = Satz abschließend) bzw. Vergangenheits- und Irrealitätspartikeln (*ná, anka*): *ná me- wɔ há* ‘PRT 1.SG-sein(LOKAT) hier ich war hier’.

### 5. Wortbildung

Zu grundlegenden Verfahren der Wortbildung im Twi gehören die Derivation, Komposition und Reduplikation.

#### 5.1. Derivation

Derivative Prozesse (s. Art. 89) spielen vor allem im nominalen Bereich eine große Rolle.

##### 5.1.1. Derivative Substantive

Als morphologische Mittel zur Bildung von Substantiven fungieren Prä- und Suffixe. (Zu den Präfixen vgl. vor allem 4.1.) Ihre derivative Funktion ist u. a. an der Bildung von Nomina actionis nachzuweisen, wozu die Präfixe *ɔ-lo-*, *a-* oder {NASAL} dienen: *hwe* ‘fallen’ > *ɔ-hwé*; *kan* ‘zählen’ > *ɔ-kán*; *pra* ‘fegen’ > *a-prá*; *wo* ‘gebären’ > *a-wó*; *san* ‘anstecken’ > *n-sán*; *kra* ‘sich verabschieden’ > *n-krá*. Verbindungen von einem Verb und einem als dessen Objekt fungierenden relationalen Substantiv tragen grundsätzlich ein nasales Präfix: *kyeré a-se* ‘zeig sg-Unterseite (erklären)’ > *n-kyere-a-sé* ‘Erklärung’; *hwehwé mu* ‘such Innenseite (nachforschen)’ > *n-hwehwe-mú* ‘Nachforschung’.

Die Suffixe stehen in der Regel in Korrelation zu bestimmten Präfixen.

(a) Suffix *-ní*, *-fó* {SINGULAR}, *-fó* {PLURAL}: Beide Suffixe treten an Verben und Substantive und bilden in erster Linie Nomina agentis oder bezeichnen die Herkunft von Personen. *-fó* ist bei deverbalen Ableitungen im Singular häufiger. Solche deverbalen Nomina weisen im {SINGULAR} Präfix *ɔ-lo-*, im {PLURAL} *a-* auf:

{SINGULAR}	{PLURAL}
<i>ɔ-tám-fo</i>	<i>a-tám-fo</i> < <i>tan</i> ‘hassen’
<i>o-suá-ni</i>	<i>a-suá-fo</i> < <i>suá</i> ‘lernen’

Auch denominale Ableitungen sind, bei Substitution des ursprünglichen Nominalpräfixes, mit diesen Suffixen verbunden: *ɔ-tóro-fo* ‘Lügner’ < *a-tóro* ‘Lüge’; *o-busu-fó* ‘Bösewicht’ < *m-musú* ‘Unfug’. Komposita, in denen Substantiv + Verbalnomen in einem Pos-

sessivverhältnis zueinander stehen, bewahren entweder das Präfix des Substantivs, wie in *n-homa-pám-fo* ‘Buchbinder’ < *n-hómá* ‘SG-Buch’ + *pam* ‘zusammenfügen’ oder substituieren es im {SINGULAR} durch *ɔ-lo-*, s. *o-dua-twá-fo* ‘Holzfäller’ < *e-duá* ‘SG-Holz’ + *twa* ‘schneiden’.

(b) Suffix *-nóm*: *-nóm* ist ein pluralisches Suffix zur Bezeichnung von Verwandten und Freunden. Das {SINGULAR}-Präfix bleibt dabei in der Regel erhalten: ‘Ehemann’ *o-kúnú*, {PLURAL} *o-kúnú-nom*; ‘Ehefrau’ *ɔ-yére*, {PLURAL} *ɔ-yére-nom*; ‘Freund’ *a-damfo*, {PLURAL} *n-namfo-nóm*.

(c) Suffix *-wá*: *-wá*, nach nasalem Auslaut *-má*, bildet Diminutiva. Es ist von *ɔ-bá* ‘SG-Kind’ herzuleiten. Die mit dem {DIMINUTIV} gebildeten Formen nehmen in der Regel im {SINGULAR} Präfix *a-*, im {PLURAL} nasales Präfix an: *ɛ-kwá* ‘Wald’ > *a-kwae-wá* ‘kleiner Wald, Unterholz’; *ɛ-hyén* ‘Schiff’ > *a-hyem-má* ‘Boot’, *a-bóá* ‘Tier’ > *a-bóá-wa* auch *a-bóá-a*. ‘kleines Tier’.

(d) Suffix *-él-í*: Dieses Suffix bildet deverbale Substantive, die vor allem den Ort und das Ergebnis eines Geschehens bzw. das Geschehen selbst, seltener das Instrument einer Handlung bezeichnen.

- Ort des Geschehens: Die so gebildeten Substantive sind in der Regel mit dem Präfix *a-* verbunden; *a-da-é* ‘Schlafplatz’ < *da* ‘schlafen’; *a-guare-é* ‘Badestelle’ < *guaré* ‘baden’.
- Ergebnis eines Geschehens: Die entsprechenden Substantive nehmen dabei ein nasales Präfix an: *n-ka-é* ‘Silbe’ < *ka* ‘sprechen’; *n-kyere-é* ‘Beispiel’ < *kyeré* ‘zeigen’.
- Instrument einer Handlung: *fita-é* ‘Fächer’ < *fitá* ‘fächeln’.

##### 5.1.2. Derivative Verben

Derivative Verben sind nur in Spuren feststellbar: *bu-á* ‘bedecken’, *bu-é* ‘aufdecken’; *so-á* ‘hochheben (auf Kopf)’, *so* ‘tragen (auf Kopf)’.

### 5.2. Komposition

#### 5.2.1. Komposition zur Bildung von Substantiven

(a) Substantiv + Substantiv: Zu unterscheiden ist zwischen einer appositionellen Fügung (*ɔsɔfopanyín* ‘Hohepriester’ < *ɔsɔfó* ‘Priester’ + *ɔpanyín* ‘Person hohen Ranges’)

und einem assoziativen Verhältnis (*ahemfi* ‘Häuptlingspalast’ < *shéne* ‘Häuptling’{GEN} + *ofi* ‘Haus’ = ‘Haus des Häuptlings’.

(b) Substantiv + Adjektiv: Adjektive können mit dem durch sie modifizierten Substantiv zu einem Kompositum verschmelzen: *shemfó* ‘neuer Häuptling’ < *shéne* ‘Häuptling’ + *sofó* ‘neu’; *anam móno* ‘rohes Fleisch’ < *enám* ‘Fleisch’ + *momonó* ‘roh’.

(c) Substantiv + Verbalsubstantiv: Beide Bestandteile stehen zueinander in einem assoziativen Verhältnis, wobei das Präfix des nominalisierten Verbs ausfällt. Ein solches Kompositum nimmt häufig Präfix *a-* an: *akatuá* ‘Bezahlung’ < *ekáw* ‘Schuld’ + *otua* ‘das Befestigen’; *afobú* ‘Verurteilung’ < *efí* ‘Schuld’ + *obú* ‘das Beugen’.

(d) Verb + Substantiv: Ein aus Verb und Substantiv bestehendes Kompositum weist in der Regel dann kein assoziatives bzw. Possessionsverhältnis auf (s.(c)), wenn die Ergänzung des Verbs durch ein relationales Substantiv, wie *ase* ‘Unterseite’, *ehó* ‘Außenseite’, repräsentiert wird: *mfiasé* ‘Beginn’ < *fi* ‘herkommen von’ + *ase* = ‘beginnen’; *nkatahó* ‘Deckel’ < *katá* ‘bedecken’ + *eho* = ‘zudecken, einhüllen’. Diese Komposita nehmen nasales Präfix an.

(e) Verb + Verb: *ɔ-dá-dwen* ‘Meditation’ < *da* ‘liegen’ + *dwen* ‘denken’; *ɔ-ká-kyere* ‘Information’ < *ka* ‘sagen’ + *kyeré* ‘zeigen’.

Die einzelnen Bestandteile der Komposita, sind, verglichen mit ihrer tonalen Charakteristik in absoluter Stellung, Tonveränderungen unterworfen. Das lexikalische Morphem des 1. Bestandteils ist in der Regel tieftonig: *afobú* ‘Verurteilung’ < *efí* ‘Schuld’ + *obú* ‘das Beugen’; *shempá* ‘guter Häuptling’ < *shéne* ‘Häuptling’ + *pá* ‘gut’; (vgl. dagegen die Komposita unter (e)).

**5.2.2. Komposition zur Bildung von Verben**  
Sie ist im wesentlichen auf zwei Varianten, in denen die jeweiligen Bestandteile diskontinuierliche Elemente darstellen, beschränkt:

- (a) Verb + Verb (= serielle Verbkonstruktion) ergeben zusammen einen neuen verbalen Begriff (vgl. 4.5.6.)
- (b) Verb + Ergänzung bilden eine semantische Einheit (vgl. 4.5.5.): *bu...fó* ‘verur-

teilen’ < *bu* ‘beugen’ + *efí* ‘Schuld’; *ma... so* ‘hochheben’ < *ma* ‘geben’ + *esó* ‘Oberseite’.

### 5.3. Reduplikation

Reduplikation (total oder partiell; s. Art. 57) ist nicht auf eine bestimmte Wortart beschränkt; sie tritt jedoch am häufigsten bei Verben auf, wobei ihr die Funktion der Pluralität zuzuschreiben ist, indem sie einerseits die Iteration oder Intensivierung eines Geschehens ausdrückt; andererseits ist Reduplikation von Verben durch pluralische Subjekte oder Objekte bedingt. In einigen Fällen kommen Bedeutungsnuancen im Verhältnis zu den jeweiligen Grundformen zum Ausdruck: *hwe* ‘betrachten’, *hwehwé* ‘suchen’; *gye* ‘nehmen’, *gyigye* ‘aufziehen’. Der Vokal der als Präfix gedeuteten 1. Silbe der reduplizierten Form ist in der Regel einer phonologischen Veränderung unterworfen; vgl. auch die folgenden Formen: *tɔ/totɔ* ‘kaufen’; *dal deda* ‘liegen’; *pe/pepe* ‘wünschen’; *paw/popaw* ‘wählen’; *nyal/nyinya* ‘erhalten’.

In einer Reihe von Fällen haben Reduplikation bzw. eine mehrfache Wiederholung der Wurzel eine grammatische Funktion, z. B. prädikatives Adjektiv (*fe* ‘schön’, *den* ‘stark’) > attributives Adjektiv (*fe fe* bzw. *dennén*) > Adverb (*fe fe fe* bzw. *dennénen*). Reduplikationen repräsentieren auch Pluralformen von im wesentlichen im Singular mit *e-/e-*-verbundenen Substantiven: *e-fí*, {PLURAL} *a-fia-fí* ‘Bund’. Die pluralischen Substantive sind dabei zusammen mit ihrem Präfix (*a-*) redupliziert.

### 6. Illustrativer Text

<i>Anna hotel-tra</i>	<i>A-kwan-tu</i>
Anna Hotel-sich.befind	sg-Weg-herauszieh'
<i>ten-ten no a-kyi no Anna</i>	
lang(RDP) DEF SG-Rücken	DEF Anna
<i>hu hotel bi wɔ</i>	
seh(HABIT) Hotel INDEF sein(LOKAT)	
<i>kurom o-bisa</i>	
Stadt.Innenseite(ZIRK) 3.SG-frag(HABIT)	
<i>krakye no a ɔ-yε</i>	
Angestellter(AKK) DEF REL 3.SG-mach(HABIT)	
<i>a-dwuma wɔ reception no</i>	
SG-Arbeit sein(LOKAT) Rezeption	DEF

se: "ɔ-dan bi da hɔ  
 sag(HABIT): SG-Haus INDEF lieg{STAT} dort  
 kwa anaa?" Krakye no  
 unbenutzt INT Angestellter DEF  
 bua no se:  
 antwort(HABIT) 3.SG(IO) sag(HABIT)  
 "yiw, a-dan wɔ ha a  
 Ja PL-Haus sein(LOKAT) hier REL  
 o-biara n-na mu.  
 SG-jemand NEG-lieg{STAT} Innenseite  
 wo-pe o-biako-fo dan  
 2.SG-woll(STAT) SG-eins-SBSTR Haus  
 anaase n-nipa baa-nu dan." Anna  
 oder PL-Mensch PRÄF-zwei Haus? Anna  
 fa o-biako-fo dan.  
 nehm(HABIT) SG-eins-SBSTR Haus.  
 o-gye ne dan safe  
 3.SG-nehm(HABIT) 3.SG.POSS Haus Schlüssel  
 na ɔ-ma krakye no  
 und 3.SG-geb(HABIT) Angestellter DEF  
 de ne n-nee-ma  
 nehm(HABIT) 3.SG.POSS PL-Ding-DIM  
 bre no wɔ ne  
 bring(HABIT) 3.SG sein(LOKAT) 3.SG.POSS  
 dan mu. ε-no a-kyi no  
 Haus Innenseite SG-PRON SG-Rücken DEF  
 Anna guare  
 Anna sich.wasch(HABIT)  
 sie-sie ne  
 sich.vorbereit-RDP(HABIT) 3.SG.POSS  
 ho. afei ɔ-bɔ  
 Außenseite nun 3.SG-schlag(HABIT)  
 m-mɔ-den se ɔ-be-fre  
 SG-schlag-Stärke daß 3.SG-FUT-ruf  
 n' a-damfo Ghana -ni no a  
 3.SG.POSS SG-Freund Ghana -SBSTR DEF REL  
 w-a-m-me-fa no wɔ  
 3.SG-PF-NEG-VEN-nehm 3.SG sein(LOKAT)  
 aluplenpak no wɔ a-homa  
 Flugplatz DEF sein(LOKAT) SG-Strick  
 so. ɔ-dare n'  
 Oberseite 3.SG-wähl(HABIT) 3.SG.POSS  
 a-damfo no number. Nyame  
 SG-Freund DEF Nummer Gott  
 n-kye a-de na n'  
 NEG-schenk(HABIT) SG-Ding FOK 3.SG.POSS  
 a-damfo no wɔ o-fie.  
 SG-Freund DEF sein(LOKAT) SG-Haus

wɔ-hye-hye ho  
 3.PL-festsetz-RDP(HABIT) Außenseite  
 se wo-be-hyia wɔ hotel no  
 daß 3.PL-FUT-treff sein(LOKAT) Hotel DEF  
 a-di-di-bea dɔn-hwerew a-bien  
 SG-essen-RDP-Ort Stunde-vergeh PL-zwei  
 a-kyi.  
 SG-Rücken.

'Annas Hotelaufenthalt: Nach der langen Reise sieht Anna ein Hotel in der Stadt. Sie fragt den Angestellten, der an der Rezeption arbeitet: "Haben Sie ein freies Zimmer?". Der Angestellte antwortet ihr: "Ja, es gibt freie Zimmer. Willst du ein Einzel- oder ein Doppelbettzimmer?". Anna nimmt ein Einzelzimmer. Sie nimmt den Schlüssel ihres Zimmers und veranlaßt den Angestellten, daß er ihr ihre Dinge in ihr Zimmer bringt. Danach wäscht sich Anna und macht sich fertig. Nun bemüht sie sich, ihren ghanaischen Freund anzurufen, der sie nicht auf dem Flugplatz abgeholt hat. Sie wählt die Nummer ihres Freundes. Gott sei Dank ist ihr Freund zu Hause. Sie verabreden, daß sie sich in zwei Stunden im Restaurant des Hotels treffen werden.'

## 7. Unübliche Abkürzungen

LOKAT	Lokale Kopula
PRÄF	Präfix
PRT	Partikel
DIM	Diminutiv
SBSTR	Substantivator

## 8. Zitierte Literatur

- Christaller, J[ohann] G. (1875), *A Grammar of the Asante and Fante Language, Called Tshi (Chwee, Twi)*. Basel: The Evangelical Missionary Society
- Christaller, J[ohann] G. (1933), *A Dictionary of the Asante and Fante Language, Called Tshi (Twi)*. Basel: The Evangelical Missionary Society [1881]
- Danso, M[ary] (1983), *The Formation of the Plural of Twi Nouns*. Arbeitsberichte aus dem Seminar für Allg. und Indogerman. Sprachwissenschaft (Kiel), 6
- Dolphyne, F[llorence Abena] (1988), *The Akan (Twi-Fante) Language: Its Sound Systems and Tonal Structure*. Accra: Ghana Universities Press
- Osam, E[mmanuel Kweku] (1994), *Aspects of Akan Grammar: A Functional Perspective*. Ph.D. dissertation, University of Oregon

- Riis, H[ans] N. (1853), *Elemente des Akwapim-Dialekts der Odschi-Sprache*. Basel
- Stewart, J[ohn] M. (1965), *The Typology of the Twi Tone System*, with comments by P[paul] Schachter and W[illiams] E. Welmers. Legon: University of Ghana

Stewart, J[ohn] M. (1989), “Kwa”. In: Bendor-Samuel, J[ohn] (Hrsg.), *The Niger-Congo Languages: A Classification and Description of Africa’s Largest Language Family*. Lanham, New York, London: University Press of America

*Brigitte Reineke, Berlin (Deutschland)*

## 141. Kinyarwanda (Bantu)

1. Introduction
2. Noun
3. Verb
4. Adjective
5. Unclassified categories
6. Tones
7. Reduplication
8. Problems in Kinyarwanda morphology
9. Illustrative text
10. Uncommon abbreviations
11. References

### 1. Introduction

Kinyarwanda, which is called *ikinyarwanda* by its native speakers and *urunyarwanda* by Rwandans in Uganda because languages that are spoken there such as Luganda, Runyankole, Runyoro, Rukiga, etc. start with the *ru*-prefix, and *Rwanda* in Coupez's writings (cf. Coupez 1980), is the national language of Rwanda and is a sister dialect of Kirundi, the national language of Burundi. Linguistically speaking, these are indeed dialects of the same language since there is total intelligibility between their respective speakers. Together with another dialect known as Giha spoken in Tanzania, it is the second largest language cluster after Kiswahili within the Bantu language group. Its speakers are perhaps more than 20 million people. Rwanda has around 8 million people right now, Burundi has around 7 million but besides the Giha speakers there are also ethnic Banyarwanda in Southern Uganda in the Kigezi district known as Bafumbira.

Other Kinyarwanda speakers are Banyamulenge in Southern Kivu and ethnic Banyarwanda in Masisi and Rutshuro in Northern Kivu in the Democratic Republic of Congo. Kinyarwanda belongs to the interlacustrine (Great Lakes) Bantu languages. Kinyarwanda has an official orthography.

Although, both short and long vowels are phonemic and are realized at the phonetic level, only short vowels are authorized in the writing system which creates a lot of homonyms (homographs). The language also has high tones but the orthography does not mark them either.

Kinyarwanda like other Bantu languages has open syllables only. It has 5 vowels namely *a*, *e*, *i*, *o*, and *u* which are realized as either short or long at both the phonemic and phonetic levels. Only *i*, *u*, and *a*, occur in (pre)prefix position and in verbal extensions. Vowels of verbal extensions can be realized as mid vowels, however, because of vowel harmony. The other similarity with other Bantu languages is the existence of prenasalized consonants, the application of vowel harmony, vowel lengthening before prenasalized consonants and after palatalized and velarized consonants and consonant mutation also known as the modified form of the verb stem due to the perfective aspect morpheme reconstructed as *-ile* and realized in Kinyarwanda as *-ye* (Kimenyi 1999).

What Kinyarwanda shares with a handful of other Bantu languages are complex consonants and Dahl's law. Complex consonants are produced by different articulators at the same time, such as labio-velarized, labio-palatalized, alveo-palatalized, alveo-velarized, etc. and different manners of production such as aspiration, nasalization, and so on. Dahl's law which is very productive in Kinyarwanda and Kikuyu, voices stop consonants or prefixes which immediately precede the word stem if the latter starts with a voiceless consonant. Like the majority of other Bantu languages, Kinyarwanda also has tone spread rule and second high tone deletion known in Bantu terminology as Meeussen rule (cf. Meeussen 1959). As far the sentence structure is concerned, Kinyarwanda like

other Bantu languages is an SVO language. Modifiers follow head nouns and complements come after the verb. The only modifier which precedes a head noun is the demonstrative. The interesting syntactic behavior of this language and many other Bantu languages are: (i) object-subject reversal, (ii) wh-questioning in situ, (iii) lack of relative pronouns, (iv) multiple auxiliary serialization, (v) the existence of ditransitive verbs, and (vi) objectivization of adjunct complements (Kimenyi 1980; 2002). The object-subject reversal which consists of interchanging the direct object and the subject has the same function as the passive construction: The focus is put on the subject and the object is considered as old information. Wh-questioning is always in-situ. The only way it can appear before the main verb is through cleft or pseudo-cleft constructions.

Nouns and verbs are indeed distinct categories in Kinyarwanda. They differ a lot as far as their form, phonetic shape, and behavior are concerned. The verbal morphology is very complex but the nominal is not.

## 2. Noun

The **noun** in Kinyarwanda consists of a stem, a prefix, and a preprefix or augment. This preprefix which is realized as a vowel, seems to be a copy of the prefix vowel. These prefixes are also known in Bantu as class markers.

### 2.1. Class markers

Kinyarwanda has 16 classes (cf. Art. 98). Modifiers (adjectives, demonstratives, numerals, possessives) agree with the head noun by taking this class marker.

In some cases, however, the modifier takes a different type of prefix depending on whether it is an adjective, a verb, an object pronoun, a quantifier or a possessive as illustrated in Tab. 141.1. The numbers 1–16 correspond to traditional conventional Bantu noun classification.

The sentence in (1) shows how this type of noun class agreement works. The head noun is *utubaati* ‘cupboards’ whose preprefix has been deleted because of the fact that it is preceded by a demonstrative.

- (1) *U-tu tu-baati tw-aa-we*  
 AUG-CL13 CL13-cupboards CL13-of-you  
*du-shyá du-tatu tu-rí muu nzu*  
 CL13-new CL13-three CL13-are on table  
*tw-aá-tu-guz-ye tw-óose ejó.*  
 we-T-them-buy-PFV CL13-all yesterday

class	noun	adj.	obj. pron.	dem.	possessive
1	<i>u-mu-</i>	<i>mu-</i>	<i>-mu-</i>	<i>u-</i>	<i>u-</i>
2	<i>a-ba-</i>	<i>ba-</i>	<i>-ba-</i>	<i>ba-</i>	<i>ba-</i>
3	<i>u-mu-</i>	<i>mu-</i>	<i>-wu-</i>	<i>u-</i>	<i>u-</i>
4	<i>i-mi-</i>	<i>mi-</i>	<i>-yi-</i>	<i>i-</i>	<i>i-</i>
5	<i>i-ri-</i>	<i>ri-</i>	<i>-ri-</i>	<i>ri-</i>	<i>ri-</i>
6	<i>a-ma-</i>	<i>ma-</i>	<i>-ya-</i>	<i>a-</i>	<i>a-</i>
7	<i>i-ki-</i>	<i>ki-</i>	<i>-ki-</i>	<i>ki-</i>	<i>ki-</i>
8	<i>i-bi</i>	<i>bi</i>	<i>-bi-</i>	<i>bi-</i>	<i>bi-</i>
9	<i>i-n-</i>	<i>n-</i>	<i>-yi-</i>	<i>i-</i>	<i>i-</i>
10	<i>i-n-</i>	<i>n-</i>	<i>-zi-</i>	<i>zi-</i>	<i>zi-</i>
11	<i>u-ru-</i>	<i>ru-</i>	<i>-ru-</i>	<i>ru-</i>	<i>ru-</i>
12	<i>a-ka-</i>	<i>ka-</i>	<i>-ka-</i>	<i>ka-</i>	<i>ka-</i>
13	<i>u-tu-</i>	<i>tu-</i>	<i>-tu-</i>	<i>tu-</i>	<i>tu-</i>
14	<i>u-bu-</i>	<i>bu-</i>	<i>-bu-</i>	<i>bu-</i>	<i>bu-</i>
15	<i>u-ku-</i>	<i>ku-</i>	<i>-ku-</i>	<i>ku-</i>	<i>ku-</i>
16	<i>a-ha-</i>	<i>ha-</i>	<i>-ha-</i>	<i>ha-</i>	<i>ha-</i>

Tab. 141.1: Noun classes in Kinyarwanda

‘These three new cupboards of yours which are on the table, we bought them all yesterday.’

### 2.2. Loan-word morphology

Most of **loan-words** are put in class 5 or 9 because there are no class markers there. These words thus take classes 6 or 10 respectively to make their plural.

- (2) *i-taábi* ‘AUG-tobacco’ >  
*a-ma-taábi* ‘AUG-CL6-tobacco (tobaccos)’/  
*i-taábi* ‘AUG-tobacco (tobaccos)’  
*i-shaáti* ‘AUG-shirt’ >  
*a-ma-shaáti* ‘AUG-CL6-shirt (shirts)’/  
*i-shaáti* ‘AUG-shirt (shirts)’  
*i-shuúri* ‘AUG-school’ >  
*a-ma-shuúri* ‘AUG-CL6-school (schools)’

There are loan-words which are found in other classes, however, because of reanalysis. They are mistaken for native words and their initial syllables are taken for class markers or prefixes (cf. 3 a–c; for further examples see Kimenyi 2002).

- (3) (a) *u-mu-sírimú*  
 ‘AUG-CL1-cultivated.person’  
 from Arabic ‘muslim’  
*a-ba-sírimú*  
 ‘AUG-CL2-cultivated.person (cultivated people)’  
 (b) *u-mu-sáriyamú* ‘AUG-CL1-muslim’  
 from Arabic ‘muslim’  
*a-ba-sáriyamú* ‘AUG-CL2-muslim  
 (muslims)’

- (c) *u-mu-notá* ‘AUG-CL3-minute  
from French *minute*  
*i-mi-notá* ‘AUG-CL4-minute (minutes)’

### 2.3. The use of the prefix

The preprefix or **augment** usually doesn't have any semantic function. It is deleted, after demonstratives, in the vocative case, and in onomastics (name creation). Within certain words, however, its absence marks definiteness and its presence indefiniteness as shown in (4).

- (4) *mu-gaanga* ‘CL1-doctor (the doctor)’ >  
*u-mu-gaanga* ‘AUG-CL1-doctor (a doctor)’  
*mu-geenzi* ‘CL1-friend (the friend)’ >  
*u-mu-geenzi* ‘AUG-CL1-friend (a friend)’

The lack of the preprefix bleeds the application of tone rhythm rules such as beat addition namely the addition of a high tone on the first mora of the noun stem if the latter has a lexical high tone at least two moras on the right as shown below (cf. 6.4.1).

- (5) *u-mw-áarimú* ‘AUG-CL1-teacher (a teacher)’  
*mw-aarimú* ‘CL1-teacher (the teacher)’

### 2.4. Grammatical categorization in Kinyarwanda: From natural categories to grammatical categories

The majority of these class markers occur in pairs 1/2; 3/4; 5/6; 7/8; 9/10; 11/10; 16 to mark singularity and plurality. The human category, class 1/2, marked by prefixes *-mu-* and *-ba-* and class 16, the place category, marked by the prefix *-ha-*, are the only ones in Kinyarwanda and in other Bantu languages, which exclude nouns referring to other semantic categories. Other classes include nouns which don't have any semantic relationship whatsoever.

### 2.5. Homonymy

Class 1 *-mu-* (singular for human category) and class 3 marker *-mu-* which pluralizes in *-mi-* look exactly the same, which results in confusing language learners and sometimes in creating regional dialects based on this homonymy.

Class 5 which is marked in the great majority by a zero marker because of the loss of the class marker *-ri-* resembles class 9 which also lost its marker *-n-*, this being phonetically conditioned. Most of the words starting with voiceless fricatives have lost this marker:

- (6) *i-sha* ‘AUG-gazelle’  
*iyi sha* ‘this (CL9) gazelle’  
*i-shámi* ‘AUG-branch’  
*iri shámi* ‘this (CL5) branch’

This derivational function of prefixes and loan-words have made them polysemous. This polysemy is responsible for the grammaticalization of the class markers in Kinyarwanda and other Bantu languages.

### 2.6. Polysemy with class markers

The same class marker or prefix can have different functions. As said earlier (cf. 2.4), most of these classes occur in pairs (singular and plural). Some of these prefixes serve as plural markers for more than one class. Thus class 6 *-ma-* serves as plural for class 5, class 9, class 14, and class 15. Class 8 whose classifier is *-bi-* serves as plural for class 7 and class 11 if the derived pejorative or augmentative nouns with *-ru-* prefix; class 10 also is the plural marker for class 9 and 11. Whereas class 12 can pluralize either in class 13 *-tu-* or class 14 *-bu-*. The following examples illustrate the use of the prefix *-ma-* (class 6 marker).

- (7) *i-táma* ‘AUG-cheek (cheek)’ >  
*a-ma-táma* ‘AUG-CL6-cheek (cheeks)’  
*i-paási* ‘AUG-iron (iron)’ >  
*a-ma-páasi* ‘AUG-CL6-iron (irons)’  
*u-bw-áato* ‘AUG-CL14-canoe (canoe)’ >  
*a-má-ato* ‘AUG-CL6-canoe (canoes)’  
*u-ku-guru* ‘AUG-CL15-leg (leg)’ >  
*a-ma-guru* ‘AUG-CL6-leg (legs)’

The loss of the nasal class marker is common to many Bantu languages spoken in East Africa including Kiswahili.

### 2.7. Metonymic derivation with the prefix

Class prefixes are used very productively to create new lexical items which have a metonymic relationship (semantic association) with the initial nouns they are derived from. This association between the basic lexical item and the new derived one can be that of cause and effect, possessor and possessed, container and content, etc. This type of derivation is also similar to metonymic “zero derivation” in non-category languages such as English (cf. (8); for further examples see Kimenyi 2002).

- (8) (a) *-íuzi*  
*u-rú-uzi* ‘AUG-CL11-river’  
*i-cy-íuzi* ‘AUG-CL7-lake’

- (b) *-kwé*  
*u-mu-kwé* ‘AUG-CL1-son-in-law’  
*u-bu-kwé* ‘AUG-CL14-wedding’
- (c) *-haánga*  
*i-gi-haánga* ‘AUG-CL7-skull’  
*u-ru-haánga* ‘AUG-CL11-forehead’  
*u-mu-haánga* ‘AUG-CL1-intelligent’
- (d) *-rími*  
*u-ru-rími* ‘AUG-CL11-tongue/language’  
*u-bu-rími* ‘AUG-CL14-stuttering’

## 2.8. Metaphoric derivation

Although, classes in Bantu languages are grammatical as pointed out earlier, it happens that the majority of nouns referring to inanimate objects are found in class 7. What is interesting in this language is that human beings who are useless, harmful, and fearful are found in this class also.

- (9) *i-ki-reémba* ‘AUG-CL7-impotent’  
*i-gi-saambo* ‘AUG-CL7-crook’  
*i-gi-pfúamatwí* ‘AUG-CL7-deaf’

This is clearly a case of metaphoric derivation.

Similarly, all praise-names are found in class 9 where the majority of names for animals are found, especially (*i-n-ká*) the cultural icon.

- (10) *i-n-twáari* ‘AUG-CL9-hero’  
*i-n-toóre* ‘AUG-CL9-cadet’  
*i-m-fúra* ‘AUG-CL9-noble’

This is again a metaphoric derivation from the cow. Cow metaphors have of course very much influenced the language especially in vocabulary (Kimenyi 1999). This is an instance of its influence on morphology as well.

## 2.9. Derivational classifiers

Some classes (6, 7, 8, 11, 12, 13, and 14) serve as expressive classes (showing the attitude and feelings of the speaker: likes, dislikes), augmentatives, diminutives, and abstract nouns. Any noun can be shifted in any one of these derivational classes.

Class 6, *-ma-*, when used as derivational morpheme shows the speaker’s indignation:

- (11) *i-gi-tabo* ‘AUG-CL7-book’ >  
*a-ma-tabo* ‘AUG-CL6-book (worthless books)’  
*u-mu-goré* ‘AUG-CL1-woman’ >  
*a-ma-goré* ‘AUG-CL6-woman (worthless women)’  
*u-mu-koóbwa* ‘AUG-CL1-girl’ >

- a-ma-koóbwa* ‘AUG-CL6-girl (worthless girl)’

Class 7, *-ki-* (plural 8, *-bi-*), as a derivational class functions as both expressive (pejorative or ameliorative) and augmentative:

- (12) *u-mu-goré* ‘AUG-CL1-woman (woman)’ >  
*i-ki-goré* ‘AUG-CL7-woman (attractive woman, disgusting woman or big woman)’  
*u-mu-koóbwa* ‘AUG-CL1-girl (girl)’ >  
*i-gi-koóbwa* ‘AUG-CL7-girl (attractive girl, disgusting girl or big girl)’

The same morpheme is also used to refer to languages or national customs:

- (13) *i-gi-faraansá* ‘AUG-CL7-French (French ways)’  
*i-ki-nyáfuriká* ‘AUG-CL7-African (African way)’  
*i-gi-kirisítu* ‘AUG-CL7-Christian (Christian way)’

This same class marker when used without a preprefix, serves as an adverb marker:

- (14) *u-mu-goré* ‘AUG-CL1-woman (woman)’ >  
*ki-goré* ‘CL7-woman (womanly)’  
*u-mw-áana* ‘AUG-CL1-child’ >  
*cy-áana* ‘CL7-child (like a child)’  
*u-mu-swa* ‘AUG-CL1-ignorant’ >  
*gi-swa* ‘CL7-ignorant (like an ignorant person)’

The other classes which are used to derive adverbs are class 6 marker *-ma-* and class 14 *-bu-*, but they are not as productive as *ki-* marker.

Adverbs in *-ma-* are used idiomatically:

- (15) (a) *guheeka ma-pfisi*  
 carry CL6-hyena  
 ‘to carry on the back upsidedown’  
 (literally ‘to carry like a hyena’)
- (b) *kwíciisha bu-gufi*  
 to.make.oneself.pass CL14-short  
 ‘to be humble’

Class 9, *-n-* prefix, is used to derive nouns which express qualities and these nouns can modify nouns in any category without changing the class marker of the head:

- (16) *i-n-kootan-yi* ‘great fighter (CL9)’  
*i-n-twáar-i* ‘courageous (CL9)’  
*i-n-guumba* ‘sterile (CL9)’

This class is sometimes used with certain nouns to show contempt:

- (17) *a-b-áana* ‘AUG-CL2-child (children)’ >  
*i-n-záana* ‘AUG-CL9-child (bad children)’  
*a-ba-sóre* ‘AUG-CL2-young.man (young men)’ >  
*i-n-sóre* ‘AUG-CL9-young.man (bad young men)’

Class 11, *-ru-* (plural 8, *-bi-*), as derivational class marker also functions as both expressive and augmentative:

- (18) *u-mu-goré* ‘AUG-CL1-woman’ >  
*u-ru-goré* ‘AUG-CL11-woman (attractive woman, disgusting woman, big woman)’  
*u-mu-koóbwa* ‘AUG-CL1-girl’ >  
*u-ru-koóbwa* ‘AUG-CL11-girl (attractive girl, disgusting girl, big girl)’

Class 12, *-ka-* (plural marker 13, *-tu-*, or 14, *-bu-*), as a derivational marker also has an expressive or diminutive function (vgl. 18):

- (19) *a-ka-goré* ‘AUG-CL12-woman (attractive woman, disgusting woman or small woman)’  
*a-ga-koóbwa* ‘AUG-CL12-girl (attractive girl, disgusting girl or small girl)’

#### 2.10. Nominal suffixes

As said earlier, nouns are poor in derivation. Only **zero-derivation** or **prefix derivation** is used mostly for metonymic derivation. In very few cases, however, noun stems are added to certain words to show gender. These noun stems are *-gabo* ‘man’; *-rume* ‘man (archaic)’, *-goré* ‘woman’.

- (20) (a) *igishéegabo* ‘virago’  
*igitsínagabo* ‘masculine’  
(b) *nyokórume* ‘your mother’s brother’  
*impweerume* ‘male dog’  
(c) *umukómagoré* ‘type of tree’  
*igitsínagoré* ‘feminine’

The noun stem *-áana* ‘child’ is also found in a handful of words to refer to a young thing.

- (21) *ikibwáana* ‘puppy’  
*umukazáana* ‘son’s wife’

The only productive suffix is *-kazi* from Proto-Bantu ‘wife’ which is added to any noun to mean female of:

- (22) *umwáarimú* ‘teacher’ >  
*umwáarimú-kazi* ‘teacher-F’  
*umunyarwaanda* ‘Rwandan’ >  
*umunyarwaanda-kazi* ‘Rwandan-F’

The only true suffix whose etymology is unknown is *-aazi* which conveys a pejorative meaning:

- (23) *ikiyogá* ‘drink’ >  
*ikiyogáazi* ‘bad drink’  
*ikiintu* ‘thing’ >  
*ikíntaázilíkíntaází* ‘a bad thing’  
*ikigoré* ‘woman’ >  
*ikigoráazi* ‘woman of low character’

#### 2.11. Denominal verbs

There are no specific rules to derive verbs from nouns. They are obtained by either putting the verb final vowel at the end of the stem or adding a prefix.

**Denominal verbs** receive the lexical tone of the noun they derive from: If the noun is toneless then the verb will be toneless as well. The lexical tone is always assigned to the first mora of the verb stem because infinite verbs carry a high tone only on the first mora of the first syllable.

- (24) *u-mu-nebwe* ‘AUG-CL1-lazy’ >  
*ku-nebw-a* ‘INF-lazy-IPFV (to become lazy)’  
*i-n-jiji* ‘AUG-CL9-stupid’ >  
*ku-jij-w-a* ‘INF-stupid-PASS-IPFV (to become stupid)’  
*u-mu-sáaza* ‘AUG-CL9-old.man’ >  
*gu-sáz-a* ‘INF-old.man-IPFV (to get old)’

Denominal verbs are also sometimes obtained by suffixation. Here, too, the lexical tone of the noun is not affected, either. These verbs are obtained by adding a variety of morphemes.

- (25) *u-ru-muri* ‘AUG-CL11-light’ >  
*ku-mur-ik-a* ‘INF-light-STAT-IPFV (to light)’  
*i-naanga* ‘AUG-harp’ >  
*ku-naang-ir-a* ‘INF-harp-APPL-IPFV (to tighten)’

#### 2.12. Compound nouns

Kinyarwanda has two types of **compound nouns**, nouns with onomastic prefixes and nouns which consist of a verb in the relative clause form and a complement.

##### 2.12.1. Nouns with onomastic prefixes

In Kinyarwanda, names are created from any noun by deleting the prefix or by adding an onomastic prefix before the noun or any syntactic construction. Full details and illustrative examples are provided in Kimenyi (1989).

- (26) *i-n-gabo* 'shield'  
*ngabo* 'Ngabo'  
*sée-ngabo* 'Sengabo'  
*rwáa-ngabo* 'Rwangabo'

The prefixes *rwáa-* and *káa-* have the high tone disappear if the head noun has also a high tone:

- (27) (a) *a-ma-nywá* 'AUG-CL6-day.times'  
*rwaamanywá* 'Rwamanywa'  
*kaamanywá* 'Kamanywa'  
*mukaámmanywá* 'Mukamanywa'  
*nyirámmanywá* 'Nyiramanywa'  
(b) *i-sibo* 'AUG-speed'  
*rwáasibo* 'Rwasibo'  
*kásibo* 'Kasibo'  
*mukásibo* 'Mukasibo'  
*nyirasibo* 'Nyirasibo'

The prefixes *sée-* and *rwáa-* are used for male names only, whereas *káa-*, *mukáa-*, and *nyirá-* are for female names.

#### 2.12.2. Verb-noun compound

The verb-noun compound is obtained by putting the verb in the headless relative clause and adding a noun without a prefix. The compound word undergoes tone rhythm rules. That is, if the head noun has a high tone, the compound noun can acquire a secondary or tertiary tone.

The rules which are responsible for compound noun creation are the noun preprefix deletion, assignment of a high tone on the first mora of the noun if has lexical high tone and another high tone on the first mora of the verb stem. High tone deletion or beat deletion or what is referred to in Bantu linguistics as Meeussen's rule takes place within the noun complement if primary beat addition is followed by a high tone on the next syllable. The maximum number of tones that a noun can have is only three.

- (28) *aba-tég-ariú-goré*  
*/a-ba-tég-a+u-ru-goré/*  
 AUG-CL2-wear-IPFV + AUG-CL11-maternity.crown  
 'those who wear the maternity crown (women)'  
*in-shóber-amá-haánga*  
*/i-n-shóber-a+a-ma-haánga/*  
 AUG-CL9-not.understand-IPFV + AUG-CL6-for-eign.country  
 'idiomatic expressions'

The noun morphology is completely different from that of the verb to which we will now turn.

### 3. Verb

The **verb morphology** is more complex than that of the noun. The verb can have multiple prefixes and multiple suffixes. Prefixes have only grammatical information whereas suffixes can be both grammatical and lexical. The verb morphology, however, is influenced by the following factors: Whether it occurs in the main clause or the subordinate clause, whether it is in the affirmative or negative form, and finally whether it occurs with or without a complement.

The obligatory morphemes are the verb stem, the subject agreement prefix, and the final vowel which most of the time is the aspect marker. This is illustrated in (29).

- (29) (a) *A-b-áana ba-som-a*  
 AUG-CL2-child they-read-IPFV  
*i-bi-tabo.*  
 AUG-CL8-book  
 'The children read books.'  
(b) *A-b-áana ba-som-ye*  
 AUG-CL2-child they-read-PFV  
*i-bi-tabo.*  
 AUG-CL8-book  
 'The children have just read books.'  
(c) *A-b-áana ba-som-e*  
 AUG-CL2-child they-read-SUBJ  
*i-bi-tabo.*  
 AUG-CL8-book  
 'The children should read books.'

In (29 a) the verb ends with the imperfective aspect marker *-a*, in (29 b) the perfective aspect morpheme *-ye*, and (29 c) the subjunctive marker *-e*.

The optional morphemes are the **proclitics** *nti-* 'not' and *ni-* 'if/when', the tense-aspect-modality morphemes, the morpheme *-na* 'also', the object pronouns which can be one or many, lexical verb extensions, grammatical suffixes, and the enclitics *-ga* and locative postclitics: *-mó*, *-hó*, and *-yó*.

#### 3.1. The proclitics *nti-* and *ni-*

The negative marker *nti-* precedes the whole verb complex and is used in all verbs in main clauses regardless of mood whereas in subordinate clauses *-ta-* which immediately follows the subject marker is used. *Nti-* is a tone neutralizing morpheme. The morpheme *ni-* is used optionally with verbs in the subjunctive mood and in realis conditionals. It can thus be translated as 'if' or 'when'.

- (30) (a) *A-b-áana ba-kor-é.*  
 AUG-CL2-child they-work-SUBJ  
 ‘The children work.’
- (b) *A-b-áana ní-ba-kór-e.*  
 AUG-CL2-child *ni*-they-work-SUBJ  
 ‘The children should work.’
- (31) *A-b-áana ní-ba-kor-á,*  
 AUG-CL2-child *ni*-they-work-IPFV  
*u-ba-heémb-e.*  
 you-them-pay-SUBJ  
 ‘If the children work pay them. / When  
 the children work pay them.’
- 3.2. The tense-aspect-modality morphemes  
 The **tense-aspect-modality** refers not only to tenses in the traditional sense but to other aspects such as mood, inchoativity, progressivity, etc.
- Semantically, Kinyarwanda has only five referential tenses, namely present, habitual, early today, past, late today, and future. Morphologically, however, the present tense, the habitual tense, and the late today tense look the same.
- (32) *A-b-áana ba-ra-som-a.*  
 AUG-CL2-child they-T-read-IPFV  
 ‘The children are reading. / The children read. / The children will be reading (today).’
- (33) *U-mu-goré a-vuz-e kó*  
 AUG-CL1-woman she-say-IPFV that  
*a-b-áana ba-som-á.*  
 AUG-CL2-child they-read-IPFV  
 ‘The woman just said that the children are reading. / The woman just said that the children read. / The woman just said that the children will be reading.’

A full discussion of the tense-aspect-morphology, how it is assigned and how it affects tone rules is discussed at great length in Kimenyi (1980; 2002). As examples in (34) show, the verb stem *-kór-* has a lexical high tone and the verb stem *-som-* is not tone marked. In some cases, tenses keep this tone contrast, in others these tones are neutralized by either assigning tones to non-toned verb stems or by deleting tones, making all types of verb stems look the same tonologically speaking.

- (34) *a-ba-gabo ba-racyáa-kór-a*  
 AUG-CL2-man they-still-work-IPFV  
 ‘the men are still working’
- a-ba-gabo ba-racyáa-som-a*  
 AUG-CL2-man they-still-read-IPFV  
 ‘the men are still reading’

- a-ba-gabo ba-a-zaa-kor-a*  
 AUG-CL2-man they-COND-FUT-work-IPFV  
 ‘the men would work’
- a-ba-gabo ba-a-zaa-som-a*  
 AUG-CL2-man they-COND-FUT-read-IPFV  
 ‘the men would read’

It is also possible to have two tenses in the same spot.

- (35) *U-mu-gabo y-a-tum-ye kó*  
 AUG-CL1-man he-PAST-send-IPFV that  
*ba-záa-z-a*  
 they-FUT-come-IPFV  
*ba-ka-zaa-som-a.*  
 they-CONSEC-FUT-read-IPFV  
 ‘The man sent a message that they will come and play.’

### 3.3. Object pronouns

Kinyarwanda is among Bantu languages which allow multiple object incorporated pronouns. Unlike some languages such as Swahili in which the pronoun is used as object agreement, in Kinyarwanda these objects are strictly used as pronouns only. Thus the whole verb phrase of the English sentence in (36) is translated into one single agglutinative verb as shown by the example in (37).

- (36) *The woman is also making us read it with them to you for me there.*  
 (The woman is also making us read the book with eyeglasses to you for me in the house).

- (37) *U-mu-goré*  
 AUG-CL1-woman  
*a-ra-na-ha-ki-zi-ba-ku*  
 she-T-also-there-it-it-them-you-  
*n-som-eesh-eesh-er-er-eza*  
 me-read-CAUS-CAUS-APPL-APPL-IPFV

The word order of multiple **object-pronouns** is dictated by either case or personal pronouns. The benefactive comes closer to the verb stem, the dative comes next, and the accusative precedes all of them.

- (38) (a) *U-mu-gabo a-ra-som-er-a*  
 AUG-CL1-man he-T-read-APPL-IPFV  
*u-mu-goré a-b-áana*  
 AUG-CL1-woman AUG-CL2-child  
*i-gi-tabo.*  
 AUG-CL7-book  
 ‘The man is reading the book to the children for the woman.’

- (b) *U-mu-gabo*  
 AUG-CL1-man  
*a-ra-ki-ba-mu-som-er-a.*  
 he-T-it-them-her-read-APPL-IPFV  
 'The man is reading it to them for her.'

If the verb has a reflexive pronoun and a first person direct object, the reflexive becomes part of the stem, followed by the first person, then the second person and the third person comes last. The structure thus becomes ambiguous because any of these pronouns can have any semantic role.

- (39) *A-ba-goré*  
 AUG-CL2-woman  
*ba-zaa-mu-kuu-ny-iy-erek-er-er-a.*  
 they-FUT-him/her-you-me-show-APPL-  
 APPL-IPFV  
 'The women themselves will show her/him to you for me. / The women themselves will show you to her/him for me. / The women themselves will show me to you for him/her. / The women themselves will show you to me for him/her. / The women themselves will show you to him/her for me.'

These complex structures are not used very often because of the difficulty in both production and processing.

### 3.4. Lexical suffixes

**Lexical suffixes** have the same shape as the verb stem namely: -VC-, -VVC-, VVNC-, -VCVC-, -VVVCVC-.

The most common lexical suffixes which are either lexical (part and parcel of the verb stem) or derivational are *-am-*, *-aang-*, *-iriz-*, *-agur-*, *-aanur-*, *-uur-/uuk-*, *-ur-/uk-*, *-iir-*, *-arar-*. It is also possible sometimes to find multiple lexical suffixes within the same verb.

The meanings conveyed by these suffixes are those of iterativity, intensification, stativity, reversal, completeness, etc. This is again discussed at great length in Kimenyi (1980; 2002).

### 3.5. Grammatical suffixes

**Grammatical suffixes** are the causative/instrumental morpheme *-iish-*, the applicative *-ir-*, the comitative/reciprocal *-an-*, and the passive morpheme *-w-*. They differ from the lexical ones not because they come last but also because they are productive. Any verb can occur with them. They are thus in a sense like inflectional morphemes.

The passive morpheme is inserted between the last consonant after consonant mutation has taken place and the final vowel.

- (40) (a) *gu-som-a* 'INF-read-IPFV (to read)' >  
*gu-som-eesh-a* 'INF-read-CAUS-IPFV  
 (to cause to read)' >  
*gu-som-er-a* 'INF-read-APPL-IPFV (to  
 read for)' >  
*gu-som-an-a* 'INF-read-REC-IPFV (to  
 read together)'  
 (b) *gu-som-w-a* 'INF-read-PASS-IPFV (to  
 be read)' >  
*gu-som-eesh-w-a* 'INF-read-CAUS-  
 PASS-IPFV (to be cause to be read)' >  
*gu-som-er-w-a* 'INF-read-APPL-PASS-  
 IPFV (to be read to)' >  
*gu-som-an-w-a* 'INF-read-REC-PASS-  
 IPFV (to be read with)'

#### 3.5.1. Multiple grammatical suffixes

The same verb can have multiple suffixes because in this language oblique objects can become structural direct objects by deleting prepositions and adding appropriate suffixes to the verb.

- (41) *U-mu-goré a-ra-som-a.*  
 AUG-CL1-woman she-T-read-IPFV  
 'The woman is reading.'  
 (42) *U-mu-goré a-ra-som-er-a*  
 AUG-CL1-woman she-T-read-APPL-IPFV  
*a-ba-koóbwa n'i-i-n-dórerwamú*  
 AUG-CL2-girl with-AUG-CL9-glass  
*z'a-a-b-áana.*  
 of-AUG-CL2-child  
 'The woman is reading to the girls with  
 glasses of the children.'  
 (43) *U-mu-goré*  
 AUG-CL1-woman  
*a-ra-som-eesh-er-er-er-er-ez-a*  
 she-T-read-CAUS-APPL-APPL-APPL-CAUS-  
 IPFV  
*a-b-áana i-n-dórerwamó*  
 AUG-CL2-child AUG-CL9-glass  
*a-ba-koóbwa.*  
 AUG-CL2-girl  
 'The woman is using the children's  
 glasses to read to the girls.'

#### 3.5.2. The repetition of the same suffix

The same suffix can be repeated once or many times because these grammatical suffixes have different semantic roles.

The suffix *-an-* indicates reciprocity, accompaniment or the lack of a dative complement as the following examples show.

- (44) (a) *a-ba-goré ba-r-éerek-a*  
 AUG-CL2-woman they-T-show-IPFV  
*a-ba-antu a-ma-shusho*  
 AUG-CL2-person AUG-CL6-picture  
 ‘the women are showing people pictures’
- (b) *a-ba-goré*  
 AUG-CL2-woman  
*ba-r-éerek-an-a*  
 they-T-show-REC-IPFV  
*a-ba-antu a-ma-shusho*  
 AUG-CL2-person AUG-CL6-picture  
 ‘the women are showing the pictures to the people together’
- (c) *a-bagoré*  
 AUG-CL2-woman  
*ba-r-éerek-an-a*  
 they-T-show-REC-IPFV  
*a-ma-shusho*  
 AUG-CL6-picture  
 ‘the women are showing each other pictures’
- (d) *a-ba-goré*  
 AUG-CL2-woman  
*ba-r-éerek-an-an-a*  
 they-T-show-COMIT-REC-IPFV  
*a-ma-shusho*  
 AUG-CL6-picture  
 ‘the women together are showing each other pictures’
- (e) *a-ba-koóbwa*  
 AUG-CL2-girl  
*ba-ra-vúg-ir-an-ir-an-a*  
 they-T-talk-APPL-REC-APPL-REC-IPFV  
 ‘the girls talk for each other’

The morpheme *-iish-* is both a causative and an instrumental marker.

- (45) (a) *u-mu-huúngu a-ra-tém-eesh-a*  
 AUG-CL1-boy he-T-cut-INST-IPFV  
*i-gi-tí i-n-tóorezo*  
 AUG-CL7-tree AUG-CL9-ax  
 ‘the boy is cutting the tree with an ax’
- (b) *u-mu-gabo a-ra-tém-eesh-a*  
 AUG-CL1-man he-T-cut-CAUS-IPFV  
*u-mu-huúngu i-gi-tí*  
 AUG-CL1-boy AUG-CL7-tree  
 ‘the man is making the boy cut the tree’
- (c) *u-mu-gabo*  
 AUG-CL1-man  
*a-ra-tém-eesh-eesh-a*  
 he-T-cut-CAUS-CAUS-IPFV

- u-mu-huúngu i-gi-tí*  
 AUG-CL1-boy AUG-CL7-tree  
*i-n-tóorezo*  
 AUG-CL9-ax  
 ‘the man is making the boy cut the tree with an ax’

Similarly, the applicative morpheme *-ir-* can act as a dummy morpheme, a carefree morpheme, a circumstantial morpheme, a possessive morpheme, or a benefactive and dative case marker (cf. Kimenyi 1980; 1995; 2002).

### 3.5.3. The word order of grammatical suffixes

The word order of grammatical suffixes mirrors syntax. Direct objects which come close to the verb have their corresponding suffixes closer to the verb stem. The applicative suffix *-ir-*, however, always follows the causative morpheme *-iish-*, and *-y-* another causative morpheme always comes last.

- (46) (a) *A-b-áana ba-ra-kór-a.*  
 AUG-CL2-child they-T-work-IPFV  
 ‘The children work.’
- (b) *A-b-áana*  
 AUG-CL2-child  
*ba-ra-kór-an-a.*  
 they-T-work-COMIT-IPFV
- (c) *U-mu-goré*  
 AUG-CL1-woman  
*a-ra-kór-an-iish-a*  
 she-T-work-COMIT-CAUS-IPFV  
*a-b-áana b'-ú-mu-gabo.*  
 AUG-CL2-child of-AUG-CL1-man  
 ‘The woman is making children of the man work together.’
- (d) *Umugore*  
*U-mu-goré*  
 AUG-CL1-woman  
*arakoraniishiriza*  
*a-ra-kór-an-iish-ir-ir-y-a*  
 she-T-work-with-CAUS-APPL-APPL-  
 CAUS-IPFV  
*umugabo abaana.*  
*u-mu-gabo a-b-áana.*  
 AUG-CL1-man AUG-CL2-child  
 ‘The woman is making the man’s children work together.’

The word order of object pronouns and grammatical suffixes is a mirror image of the sentence structure of the language. First, Kinyarwanda allows two inherent direct objects and multiple direct objects through the process of objectivization of adjunct noun phrases by suffixation. The word order of

these objectivized objects is not random. There seems to be a hierarchy of semantic roles. The agent seems to come first, followed by the possessor or the benefactive. The accusative always comes last before the dative.

Many Kinyarwanda verbs allow two direct objects (inherent accusatives) and several direct objects (structural accusatives) through the process of objectivization.

### 3.6. The aspect marker

Kinyarwanda has two **aspect markers**: the imperfective marked by *-a* and the perfective marked by *-ye*.

The imperfective is found with habitual tenses, future tenses, progressive tenses, “not yet” tenses because it shows an event or an action not yet over. Interestingly, however, stative verbs have to take the perfective marker to show that the state has not changed. When suffixed with the imperfective aspect morpheme, stative verbs have a habitual reading only. The verbs *-som-* ‘read’ and *-rwáar-* ‘be sick’, active and stative verbs respectively are used below for illustration.

- (47) (a) *A-b-áana ba-ra-som-a.*  
AUG-CL2-child they-T-read-IPFV  
‘The children are reading.’
- (b) *A-b-áana ba-ra-som-ye.*  
AUG-CL2-child they-T-read-PFV  
‘The children have just read.’
- (48) (a) *A-b-áana ba-ra-rwáar-a.*  
AUG-CL2-child they-T-be.sick-IPFV  
‘The children get sick.’
- (b) *A-b-áana ba-ra-rwáa-ye.*  
AUG-CL2-child they-T-be.sick-PFV  
‘The children are sick.’

*-ga* which is used in past tenses to mark imperfective aspect is really a clitic and not part of the verb stem, first because it comes at the end, secondly because Kirundi doesn’t have it and lastly because vowel lengthening doesn’t take place when the verb ends in *w* as part of the stem or as a passive marker.

### 3.7. The locative postsuffixes

The **locative enclitics**, *hó*, *mó*, and *yó* correspond respectively to the prepositions *ku*, *mu*, and *i*. They are not yet completely part of the verb morphology. They don’t affect its phonology, segmentally or suprasegmentally. Native speakers are always aware of this because in orthography there are those who use them as part of the verb and those who write

them separately. Nothing, however, can be inserted between them and the verb.

- (49) *gu-sig-a-hó* ‘INF-stop-IPFV-LOC<sub>1</sub> (to stop)’
- kw-iigir-a-yó* ‘INF-stop-IPFV-LOC<sub>2</sub> (to push)’
- gu-hít-a-mó* ‘INF-stop-IPFV-LOC<sub>3</sub> (to choose)’

### 3.8. Deverbal Nouns

**Deverbal nouns** and denominal verbs are formally distinguishable in Kinyarwanda. Denominal verbs are assigned the lexical tone of the noun, a high tone on the stem first mora if the original noun has a high tone but toneless if the original noun doesn’t have any. Deverbal nouns, on the other hand, are assigned a high tone either on the second mora of the stem or on the final vowel.

Nominal stems don’t have derivational suffixes except few such as *-kazi* and *-aazi* as seen in the section on noun morphology (see 2.10).

There are many verbal nominal suffixes, however, namely nominalizer *-yi*, the objective *-o*, the manner *-ré*, the stative *-e* as examples below indicate.

- (50) (a) *-yi:*  
*gu-híng-a* ‘INF-farm-IPFV (to farm)’ >  
*u-mu-híinz-yi* ‘AUG-CL1-farm-NR<sub>1</sub>  
(farmer)’
- (b) *-o:*  
*gu-téer-a* ‘INF-attack-IPFV (to attack)’ >  
*i-gi-téer-o* ‘AUG-CL7-attack-NR<sub>2</sub> (an attack)’
- (c) *-ré:*  
*ku-íi-twaar-a* ‘INF-REFL-carry-IPFV  
(to behave)’ >  
*i-my-íi-twaar-re* ‘AUG-CL4-REFL-  
carry-MANR (behavior)’

The suffixes *-yi* and *-o* can occur with verb stems with prefixes of any noun class but *-ré* always requires the prefix *-mi-* of class 4.

There are other nouns also which are derived by tone, by zero affixation or by syllable truncation.

#### 3.8.1. Derivation by tone

The final vowel of deverbal nouns is either the one found in all infinitive forms *-a* or the one found in subjunctive forms *-e*. These nouns are also found in both the active or passive forms, the latter being marker by the

morpheme *-w-* which appears before the final vowel (cf. 3.5).

- (51) *i-hené* ‘AUG-goat’ <  
*gu-hen-a* ‘INF-goat-IPFV (to show the anus’)  
*u-mu-kené* ‘AUG-CL1-poor’ <  
*gu-ken-a* ‘INF-poor-IPFV (to be poor’)  
*i-temé* ‘AUG-bridge’ <  
*gu-tém-a* ‘INF-bridge-IPFV (to cut’)

**3.8.2. Derivation by last syllable truncation**  
Kinyarwanda uses very productively, last syllable **truncation** process to create deverbal nouns. The second mora high tone assignment is of course present all the time as the following examples show.

- (52) *i-cy-aámbu* ‘AUG-CL7-bridge’ <  
*kw-aambuk-a* ‘INF-cross-IPFV’  
*u-mu-hemú* ‘AUG-CL1-traitor’ <  
*gu-hemuk-a* ‘INF-betray-IPFV’  
*i-gi-kanú* ‘AUG-CL7-neck’ <  
*gu-kanur-a* ‘INF-open.eyes-IPFV’

Last syllable truncation is a case of markedness reversal. It becomes a marker also.

#### 4. Adjective

The number of **adjectives** is very much limited. Kinyarwanda doesn’t have any color adjective for instance. They are all nouns and they all occur in the possessive construction as the following color names illustrate:

- (53) (a) ‘green’ < *i-cy-aátsi kibísi* ‘raw grass’;  
(b) ‘red’ < *u-ru-gina* ‘anthill’;  
(c) ‘black’ < *u-mu-kara* ‘otter/marten’;  
(d) ‘light skin’ <*i-n-zóbe* ‘swamp antelope’;  
(e) ‘white’ < *u-ru-táre* ‘rock’;  
(f) ‘yellow’ < *u-mu-hoondo* ‘milk from a cow which has just mated’; etc.

Words which have adjectival properties, thus agreeing with the head noun are the following:

- (a) Numerals referring to one through seven  
(54) *-mwé* ‘one’, *-biri* ‘two’, *-tatu* ‘three’, *-né* ‘four’, *-taanu* ‘five’, *-taandátu* ‘six’ and *-riindwi* ‘seven’.

Other numerals occur in the apposition construction and ordinal numbers even the ones

referring to one through seven occur in the possessive construction (N of N)

(b) Quantifiers

- (55) *-óombí* ‘both’, *-óosel-éese* ‘all’, *-ndi* ‘other’, *-vvhe?* ‘which’.

(c) Qualitative adjectives

- (56) *-bísi* ‘raw’, *-iizá* ‘beautiful/good’, *-bí* ‘bad’, *-tó* ‘young’, *-gufí* ‘short’, *-ree-re* ‘long/tall/deep’, *-shyá* ‘new’, *-zima* ‘healthy/alive’, *-kurú* ‘old’, *-taraga* ‘alive’, *-ké* ‘few’, *-iinshi* ‘many’, *-gari* ‘large’, *-nini* ‘big’, *-hiire* ‘happy’, *-nziinyá* ‘tiny’.

(d) Demonstratives

There are four types of demonstratives: near speaker which is realized as the copy of both the preprefix and prefix of the head noun; near hearer realized as *-o*; far from both hearer and speaker marked by *-riiyá*, and *-áa* and *-ryá*, which refers to antecedents distant in time.

- (57) *iki gitabo* ‘this book’  
*icy-o gitabo* ‘that book (next to you)’  
*ki-riyya gitabo* ‘that (far away) book’  
*cy-áa gitabo* ‘that book (remote past)’  
*ki-rya gitabo* ‘that book (remote past)’

The *-o* demonstrative is also used to refer to something that has just been mentioned in the conversation.

Modification is accomplished by other devices such as the use of apposition (see 58), the possessive construction (see 59), or the relative clause (see 60).

- (58) *a-b-áana i-sokó*  
AUG-CL2-child AUG-flea.market  
‘a lot of children’  
(59) *u-mu-gabo w'u-mu-sáaza*  
AUG-CL1-man of-AUG-CL1-old  
‘an old man’  
(60) *u-mw-áana u-rwáa-ye*  
AUG-CL1-child he.who-be.sick-PFV  
‘sick child’

Possessive adjectives don’t exist in Kinyarwanda either. The possessive construction (such ‘of me’, ‘of you’, ‘of her’, etc.) is used instead.

#### 5. Unclassified categories

Bantu languages are also characterized by the existence of a linguistic category known as “ideophones” as well as the lack of “function words”.

### 5.1. Ideophones

**Ideophones**, a concept studied by Welmers (1973) in different African languages, is a term coined by C. M. Doke (1935) to mean a “vivid representation of an idea in sound”. This linguistic item which conveys “a pictur-esque connotation” or “a word which describes a predicate, qualificative or adverb in respect to manner, color, action, state or intensity”, is not only found in Bantu languages but in the whole Niger-Congo family.

Some of the formal properties of ideophones is the lack of class markers and the excessive use of reduplication, triplication or quadruplication.

- (61) *hó* ‘sound of barking dog’;  
*pyii* ‘sound of hot object in contact with a liquid’;  
*yuguyugu* ‘very fast’;  
*tsiritsiri* ‘very dark’;  
*yíuyíuyíuyíú* ‘sound which expresses pain or astonishment’;  
*póopóopóopóó* ‘sound expressing astonishment’.

Ideophones are expressive and can have multiple interpretations. It is thus difficult to get the proper meaning out of the proper context. I have argued (Kimenyi 1997) that ideophones have to be examined at four separate levels namely: form, function, content, and context.

### 5.2. Function words

Kinyarwanda and other Bantu languages don't have **function words**, such as prepositions, conjunctions, subordinators, auxiliaries. Most of the prepositions, for instance, are expressed by prepositional noun phrases referring to body parts: *in front* ‘in the eyes of’, *behind* ‘in the back of’, *next* ‘in the ribs of’, *inside* ‘in the stomach of’, *over/above/on* ‘the sky of’, *under* ‘the earth of’.

- (62) (a) *h-eejuru y'áa-m-éezá*  
CL16-sky of:AUG-CL6-table  
‘on the top of the table’  
(b) *muunsi y'áa-m-éezá*  
earth of:AUG-CL6-table  
‘under the table’

In some instances this category is being grammaticalized but in the majority of cases, function words become so because of the syntagmatic derivation. It is the structure that is its position within the sentence which determines whether a noun or verb is a function word.

### 5.3. Serial construction of auxiliaries

Another interesting feature of Bantu languages is the multiple **serial construction** of auxiliary verbs as seen in (63). These verbs, as their translation shows, become auxiliary by virtue of their syntagmatic position. Otherwise they can also occur as independent content verbs.

- (63) *Aba b-áana ba-a-ri*  
these CL2-child they-PAST-be  
*bá-tuu-ye bá-saanz-w-e*  
they-live-PFV they-find-PASS-IPFV  
*bá-jy-a bá-kuund-a gu-pf-á*  
they-go-IPFV they-like-IPFV INF-die-IPFV  
*kuu-z-a ku-tu-reeb-a byíburá*  
INF-come-IPFV INF-us-see-IPFV at.least  
*rimwé mu kwéezi.*  
once in month  
‘These children usually came to see us at least once a month.’

## 6. Tones

### 6.1. Role of tones

**Tones** in Kinyarwanda play different roles: they are lexical, morphological, and syntactic (cf. Kimenyi 1979; Sibomana 1974).

They differentiate words which are segmentally the same, they mark tenses and moods or types of phrases, they derive new words or mark syntactic constructions. Examples in (64) show words which are differentiated only by tones. Examples in (65) show that relative clauses are marked by tones, in (66) we have deverbal nouns marked by tones, in (67) we see compound words which have a different tone output than the words they are derived from and finally, the last examples in (68) show that verbal tones are marked or not marked depending on whether there is a complement or not.

- (64) *u-mu-ryaango* – *u-mu-ryáango*  
‘AUG-CL1-family’ ‘AUG-CL1-door’  
*ino* – *inó*  
‘toe’ ‘here’
- (65) (a) *a-bá-ana ba-som-a*  
AUG-CL2-child they-read-IPFV  
*i-bi-tabo*  
AUG-CL8-book  
‘the children read books’  
(b) *a-bá-ana ba-som-á*  
AUG-CL2-child they-read-IPFV.REL  
*i-bi-tabo*  
AUG-CL8-book  
‘the children who read books’

- (66) *gu-kór-w-a* ‘INF-do-PASS-IPFV (to be done)’ – *i-bi-kor-w-á* ‘deeds’  
*gu-kórор-a* ‘to cough’ –  
*i-n-kórор-á* ‘a cough’
- (67) (a) *nyírí izína* ‘the owner of the name’ – *nyírína* ‘in person’ (owner name)  
(b) *amayíra abíri* ‘two paths’ – *amayiráabíri* ‘crossroads’  
(c) *wé nyíne* ‘s/he indeed’ – *wéenyíne* ‘(s/he) alone’
- (68) (a) *A-ba-gabo ba-raa-z-a*  
AUG-CL2-man they-PRES-come-IPFV  
*ba-ga-kór-a*.  
they-CONSEC-work-IPFV  
‘The men come and work.’
- (b) *A-ba-gabo ba-raa-z-a*  
AUG-CL2-man they-PRES-come-IPFV  
*ba-ga-kor-a i-sáahá.*  
they-CONSEC-work-IPFV AUG-hour  
‘The men come and work for an hour.’

## 6.2. Tone assignment

Tone assignment takes into account word category. In nouns, lexical tones can be assigned anywhere in the word stem except the prefix and the preprefix. The noun can have “stray tones”, however. Stray tones don’t affect tone rules. In verbs, tones are assigned either on the first mora of the verb stem or on the first mora of the second syllable of the verb stem. The assignment of a high tone on the first mora of the verb stem is referred to as lexical tone and the assignment of high tone on the first mora of the verb stem second syllable as grammatical tone because low verb stems get it as well. Each morpheme is assigned only one high tone.

- (68) (a) *gu-kór-a* ‘to work’  
*gu-kin-a* ‘to play’
- (b) *a-ba-antu ba-kor-á*  
AUG-CL2-person they-work-IPFV.REL  
‘people who work’
- (c) *a-ba-antu ba-kin-á*  
AUG-CL2-person they-play-IPFV.REL  
‘people who play’

## 6.3. Floating tones

There are some tones in Kinyarwanda which are phonetically manifested only when a full word is formed or are only realized in phrases or sentences. Tones which require these specific structures to appear phonetically are referred to as **floating tones**.

The following words, morphemes and phrase types acquire a high tone only when they occur in the middle of the sentence or if another word precedes: *ni*, *si*, *nta*, *nga*, imperative verb form, negative imperative marker *-ii-*, conditional *ni-*, negative *nti-* and *si-*.

- (69) (a) *ni-i-bi-tabo*  
it.is-AUG-CL8-book  
‘it is books’
- (b) *ibyo ní-i-bi-tabo*  
those it.is-AUG-CL8-book  
‘those are books’
- (70) (a) *b-ii-kor-á*  
they-NEG-work-IPFV  
‘they should not work’
- (b) *ubu b-ii-kor-á*  
now they-NEG-work-IPFV  
‘they should not work’

Some noun and verb stems also acquire a high tone only when a prefix precedes.

- (71) *zi* ‘know’: *ba-ra-zi* > *ba-ráa-zi* ‘they-T-know (they know)’  
*tu-ra-ku-zi* > *tu-ra-kíuu-zi* ‘we-T-you-know (we know you)’

The presence of these tones only when a full word or sentence are formed strongly supports the existence of floating tones.

## 6.4. Tone rules

Using a metrical approach (Kimenyi 2002), Kinyarwanda tone rules are very simple. There are five types of them, namely: **beat movement**, **beat addition** and **beat deletion**, **iambic reversal**, and **weak beat alternation**.

### 6.4.1. Beat addition

In nouns, a second high tone (first beat addition) is assigned on the first mora of the stem if there are at least two moras between the mora which carries the lexical tone and the stem first syllable.

- (72) *u-mu-eendá* > *umwéendá* ‘AUG-CL3-debt’  
*i-ki-kookó* > *igikóokó* ‘AUG-CL7-animal’

Second beat addition occurs in compounds only. For verbs, the second beat is added on the first mora of the verb stem (cf. 73 a–b) and for the onomastic prefix *nyira-*, it is added on its second syllable (cf. 73 c).

- (73) (a) *i-n-shober-a+a-ma-haánga*  
> *inshóberamáhaánga*

- AUG-CL9-not.know-IPFV + AUG-CL6-for-  
eign.country  
'idiomatic expressions'
- (b) *ga-shaak-a+u-bu-haké*  
> *gasháakabúhaké*  
CL12-want-IPFV + AUG-CL14-servi-  
tude  
'colonialist'
- (c) *nyirá-umu-baandé* > *nyirámubáandé*  
OP-CL1-valley  
'valley'

#### 6.4.2. Beat deletion

When two high tones occur in adjacent moras because of high tone addition, the high tone on the right deletes. This beat deletion is known as Meeussen rule.

- (74) *i-n-kor-a + u-mu-tíma* > *inkóramútima*  
AUG-CL9-touch-IPFV + AUG-CL3-heart  
'conscientious'  
*nka kiményi* > *nka kímenyi*  
'like Kimenyi'

#### 6.4.3. Beat movement

Beat movement occurs in verbs only. As noted earlier, verbs in Kinyarwanda can have multiple direct object pronouns. When the verb stem has a high tone, it spreads to all vowels of these pronouns. This area is referred to as the "prosodic domain" or the "suprastem". When it reaches the leftmost mora of the prosodic domain it makes a U-tone creating a Left-to-Right Beat Alternation. That is strong beats switch evenly from strong to weak from the leftmost to the rightmost.

- (75) *ba-ra-kór-a* > *barakóra*  
they-T-work-IPFV  
'they work'  
*ba-ra-na-ha-bi-ba-mu-kór-eesh-erez-a*  
*baranáhabíbamúkoreeshereza*  
they-T-also-there-it-them-him/her-do-  
CAUS-APPL-IPFV  
'they also make them do it for them  
there'

#### 6.4.4. Iambic reversal

Some monosyllabic vowel prefixes or associative prefixes have their high tone shifted to the next mora on the right.

- (76) *ú-no* > *unó* 'this'  
*bá-no* > *báno* 'these'  
*á-te* > *até* 'how s/he?'  
*mú-te* > *múte* 'how you?'

#### 6.4.5. Weak beat insertion

The associative morphemes *na* 'with', *-a* 'of', *nka* 'like', the locative postsuffixes *-mó*, *-hó*, and *-yó* as well as the complementizers *kó*, *ahó*, *iyó*, *ubwó*, etc. undergo Weak Beat Addition.

- (77) *nká + inká* > *nkíinká* 'like + cow (like  
a cow)'  
*kirimó + amáazi* > *kirim(w)áamáazi*  
there.is.in + water  
'there is water in it'

### 7. Reduplication

**Reduplication** (cf. Art. 57) is very active in Kinyarwanda. There are three types of reduplication in this language: stem first syllable reduplication, stem reduplication, and whole word reduplication. Reduplication is both lexical and grammatical. Many words appear in the dictionary already reduplicated and others are formed from existing words. Reduplication expresses either plurality, iterativity, intensification, completeness, etc. As far as verbs are concerned only bisyllabic stems undergo reduplication. Monosyllabic and polysyllabic don't. Only adjectives allow full word reduplication. Stem first syllable reduplication is mostly lexical.

#### 7.1. First syllable reduplication

If the verb stem has a lexical high tone, this tone is only realized on the first mora of the reduplicated stem. The reduplicated syllable is also long or short depending on whether the stem first syllable is long or short.

- (78) *gu-cyoo-cyoor-a*  
'INF-RDP-STEM-IPFV (to ask nutty ques-  
tions)'  
*ku-jíuu-juuby-a*  
'INF-RDP-STEM-IPFV (to harass)'  
*ku-nyíuu-nyuuz-a*  
'INF-RDP-STEM-IPFV (to suck)'

#### 7.2. Verb stem reduplication

When the verb stem is reduplicated, both lexical tones and grammatical tones are only realized on the leftmost reduplicated stem. In some cases, the last vowel of the stem on the left is lengthened.

- (79) *ku-geend-a* 'INF-walk-IPFV'  
*ku-geend-a-geend-a* 'INF-walk-IPFV-  
walk-IPFV (to walk around)'  
*ku-ruund-a* 'INF-pile-IPFV'

- ku-ruund-a-ruund-a* ‘INF-pile-IPFV-pile-IPFV (to pile)’  
*gu-shim-a* ‘INF-scratch-IPFV’  
*gu-shim-a(a)-shim-a* ‘INF-scratch-IPFV-scratch-IPFV (to scratch repetitively)’

### 7.3. Noun reduplication

In nouns, if the stem has a high tone on the last syllable, it is not repeated on the second part of the stem.

- (80) *i-hené* ‘AUG-goat’  
*a-ma-hené-hene* ‘AUG-CL6-goat-RDP  
 (goat milk)’

### 7.4. Adjective reduplication

Polysyllabic adjectives get a high tone on the prefix of the stem on the right and monosyllabic stems have the last vowel of the stem on the left lengthened.

- (81) *-nini* ‘big’ *abaantu* ‘people’  
*banini* *banínibánini*  
*-tatu* ‘three’ *inká* ‘cows’  
*zitatu* *zitatuzítatu*

Noun stems’ last syllable lengthen also when reduplication takes place.

## 8. Problems in Kinyarwanda morphology

The following Kinyarwanda phenomena pose a challenge to morphological studies: (a) hundreds of words have different phonetic variations, (b) a larger number of words have an identical vowel in all stem syllables, (c) tones have morphological functions, and (d) the existence of dummy morphemes.

### 8.1. Words with phonetic variations

Kinyarwanda has thousands of words which have alternative different phonetic forms without any change in meaning or register. Look at the Kinyarwanda equivalent English adverbs ‘or’ and ‘at least’ below:

- (82) (a) ‘or’: *cyángwá*, *byáangwá*, *cyáangá*,  
*cyáangwá*, *náangwá*, *yáangwá*.  
 (b) ‘at least’: *byíiburá*, *byíibuzé*, *cyáaburá*,  
*cyáabuzé*, *cyíiburá*, *cyíibuzé*, *níiburá*, etc.  
 (c) ‘diploma’: *dipóroómi*, *dipóromá*, *dipóromó*,  
*dipóromú*, *dipóróoma*, *dipóróome*, etc.

All dictionaries list all these phonetic realizations of the same word as independent lexical

entries making the dictionary voluminous and superfluous. These clearly have one underlying morphological representation. The problem is to decide among all these different phonetic representations, which one is closer to the abstract one.

### 8.2. Suprasegmental representations or non-lineal morphology

It was shown that tones in this language have a morphological function as well (see 6.1). It was also indicated that some phonological tones “float” and are realized phonetically when they find a vowel to attach to (see 6.3). Grammatical tones are assigned to specific moras or syllables in the metrical domain. It is now accepted in phonology that indeed tones and segments belong to two distinct different tiers. As was pointed out earlier also, vowels or the noun prefix and preprefix (augment) are identical: this is clearly a vowel copy. The majority of Kinyarwanda words have an identical vowel in the stem:

- (83) *u-mu-biri* ‘AUG-CL1-body’, *i-ki-biíndi* ‘AUG-CL7-clay.jar’, *u-ru-háanga* ‘AUG-CL11-forehead’, *a-ma-táma* ‘AUG-CL6-cheeks’, *u-ku-bóko* ‘AUG-CL15-arm’.

The fact that loan-words from languages that have closed syllables have the vowel of the open syllable repeated in the closed syllables when they enter Kinyarwanda or are assigned default vowels, namely *u* for bilabial consonants and *i* elsewhere (Kimenyi 2002), lends support to the idea of having two distinct tiers for vowels and consonants in the morphological representation.

### 8.3. Dummy Morphemes

There are morphemes which are added to the verb structure without any semantic or grammatical functions. These are the morphemes *-ij-* in (84), the applicative morpheme *-ir-* in (85), the infinitive morpheme *-ku-* in (86) and the 10 class marker *-zi-* in (87).

- (84) (a) *A-ba-gabo*  
 AUG-CL2-men  
*ba-ra-hiing-iish-a*  
 they-T-cultivate-CAUS-IPFV  
*i-mi-rímá*      *a-ma-súka*.  
 AUG-CL4-field AUG-CL6-hoe  
 ‘The men are using hoes to cultivate the fields.’

- (b) *A-ba-gabo*  
 AUG-CL2-man  
*ba-hiing-iish-ij-e*  
 they-cultivate-CAUS-APPL-PFV  
*i-mi-rimá a-ma-súka.*  
 AUG-CL4-field AUG-CL6-hoe  
 ‘The men have just cultivated the fields with the hoes.’
- (85) (a) *U-mu-goré*  
 AUG-CL1-woman  
*a-ra-som-eesh-a i-n-dórerwamó*  
 she-T-read-CAUS-IPFV AUG-CL9-glass  
*z'uu-m-wáana.*  
 of:AUG-CL1-child  
 ‘The woman is reading with the child’s glasses.’
- (b) *U-mu-goré*  
 AUG-CL1-woman  
*a-ra-som-eesh-er-ez-a*  
*/a-ra-som-iish-ir-ir-y-a/*  
 she-T-read-CAUS-APPL-APPL-CAUS-  
 IPFV  
*u-mw-áana i-n-dórerwamó.*  
 AUG-CL1-child AUG-CL10-glass  
 ‘The woman is reading with the child’s glasses.’
- (86) *mu-i-andik-a > mu-i-ku-andik-a /mwiik-waandikal* ‘you-NEG-INF-write-IPFV’
- (87) *i-n-áana > inzáana* ‘naughty children’  
 comp. *u-mw-áana* ‘AUG-CL1-child’  
*i-n-áagwá > inzáagwá* ‘banana wines’  
 comp. *u-rw-áagwá* ‘AUG-CL11-banana wine’

The insertion of the morphemes in the examples above is historically motivated. These are not random or arbitrary morphemes. The insertion of the infinitive marker in finite verbs is also found with monosyllabic verbs in Kiswahili with certain tense markers. The morpheme *-zi-* class 10 marker goes back to noun stems which lost it. The applicative morpheme *-ir-* also happens to be polyfunctional as it was pointed out earlier.

#### 8.4. Discontinuous Morphemes

In Kinyarwanda and many other Bantu languages, the perfective aspect *-ye* modifies phonetically the verb stem, in Kinyarwanda this modification consists of either making the last consonant palatal and an alveolar sibilant. As the following examples illustrate, when the passive morpheme is added to the

verb stem, consonant mutation applies first and the passive morpheme is inserted between the mutated consonant and the vowel *e* of the aspect morpheme *-ye*. Another way of explaining this phenomenon is to postulate that the passive morpheme *-w-* is inserted inside the morpheme *-ye* before consonant mutation.

- (88) *ba-ra-kúbit-a ba-ra-kúbis-e*  
*/ba-ra-kúbit-a/ /ba-ra-kúbit-ye/*  
 they-T-beat-IPFV they-T-beat-PFV  
 ‘they beat’ ‘they’ve just beat’  
*ba-ra-kúbis-w-e*  
*/ba-ra-kúbit-y<w>e/*  
 they-T-beat-*< PASS >* PFV  
 ‘they just got beaten’

#### 9. Illustrative text

The first line of the text illustrates the official Kinyarwanda orthography.

- (89) *Urwanda ni igihugu*  
*U-rwaanda ní i-gi-húgu*  
 AUG-rwanda ís AUG-CL7-country  
*kiba mu karere*  
*ki-bá mu ka-réré*  
 CL7-be in CL12-region  
*k'ibiyaga binini by'Afurika*  
*k'i-bi-yága bi-níni by'Aafuriká*  
 of:AUG-CL8-lake CL8-big of:Africa  
*yo hagati.*  
*yó ha-gatí.*  
 of CL16-middle
- (90) *Rukikijwe*  
*Ru-kikij-w-e*  
 CL11-surround-PASS-PFV  
*n'Ubugande, Tanzaniya,*  
*n'Ú-bu-gaande, Taanzaniyá,*  
 by:AUG-CL14-Uganda Tanzania,  
*Uburundi na Kongo.*  
*Ü-bu-ruúndi na Koongó.*  
 AUG-CL14-Burundi and Congo
- (91) *Iki gihugu*  
*I-ki gi-húgu*  
 AUG-this CL7-country  
*kirangwa*  
*ki-raang-w-a*  
 CL7-advertise-PASS-IPFV  
*n'imisozi myinshi*  
*n'i-mi-sózi my-iínshi*  
 by:AUG-CL4-hill CL4-many

- miremire                  *hamwe*  
*mi-ree-mi-re,*                  *hamwé*  
CL4-tall-CL4-RDP together  
*n'imigezi*                  *n'ibibaya.*  
*n'l-mi-gezi*                  *n'l-bi-baya.*  
with:AUG-CL4-river and:AUG-CL8-valley
- (92) *Umrurwa*                  *warwo*  
*U-mu-rwa*                  *w-áa-rw-o*  
AUG-CL3-capital CL3-of-CL11-it  
*witwa*                  *Kigali.*  
*w-iit-w-a*                  *Kigali.*  
CL3-call-PASS-IPFV Kigali
- (93) *Indimi*  
*I-n-dimi*  
AUG-CL10-language  
*zikoreshwa*                  *mu*  
*zí-kór-eesh-w-a*                  *mu*  
CL10-use-CAUS-PASS-IPFV in  
*mashuri,*                  *mu kazi*  
*ma-shuúri,*                  *a-ka-zi*  
CL6-school AUG-CL12-work  
*n'inyandiko*                  *za*  
*n'i-ny-andik-o*                  *zaa*  
and:AUG-CL10-write-NR of  
*leta*                  *ni ikinyarwanda,*  
*leeta*                  *ni i-ki-nya-rwaanda,*  
government are AUG-CL7-OP-rwanda  
*igifaransa*  
*i-gi-faraansá*  
AUG-CL7-France  
*n'icyongereza.*  
*n'ii-cy-óongerezá.*  
and:AUG-CL7-England
- (94) *Mu ibarura*  
*Mu i-bar-uur-a*  
in AUG-count-REVERSIVE-NR  
*ríherutse,*                  *Leta*  
*ri-her-íuuts-e,*                  *leeta*  
CL5-end-REVERSIVE-PFV government  
*ivuga*                  *ko abaturage*  
*i-vug-a*                  *kó a-ba-tuúr-age*  
AUG-say-IPFV that AUG-CL2-inhabit-NR  
*b'igihugu*                  *bamaze*  
*b'i-gi-húgu*                  *ba-maz-e*  
of:AUG-CL7-country CL2-finish-PFV  
*kugera*                  *kuri miriyoni*  
*ku-ger-a*                  *kuri míryooni*  
INF-arrive-IPFV to million  
*umunani*                  *n'igice.*  
*u-mu-naáni*                  *n'i-gi-cé.*  
AUG-CL3-eight and:AUG-CL7-half
- (95) *Ntibyumvikana*                  *cyakora*  
*Nti-by-uumv-iik-an-a*                  *cyakorá*  
NEG-CL8-hear-STAT-REC-IPFV but
- kubera*                  *ko mu mwaka*  
*ku-b-éer-a*                  *kó mu mw-áaka*  
INF-be-APPL-IPFV that in CL3-year  
*igihumbi*                  *kimwe*  
*i-gi-huumbi*                  *ki-mwé*  
AUG-CL7-thousand CL7-one  
*magana*                  *cyenda na mirongo*  
*ma-gana*                  *cy-eendá na mi-roongo*  
CL6-hundred CL7-nine and CL4-line  
*cyenda*                  *na kane,*  
*cy-eendá*                  *na ká-ne,*  
CL7-nine and CL12-four  
*habaye*  
*ha-d-baa-ye*  
CL16-PAST-be-IPFV  
*itsembabwoko*  
*i-tséemb-a+bw-óoko*  
AUG-exterminate-NR+CL14-race  
*ry'abatutsi,*                  *hagapfa*  
*ry'á-ba-tuutsi*                  *ha-ga-pf-a*  
of:AUG-CL2-tutsi CL16-CONSEC-die-IPFV  
*abantu*                  *barenze*  
*a-baa-ntu*                  *ba-reenz-é*  
AUG-CL2-person CL2-go.beyond-IPFV  
*miriyoni.*  
*miríyooni.*  
million.
- (96) *Ubu igihugu*                  *cyuzuye*  
*Ubu i-gi-húgu*                  *cy-uuzu-ye*  
now AUG-CL7-country CL7-be.full-IPFV  
*imfubyi*  
*i-m-fiubyi*  
AUG-CL19-orphan  
*n'abapfakazi.*  
*n'á-ba-pf-áakaz-i.*  
and:AUG-CL2-die-STAT-NR
- (97) *Abarokotse*  
*A-ba-á-rokots-e*  
CL2/REL-CL2-PAST-survive-PFV  
*bafite*                  *ibibazo*  
*ba-fit-e*                  *i-bi-báz-o*  
CL2-have-PFV AUG-CL8-ask-NR  
*ishyano*                  *ryose:*  
*i-shyano*                  *ry-óose:*  
AUG-calamity CL5-all  
*ubukene,*  
*u-bu-ken-é,*  
AUG-CL16-be.poor-NR  
*kutagira*                  *aho*  
*ku-tá-gir-á*                  *a-h-ó*  
INF-NEG-have-IPFV AUG-CL16-it  
*baba,*                  *ubafasha,*  
*ba-bá,*                  *u-ú-ba-fash-a*  
CL2-be CL1/REL-CL1-them-help-IPFV

<i>n'ubwoba</i>	<i>ko</i>
<i>n'úu-bw-óoba</i>	<i>kó</i>
and:AUG-CL14-fear	that
<i>bashobora</i>	<i>kwongera</i>
<i>ba-shob-ór-a</i>	<i>kw-oonger-a</i>
CL2-may-SUF-IPFV	INF-repeat-IPFV
<i>kwicwa.</i>	
<i>kw-íc-w-a</i>	
INF-kill-PASS-IPFV	
(98) <i>Vuba aha Leta</i>	
<i>Vubá aha leeta</i>	
Fast here government	
<i>yarekuye</i>	
<i>y-a-rék-u-ye</i>	
CL5-PAST-let-REVERSIVE-PFV	
<i>abicanyi</i>	<i>ibihumbi</i>
<i>a-b-iic-an-yi</i>	<i>i-bi-huumbi</i>
AUG-CL2-kill-REC-NR	AUG-CL8-thousand
<i>mirongo</i>	<i>ine</i>
<i>mi-roongo</i>	<i>kubera</i>
CL4-line	AUG-four
<i>gusa</i>	<i>ko bemeye</i>
<i>gu-sa</i>	<i>kó b-eéme-ye</i>
CL15-alone	that CL2-accept-PFV
<i>ubugome</i>	<i>bwabo</i>
<i>u-bu-gom-é</i>	<i>bw-áa-b-o</i>
AUG-CL14-rebel-NR	CL14-of-CL2-it
<i>burenze</i>	<i>kamere</i>
<i>bu-reenz-é</i>	<i>ka-mer-é</i>
CL14-go.beyond-PFV	CL12-grow-NR
<i>kandi</i>	<i>batarigeze</i>
<i>kaá-ndi</i>	<i>bá-ta-ar-iigez-e</i>
CL12-other	CL2-NEG-PAST-try-PFV
<i>bicuza</i>	<i>cyangwa ngo</i>
<i>b-iicuz-a</i>	<i>cyángwá</i>
CL2-repent-IPFV	or
<i>basabe</i>	<i>imbabazi</i>
<i>bá-sab-a</i>	<i>i-m-babáz-i</i>
CL2-ask-IPFV	AUG-CL10-suffer-NR
<i>abo</i>	<i>bahemukiye.</i>
<i>a-b-ó</i>	<i>ba-á-hemuk-i-ye.</i>
AUG-CL2-it	CL2-PAST-harm-APPL-PFV

(89) Rwanda is a country of the Great Lakes Region in Central Africa. (90) It is surrounded by Uganda, Tanzania, Burundi and Congo. (91) This country is characterized by many tall hills, rivers, and valleys. (92) Its capital is called Kigali. (93) Languages that are used in schools, in works and writings of the government are Kinyarwanda, French and English. (94) In the latest census, the government says that the population has just reached eight million and half. (95) It is not understand-

able, however, because in the year 1994, there was genocide of Tutsi and then died more than a million people. (96) Now the country is full of orphans and widows. (97) Those who survived have a lot of problems, poverty, not to have who can help them, and the fear that they may be killed again. (98) Recently the government released 40,000 murderers because they admitted that they killed even though they didn't try to repent or ask for forgiveness.'

## 10. Uncommon abbreviations

AUG	augment
OP	onomastic prefix
T	present, habitual, and late today tense

## 11. References

- Coupez, André (1980), *Abégé de grammaire rwanda*, Vol. 1–2. Butare: Institut national pour la recherche scientifique
- Doke, Clement Martyn (1935), *Bantu Linguistic Terminology*. London: Longmans
- Kimenyi, Alexandre (1979), *Studies in Kinyarwanda and Bantu Phonology*. Edmonton, Alberta: Linguistic Research
- Kimenyi, Alexandre (1980), *A Relational Grammar of Kinyarwanda*. Berkeley, Los Angeles: University of California Press
- Kimenyi, Alexandre (1989), *Kinyarwanda and Kirundi Names: A Semiolinguistic Analysis of Bantu Onomastics*. Lewiston: The Edwin Mellen Press
- Kimenyi, Alexandre (1995), "Kinyarwanda Applicatives Revisited". Paper presented at the 8<sup>th</sup> Niger-Congo Syntax and Semantics Workshop, Boston University, 25 October
- Kimenyi, Alexandre (1997), "Kinyarwanda Ideophones: Form, Function, Content and Context". *Wihogora Rwanda* (Sacramento) 3
- Kimenyi, Alexandre (1999), "Cow metaphors". *Wihogora Rwanda* 5, 78–85
- Kimenyi, Alexandre (2002), *A Tonal Grammar of Kinyarwanda: An Autosegmental and Metrical Analysis*. Lewiston: The Edwin Mellen Press
- Kimenyi, Alexandre (2003), "The perfective aspect and consonant mutation in Bantu languages". In:

Githora, Chege & Littlefield, Heather & Manfredi, Victor (eds.), *Trends in African Linguistics, Vol. V*. Lawrenceville/NJ: Africa World Press, 135–47  
 Meeussen, Achiel Emiel (1959), *Essai de grammaire rundi*. Tervuren: Musée du Congo Belge (Annales; Sciences de l'Homme; Linguistique 24)

Sibomana, Léonidas (1974), *Deskriptive Tonologie des Kinyarwanda*. Hamburg: Buske

Welmers, William E. (1973), *African language structures*. Berkeley: University of California Press

*Alexandre Kimenyi, Sacramento (U.S.A.)*

## 142. Vietnamesisch (Viet-Muong)

1. Allgemeine Angaben
2. Abgeleitetes Wort
3. Zusammengesetztes Wort
4. Reduplikation
5. Iteration
6. Von der syntaktischen Fügung zum Wort
7. Wortklassen
8. Illustrativer Text
9. Zitierte Literatur

### 1. Allgemeine Angaben

#### 1.1. Geographische Verbreitung. Soziolinguistischer Status

Die vietnamesische Sprache ist die Staats-sprache der im östlichen Teil Hinterindiens gelegenen Sozialistischen Republik Vietnam (Fläche 331.700 km<sup>2</sup>, Bevölkerung über 60 Mio. Einwohner). Der Gebrauch der vietnamesischen Sprache erstreckt sich auf die Bereiche der Presse, des Rundfunks und Fernsehens, der Wissenschaft und Kultur, des Bildungswesens und des überregionalen Verkehrs.

Neben den Vietnamesen als den Sprachträgern, die 80 Prozent der Gesamtbevölkerung ausmachen und im wesentlichen die Küste und die Flachlandregionen besiedeln, leben in Vietnam etwa 60 Völkerschaften und ethnische Gruppen, darunter die Muong, die den Vietnamesen sprachlich und ethnisch am nächsten stehen.

#### 1.2. Genetische Beziehungen

Die zu Beginn des 20. Jh. vertretene Auffassung von der Verwandtschaft des Vietnamesischen mit den Thai-Sprachen (Maspero 1912) wurde in der Folge einer kritischen Überprüfung unterzogen (Przyluski 1924). Gemäß der heute vorherrschenden Ansicht gehört das Vietnamesische zur Gruppe der Mon-Khmer-Sprachen, die ihrerseits ein Bestandteil der austroasiatischen Sprachfamilie ist (Haudri-court 1953). Der langanhaltende kulturelle

Einfluß Chinas auf Vietnam, der sich vom 1. Jh. v. u. Z. bis ins 20. Jh. verfolgen lässt, hatte eine Fülle von chinesischen Entlehnun-gen im Vietnamesischen zum Ergebnis, je-doch entbehrt dieser Umstand, wie auch die typologische Ähnlichkeit der vietnamesischen und der chinesischen Sprache, einer geneti-schen Grundlage (vgl. Hồ Lê 1971).

### 1.3. Schrift

Das chinesische Hieroglyphensystem (*chữ Hán* ‘Han Schriftzeichen’), das in der vietna-mesischen Kultur etwa vom Beginn unserer Zeitrechnung bis zum 13. Jh. uneinge-schränkt dominierend war, besteht neben dem auf seiner Grundlage geschaffenen vietnamesischen Hieroglyphensystem (*chữ Nôm* ‘südliche Schriftzeichen’) fort. Mit Ausnahme einzelner kürzerer historischer Perioden be-wahrte das chinesische Hieroglyphensystem bis zum 20. Jh. den Status der offiziellen Schriftsprache in der Sphäre des Verwal-tungs- und Bildungswesens.

Das moderne, auf dem Lateinischen basie-rende vietnamesische Alphabet (*quốc ngữ* ‘nationale Sprache/Schrift’) wurde im 17. Jh. von portugiesischen, italienischen und fran-zösischen Missionaren ausgearbeitet. Es stellt eine Variante der phonematischen Schrei-bung dar und wird seit 1910 im offiziell-am-tlichen Verkehr verwendet. Das Alphabet sieht wie folgt aus:

<i>a</i>	[a:]	<i>d̥</i>	[d]	<i>k</i>	[k]	<i>o'</i>	[ø:]	<i>u</i>	[u]
<i>ă</i>	[a]	<i>e</i>	[ɛ]	<i>l̥</i>	[l]	<i>p</i>	[p]	<i>u'</i>	[w]
<i>â</i>	[ə]	<i>ê</i>	[e]	<i>m̥</i>	[m]	<i>q</i>	[k]	<i>v</i>	[v]
<i>b</i>	[b]	<i>g</i>	[β]	<i>n̥</i>	[n]	<i>r̥</i>	[ʒ]	<i>x</i>	[s]
<i>c</i>	[k]	<i>h</i>	[h]	<i>o̥</i>	[ɔ]	<i>s̥</i>	[ʃ]	<i>y</i>	[i]
<i>d</i>	[z]	<i>i̥</i>	[i]	<i>ô̥</i>	[o]	<i>t̥</i>	[t]		

Tab. 142.1: *quốc ngữ*

Einige Laute werden durch Digraphen wiedergegeben (s. Tab. 142.2).

<i>ph</i> [f]	<i>ch</i> [t <sup>j</sup> ]	<i>tr</i> [tʃ]	<i>kh</i> [χ]	<i>ng</i> [ɲ]
<i>th</i> [t <sup>h</sup> ]	<i>gi</i> [ʒ <sup>t</sup> ]	<i>nh</i> [n]		

Tab. 142.2: Digraphen

Im Alphabet *quốc ngữ* sind für die Bezeichnung der vietnamesischen Töne spezielle diakritische Zeichen gebräuchlich. Das Fehlen eines diakritischen Zeichens entspricht dem hohen Level-Ton: *ma* ‘Gespenst’. Weiter folgen Zeichen, die in der Regel bei dem silbenbildenden Vokal angeordnet sind: der Gravis zur Wiedergabe des oberen fallenden (zweiten) Tons (*mà* ‘aber’); die Tilde zur Wiedergabe des fallend-steigenden diskontinuierlichen (dritten) Tons (*mā* ‘Pferd im Schachspiel’); die “Frage” zur Wiedergabe des fragenden fallend-steigenden kontinuierlichen (vierten) Tons (*má* ‘Grab’); der Akut zur Wiedergabe des steigenden gespannten (fünften) Tons (*má* ‘Wange’) und schließlich der Punkt in der Unterzeile zur Wiedergabe des abrupt-fallenden schweren (sechsten) Tons (*mq* ‘Setzung(e) der Reispflanze’).

#### 1.4. Literatursprache und Dialekte

Die vietnamesische Sprache der Gegenwart weist drei Dialekte auf:

- (a) den nördlichen Dialekt (hauptsächlich gesprochen im Gebiet von Hanoi),
- (b) den zentralen Dialekt (hauptsächlich gesprochen in den Städten Vinh und Hué) und
- (c) den südlichen Dialekt (hauptsächlich gesprochen in Saigon).

Die Dialekte unterscheiden sich vor allem in phonetischer Hinsicht (vgl. Thompson 1967: 18–51, 78–97).

Die moderne Literatursprache kann zum gegenwärtigen Zeitpunkt noch nicht als gefestigt angesehen werden. Ihre schriftliche Variante, die den Gegenstand der weiterführenden Betrachtung bildet, orientiert sich am Lautbestand des zentralen Dialekts und am Tonsystem des nördlichen Dialekts, im mündlichen Sprachgebrauch wird dieses ideale Schema jedoch selten durchgängig eingehalten.

## 2. Abgeleitetes Wort

### 2.1. Formale und semantische Eigenschaften der Derivation

Das **abgeleitete Wort** ist ein Wort, welches von einem anderen Wort mit Hilfe eines Derivationsformanten gemäß den in der Sprache

wirkenden Wortbildungsmustern gebildet wurde (s. Art. 89). Die Wortbildungsanalyse muß auf die formal-semantische Beziehung zwischen Ausgangswort und Ableitung gerichtet sein: Man kann nicht davon ausgehen, daß engl. *tailor* durch das Wort *tail* motiviert ist (der Zusammenhang ist rein formal), ebenso kann engl. *fear* nicht auf *awe* zurückgeführt werden (die Beziehung ist rein semantisch). Möglich sind indessen Fälle, in denen die semantischen und die formalen Eigenschaften der Derivation nicht übereinstimmen (s. Art. 81): russ. *kačat'* ‘jmdn. schaukeln’ ist semantisch komplexer als *kačat'sja* ‘schaukeln’, in morphologischer Hinsicht verhält es sich umgekehrt (Apresjan 1974: 277), dt. *Schäferei* ist semantisch von *Schaf* abgeleitet, formal hingegen von *Schäfer* (Murjasov 1977: 120). In der vietnamesischen Sprache sind Beispiele, wo die semantischen und die formalen Eigenschaften der Derivation nicht übereinstimmen, unbekannt.

#### 2.2. Derivationsformant

Der **Derivationsformant** zeichnet sich durch die Besonderheit seiner Position im Wort und die Fähigkeit aus, mehr oder weniger großen Klassen von Wörtern als Wortbildungsmittel zu dienen. In semantischer Hinsicht ist wesentlich, daß die Bedeutung des abgeleiteten Wortes sich nicht nach der Formel “Bedeutung des Ausgangswortes, verbunden mit einer gewissen Einschränkung” erklären läßt, auch wenn der Derivationsformant eine einheitliche Veränderung der lexikalischen Bedeutungen der Ausgangswörter bewirkt, aus der sich die Wortbildungsbedeutung des entsprechenden Modells ergibt.

#### 2.3. Vietnamesische Derivationsformanten

Die Zahl der vietnamesischen Derivationsformanten beläuft sich auf einige wenige Einheiten. Das Präfix *nhà*, angefügt an Substantive, bildet Nomen mit Agensbedeutung:

- |                      |                        |
|----------------------|------------------------|
| (1) <i>chính trị</i> | → <i>nhà chính trị</i> |
| ‘Politik’            | ‘Politiker’            |
| <i>khoa học</i>      | → <i>nhà khoa học</i>  |
| ‘Wissenschaft’       | ‘Wissenschaftler’      |
| <i>kiến trúc</i>     | → <i>nhà kiến trúc</i> |
| ‘Architektur’        | ‘Architekt’            |

Das Suffix *hóa*, angefügt an Substantive und Adjektive, bildet Nomen mit kausativer Semantik:

- |                        |  |
|------------------------|--|
| (2) <i>công nghiệp</i> | → <i>công nghiệp hóa</i>                   |
| ‘Industrie’            | ‘industrialisieren/<br>Industrialisierung’ |

<i>tự động</i>	→ <i>tự động hóa</i>
‘Automat’	‘automatisieren/ Automatisierung’
<i>quân sự</i>	→ <i>quân sự hóa</i>
‘militärisch’	‘militarisieren/ Militarisierung’

Zu den genannten Formanten kommen vielleicht noch zwei hinzu – die Präfixe *dánh* und *làm*, die von intransitiven Verben und Adjektiven transitive Verben bilden:

(3) <i>thúc</i>	→ <i>dánh thúc</i>
‘wachen’	‘wecken’
<i>rơi</i>	→ <i>dánh rỗi</i>
‘fallen’	‘fallen lassen’
<i>bẩn</i>	→ <i>dánh bẩn</i>
‘schmutzig’	‘verschmutzen’
<i>hở</i>	→ <i>làm hở</i>
‘kaputt’	‘kaputt machen’
<i>khổ</i>	→ <i>làm khổ</i>
‘qualvoll’	‘quälen’

Das Präfix *dánh* verbindet sich nur mit einer normativ beschränkten Klasse von Wörtern und läßt in allen Fällen seiner Verwendung den Ersatz durch den produktiveren Formanten *làm* zu, währenddessen ein Ersatz in umgekehrter Richtung bei weitem nicht immer möglich ist. Das Element *làm* ist noch insoweit von Interesse, als es ein selbständiges Wort (Verb) mit derselben kausativen Bedeutung zu repräsentieren vermag. Deshalb kann in einer Reihe von Fällen ein und dieselbe Semantik entweder durch eine kausative Konstruktion oder durch ein transitives Verb mit dem Derivationsformanten *làm* ausgedrückt werden: *làm tôi nhục* ‘mich erniedrigen’ (mich zum Erniedrigten machen); *làm nhục chúng ta* ‘uns erniedrigen’ (Nguyễn Kim Thành 1977: 55).

### 3. Zusammengesetztes Wort

#### 3.1. Zusammengesetztes Wort und syntaktische Fügung

Das **zusammengesetzte Wort** ist ein Wort, das durch die Juxtaposition von mindestens zwei Wurzelmorphemen ohne Beteiligung von Derivationsformanten gebildet wurde (s. Art. 87). In Verbindung mit der Analyse des Kompositums entsteht die Frage nach der Möglichkeit des Vergleichs mit der syntaktischen Fügung. Im Rahmen der synchronen Wortbildung ist ein solcher Vergleich auf die Untersuchung der strukturell-semantischen Be-

sonderheiten der Komposita ohne Bezug zu ihrer Entstehungsgeschichte gerichtet. Deshalb setzt die Anwendung dieses Untersuchungsverfahrens nicht die Abstammung des Wortes von einer mit ihm korrelierenden syntaktischen Fügung voraus.

Ein Extrem bei der Untersuchung der Relation zwischen Kompositum und syntaktischer Fügung läuft darauf hinaus, das Kompositum als Korrelat zur syntaktischen Fügung zu definieren, so daß die Wurzelmorpheme des Kompositums sich in den Komponenten der syntaktischen Fügung wiederholen. Dem entgegengesetzten Standpunkt folge darf das Kompositum auch dann nicht auf einer Ebene mit der syntaktischen Fügung betrachtet werden, wenn es durch eine solche wiedergegeben werden kann. Eine Klassifikation der Komposita nach der assoziativen Zuordnung ihrer Wurzelbestandteile zu der einen oder anderen Wortart überträgt die Verhältnisse der Syntax in ungerechtfertigter Weise auf die Wortbildung.

Wie sonderbar es auch scheinen mag, es müssen beide Extrempositionen Berücksichtigung finden: Bei allen zweifellos bestehenden Unterschieden zwischen Wort und syntaktischer Fügung lassen sich, insbesondere im Vietnamesischen, einige Varianten von Wörtern nicht ohne Probleme von syntaktischen Fügungen abgrenzen. So werden z. B. Bildungen des Typs *dánh chét* ‘schlag sterb (zu Tode prügeln)’, *tìm thấy* ‘such seh (finden)’ von den einen Autoren als Wörter, von den anderen als syntaktische Fügungen angesehen. Offensichtlich ist, daß die Komposita einen unterschiedlichen Grad an Nähe zu den syntaktischen Fügungen aufweisen, angefangen bei einer Nullstufe, die einen Vergleich dieser Einheiten ausschließt, bis hin zu einer maximalen Stufe, die den Vergleich unausweichlich macht. Unbedingt ist jedoch zu berücksichtigen, daß nur solche Wörter als Gegenstand des Vergleichs mit der syntaktischen Fügung in Betracht kommen, deren Wurzelmorpheme als selbständiges Wort verwendet werden können, z. B. *máy bay* ‘Maschine fliegt (Flugzeug)’. Die Möglichkeit des Vergleichs entfällt, wenn auch nur eines der Wurzelmorpheme nicht über diese Eigenschaft verfügt (wenn es, fachterminologisch gesprochen, ein ‘gebundenes Morphem’ ist). Das ist z. B. der Fall bei *súng trù'ó'ng* ‘Gewehr lang (Gewehr/langes Gewehr)’, wo die erste Komponente ein selbständiges Wort sein kann, während die zweite nur als Wortbestandteil gebraucht wird.

### 3.2. Nähe des Kompositums zur syntaktischen Fügung

Es erscheint zweckmäßig, die Nähe des Kompositums zur syntaktischen Fügung nach der folgenden dreistufigen Skala einzuschätzen:

- (a) Minimale Nähe. Die Wortkomponenten lassen sich in der Terminologie der Wortarten natürlich charakterisieren, jedoch erlaubt ihre Relation weder eine semantische noch eine formale Gegenüberstellung mit der syntaktischen Fügung: *trai tré* ‘Bursche jung (jugendlich)’ – Juxtaposition einer Substantiv- und einer Adjektivwurzel, die sich unter Beachtung der Bedeutung des Ganzen nicht als attributive Zusammensetzung interpretieren lässt.
- (b) Mittlere Nähe. Die Wortkomponenten lassen sich in der Terminologie der Wortarten charakterisieren, die innere Form (Motivierung) des Wortes ist mittels einer für die syntaktische Fügung typischen semantischen Relation interpretierbar, wobei die formalen Merkmale der entsprechenden syntaktischen Fügung fehlen: *cà chua* ‘Aubergine sauer (Tomate)’ – semantisch attributive Verbindung, die in

keinem Kontext die Weglassung des “Attributs” gestattet.

- (c) Maximale Nähe. Das Wort lässt sich nach der Zugehörigkeit seiner Komponenten zu der einen oder anderen Wortart sowie nach der syntaktischen Bedeutung der Relation zwischen den Komponenten charakterisieren, wobei eine formale Gegenüberstellung mit der syntaktischen Fügung in der einen oder anderen Weise möglich ist: *xe dép* ‘Fahrzeug tretr (Fahrrad)’ – ein nach dem attributiven Modell aufgebautes Wort, das in einem eindeutigen Kontext auf die erste Silbe verkürzt wird; *làm việc* ‘mach Sache (arbeiten)’ – ein nach dem Objekt-(Komplement-)Modell aufgebautes Wort, das die Umwandlung der zweiten Komponente in ein vollgültiges Objekt erlaubt (*làm một việc nặng* ‘eine schwere Arbeit verrichten’).

### 3.3. Modelle zusammengesetzter Wörter im Vietnamesischen

Bei den Komposita (mit Ausnahme der Wörter mit minimaler Nähe zur syntaktischen Fügung) ist eine koordinative oder subordinative Relation zwischen den Komponenten zu beobachten. Gehört das Wort bei koordi-

	Mittlere Nähe	Maximale Nähe
Subst. + Subst.	<i>sát</i> <i>dá</i> Eisen Stein ‘fest’	<i>áo</i> <i>quần</i> Hemd Hose ‘Bekleidung’
	<i>trò</i> <i>i</i> <i>bé</i> Himmel Meer ‘unendlich’	<i>bút</i> <i>mực</i> Pinsel Tusche ‘Schreibzubehör’
	<i>gan</i> <i>da</i> Leber Magen ‘kühn’	<i>giày</i> <i>dép</i> Schuhe Sandalen ‘Schuhwerk’
Verb + Verb	<i>sói</i> <i>nói</i> brodel auftret ‘lebhaft’	<i>thu'o'ng</i> <i>yêu</i> Mitleid.hab lieb ‘lieben’
	<i>đi</i> <i>đi</i> <i>ng</i> geh steh ‘Gang(art)’	<i>nhảy</i> <i>múa</i> tanz gestikulier ‘tanzen’
	<i>chǎn</i> <i>nuôi</i> sich.sorg aufzieh ‘Viehzucht’	<i>cày</i> <i>cây</i> pflüg Setzlinge.umpflanz ‘Ackerbau betreiben’
Adj. + Adj.	<i>xa</i> <i>gắn</i> weit nah ‘überall’	<i>cô</i> <i>xưa</i> alt vergangen ‘alt’
	<i>đỏ</i> <i>đen</i> rot schwarz	<i>già</i> <i>yếu</i> alt schwach ‘gebrechlich’
	‘Partie im Kartenspiel’	<i>mạnh</i> <i>khỏe</i>
	<i>đen</i> <i>tối</i> schwarz dunkel ‘verbrecherisch’	stark gesund ‘gesund’

Tab. 142.3: Komposita mit koordinativer Relation der Wortkomponenten

	Mittlere Nähe		Maximale Nähe	
Attributives Modell	<i>máy</i>	<i>bay</i>	<i>xe</i>	<i>hṝa</i>
	Maschine	flieg ‘Flugzeug’	Fahrzeug	Feuer ‘Zug’
	<i>tháng</i>	<i>hai</i>	<i>cá</i>	<i>vàng</i>
	Monat	zwei ‘Februar’	Fisch	Gold ‘Goldfisch’
	<i>nhà</i>	<i>máy</i>	<i>áo</i>	<i>mura</i>
	Haus	Maschine ‘Betrieb’	Bekleidung	Regen ‘Regenmantel’
Komplement-Modell	<i>trả</i>	<i>lời</i>	<i>đán</i>	<i>co'm</i>
	erwider	Worte ‘antworten’	ess	Reis ‘essen’
	<i>làm</i>	<i>khách</i>	<i>chụp</i>	<i>ánh</i>
	mach	Gast ‘Umstände machen’	fotographier	Aufnahme ‘fotographieren’
	<i>theo</i>	<i>gót</i>	<i>nói</i>	<i>chuyện</i>
	verfolg	Absatz ‘verfolgen’	sprech	Ereignis ‘sich unterhalten’

Tab. 142.4: Komposita mit subordinativer Relation der Wortkomponenten

nativer Relation nicht zu derselben Wortart wie seine Bestandteile, so ist für ein solches Wort nur eine mittlere Nähe zur syntaktischen Fügung möglich. In den übrigen Fällen, in denen das Wort und seine Bestandteile derselben Wortart zugerechnet werden können, ist gleichermaßen eine mittlere wie eine maximale Nähe zur syntaktischen Fügung möglich.

In Tab. 142.3 werden die Grundmodelle der Komposita mit koordinativer Relation der Wortkomponenten angeführt.

In den Modellen mit subordinativer Relation lassen sich ebenfalls unterschiedliche Stufen der Nähe des Kompositums zur syntaktischen Fügung ausmachen. Wir veranschaulichen das hier am Beispiel des attributiven und des Komplement-Modells.

#### 4. Reduplikation

##### 4.1. Reduplikation und Ausgangswort

**Reduplikationen** sind Wörter des Typs *đẹp đẽ* ‘schön’, die aus einem Wurzelmorphem (*đẹp* ‘schön’) und einer bedeutungsleeren phonetisch modifizierten Variante desselben bestehen (s. Art. 57). Die asemantische Komponente der Reduplikation nennen wir Reduplikator. Wörter, die allein in iterativer Form existieren (*ba ba* ‘Schildkröte’, *đu đú* ‘Papaya’), werden nicht zu den Reduplikationen gerechnet. Sie verfügen über keine korrelierende Ausgangseinheit und fallen insofern nicht in die Zuständigkeit der synchronen Wortbildung. Die semantische Korrelation mit dem Ausgangswort ist neben der formalen Korrelation ein wesentliches Kriterium

der Wortbildungsanalyse sekundärer Einheiten. Für die Reduplikation ist die Verbindung zum Ausgangswort solange gegeben, wie dieses an der Bedeutungserklärung der Reduplikation teilnehmen kann, z. B.: *nhỏ* ‘klein’ → *nhỏ nhán* ‘winzig’ (*nhỏ* và *xinh* ‘klein und angenehm’). Sobald die zu untersuchende Einheit nur noch in phonetischer Hinsicht Merkmale einer Reduplikation aufweist und sich in semantischer Hinsicht so weit vom Ausgangswort entfernt hat, daß es nicht mehr in die Bedeutungserklärung dieser Einheit einzubezogen werden kann, wird sie zu einem einfachen Wort: *nhỏ nhén* ‘gemein’.

Im Unterschied zur Ableitung als Verbindung eines Wurzelmorphems mit einem Derivationsformanten, der die reguläre Bedeutung des entsprechenden Wortbildungsmodells wiedergibt, stellt sich die Reduplikation als Verbindung einer Wurzel mit einem unikalen Wortbildungsmittel dar, wobei eine streng individuelle Korrelation mit dem Ausgangswort vorliegt.

##### 4.2. Reduplikator

Der Reduplikator wird gewöhnlich nach der Wurzel angeordnet, es sind aber auch Fälle einer Voranstellung des Reduplikators anzutreffen: *toi* ‘verstreut, verschüttet’ → *tá to'i* ‘zerstreut’, *trộn* ‘mischen’ → *trà trộn* ‘sich unter die Menge mischen’ (Lê Văn Ly 1981: 58). Nach der Form des Reduplikators, die zu einem gewissen Grade von der Form der Wurzel bestimmt wird, lassen sich zwei Arten von Reduplikationen unterscheiden: initiale Reduplikationen – im Reduplikator wiederholt sich der Anlaut der Wurzelsilbe (*lám* ‘machen’ → *lám lỵng* ‘arbeiten’) – und finale Reduplikationen – im Reduplikator wiederholt

sich der Auslaut der Wurzelsilbe (*ké* 'erzählen' → *ké lè* 'faseln'). Der Ton des Reduplikators wird auch zu einem gewissen Grade vom Ton der Wurzel bestimmt (vgl. Glebova & Sitnikova 1980: 66 f.), aber obwohl die lautive Form des Reduplikators nicht beliebig ist, ist es nicht möglich, sie in jedem konkreten Fall völlig eindeutig vorherzusagen.

Die Anfügung eines Reduplikators an eine Ausgangseinheit signalisiert die Bildung eines neuen, vom Ausgangswort verschiedenen Wortes. Ganz allgemein bestehen die grammatischen Unterschiede zwischen Reduplikation und Ausgangswort in folgendem: bei Substantiven – Übergang in die Unterkategorie der Kollektiva (*máy* 'Maschine' → *máy móc* 'verschiedene Maschinen und Apparate'); bei Verben – Übergang in die Unterkategorie der Intransitiva (*làm* 'machen' → *làm lung* 'arbeiten'); bei Adjektiven – Verlust der Möglichkeit einer substantivischen Ergänzung: *đẹp* 'schön', wo *đẹp măt* 'schön Auge (angenehm fürs Auge)', *đẹp ngu'o'i* 'schön Körper (äußerlich schön)' möglich ist, gegenüber *đẹp dẽ* 'schön', wo eine nominale Ergänzung unmöglich ist (Đái Xuân Ninh 1978: 203). Neben den grammatischen Unterschieden gibt es auch semantische. Zum Beispiel wird *lanh* 'kalt' sowohl im wörtlichen als auch im übertragenen Sinne gebraucht, während *lanh léo* 'kalt' nur im wörtlichen Sinne und *lanh lung* 'kalt' nur im übertragenen Sinne verwendet wird: *một gian phòng lanh léo* 'ein kaltes Zimmer'; *một thái độ lanh lung* 'ein kühles Verhältnis' (Hô Lê 1976: 285; Sitnikova 1980: 245 f., 254–256).

## 5. Iteration

### 5.1. Wurzeliteration

Die **Iteration** ist der einzige Bereich der Formenbildung im Vietnamesischen, berücksichtigt man die Regularität sowie den rein quantitativen Charakter der zusätzlichen Bedeutung, die der Ausgangseinheit verliehen wird. Bei den Wurzeliterationen wird das Wurzelmorphem der Ausgangseinheit dem Verfahren der Iteration unterzogen. Es gibt vollständige und unvollständige (abweichende) Wurzeliterationen. Bei ersteren wiederholt sich die Wurzel in unveränderter Form (*đẹp* 'schön' → *đẹp đẹp* 'sehr schön'), bei letzteren erfolgt die Wiederholung mit einigen phonetischen Modifikationen. Im Ergebnis dessen wird der Wurzel eine bedeutungsleere Einheit hinzugefügt: *đẹp* 'schön' → *dém đẹp* 'ziemlich schön' (Glebova & Sitnikova 1980: 56–58).

Substantive bilden nur vollständige Iterationen mit der Bedeutung der distributiven Pluralität: *ngu'o'i* 'Mensch' → *ngu'o'i ngu'o'i* 'jeder Mensch, Leute', *nhà* 'Haus' → *nhà nhà* 'jedes Haus'. Einige semantische Klassen von Handlungsverben lassen ausschließlich vollständige Iterationen mit der Bedeutung der Iterativität der Handlung zu: *vuốt* 'streichen' → *vuót vuót* 'ab und zu streichen', *gãi* 'kratzen' → *gãi gãi* 'ab und zu kratzen'. Nicht-handlungsverben bilden in Abhängigkeit von ihrer phonetischen Gestalt vollständige oder abweichende Iterationen (vgl. Glebova & Sitnikova 1980: 57) mit der Bedeutung der Abschwächung des Merkmals: *lo* 'sich beunruhigen' → *lo lo* 'sich leicht beunruhigen', *sợ* 'sich ängstigen' → *sợ sợ* 'sich etwas ängstigen'. Komplizierter verhält es sich mit den Adjektiven. Adjektive im hohen Level oder oberen fallenden Ton können nur vollständige Iterationen bilden, die in Abhängigkeit vom Kontext die Abschwächung oder die Verstärkung des Merkmals wiedergeben: *hay* 'interessant' → *hay hay* 'ziemlich/sehr interessant'. In allen anderen Tönen bezeichnet die vollständige Iteration die Verstärkung des Merkmals, während die Semantik abweichender Iterationen zu einem gewissen Grade von der Position des bedeutungsleeren Segments in Relation zur Wurzel abhängt: die Bedeutung der Abschwächung des Merkmals wird nur durch die Voranstellung der bedeutungsleeren Komponente wiedergegeben (*héo* 'verwelkt' → *heo héo* 'leicht verwelkt'); die Bedeutung der Verstärkung des Merkmals ist nicht fest an eine bestimmte Position der bedeutungsleeren Komponente gebunden (vgl. *còn* 'klein' → *còn còn* 'sehr klein', *héo* 'verwelkt' → *héo héo* 'ganz verwelkt') (Nguyễn Phú Phong 1977: 62). Außerdem sei angemerkt, daß nur abweichende Iterationen mit der Bedeutung der Merkmalsverstärkung mehr als ein bedeutungsleeres Segment enthalten können: *héo* 'verwelkt' → *héo héo* 'ganz verwelkt' → *héo héo héo* 'ganz ganz verwelkt' → *héo héo héo héo* 'ganz ganz ganz ganz verwelkt' (Nguyễn Phú Phong 1977: 62).

### 5.2. Silbeniteration

Neben der Wurzeliteration gibt es im Vietnamesischen auch die Silbeniteration, bei der die asemantische Silbe der Ausgangseinheit dem Verfahren der Iteration unterzogen wird. Silbeniterationen lassen sich ebenfalls in vollständige und abweichende unterteilen. So kann z. B. vom Adjektiv *đen* 'schwarz' durch Hinzufügung der bedeutungsleeren Silbe *sì*

das Wort *đen sì* ‘schmutzig’ gebildet werden, dessen vollständige Silbeniteration die Einheit *đen sì sì* ‘sehr schmutzig’ ergibt. Vom Adjektiv *tho;m* ‘wohlriechend’ läßt sich durch Ergänzung der bedeutungsleeren Silbe *phú;c* das Wort *tho;m phú;c* ‘aromatisch’ bilden, dessen abweichende Silbeniteration wiederum die Einheit *tho;m phúng phú;c* ‘sehr aromatisch’ ergibt (vgl. Đái Xuân Ninh 1978: 186–190).

## 6. Von der syntaktischen Fügung zum Wort

### 6.1. Silbenschwund

Möglich ist eine “Kompression” der syntaktischen Fügung im Text, die unter günstigen Bedingungen zu einem vollwertigen Ergebnis aus Sicht der Wortbildung führt (s. Art. 60). So scheint es aus rhythmisch-phonetischen Erwägungen angebracht, die dreisilbige Wortfügung *khoa triết học* ‘philosophische Fakultät’ auf die zweisilbige Konstruktion *khoa triết* mit derselben Bedeutung, jedoch mit den Formmerkmalen eines Wortes zu verkürzen (vgl. russ. *biologičeskij fakul'tet* → *biofak*). Anstelle von *xír Nghé An* ‘Provinz Nghean’ wird häufiger *xí Nghé* gebraucht, anstelle von *tàu Hải Phòng* ‘Haiphonger Zug’ – *tàu Phòng*. Da das letztere Beispiel aber den Einschub eines autosemantischen Wortes gestattet (*tàu đi Phòng* ‘Zug, der nach Haiphong fährt’), scheint es eher am Platz, hier von der Bildung einer einsilbigen, funktional eingeschränkten Variante des Wortes *Hải Phòng* zu sprechen. Eine einsilbige Variante des Wortes *Hải Phòng* ist z. B. nicht möglich in der Fügung *học sinh Hải Phòng* ‘Haiphonger Schüler’ (Tru'o'ng Đông San 1976: 10).

Bisweilen wird der Schwund einer oder mehrerer Silben von einer Veränderung der Stellung der verbliebenen Silben zueinander begleitet, wodurch der Wortstatus der neu entstandenen Bildung sowohl grammatisch als auch lexikologisch noch überzeugender ausfällt: *vận động thanh niên* ‘Bewegung Jugend’ → *thanh vận* ‘Jugendbewegung’, *phòng thủ biên giới* ‘Schutz Grenze’ → *biên phòng* ‘Grenzsicherung’ (Lê Xuân Thái 1982: 53).

### 6.2. Silbenausklammerung

Das Verfahren der Silbenausklammerung kommt in Betracht für das gemeinsame bedeutungstragende Element von zwei oder drei zweisilbigen Wörtern, die ein parataktisches Gefüge bilden: *công nghiệp*, *nông nghiệp*

→ *công nông nghiệp* ‘Industrie (und) Landwirtschaft’, *thủy quân, lục quân, không quân* → *thủy, lục, không quân* ‘Flotte, Infanterie, Luftfahrt’. Silbenausklammerung in Gefügen mit mehr als drei Wörtern tritt nur in Einzelfällen auf: *nông nghiệp, công nghiệp, lâm nghiệp, ngư nghiệp* → *nông, công, lâm, ngư nghiệp* ‘Landwirtschaft, Industrie, Forstwirtschaft, Fischerei’ (Nguyễn Huy Cần 1981: 66).

Innerhalb des Komplexes, der durch Silbenausklammerung gebildet wurde, haben die einzelnen Elemente eine feste, normativ vorgegebene Anordnung. Der Komplex selbst weist Merkmale eines Wortes auf. Einige Bildungen dieser Art sind lexikalisiert und werden in Wörterbücher aufgenommen: *bi hài kịch* ‘Tragikomödie’ (*bi kịch* ‘Tragödie’ + *hài kịch* ‘Komödie’), *âm du'o'ng lịch* ‘Sonnen- und Mondkalender’ (*âm lịch* ‘Mondkalender’ + *du'o'ng lịch* ‘Sonnenkalender’) (Nguyễn Huy Cần 1981: 68).

### 6.3. Silbenüberschneidung

Bei Kontaktposition eines ein- und eines zweisilbigen Wortes kann das erstere innerhalb des zweiten “wiederholt” werden: *nghĩ vo* → *nghĩ vân* ‘zerstreut (über etw.) nachdenken’. Bei Kontaktposition zweisilbiger Wörter ist eine wechselseitige Überschneidung ihrer Silben nach dem Muster ABCD → ACBD möglich: *thay đổi lòng dạ* → *thay lòng đổi dạ* ‘die Absicht ändern’.

Die Silbenüberschneidung ist – ungeachtet ihrer Verbreitung im vietnamesischen Text, insbesondere im künstlerischen Text, wo ihr eine wichtige stilistische Funktion zukommt – ebenfalls exquisiten normativen Einschränkungen unterworfen. Bekannt sind auch Fälle einer Lexikalisierung von Überschneidungsbildungen, wo die Möglichkeit einer Gegenüberstellung mit den (normativ nicht zulässigen) Ausgangsvarianten lediglich theoretisch besteht. Derartige Fälle sind z. B. die im Wörterbuch der festen Wortverbindungen angeführten Ausdrücke *bu'o'm chán ong chu'o'ng* ‘Schmetterling sich.übersättig Biene überdrüssig.werd (verstoßen (eine Frau))’, *cơm bưng nu'o'c rót* ‘Reis (dar)reich Tee eingieß (mit Kost und Logis)’.

Die Komponenten wenigstens eines der sich überschneidenden Wörter müssen semantisch vollwertig sein, während dieses Moment für die Komponenten des anderen Wortes nicht wesentlich ist. Unter diesem Gesichtspunkt betrachtet, heißt das für die Wortüberschneidung, daß sich ein Wortbe-

standteil (der bedeutungstragende oder nichtbedeutungstragende) in einem gewissen engen formalen Rahmen wie das Wortganze verhält.

## 7. Wortklassen

### 7.1. Morphologischer und syntaktischer Ansatz

Die Erörterung des Problems der **grammatischen Wortklassen** (Wortarten) im Vietnamesischen, wo morphologische Formenbildung vollständig und morphologische Wortbildung nahezu vollständig fehlt, macht deutlich, daß für die Wortartenunterscheidung nicht allein vom morphologischen Ansatz ausgegangen werden kann (s. Art. 70–72). Die ausschließliche Berücksichtigung des morphologischen Aspekts führt dazu, die Existenz von Wortarten im Vietnamesischen zu negieren (Grammont & Lê Quang Trinh 1911–1912). Morphologische Wortartklassifikationen rufen selbst in bezug auf Sprachen mit entwickeltem Formensystem Einwände hervor. Es wird z. B. darauf hingewiesen, daß die Deklinierbarkeit nicht als Merkmal der russischen Substantive angesehen werden kann, weil es in der russischen Sprache viele indeklinable Substantive gibt. Beim morphologischen Ansatz erfolgt die Subklassifikation der flektierbaren und der nichtflektierbaren Wörter anhand verschiedener Kriterien. Im weiten typologischen Kontext erscheinen die morphologischen Wortartindikatoren als rein äußerliche, im Grunde fakultative Marker, durch die tiefergehende, vorwiegend syntaktische Unterschiede fixiert werden. Ohne die morphologischen Indikatoren büßt die Wortartenunterscheidung an formaler Einfachheit ein, sie kehrt aber zugleich zu ihrer realen Basis – zur Syntax – zurück.

### 7.2. Klassifikation nach kombinatorischen Merkmalen

Der erste Versuch einer Wortartklassifikation nach kombinatorischen Merkmalen wurde von Lê Vǎn Ly unternommen, in dessen Arbeit die vietnamesischen Wörter nach ihrer Kombinierbarkeit mit sog. "Zeugenwörtern" (*mots temoins*) in einige größere Klassen unterteilt werden. Die Prinzipien der Herausgliederung der Zeugenwörter, zu denen die Marker von grammatischen Kategorien und einige synsemantische Wörter gehören, werden nicht diskutiert, und die Zeugenwörter selbst werden gleichsam a priori in einer Liste

angegeben. So sind etwa Zeugenwörter für Substantive die Pluralmarker *nhung*, *các*, die Wörter *máy* 'einige', *mọi* 'alle', *nhiều* 'viele', *bao nhiêu* 'wieviele' und eine Reihe weiterer (Lê Vǎn Ly 1948: 173–197). Der vorliegende Ansatz ist interessant als grundsätzlich richtiger Versuch einer Wortartklassifikation nach kombinatorischen Merkmalen. Überzeugender ist dieser Ansatz in Klassifikationen vertreten, die die Tradition von Lê Vǎn Ly fortsetzen, aber seine sehr umfangreichen Listen von Zeugenwörtern vereinfachen (Martini 1950: 186–190, Honey 1956: 537–539).

### 7.3. Funktionale Klassifikation

Der funktionale Ansatz stützt sich auf die unterschiedliche Verwendbarkeit der Wörter in verschiedenen syntaktischen Funktionen. Genauer gesagt geht es nicht darum, in welcher Funktion ein Wort in einem konkreten Beispiel auftritt, sondern um das funktionale Spektrum an syntaktischen Möglichkeiten des Wortes. Ebenso wird ein morphologisches Merkmal nicht durch irgendeine einzelne Form, sondern durch ein System von Formen repräsentiert (Jachontov 1968: 73 f.). Das sind die allgemeinsten theoretischen Ausgangspositionen für die funktionale Klassifikation, deren Ausarbeitung bestimmte Schwierigkeiten bereitet. Erforderlich ist die theoretisch überzeugende und praktisch wirkungsvolle Bestimmung des Funktionsbegriffs, die Abgrenzung unterschiedlicher Typen von Funktionen, die Beachtung möglicher Zusammenhänge zwischen den Funktionen sowie die Lösung einer Reihe weiterer komplizierter Fragen. Diese Schwierigkeiten lassen sich jedoch überwinden. Mit Hilfe des funktionalen Ansatzes – das scheint besonders wichtig – wird es möglich sein, eine Klassifikation des gesamten Materials auf einheitlicher Grundlage vorzunehmen (Panfilov 1990).

## 8. Illustrativer Text

In der vietnamesischen Orthographie wird jede Silbe getrennt geschrieben, unabhängig davon, ob sie ein Wort oder einen Wortbestandteil repräsentiert.

*Quá trình hình thành tiêu thuyết ở  
Prozeß sich.bild Roman in  
Tây Âu so với Trung Quốc  
westeuropäisch vergleich mit China  
và Việt Nam tuy có một số điểm  
und Vietnam obwohl EXIST einige Punkt*

giống nhau, nhưng lại có  
ähnlich zusammen aber trotzdem EXIST  
nhưng mặt rất khác nhau.  
PL Gesicht sehr unterschied zusammen  
Phân tích nguồn gốc phát sinh và quá trình  
analysier Quelle entsteh und Prozeß  
hình thành của tiểu thuyết, chúng ta nhầm  
sich.bild KONN Roman wir hinziel  
mục đích bu'ó'c đầu rút ra một số  
Ziel Schritt erst herauszieh einige  
đặc trưng thâm mỹ, một số đặc điểm  
Besonderheit ästhetisch einige Kennzeichen  
truyền thống của thể loại. Trong những  
Tradition KONN Genre innerhalb PL  
cuốn tiểu thuyết đầu tiên, những  
Papierrolle Roman Anfangs- PL  
đặc trưng thâm mỹ của thể loại  
Besonderheit ästhetisch KONN Genre  
hiện ra một cách đơn giản nhất,  
erschein ein Verfahren einfach selbst  
do đó cũng rõ ràng nhất, dễ thấy  
deshalb dieses auch klar selbst, leicht  
nhất.  
selbst

‘Obwohl der Entstehungsprozeß des westeuropäischen Romans in bestimmter Hinsicht an analoge Prozesse in China und Vietnam erinnert, gibt es auch wesentliche Unterschiede. Im Zusammenhang mit der Untersuchung der Entstehung und Entwicklung des Romans setzen wir uns zuerst das Ziel, eine Reihe ästhetischer und traditioneller Besonderheiten des Genres aufzudecken. In den ersten Werken treten die ästhetischen Besonderheiten des Genres in sehr einfacher und damit in sehr klarer, unvermittelte Form hervor.’

## 9. Zitierte Literatur

- Apresjan, J[urij] D. (1974), *Leksičeskaja semantika*. Moskva: Nauka
- Đái Xuân Ninh (1978), *Hoạt động của từ tiếng Việt*. Hanoi: Khoa học xã hội
- Glebova, J[vetta] J. & Sitnikova A[ntonina] N. (1980), “Grammatičeskaja klassifikacija neosložnennych povtorov v sovremennom v'etnamskom jazyke”. In: Alieva, N. F. (Hrsg.), *Jazyki Jugo-Vostočnoj Azii. Problemy povtorov*. Moskva: Nauka, 48–75

*stočnoi Azii. Problemy povtorov*. Moskva: Nauka, 48–75

Grammont, M[aurice] & Lê Quang Trinh (1911–1912), “Etudes sur la langue annamite”. *Mémoires de la Société de Linguistique de Paris* 17, 201–241, 295–310

Haudricourt, A[ndre] G. (1953), “La place du vietnamien dans les langues austroasiatiques”. *Bulletin de la Société de Linguistique de Paris*, 4.91, 122–128

Honey, P[atric] J. (1956), “Word Classes in Vietnamese”. *Bulletin of the School of Oriental and African Studies* XVIII.3, 534–544

Hồ Lê (1971), “Những su nghien cuu va tranh luan cua nuoc ngoai xung quanh van de nguon goc tieng Viet”. *Ngon Ngu* 4, 42–49

Hồ Lê (1976), *Van de cau tao tu cua tieng Viet hien dai*. Hanoi: Khoa hoc xa hoi

Jachontov, S[ergej] E. (1968), “Ponjatie častej reči v obščem i kitajskom jazykoznanii”. In: Zirmunkij, V. M. & Sunik, O. P. (Hrsg.), *Voprosy teorii častej reči na materiale jazikov različnykh tipov*. Leningrad: Nauka, 70–90

Lê Văn Ly (1948), *Le parler vietnamien*. Paris: Huong Anh

Lê Văn Ly (1981), “Cach thuc cau tao va to hop cua tu ngu Viet Nam”. In: *Giu gin su trong sang cua tieng Viet ve mat tu ngu*. Hanoi: Khoa hoc xa hoi, 56–63

Lê Xuân Thái (1982), “Ve cac dong tu va tinh tu khong truc tiep lam vi ngu trong tieng Viet”. *Ngon ngu* 1, 48–53

Martini, Francois (1950), “L’opposition nom et verbe en vietnamien et en siamois”. *Bulletin de la Société de Linguistique de Paris* XLVI.1, 183–196

Maspero, Henri (1912), “Etudes sur la phonétique historique de la langue annamite: Les initiales”. *Bulletin de l’Ecole Française d’Extrême Orient* XII.1, 1–127

Murjasov, R[achim] Z. (1977), “O napravlenii prizvodnosti i toždestve derivacionnykh morfem”. *Voprosy jazykoznanija* 6, 119–125

Nguyễn Huy Cận (1981), “Hien tuong gop tieng cua cac chuoii tu ghep phu nghia”. In: *Giu gin su trong sang cua tieng Viet ve mat tu nghia*. Ha Noi: Khoa hoc xa hoi, 64–70

Nguyễn Kim Thành (1977), *Dong tu trong tieng Việt*. Ha Noi: Khoa học xã hội

Nguyễn Phú Phong (1977), “Ve van de lay tu trong tieng Viet”. *Ngon ngu* 2, 61–68

Panfilov, V[alerij] S. (1990), "Klassy slov (časti reči) vo v'etnamskom jazyke". *Voprosy jazykoznanija* 5, 74–89

Przyluski, J[ean] (1924), "L'annamite". In: Meillet, A. & Cohen, M. (Hrsg.), *Les langues du monde*. Paris: Champion, 395–398

Sitnikova, A[ntonina] N. (1980), "K charakteristike značenij častičnyx motivirovannyx povtorov v sovremenном v'etnamskom jazyke". In: Alieva, N. F. (Hrsg.), *Jazyky Jugo-Vostocnoj Azii. Problemi povtorov*. Moskva: Nauka, 240–258

Thompson, Laurence C. (?1967), *A Vietnamese Grammar*. Seattle, London: Univ. of Washington Press [?1965]

Tru'o'ng Đōng San (1976), "Cac bien the cua mot tu va mot cum tu co dinh". *Ngon ngu* 2, 9–14

*Valerij S. Panfilov, St. Petersburg (Rußland)*  
Übersetzung:

*Holger Baumann, Erfurt (Deutschland)*  
*Martin Haspelmath, Leipzig (Deutschland)*

## 143. Deutsche Gebärdensprache

1. Einleitung
2. Problem der Entwicklung von Beschreibungskategorien
3. Aspekte der Morphologie in Gebärdensprachen
4. Modell der Morphologie der Bewegungsverben
5. Serialverb-Konstruktionen
6. Gegenpositionen
7. Weiterentwicklung
8. Zitierte Literatur

### 1. Einleitung

Im folgenden geht es vornehmlich um die Problematik der morphologischen Beschreibung einer nicht primär auditiv, sondern primär visuell wahrnehmbaren Sprache, der Deutschen Gebärdensprache (DGS), stellvertretend für die verschiedenen existierenden Gebärdensprachen der Gehörlosen. "SprecherInnen" dieser Sprache (in der Linguistik sind noch keine auf Gebärdensprachen bezogenen Ersatztermini für "SprecherIn – HörerIn" etabliert) sind in Deutschland schätzungsweise 80.000 prälingual Gehörlose sowie eine Anzahl schwerhöriger und hörender Personen wie hörende Kinder gehörloser Eltern oder GebärdensprachdolmetscherInnen. Die Weltgesundheitsorganisation geht davon aus, daß weltweit 0,08–0,1 Prozent der Bevölkerung prälingual gehörlos sind.

Entgegen dem landläufigen Glauben gibt es keine internationale natürliche Gebärdensprache, sondern untereinander nur begrenzt verständliche (nationale) Gebärdensprachen, die in bestimmten Bereichen strukturelle Gemeinsamkeiten haben. Dabei sind die jeweili-

gen Gebärdensprachen keine Erfindung der hörenden GehörlosenlehrerInnen zum Zweck der sogenannten Sprachanbahnung, sondern von den Gehörlosen-Gemeinschaften in der Alltagskommunikation selbst entwickelte Sprachen. Für die Sprachgeschichte können wir davon ausgehen, daß, wo immer Gehörlose zusammentreffen, zunächst in der Familie, später verstärkt durch die Bildung größerer Gehörlosen-Gemeinschaften in Städten, dann durch das "institutionalisierte Taubstummenwesen" (in Deutschland ab dem Ende des 18. Jahrhunderts durch Samuel Heinicke), gebärdete Kommunikationsformen ausgebildet werden, die u. a. durch Konventionalisierung über Generationen hinweg zu Gebärdensprachen entwickelt werden (wir wollen mit dieser Aussage Gebärdensprachentwicklung nicht an Schulgründungen binden – für Paris ist z. B. belegt, daß es dort eine Gehörlosengemeinschaft unabhängig von der Gründung der Pariser Schule gab; vgl. Desloges 1779). Die sprachlichen Minderheiten der Gehörlosen sind soziolinguistisch insofern eine Besonderheit, als in der Regel ein Generationenbruch vorliegt. Da nur schätzungsweise 10 Prozent der Gehörlosen auch gehörlose Eltern haben, lernen die meisten Gehörlosen Gebärdensprache nicht von ihren Eltern, sondern in der *peer group* resp. von älteren SchülerInnen, vorzugsweise von gehörlosen Kindern gehörloser Eltern.

Es handelt sich bei der Deutschen Gebärdensprache (wie bei jeder Gebärdensprache Gehörloser) um eine Sprache mit eigener Grammatik und Lexik. Damit nicht identisch sind andere visuelle Kommunikationsformen

wie z. B. von Gebärdensprache begleitetes Deutsch ("lautsprachbegleitendes Gebärdensprache") oder das Fingeralphabet (mit welchem deutsche Wörter "in der Luft buchstabiert" werden), auch wenn einige dieser Elemente für bestimmte Teifunktionen in die Deutsche Gebärdensprache integriert sind.

Die Deutsche Gebärdensprache wird im deutschsprachigen Raum benutzt. Hier wie auch für andere Gebärdensprachen ist beim jetzigen Stand der Dialektforschung eine Bezeichnung wie "die" Deutsche Gebärdensprache allerdings eher eine Notlösung. Es gibt bisher weltweit keine normierten Gebärdensprachen Gehörloser; die Gebärdensprachdialektologie steckt noch in den Anfängen. Mehr als in sonstigen soziolinguistischen Fragen spielen bei Gebärdensprachdialekten neben geographischen Verhältnissen die jeweiligen Gehörlosenschulen eine Rolle; auch in der Deutschen Gebärdensprache erkennen Gehörlose sich z. T. als (ehemalige) Angehörige unterschiedlicher Schulen/Internate aufgrund des jeweiligen Gebärdensprachdialekts, obwohl Gebärdensprache in den Schulen bisher kein reguläres Schulfach ist, sondern allenfalls im Rahmen von Modellversuchen oder Lerngruppen unterrichtet wird.

Keine Gebärdensprache Gehörloser hat eine Gebrauchsschrift. Für die Deutsche Gebärdensprache wie auch international gibt es keine gebräuchlichen Schriftsysteme, sondern allenfalls miteinander konkurrierende, wenig verbreitete Gebärdensprach-Verschriftungsvorschläge oder Gebärdensprach-Notationen für den wissenschaftlichen Gebrauch. Die erste Gebärdensprachsschrift, die ausdrücklich Alltags- wie gelehrt Zwecken dienen sollte, wurde zu Beginn des 19. Jahrhunderts in Paris entwickelt (Bébian 1825; vgl. Fischer 1995); nicht zuletzt aus sprachpolitischen Gründen wurde ihre Verwendung an der Pariser Schule regelrecht verhindert.

Folglich gibt es im strengen Sinne keine historischen Quellen der Deutschen Gebärdensprache – und dies, obwohl z. B. Graßhoff in Anerkennung eines leistungsfähigen Kommunikationsmediums für eine "Lebens-Erliechterung der Taubstummen durch Gründung einer Taubstummengemeinde" (Graßhoff 1820: Titel) plädierte, in der sie – wie man heute sagen würde – als sprachliche Minderheit "die ihnen allen gemeinsame Gebärdensprache" (1820: 6) verwenden würden. Zwar gibt es etliche Quellen, die vor allem Einzelgebärden aus dem deutschen Sprachraum beschreiben oder sie in Zeichnungen

darstellen (vgl. Fischer 1996); als frühe Zeugnisse der Deutschen Gebärdensprache können sie aber unseres Wissens insofern nur eingeschränkt gelten, als

- sie (bis auf zeichnerische Darstellungen) lautsprachvermittelt (d. h. in Lautsprache beschriebene Gebärdensprache) sind, also genau genommen keine Primärquellen darstellen;
- sie ausgewählte Einzelgebärdenbeschreibungen darstellen, Angaben z. B. über Syntax oder auch nur über Gebärdensprachverknüpfung weitgehend fehlen;
- sie z. T. keine Gebärdensprache Gehörloser, sondern z. B. Mönchsgebärden (Leibniz 1717) sind;
- fast alle Gebärdensprachbeschreibungen von Hörenden stammen, die sie für bestimmte (pädagogische u. a.) Zwecke formulieren; d. h. diese Gebärdensprachbeschreibungen sind in Auswahl und Form abhängig vom pädagogischen Zweck, also nicht ohne weiteres als authentische Gebärdensprache Gehörloser zu bezeichnen (Czech 1836; Reich 1834; Gebärdensprachbeispiele aus beiden Werken sind von R. Fischer in Reich 1991 zusammengestellt worden).

Der wohl wichtigste Aspekt noch heute sowohl für die GebärdensprachverwenderInnen wie auch für die Gebärdensprachforschung ist die Tatsache, daß Gebärdensprachen unterdrückte Sprachen sind – im Fall der Deutschen Gebärdensprache eine fast mit allen Mitteln unterdrückte Sprache, denkt man an die Gesetzgebung zur "Verhinderung erbkranken Nachwuchses" im Dritten Reich mit Zwangssterilisation und Zwangsausbildung mit dem Ziel der Ausrottung von (unter anderem) erblicher Gehörlosigkeit (Biesold 1988; Fischer et al. 1997). Erst das Sozialgesetzbuch IX (2001) und das Bundes-Behindertengleichstellungsgesetz (2002) haben eine neue gesetzliche Grundlage in Richtung auf kommunikative Barrierefreiheit geschaffen.

## 2. Problem der Entwicklung von Beschreibungskategorien

Auf internationaler Ebene setzt man den Beginn der Gebärdensprachforschung nach 1960 mit der verspätet einsetzenden Rezeption der Arbeiten Stokoes an (beispielsweise Stokoe 1960). Eine solche Einordnung läßt allerdings außer acht, daß im 19. Jahrhundert in verschiedenen europäischen Ländern eine

auch sprachwissenschaftlich orientierte Beschäftigung mit Gebärdensprache, teilweise durch Gehörlose selbst, bereits möglich war, bevor in der zweiten Hälfte des 19. Jahrhunderts länderübergreifend die Unterdrückung der Gebärdensprachen (erneut) durchgesetzt wurde. Die 40 Jahre neuerer linguistischer Forschungstätigkeit zu Gebärdensprachen haben (durchaus politisch mitbedingt) Arbeiten insbesondere über die Amerikanische (*American Sign Language*, ASL), die Australische (*Auslan*) und die Britische Gebärdensprache (*British Sign Language*, BSL) sowie über skandinavische Gebärdensprachen hervorgebracht. Für die Deutsche Gebärdensprache gibt es bisher weder eine präskriptive noch eine deskriptive Grammatik. Interessierte können einen ersten Eindruck von Strukturbeschreibungen aus Boyes Braem (¹995) gewinnen; die Verfasserin geht jedoch ausdrücklich nicht speziell auf die Deutsche Gebärdensprache ein, sondern gibt als Ziel an, „die linguistischen Prinzipien (zu) illustrieren, die allen Gebärdensprachen gemeinsam sind“ (Boyes Braem ¹995: 15).

In gewisser Weise ist dieser Stand Spiegel der oben kurz angedeuteten Situation der Unterdrückung der Gebärdensprachen, zu der auch die jahrhundertealte Fixiertheit auf Lautsprache als vermeintliches Definiens des Menschen gehört.

Diese Umstände wirken sich in der Gebärdensprachforschung inhaltlich aus. Die Gebärdensprachforschung ist überwiegend orientiert an Lautsprachlinguistik, indem sie ganz grundsätzlich auf deren Kategorien aufbaut. Dies geschieht als zunehmend reflektierter Rückgriff auf bereitliegende wissenschaftliche Begrifflichkeit oder ganz offen in der Absicht, Gebärdensprachen sprach- und gesellschaftspolitisch zur Anerkennung zu verhelfen. Comrie notierte als einer der ersten UniversalienforscherInnen als Desiderat typologischer Forschung, daß Gebärdensprachen zu beachten seien; er tat dies mit dem ausdrücklichen Hinweis, daß sich viele der heute noch als Universalien geltenden Aspekte bei Beachtung von Gebärdensprachen schlicht als Eigenheiten des Mediums, nicht aber als Charakteristikum einer „Sprachfähigkeit“ herausstellen könnten (Comrie 1981: 221).

### 3. Aspekte der Morphologie in Gebärdensprachen

Wir greifen zunächst, um einen Eindruck zu vermitteln, auf Arbeiten zu morphologischen Teilbereichen verschiedener Gebärdenspra-

chen zurück; dabei verzichten wir weitestgehend auf die u. E. ansonsten erforderliche Spezifizierung der jeweiligen Gebärdensprache.

Zentrale Bedeutung für Gebärdensprachverwendung und -beschreibung hat der Umgang mit dem dreidimensionalen **Gebärdensraum**, der sich vor Oberkörper und Kopf der Gebärdenden befindet. In ihm kommen insbesondere die Hände und Arme, aber auch der Oberkörper insgesamt sowie der Kopf, Mimik und insbesondere auch der Mund (Mundbild, Mundgestik u. ä.) zum Einsatz. Die in diesem Gebärdensraum geleistete sprachliche Lokalisierung und Repräsentation von „Objekten“ oder Vorgängen mutet zunächst so anders als die Grammatik von Lautsprachen an, daß in früheren Phasen der Gebärdensprachforschung des 20. Jahrhunderts durchaus noch Formulierungen benutzt wurden, die die heute übliche Verwendung von Kategorien der Lautsprachlinguistik als nicht zwingend erweisen. So sagt beispielsweise Deuchar:

“(… ) we find that the nearest approximation to inflection for case is spatial modification (… ).” (Deuchar 1978: 80)

“So in a limited class of BSL verbs the direction of movement, and in some verbs the orientation also, indicate the case roles associated with those verbs. This may be compared to inflection in spoken languages, except that it is more restricted.” (Deuchar 1978: 82)

Unseres Wissens ist bisher keine Gebärdensprachmorphologie in ihrer Gesamtheit beschrieben und typologisch eingeordnet worden. Man ist lange davon ausgegangen, daß **Gebärden** in nicht unerheblichem Maß quasi isoliert als Einzelgebärden vorkommen. Soffern man sich für diese Einschätzung auf die Ebene der Hände beschränkt und nicht-manuelle Elemente wie Mimik oder Körperhaltung unberücksichtigt läßt, mag eine derartige Aussage für etliche Gebärden vertretbar scheinen. Daneben sind schon früh Gebärden aufgefallen, die sich einer derartigen Einstufung vollkommen entziehen und die man z. B. aufgrund des Eindrucks, daß sie reale Abläufe direkt nachzuahmen scheinen, anfangs noch für eher pantomimische, i. S. von nicht-analysierbare und „nicht-sprachliche“ Elemente hielt. Im Zuge der weiteren Gebärdensprachforschung sind nun diese komplexen Gebärden mit ihrem Flexionsreichtum und ihrer polysynthetischen Gebärdensstruktur zu einem viel bearbeiteten Gegenstand ge-

worden. Das Vorliegen oder Nicht-Vorliegen morphologischer Komplexität wird zur Unterscheidung von Gebärdentypen verwendet (beispielsweise die Verbklassen "inflecting verbs – plain verbs – spatial verbs" seit Padden 1988).

Für die Amerikanische Gebärdensprache ist eine Wortartenunterscheidung in Verben und Substantive postuliert worden u. a. der Art, daß die jeweilige Nominalgebärde in der Bewegung kürzer sei und als solche wiederholt werde (Supalla & Newport 1978).

Für die semantische Kategorisierung wurden etliche Ausdifferenzierungen vorgeschlagen. Es scheint beispielsweise relativ unstrittig, daß in Gebärdensprachen Negationsinkorporation vorliegt; daß Tempus wohl keine morphologische Kategorie darstellt, wohingegen es diverse Aspektunterscheidungen auch mit Mitteln der Mimik zu geben scheint; daß Numerus morphologisch beschreibbar ist, nicht aber Genus (Smith 1990 legt nahe, daß sich in der Taiwanischen Gebärdensprache als bisher erster diese Markierung doch finden läßt). Über das Vorkommen von "Person"-Markierung ist z. Zt. eine Diskussion im Gange. Generizität werde morphologisch realisiert u. a. durch Wegfall bestimmter Kongruenzmorpheme. In formaler Hinsicht überwiegen diversen Darstellungen zufolge Inkorporation und Flexion als morphologische Prozesse.

Angesichts letzterer scheint für die strukturelle Bearbeitung von Gebärdensprachen die Morphonologie von besonderer Bedeutung zu sein. Das meist verwendete, von Stokoe (1960) begründete, später erweiterte phonologische Modell für Gebärdensprachen sieht die Gebärdensbeschreibung mithilfe der phonologischen (cherologischen) Merkmale Handform, Handstellung, Ausführungsstelle und Bewegung (mit jeweils einer Anzahl von Unteraspekten) vor. Diesem Modell liegt zentral die Annahme der Simultaneität der verschiedenen Merkmale zugrunde; ein anderes Modell betont demgegenüber mit seiner Unterscheidung von Bewegungs- und Halte-Segmenten Sequentialität (Liddell & Johnson 1989) und ermöglicht darüber hinaus die morphonologische Beschreibung von Gebärdensstrukturen über Leerstellen.

Mit der auf Stokoe basierenden phonologischen Beschreibung lassen sich Gebärdeminimalpaare aufstellen und einordnen; für die Deutsche Gebärdensprache z. B. RUSSLAND vs. KUH, wo dem obigen Modell zufolge der Aspekt der Ausführungsstelle (Kinn

vs. Stirnseite) unter Beibehaltung aller anderen Merkmale bedeutungsentscheidend wirkt. Variiert man Merkmale, so erhält man nicht immer Minimalpaare oder Nonsensegebärden, sondern in vielen Fällen eine grammatische Modifikation (Flexion, Inkorporation) des ursprünglichen Gebärdentokens.

#### 4. Modell der Morphologie der Bewegungsverben

In verschiedenen Gebärdensprachen gelten **Bewegungsverben** aus der Gruppe der "spatial verbs" als ein Bereich komplexer morphologischer Verhältnisse; man zählt sie zu den polysynthetischen Konstruktionen. An ihnen möchten wir den von KritikerInnen so genannten "traditionellen" Ansatz der Gebärdensprachlinguistik (d. h. denjenigen, der Kategorien der Lautsprachlinguistik übernimmt) exemplifizieren. Wir tun dies anhand einer frühen Analyse, die für diesen Bereich wohl als die ausgearbeitetste gelten durfte, nämlich die von Supalla (1978; 1982; 1986; 1990) für die Amerikanische Gebärdensprache vorgestellte. Bei dieser morphologischen Analyse sind zunächst zu trennen 1) der Stamm und 2) fakultative Affixe wie dasjenige für Aspekt. Die Morphologie des Verbstammes selbst wird beschrieben als eine Kombination von a) obligatorischen und ggf. zusätzlichen fakultativen Elementen aus dem Bereich **Bewegung** mit b) anderen obligatorischen Affixen, die über den die Bewegung ausführenden **Körperteil** zu beschreiben sind.

- zu a): Die jeweilige Bewegung stellt in den meisten Fällen ihrerseits eine Morphemkombination aus obligatorischer Wurzel und fakultativen anderen Bewegungsaffixen dar. Dabei entspricht die Wurzel einer der Basisprädikationen für die dazugehörigen Nomina: 'existence' – 'location' – 'motion'. Es gibt nur eine begrenzte Anzahl Wurzeln, die ihrerseits aus einer noch geringeren Anzahl von Bewegungsmerkmalen wie 'stative' oder 'linear' zusammengesetzt und zu komplexen Bewegungsfolgen kombinierbar sind. Fakultativ können hinzukommen sekundäre Bewegungsaffixe für aa) 'degree of change', bb) 'directionality of change' und cc) 'frequency of change' (verstanden als Art der Bewegung), wobei es jedoch für die Kombination mit den Wurzeln verschiedene Restriktionen gibt.

	BEWEGUNGSKOMPONENTE		KLASSIFIKATOR	
	{ART}	{RICHTUNG/PFAD}	{KÖRPER(TEIL)}	semantisch
1. Sequenzteil	+	-	+	-
2. Sequenzteil	-	+	-	+

Tab. 143.1: Morphemkombinationsrestriktionen

– zu b): Der die Bewegung realisierende Körperteil verweist in Bewegungsverben auf (semantische) Merkmale und ggf. Lokalisation des oder der betreffenden Referenten und wird von Supalla in Analogie zu einigen Lautsprachen als Klassifikator eingeordnet (s. Art. 97). Die mit Bewegungsverben kombinierten Klassifikatortypen sind diejenigen der semantischen Klassifikatoren, bei denen sogenannte abstrakte semantische Merkmale oder Merkmalskombinationen wie ‘vertikal’ oder ‘mit zwei Beinen’ zum Tragen kommen, und diejenigen der Körper(teil)-Klassifikatoren, bei denen der Körper oder Körperteile auf den-/dieselben oder auf andere, über eine Ähnlichkeitsrelation verbundene Körperteile (Hände für Füße u. ä.) des jeweiligen Referenten(nomens) verweisen. Die Körper(teil)-Klassifikatoren werden von Supalla nicht in den phonologischen Teil der Analyse einbezogen; die jeweils wechselnde Handform gilt somit als morphonologisches Merkmal des betreffenden Klassifikators. Im Unterschied zu anderen, als multi-morphemisch analysierten Klassifikatortypen (z. B. den “size-and-shape-specifiers”) sieht Supalla sowohl die semantischen wie die Körper(teil)-Klassifikatoren als je ein Morphem an.

Semantische Klassifikatoren haben eine morphonologische Markierung der Hand als ausführendem Körperteil nicht nur über die Form, sondern Supallas Analysen zufolge auch über Orientierung und Plazierung der Hand; hierbei werden Aspekte aus dem Bereich “reference scale” morphologisch gefasst.

## 5. Serialverb-Konstruktionen

Die unter 4. skizzierte Morphonologie der Bewegungsverben basiert auf der Annahme, daß alle Morpheme simultan vorkommen. Supalla (1990) sieht jedoch eine Modifikation insofern für angebracht an, als er einige

grammatische Restriktionen ausgemacht hat, die die Sequentialität von Morphemgruppen erzwingen und dadurch eine spezielle Konstruktion, von Supalla in Anlehnung an Lautsprachen **serial verbs** genannt, hervorbringen. Hierbei werden, anders als in den simultan ausgeführten Morphemkombinationen, die verschiedenen Komponenten eines einzigen Geschehens in der Versprachlichung nicht in einem einzigen Verbkomplex verknüpft, sondern auf eine Gebärdensequenz in charakteristischer Weise verteilt.

Beide Teile der Sequenz bestehen wie jedes “normale” Bewegungsverb mindestens aus einem Bewegungskomplex und einem Morphem für den Klassifikator. Jedoch sind erster und zweiter Verbalteil unterschiedlich spezifisch in Hinblick auf Bewegung und Klassifikator. Der erste Sequenzteil enthält für die Bewegungskomponente ein Morphem {ART DER BEWEGUNG}; der zweite hat mindestens ein Morphem {RICHTUNG UND PFAD DER BEWEGUNG}. Obwohl der Referenzausdruck der gleiche ist, fällt das Morphem dafür im ersten Sequenzteil anders aus als im zweiten: {ART DER BEWEGUNG} wird verbunden mit einem Morphem, das für Supalla zur ikonischen Gruppe der so genannten Körper(teil)-Klassifikatoren gehört; die übrigen Bewegungsmorphe werden im zweiten Teil mit einem Morphem der “abstrakteren” Klasse der “semantischen” Klassifikatoren verbunden.

Supalla führt als Erklärungen für die “serial verbs”-Sequentialität an:

- schwierig auszuführende Bewegungen (“physical restrictions on simultaneity”; Supalla 1990: 131 f.);
- “arbitrary grammatical restrictions” (Supalla 1990: 132), zu denen essentiell die morphonologischen Kombinationsbeschränkungen gehören, die die Verbindbarkeit a) bestimmter Bewegungsmorphe untereinander, b) bestimmter Bewegungsmorphe mit bestimmten Klassifikatortypen betreffen.

S. Fischer & Janis (1990: 287) erwägen darüberhinaus das Kriterium "heaviness": daß es "too many inflections" an einer Verbwurzel geben könnte.

## 6. Gegenpositionen

Supallas morphonologische Analysen sind nicht unwidersprochen geblieben. Wir möchten zur Illustration der grundsätzlichen Problematik einige Gegenvorschläge erwähnen.

Noch im Rahmen morphologischer Analysen bleiben die Ansätze z. B. von Schick und Engberg-Pedersen. Schick (1990) kritisiert an Supalla, seine morphonologische Analyse sei mit einem wichtigen Charakteristikum gebärdensprachlicher Klassifikatorprädikate unvereinbar, nämlich mit ihrer **Ikonizität** (s. Art. 30). Daher möchte sie – neben anderen Abweichungen von Supallas Analyse – für die Bewegungsmorphe in diesen Prädikaten folgende drei Klassen, die der Ikonizität gerecht würden, zugrundelegen: "movement through space (MOV), a stylised imitation of real-world action (IMIT), and a single point in space (DOT)" (Schick 1990: 17). Im Vergleich mit Supallas erscheint ihre morphologische Analyse entsprechend "gröber".

Ebenfalls auf der Basis morphologischer Analysen arbeitet Engberg-Pedersen (insbesondere Engberg-Pedersen 1993). Neben einer anders vorgenommenen morphologischen Analyse und der Ablehnung der Klassifikator-Interpretation (Engberg-Pedersen 1989; Engberg-Pedersen & Pedersen 1985) ist die bei ihr vorfindbare Abweichung von Supalla eine inhaltlich-funktionale Interpretation. So findet sich in ihren Analysen eine Beachtung lange vernachlässigter, auch für morphologische Analysen relevanter Aspekte, insbesondere bezüglich der Gebärdensprachenverwendung im Diskurs, womit sie über Einzelgebärden oder Beispielsätze hinausgeht. Sie beschäftigt sich mit der gleichzeitigen Verwendung beider Hände außerhalb von Zwei-Hand-Gebärden, mit Fragen von Figur-Grund-Realisierungen, mit der Variation bei der Artikulatorenwahl für identische Referenten.

S. Fischer & Janis (1990), die bei einem Nachvollzug die Analysen Supallas betreffend Serialverben experimentell nicht durchgehend bestätigen konnten, bringen die von Supalla beschriebenen Konstruktionen in Verbindung mit Satzkonstruktionen, die ebenfalls bestimmte Formen der Sequentiali-

tät incl. Teilwiederholungen beinhalten (vorläufig "verb sandwiches" genannt). So gesehen wären Supallas Serialverb-Konstruktionen weniger morphonologisch als (morpho-)syntaktisch, dann auch pragmatisch zu analysieren.

Andere Ansätze (bisher in geringerer Zahl) stellen dagegen die an Lautsprach-Morphologie orientierte Analyse an sich in Frage mit der Begründung, sie sei in verschiedener Hinsicht für Gebärdensprachen nicht adäquat (daß Lautsprach-Linguistik auch ihren Gegenstand Lautsprache reduziert, ihm nicht in seiner Komplexität gerecht wird, wird dabei bisher nicht eigentlich in Betracht gezogen). Die grundsätzliche Frage, was in Gebärdensprachen überhaupt als Morphem gelten kann, muß als nicht geklärt angesehen werden, trotz des Vorliegens einer Reihe morphologischer Detailanalysen. So ist beispielsweise Supallas detailliertes morphologisches Modell nur haltbar, wenn die differenzierte Analyse auf die Hände beschränkt wird. Anderweitige Untersuchungen haben aber inzwischen erwiesen, daß die grammatische Strukturierung in Gebärdensprachen über eine Interaktion der diversen körperlichen Artikulatoren bewerkstelligt wird, die nicht einfach eine Addition isoliert bearbeiteter Artikulatorenmorphologien o. ä. darstellt – eine derartige Komplexität ist linguistischerseits bisher überhaupt nicht beschreibbar.

Zu den frühen Gegenvorschlägen zählen die Arbeiten Jouisons, die durch seinen Tod 1991 ein Ende gefunden haben. Jouison hat ebenfalls grundsätzliche Kritik geäußert und bereits Einzelanalysen, die u. a. das Klassifikatorproblem in polysynthetischen Verben berühren, vorgelegt (u. a. Jouison 1989). Charakteristikum seines Ansatzes ist, der "Faszination der Hände" entgehen zu wollen und stattdessen das körperliche Gesamtgeschehen im gebärdensprachlichen Diskurs zu beachten (Jouison 1992). Er nimmt eine doppelte Gliederung der Gebärdensprachen an; bereits bei der Bearbeitung der *signifiant*-Ebene verzichtet er auf eine Reduktion auf die manuelle Ebene. Dies wird in Verbindung mit der Grundlage Ikonizität zum Bestimmungsfaktor seiner Systemskizze: Ausgehend von der Ähnlichkeitsbeziehung des körpersprachlich Geformten mit dem außersprachlichen Gemeinten ("Objekt" oder Handlung), sucht er die von dieser ikonischen Beziehung implizierten Relationen der ausführenden Artikulationsorgane untereinander und gelangt so zu Hypothesen über phonologische und

morphologische Inventare, die gleichermaßen einem systematischen, nicht-pantomimischen Charakter der (Französischen) Gebärdensprache wie auch ihrer offensichtlichen Ikonizität Rechnung tragen sollen.

Für den deutschsprachigen Raum verweisen wir auf die Kritik des "traditionellen" Ansatzes bei Ebbinghaus & Heßmann (1991). Sie greifen auf "vergessene" frühere Vorschläge (insbesondere von DeMatteo in den 70er Jahren) zurück und kritisieren die vorherrschende Tendenz, "ausschließlich auf diskretisierbare Einheiten Bezug (zu nehmen)" und dabei "die Möglichkeit der Dynamisierung des ikonischen Potentials" als "eine der wesentlichsten Eigenschaften gebärdensprachlicher Zeichengenerierung" unberücksichtigt zu lassen (Ebbinghaus & Heßmann 1991: 65).

## 7. Weiterentwicklung

Seit 1992 hat sich international und auch bezüglich der Deutschen Gebärdensprache gebärdensprachlinguistisch sehr viel getan, so daß unser Beitrag auf grundsätzliche Fragen bezüglich gebärdensprachlicher Morphologie verweist statt den aktuellen Forschungsstand wiederzugeben. Bezüglich der Klassifikatorkonstruktionen, für die wir Supallas klassische morphologische Analyse vorgestellt haben, bietet Emmorey (2003, ed.) einen Einblick in derzeit diskutierte Forschungsperspektiven.

Zu diesen grundsätzlichen Fragen zählt weiterhin die Berücksichtigung beziehungsweise linguistische Identifizierung ikonischer/motivierter Gebärdanteile. So werden zum Beispiel wechselnde Ausführungsstellen bei der Gebärdengruppe der Kongruenzverben/Richtungsgebärden als grammatischer Bestandteil (Kongruenz) einer polymorphemischen Gebärde resp. als ein durch räumliche Faktoren motivierter gestischer Gebärdanteil interpretiert (vgl. Liddell 2000). Keller (1998) diskutiert eine Anzahl morphosyntaktischer Phänomene, u. a. Kongruenz, umfassend vor dem Hintergrund der grammatisierten Raumnutzung in der Deutschen Gebärdensprache.

Bis heute ist die morphologische Beschreibung der manuellen Ebene fortgeschritten gegenüber derjenigen der non-manuellen Komponenten, und um so mehr fehlt es an Vorschlägen für deren gemeinsame morphologische Betrachtung. Mit Ebbinghaus (1996)

liegt ein Beispiel für die Leistung der non-manuellen Komponente Blick(-richtung) für Referenz in Deutscher Gebärdensprache vor. Fachgebärdensleksikographische Arbeiten haben zunächst praktische Probleme identifiziert und dann zur vertieften Auseinandersetzung mit der "Wortstruktur" in der Deutschen Gebärdensprache geführt; die Schwierigkeiten, ein Lemma angesichts einer durch ggf. zahlreiche grammatische Leerstellen charakterisierten Gebärdensstruktur zu bestimmen, ähneln denjenigen, mit denen "die" (Lautsprach-) Linguistik bei so genannten kleinen Sprachen außerhalb des indo-europäischen Sprachraums konfrontiert ist. Damit hängt ferner auch das Problem zusammen, ein Lemma X als einen Gebärdentyp semantisch und morphologisch von Tokens und anderen Types zu unterscheiden (Konrad et al. 2001; vgl. auch Ebbinghaus & Heßmann 2000). Linguistische Identifizierung und Terminologiearbeit gehören entsprechend weiterhin notwendig zur Gebärdensprachforschung.

Mit der Gebärdensstruktur im Sinne der Wortartenbeschreibung und Wortbildung beschäftigen sich für die Deutsche Gebärdensprache zum Beispiel Becker (2001) und Erlenkamp (2000). Morphologische Aspekte berücksichtigen auch die Studien zu Vergebärdlern ("Versprechern") in Deutscher Gebärdensprache (Happ & Hohenberger 2001).

Angewandte Linguistik als kontrastive Grammatik für Deutsch und Deutsche Gebärdensprache ist ebenfalls ein neuer gebärdensprachlinguistischer Bereich, in dem für die Numeruskongruenz didaktisch aufbereitete Materialien zur Verfügung stehen (Fischer et al. 2000).

Die vorherrschenden Untersuchungsebenen sind weiterhin die Einzelgebärde und der Beispielsatz, textgrammatische Forschungen dagegen noch der Einzelfall.

Dies gilt auch für die textlinguistische Relevanz von Perspektivierung und Rollenwechsel, die zu den grundlegenden Charakteristika von Texten der Deutschen Gebärdensprache gehören.

Insgesamt muß trotz aller in den vergangenen Jahren erreichten Fortschritte in der Gebärdensprachforschung festgestellt werden, daß viele theoretische Fragen noch ungeklärt sind und darüber hinaus so wichtige Bereiche wie die Lehre der Deutschen Gebärdensprache und die GebärdensprachdolmetscherInnenbildung hohe Anforderungen stellen (vgl. Johnston 2000; Kollien 2000; Leven

2001), denen auch in der Zukunft von Seiten der Gebärdensprachforschung begegnet werden muß.

## 8. Zitierte Literatur

- Bébian, Roch-Ambroise Auguste (1825), *Mimographie ou Essai d'écriture mimique, propre à régulariser le langage des sourds-muets*. Paris: Colas
- Becker, Claudia (2001): "Gebärdensprachprozesse in der Deutschen Gebärdensprache: Zur Rolle von Komposita". In: Leuninger & Wempe (Hrsg.), 147–167
- Biesold, Horst (1988), *Klagende Hände*. Solms: Oberbiel
- Boyes Braem, Penny (1995), *Einführung in die Gebärdensprache und ihre Erforschung*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 11)
- Comrie, Bernard (1981), *Language Universals and Linguistic Typology: Syntax and Morphology*. Oxford: Blackwell
- Czech, Franz Herrmann (1836), *Versinnbildlichte Denk- und Sprachlehre, mit Anwendung auf die Religions- und Sittenlehre und auf das Leben*. Wien: Mönchianen-Congregations Buchh.
- Desloges, Pierre (1779), *Observations d'un sourd et muet, sur un cours élémentaire d'éducation des sourds et muets*, publié en 1779 par M. l'Abbé Deschamps, chapelain de l'église d'Orléans. Amsterdam, Paris: Morin
- Deuchar, Margaret (1978), "Doing without Word Order and Inflections: The Case of British Sign Language". *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 4, 78–85
- Ebbinghaus, Horst (1996), "Direktionale Verben in gebärdensprachlichen Handlungsszenarien: Zum Verhältnis von Syntax, Topographie und szenischer Gehalt gebärdensprachlicher Räume". *Das Zeichen* 38, 544–553
- Ebbinghaus, Horst & Heßmann, Jens (1991), "Gebärdensprachen Gehörloser als Gegenstand der Semiotik: Von den Grenzen der Nützlichkeit linguistischer Beschreibungsmodelle". *Das Zeichen* 15, 55–66
- Ebbinghaus, Horst & Heßmann, Jens (2000), "Leben im Kommunikationskonflikt: Zur Ungleichsprachigkeit Hörender und Gehörloser". In: Hess-Lüttich, Ernest W. B. & Schmitz, H. Walter (Hrsg.), *Botschaften verstehen: Kommunikationstheorie und Zeichenpraxis. Festschrift für Helmut Richter*. Frankfurt/M. u. a.: Peter Lang, 47–66
- Emmorey, Karen (ed.) (2003), *Perspectives on Classifier Constructions in Sign Languages*. Mahwah/NJ & London: Lawrence Erlbaum Publ.
- Engberg-Pedersen, Elisabeth (1989), "Proformes en morphologie, syntaxe et discours". In: Quertinmont & Loncke (Hrsg.), 35–52
- Engberg-Pedersen, Elisabeth (1993), *Space in Danish Sign Language: The Semantics and Morphosyntax of the Use of Space in a Visual Language*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 19)
- Engberg-Pedersen, Elisabeth & Pedersen, Annegrete (1985), "Proforms in Danish Sign Language: Their Use in Figurative Signing". In: Stokoe, William C. & Volterra, Virginia (Hrsg.), *Sign Language Research '83. Proceedings of the 3rd International Symposium, Rome, June 22–26 1983*. Silver Spring/MD: Linstok; Roma: Istituto di Psicologia CNR, 202–209
- Erlenkamp, Sonja (2000), *Syntaktische Kategorien und lexikalische Klassen: Typologische Aspekte der Deutschen Gebärdensprache*. München usw.: Lincom Europa (Linguistic Research Forum 5)
- Fischer, Renate (1995): "Zur Verschriftung der Gebärdensprachen Gehörloser". *KODIKAS/CODE. Ars semiotica* 18.1–3, 63–74
- Fischer, Renate (1996): "Historische Gebärdendixika". *Das Zeichen* 36, 142–157
- Fischer, Renate & Abel, Saskia & Borth, Brigitte & Drewes, Gesine & Lüdeke, Ulrike (1997), "<mir mußten dann die Flügel abgeschnitten werden>. Hörgeschädigte in einer Hamburger 'Heil- und Pflegeanstalt' in der ersten Hälfte des 20. Jahrhunderts". *Das Zeichen* 39, 20–33
- Fischer, Renate & Kollien, Simon & Poppendieker, Renate & Vaupel, Meike & Weinmeister, Knut (2000), *Materialien zur kontrastiven Grammatik DGS – Deutsch I: Singular-/Plural-Übereinstimmung bei Verben*. Münster etc.: LIT Verlag (BiBi-Studien 1)
- Fischer, Susan D. & Janis, Wynne (1990), "Verb Sandwiches in American Sign Language". In: Prillwitz, Siegmund & Vollhaber, Tomas (Hrsg.), *Current Trends in European Sign Language Research. Proceedings of the 3rd European Congress on Sign Language Research, Hamburg, July 26–29, 1989*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 9), 279–294
- Fischer, Susan D. & Siple, Patricia (1990, Hrsg.), *Theoretical Issues in Sign Language Research. Bd. I: Linguistics*. Chicago/IL: University of Chicago Press

- Graßhoff, Ludwig (1820), *Beitrag zur Lebens-Erlichterung der Taubstummen durch Gründung einer Taubstummengemeinde*. Berlin: Amelang
- Happ, Daniela & Hohenberger, Annette (2001), "DFG-Projekt: Vergebärdler. Phonologische und morphologische Aspekte der Sprachproduktion in Deutscher Gebärdensprache (DGS)". In: Leuninger & Wempe (Hrsg.), 217–240
- Johnston, Trevor (2000), "Die Auslan-Forschung: Vergangenheit, Gegenwart und Zukunft. Ein offener Brief an alle Auslan-Aktionäre". *Das Zeichen* 54, 640–653
- Jouison, Paul (1989), "Iconicité et double articulation dans la langue des signes". In: Quertinmont & Loncke (Hrsg.), 73–107
- Jouison, Paul (1992), "Reise um die Gebärdensprache (Teil I)", übers. von Renate Fischer und Reiner Konrad. *Das Zeichen* 19, 57–65
- Keller, Jörg (1998), *Aspekte der Raumnutzung in der Deutschen Gebärdensprache*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 34)
- Kollien, Simon (2000), "Workshop für gehörlose Gebärdensprachdozenten und Linguisten". *Das Zeichen* 54, 680–682
- Konrad, Reiner & Schwarz, Arvid & Prillwitz, Siegmund (2001): "Fachgebärdenlexika". In: Leuninger & Wempe (Hrsg.), 193–216
- Leibniz, Gottfried Wilhelm von (1717), *Collectanea etymologica, Illustrationi linguarum, veteris celticae, germanicae, gallicae, aliarumque inservientia*. Hannover: Foerster
- Leuninger, Helen & Wempe, Karin (2001, Hrsg.), *Gebärdensprachlinguistik 2000 – Theorie und Anwendung. Vorträge vom Symposium "Gebärdensprachforschung im deutschsprachigen Raum", Frankfurt a. M., 11.–13. Juni 1999*. Hamburg: Signum (Internationale Arbeiten zur Gebärdensprache und Kommunikation Gehörloser 37)
- Leven, Regina (2001): "Kriterien für die Beurteilung von Dolmetscherleistungen: Ein Leitfaden für die Vorbereitung auf Dolmetscherprüfungen". *Das Zeichen* 55, 140–153
- Liddell, Scott (2000): "Blended Spaces and Deixis in Sign Language Discourse". In: McNeill, David (Hrsg.): *Language and Gesture*. Cambridge: University Press, 331–357
- Liddell, Scott K. & Johnson, Robert E. (1989), "American Sign Language: The Phonological Base". *Sign Language Studies* 64, 195–277
- Padden, Carol (1988), *Interaction of Morphology and Syntax in American Sign Language*. New York: Garland
- Quertinmont, Sophie & Loncke, Filip (1989, Hrsg.), *Études européennes en langues des signes*. Bruxelles: EDIRSA (Collection Surdité 2)
- Reich, Carl Gottlob (1834), *Der erste Unterricht des Taubstummen mit angefügten Declinations-, Conjugationstabellen und einer Zeittafel*. Leipzig: Voß
- Reich, Carl Gottlob (1991), "Die natürliche Zeichensprache", mit einer Vorbemerkung von Renate Fischer. *Das Zeichen* 16, 135–142
- Schick, Brenda (1990), "Classifier Predicates in American Sign Language". *International Journal of Sign Linguistics* 1, 15–40
- Smith, Wayne H. (1990), "Evidence for Auxiliaries in Taiwan Sign Language". In: Fischer & Siple (Hrsg.), 211–228
- Stokoe, William C. (1960), *Sign Language Structure: An Outline of the Visual Communication System of the American Deaf*. Buffalo/NY: University of Buffalo (Studies in Linguistics, Occasional Papers 8)
- Supalla, Ted (1978), "Morphology of Verbs of Motion and Location in American Sign Language". In: Caccamise, Frank (Hrsg.), *American Sign Language in a Bilingual, Bicultural Context. Proceedings of the 2nd National Symposium on Sign Language Research and Teaching, Coronado/Cal., October 15–19, 1978*. o.O., 27–46
- Supalla, Ted (1982), *Structure and Acquisition of Verbs of Motion and Location in American Sign Language*. Ph.D.Diss. University of California, San Diego
- Supalla, Ted (1986), "The Classifier System in American Sign Language". In: Craig, Colette (Hrsg.), *Noun Classes and Categorization. Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon, October 1983*. Amsterdam/Phil.: John Benjamins, 181–214
- Supalla, Ted (1990), "Serial Verbs of Motion in American Sign Language". In: Fischer & Siple (Hrsg.), 127–152
- Supalla, Ted & Newport, Elissa (1978), "How Many Seats in a Chair? The Derivation of Nouns and Verbs in American Sign Language". In: Siple, Patricia (Hrsg.), *Understanding Language through Sign Language Research*. New York: Academic Press, 91–132

Renate Fischer, Hamburg (Deutschland)  
Regina Leven, Magdeburg (Deutschland)

## 144. Plansprachen

1. Zum Begriff "Plansprache"
2. Volapük, Esperanto und Interlingua
3. Herkunft der Morpheme
4. Wortbildung
5. Grammatische Kategorien und Paradigmen
6. Linguistische Bedeutung
7. Illustrative Texte
8. Zitierte Literatur

### 1. Zum Begriff "Plansprache"

**Plansprachen** sind nach bestimmten Kriterien bewußt geschaffene sprachliche Systeme von meist regelmäßigerer Struktur als Ethnospalten. Sie sollen vor allem der Erleichterung der internationalen sprachlichen Kommunikation dienen. Sie folgen ethnospachlichen Vorbildern kaum oder gar nicht (apriorische Systeme) oder lehnen sich an solche in mehr oder weniger starkem Maße an (aposteriose Systeme). Ihre Strukturen folgen entweder stärker genetischen oder funktionalen Gesichtspunkten (vgl. Back 1996: 883 f.; Blanke 1985: 99–110; Couturat & Leau 2001; Large 1985). Der Terminus wird in der interlinguistischen Fachliteratur zunehmend synonym für **Universalsprache** (universal language), **internationale Kunstsprache/künstliche Sprache** (international artificial language) und **Welthilfssprache** (auxiliary world language) verwendet (vgl. Blanke 1997).

Plansprachen werden von der **Interlinguistik** untersucht, deren Gegenstand einige Autoren nicht auf Plansprachen begrenzen (vgl. Blanke 1998 a: 47–58; Sakaguchi 1998: 309–322; Schubert 1989: 7–18). Bis 1973 wurden 912 Plansprachensysteme bibliographisch erfaßt (vgl. Duličenko 1990: 13). Ihre Zahl liegt inzwischen bei weit über 1000. Zahlreiche Autoren veröffentlichen ihre neuen Projekte im Internet (vgl. Becker 2001).

Die meisten Autoren woll(t)en durch ihre Projekte einen Beitrag zur Völkerverständigung und für eine demokratischere und effektivere internationale Kommunikation leisten. In Abhängigkeit von philosophischen, sprachbezogenen und auch politischen Grundauffassungen wird dieses Hauptmotiv jedoch in den einzelnen Jahrhunderten durch weitere Motive in unterschiedlichem Maße ergänzt und modifiziert. Hinzu kommt die Freude am spielerischen Umgang mit Sprache. Diese Motive (vgl. Blanke 2000: 43–52; Sakaguchi 1998: 36–53) und die angestrebten Zielgrup-

pen (z. B. gebildete Elite vs. breitere Schichten, Weltregionen vs. gesamte Welt) haben die morphologische Struktur der einzelnen Plansprachensysteme mitgeprägt.

### 2. Volapük, Esperanto und Interlingua

Nur wenige Plansprachensysteme konnten sich vom Projekt zu einer in der Praxis funktionierenden Sprache entwickeln, eine Unterscheidung, die oft übersehen wird (vgl. Blanke 2000: 51–56). Auf Volapük, Esperanto und Interlingua wird genauer eingegangen.

Volapük wurde 1879 von Johann Martin Schleyer (1831–1912) veröffentlicht und basiert vor allem auf stark modifizierter germanischer, französischer und lateinischer Lexik. Volapük verfügt über ausgeprägte Wortbildungsmöglichkeiten und eine regelmäßige Grammatik. Es ist die erste Plansprache, die in gewissem Maße in der Praxis funktioniert hat. Sie war besonders Ende des 19. Jahrhunderts bekannt. 1889 gab es weltweit ca. 400 Volapük-Klubs und 18 nationale und internationale Verbände. Es erschienen über 120 Zeitschriften, zahlreiche Lehrmittel, Original- und Übersetzungsleiteratur (vgl. Blanke 1985: 204–218, 335 f.; Haupenthal 1982). In den 20er Jahren wurden in Volapük Reformen vorgenommen (vgl. Bishop 1983). Dieses Volapük, das selbst Anfang des 21. Jahrhunderts noch einige Anhänger hat, wird hier behandelt (vgl. de Jong 1931; Schmidt 1933; Cherpillod 1995).

Interlingua, auf der Grundlage der Arbeiten der **International Auxiliary Language Association**, IALA (vgl. Blanke 1985: 167–173, 329 f.) maßgeblich von Alexander Gode (1907–1970) ausgearbeitet, erschien 1951 (vgl. Gode 1971; Gode & Blair 1971). Interlingua imitiert die romanischen Sprachen und ist ein Versuch, im Sinne von Benjamin L. Whorf das **Standard Average European** (SAE) zu modellieren (vgl. Gode 1953: 90). Interlingua hat eine relativ regelmäßige Grammatik. Seine Wortbildung orientiert sich völlig an den romanischen Vorbildsprachen. Die Sprache hat, vor allem in Europa, einige hundert Anhänger, die jährliche Treffen durchführen. In Interlingua erschienen im Jahr 2002 acht Zeitschriften. Es werden pro Jahr 10–15 Buchtitel, vor allem Lehrmittel, veröffent-

licht. Interlingua wird gelegentlich als Alternative zum Esperanto gesehen (vgl. Gopsill 1990: 107 f.).

Esperanto geht auf eine Projektskizze von Ludwig L. Zamenhof (1859–1917) zurück (vgl. Dr. Esperanto 1887) und ist eine voll entwickelte Sprache mit vorwiegend romanischer Lexik, einer produktiven Wortbildung und regelmäßiger Grammatik (vgl. Gledhill<sup>2</sup> 2000). Es verfügt über eine wachsende, sozial und fachlich differenzierte, in über 100 Ländern repräsentierte aktive Sprachgemeinschaft, die über eine Infrastruktur mit Buchproduktion (jährlich 200–250 Titel), Presse, Kulturzentren, Bildungseinrichtungen, Radiosendungen und zahlreichen Veranstaltungen verfügt (vgl. Lapenna et al. 1974; Blanke 2000: 59–69). Esperanto wird für diverse Kommunikationsbereiche verwendet und hat entsprechende Funktionalstile entwickelt (so u. a. als Umgangs-, Literatur- und Fachsprache). Diese Plansprache ist Gegenstand linguistischer Forschung (vgl. Blanke 2003; Tonkin & Fettes 1996: 9–15). Die Zahl der Esperanto-Kenner kann auf einige Hunderttausend geschätzt werden.

Die nur geringe Praxis von Volapük und von Interlingua haben kaum zu einem relevanten Sprachwandel geführt. Er beschränkt sich auf Neologismen in der Lexik. In Esperanto jedoch auf Grund seiner längeren Tradition und umfangreichen Praxis ist er konstatierbar, so u. a.: Herausbildung und Stabilisierung einer Aussprachenorm, Entlehnungen von Grundmorphemen als Basis für neue Wortbildungen, Entwicklung von Polysemie, Synonymie und (weniger) Homonymie, Entwicklung von Phraseologismen, stärkere Aus schöpfung der Wortbildungsmöglichkeiten, Entwicklung von Funktionalstilen (vgl. Philippe 1991). Die hier nur skizzenhaft mögliche morphologische Darstellung der drei Plansprachen geht von geschriebenen Texten aus.

### 3. Herkunft der Morpheme

Volapük, Esperanto und Interlingua verwenden das lateinische **Alphabet**. Volapük kennt auch die deutschen Umlaute <ä>, <ö> und <ü>. Es fehlen die **Grapheme** <q>, <w>. In Esperanto fehlen die Umlaute sowie <q>, <w>, und <y>. Es benutzt einige diakritische Zeichen: <ê> [ɛ], <î> [dʒ], <ô> [χ], <j> [ʒ], <ñ> [ʃ] und <û> [w]. In Esperanto und Volapük ist die Orthographie phonema-

tisch, in Interlingua dagegen historisch orientiert.

Die **Morpheme** der drei Plansprachen sind den europäischen Hauptsprachen entnommen und in unterschiedlichem Maße verändert worden.

Die Morpheme des Volapük stammen vor allem aus dem Englischen, Deutschen und Französischen sowie aus dem Latein. Sie wurden meist bis zur Unkenntlichkeit verändert:

- (1) engl. *knowledge* → *nol* ‘Wissen’, *speak* → *pük* ‘Sprache’; dt. *Mund* → *mud* ‘Mund’, *Dank* → *dan* ‘Dank’; lat. *animal* → *nim* ‘Tier’, *angulus* → *gul* ‘Ecke’; frz. *université* → *niver* ‘Universität’, *compliment* → *plim* ‘Kompliment’.

Die Morpheme des Esperanto der ersten zwei bis drei Jahrzehnte seiner Existenz entstammen zu ca. 75 Prozent latiniden Sprachen, zu 20 Prozent germanischen und zu 5 Prozent anderen (vgl. Janton 1993: 51). Sie wurden nach graphematischen, phonematischen und semantischen Gesichtspunkten ausgewählt und z. T. verändert (vgl. Szerdahelyi 1983: 285–290). Dabei geht es um Morpheme als Teile der synchronen Wortbildung:

- (2) lat. *augurare* → *augur-i* ‘wahrsgag-INF’, *certus* → *cert-a* ‘sicher-ADJR’; frz. *accoucher* → *akuš-i* ‘entbind-INF’, *crayon* → *krajon-o* ‘Bleistift-NR’; engl. *in spite of* → *spit-e* ‘trotz-ADVR’, *team* → *team-o* ‘Mannschaft-NR’; ital. *ancora* → *ankorau* ‘noch’, *autunno* → *aŭtun-o* ‘Herbst-NR’; dt. *Bild* → *bild-o* ‘Bild-NR’; russ. *kolbasa* → *kolbas-o* ‘Wurst-NR’, *krome* → *krom* ‘außer’.

Die Morpheme in Interlingua sind Durchschnittsformen des Italienischen, Spanischen, Portugiesischen und Französischen. Fehlt Übereinstimmung in mehr als einer dieser Sprachen, die als Kontrollsprachen gelten, so werden die Morpheme aus dem Lateinischen entnommen (vgl. Gode<sup>2</sup> 1971: XXVI). Die Äquivalente des Beispiels (1) lauten in Interlingua:

- (3) *cognoscentia* (*saper*), *lingua*, *bucca* (*ore*), *regratiamento*, *animal*, *angulo*, *universitate*, *complimento*.

Die Interlingua-Äquivalente für (2) lauten:

- (4) *auguar*, *certe*, *assister in le parto*, *stilo de graphite*, *malgrado*, *equipa*, *anque (etiam)*, *autumno*, *pictura*, *salsicia*, *foris de*.

**Wortstammvarianten** fehlen in Volapük und Esperanto, sind aber in Interlingua häufig, da

diese Plansprache das Wortbildungssystem ihrer Quellsprachen übernommen hat (z. B. zu *face-r* ‘mach<sub>1</sub>-INF (machen)’, *fact-o* ‘mach<sub>2</sub>-NR (Tatsache)’, *effect-o* ‘wirk<sub>1</sub>-NR (Wirkung)’, *effic-ace*, ‘wirk<sub>2</sub>-ADJR<sub>1</sub> (wirksam)’, *effect-ive*, ‘wirk<sub>1</sub>-ADJR<sub>2</sub> (wirklich)'). Nur in diachroner Sicht auf die Quellsprachen gibt es auch im Esperanto solche Varianten (*far-i* ‘mach<sub>1</sub>-INF (machen)’, *fakt-o* ‘mach<sub>2</sub>-NR (Tatsache)’, *efekt-o* ‘wirk<sub>1</sub>-NR (Wirkung)’, *efik-a* ‘wirk<sub>2</sub>-ADJR (wirksam)’, *efektiv-e* ‘wirk<sub>3</sub>-ADVR (wirklich)'). Sie sind synchron jedoch selbständige Basen für weitere Wortbildungen.

#### 4. Wortbildung

##### 4.1. Kompatibilität der Morpheme

Die Morpheme in Volapük und Esperanto zeichnet eine hohe **Kompatibilität** aus. Die **syntagma-internen Beziehungen** der unmittelbaren Konstituenten einer Konstruktion werden durch die Faktoren der phonologisch-morphologischen und der semantischen inneren **Valenz** gesteuert (vgl. Stepanova 1971). Das Wirken der semantischen Faktoren ist bei Volapük und Esperanto stärker möglich als bei Interlingua, deren Morphemkonstruktionen ihren Prototypsprachen folgen müssen und durch zahlreiche Ausnahmen gekennzeichnet sind. Volapük und insbesondere Esperanto zeichnen sich somit durch eine leistungsfähige autonome Wortbildung aus, die natürlich auch durch ethnopsprachige Einflüsse, Analogiebildung, die Morphemstruktur, Wirkungen der Sprachtradition und andere Faktoren beeinflußt wird. Die wichtigsten morphologischen Eigenschaften des Esperanto sind durch Unveränderbarkeit, hohe Kompatibilität und eine relativ einfache Struktur der vorwiegend romanischen Morpheme umrissen. Das unterstreicht das folgende Zitat des Begründers des Esperanto:

“Ich habe die bestehenden Begriffe vollständig *zergliedert*, so dass die ganze Sprache nur aus *unveränderlichen Wörtern* besteht, anstatt aus Wörtern mit einer unendlichen Anzahl von grammatischen Formen ... Verschiedene grammatischen Formen aber, so wie die gegenseitigen Beziehungen zwischen den Wörtern u. dergl., werden durch Anreihung unveränderlicher Wörter ausgedrückt ... er [der Lernende-DB] ahnt sogar nicht, dass das, was er Endung, Präfix oder Suffix nennt, ein völlig selbständiges Wort ist, das immer dieselbe Bedeutung behält, sei es am Anfang oder am Ende eines anderen Wortes, oder auch als selbständiges Wort; dass jedes Wort mit gleichem Rechte als Stammwort

oder als grammatisches Anhängsel gebraucht werden kann.” (Dr. Esperanto 1887: 13)

Zamenhofs Auffassung von der Zergliederung der Ideen in selbständige Elemente impliziert das Konzept vom lexikalisch-kategorialen Charakter der autosemantischen Grundmorpheme (vgl. Kalocsay & Warinighien <sup>5</sup>1985: 375–390), wonach diese substantivische, adjektivische oder verbale Grundbedeutung haben und erst durch Anfügen der wortkategorialen Suffixe auf die Wortebene gehoben werden. So hat das substantivische wortkategoriale Morphem {-o}, angefügt an das substantivische Morphem {film} (*film-o*, ‘Film-NR’), in lexikalischer Hinsicht pleonastische Bedeutung, trägt aber grammatische Information (Singular). An das verbale Grundmorphem {skrib} (*skrib-i* ‘schreib-INF’) angefügt, substantiviert es dieses (*skrib-o*, ‘Schrift’ oder ‘Schreiben’, in Abhängigkeit vom Kontext). Das bedeutet, daß in Esperanto Komposition als Wortbildungsvorfahren vorherrscht. Diese Theorie wurde durch den Bruder von Ferdinand de Saussure begründet (vgl. René de Saussure 1910; Coutrat 1907). Sie hat sich u. a. für kontrastiv-linguistische Untersuchungen (vgl. Blanke 1981) und für Zwecke der automatischen Übersetzung (vgl. Schubert 1997: 129 f.) bewährt.

##### 4.2. Einige Wortbildungstypen und -prozesse

Volapük, Esperanto und Interlingua unterscheiden sämtliche traditionellen **Wortarten**. In Esperanto können substantivische, adjektivische und verbale **Grundmorpheme** (Träger des Wortbedeutungskerns und Ausgangspunkt der Wortbildung) auf der Wortebene nur als gebundene Morpheme auftreten. Die **Autosemantika** sind morphologisch gekennzeichnet. Das betrifft in Volapük nur Adjektiv, Verb und abgeleitetes Adverb und in Interlingua ebenfalls Verb und abgeleitetes Adverb. Die folgende Tabelle macht das für drei Wortarten deutlich:

Abgeleitete Adverbien werden in Esperanto durch {e}, in Volapük durch {o} und in Interlingua durch {(a)mēte} gekennzeichnet: Esperanto *rapid-e*, Interlingua *rapidamente*, Volapük *vif-o* ‘Schnelligkeit-ADVR (schnell)’.

Da Interlingua den Wortbildungsmustern der romanischen Sprachen folgt, macht es dem romanisch Gebildeten – gemessen an seinen Sprachgewohnheiten – einen natürl-

	Substan-	Adjektive	Verben
	tive	(Inf.)	
Volapük	<i>böd</i>	<i>gret-ik</i>	<i>vob-ön</i>
	<i>plan</i>	<i>ful-ik</i>	<i>vok-ön</i>
	<i>nam</i>	<i>gol-ik</i>	<i>pön-ön</i>
Interlingua	<i>ave</i>	<i>grande</i>	<i>labora-r</i>
	<i>planta</i>	<i>plen</i>	<i>puni-r</i>
	<i>mano</i>	<i>prematur</i>	<i>voce-r</i>
Esperanto	<i>bird-o</i>	<i>grand-a</i>	<i>labor-i</i>
	<i>plant-o</i>	<i>plen-a</i>	<i>vok-i</i>
	<i>man-o</i>	<i>fru-a</i>	<i>pun-i</i>
(Deutsch)	<i>Vogel</i>	<i>groß</i>	<i>arbeiten</i>
	<i>Pflanze</i>	<i>voll</i>	<i>rufen</i>
	<i>Hand</i>	<i>früh</i>	<i>strafen</i> )

Tab. 144.1: Substantive, Adjektive und Verben in Volapük, Interlingua und Esperanto

cheren Eindruck als Esperanto und insbesondere als Volapük. Diese "Natürlichkeit" erhöht die passive Verständlichkeit der Texte, geht aber zu Lasten der Regelmäßigkeit in der Wortbildung. Diese wiederum erleichtert die aktive Beherrschung und von Vorbildsprachen unabhängige Abbildung von Begriffen, wie die folgende Tabelle für die von Substantiven abgeleiteten Adjektive zeigt:

In Esperanto ist das Hauptverfahren die **Komposition**. **Denominale Komposita** sind häufig:

- (5) *dom-o + pord-o → dom-pord-o*  
‘Haus-Tür-NR (Haustür’)

Weniger häufig sind **Deadjektiva**:

- (6) *rapid-a + trajn-o → rapid-trajn-o*  
‘schnell-Zug-NR (Eilzug’)

und **deverbale Komposita**:

- (7) *plor-i + lament-i → plor-lament-i*  
‘wein-klag-INF (wehklagen’)

Diese und auch mit Affixoiden gebildete Komposita:

- (8) *skrib-i + il-o* ‘schreib-INF + Instrument-NR (Schreibstift’); *mis-skrib-i* ‘NEG-schreib-INF (verschreiben’)

sind syntaktisch auflösbar, wie die auf (5)–(8) bezogenen Beispiele zeigen:

- (9) *pord-o de dom-o* ‘Tür-NR von Haus-NR’, *rapid-a trajn-o* ‘schnell-ADJR Zug-NR’, *plor-(ant-)e lament-i* ‘wein-(PART.AKT.PRÄS)-ADVR klag-INF’, *il-o por skrib-i* ‘Instrument-NR für schreib-INF’, *mis-e skrib-i* ‘NEG-ADVR schreib-INF’.

Esperanto verfügt traditionell über 10 Präfixe und 31 Suffixe, von denen der größte Teil als **Affixoide** angesehen werden kann, da sie auch als selbständige Basen für Wortbildungsprozesse funktionieren. Das ist in Volapük selten der Fall und fehlt in Interlingua. Konstruktionen nur aus Affixoiden sind häufig:

- (10) *estr-ar-an-o* ‘Leiter-Menge-Mitglied-NR (Vorstandsmitglied’); *re-dis-ig-i* ‘wieder-auseinander-INCH-INF (sich wieder trennen’).

Aber auch Adjektive, Präpositionen und andere Wortarten können den Status eines Affixoids erhalten, wenn sie reihenbildend sind:

- (11) *plen-a* ‘voll-ADJR’, *dolor-o* ‘schmerz-NR’ → *dolor-plen-a* ‘schmerz-voll-ADJR (schmerzvoll’), *labor-plen-a* ‘arbeit-voll-ADJR (arbeitsvoll, arbeitsreich’); *en* ‘in’, *ir-i* ‘geh-INF (gehen’ → *en-ir-i* ‘in-geh-INF (hineingehen’), *jet-i* ‘werf-INF (werfen’), *en-jet-i* ‘in-werf-INF (hineinwerfen’).

**Explizite Derivativa** entstehen durch Kombinationen von Grundmorphemen (*labor-o + temp-o → labor-temp-o* ‘arbeit-Zeit-NR (Ar-

Interlingua		Esperanto		Volapük	
<i>Nation</i>	<i>nation-a</i>	<i>naci-o</i>	<i>naci-a</i>	<i>net</i>	<i>net-ik</i>
<i>lun-a</i>	<i>lun-ar</i>	<i>lun-o</i>	<i>lun-a</i>	<i>mun</i>	<i>mun-ik</i>
<i>mar</i>	<i>mar-in</i>	<i>mar-o</i>	<i>mar-a</i>	<i>mel</i>	<i>mel-ik</i>
<i>system-a</i>	<i>system-atic</i>	<i>sistem-o</i>	<i>sistem-a</i>	<i>sit</i>	<i>sit-ik</i>
<i>fratr-e</i>	<i>frat-erne</i>	<i>frat-o</i>	<i>frat-a</i>	<i>blod</i>	<i>blod-ik</i>

Tab. 144.2: Adjektivableitungen von Substantiven in Interlingua, Esperanto und Volapük

beitszeit') oder durch Konstruktionen mit wortkat. Suffixen (*blu-a + okul-o + -a → bluokul-a* 'blau-Augen-ADJR (blauäugig)').

Die hohe Kompatibilität der Morpheme des Esperanto erlaubt oft eine weitestgehende Auffüllung von Modellen der Wortbildung. Es sei z. B. das Modell Präposition + Substantiv + faktitives Suffix *-ig-* angeführt:

- (12) *sur-tabl-ig-i* 'auf-Tisch-FAKT-INF (auf den Tisch bringen, auftischen)', *sub-tabl-ig-i* 'unter-Tisch-FAKT-INF (unter den Tisch bringen)', *sur-ter-ig-i* 'auf-Erde-FAKT-INF (auf die Erde bringen, landen)', *en-ter-ig-i* 'in-Erde-FAKT-INF (in die Erde bringen, begraben)', *el-ter-ig-i* 'aus-Erde-FAKT-INF (aus der Erde bringen, ausgraben)'.

Diese Konstruktionen sind dann wieder Basen für weitere Wortbildungen, z. B. für Adjektive und Substantive:

- (13) Adjektive: *sur-tabl-ig-a* 'auf-Tisch-FAKT-ADJR', *sub-tabl-ig-a* 'unter-Tisch-FAKT-ADJR', *sur-ter-ig-a* 'auf-Erde-FAKT-ADJR', *en-ter-ig-a* 'in-Erde-FAKT-ADJR', *el-ter-ig-a* 'aus-Erde-FAKT-ADJR';  
Substantive: *sur-tabl-ig-o* 'auf-Tisch-FAKT-NR', *sub-tabl-ig-o* 'unter-Tisch-FAKT-NR', *sur-ter-ig-o* 'auf-Erde-FAKT-NR', *en-ter-ig-o* 'in-Erde-FAKT-NR', *el-ter-ig-o* 'aus-Erde-FAKT-NR' usw.

Die Faktoren der **Sprachökonomie** bei Wortbildungsprozessen in Esperanto wirken durch die Prinzipien der **Sprachbequemlichkeit** und der **Redundanz** (s. Art. 31). Sie regeln, daß in ein Wortsyntagma alle die Morpheme einge-fügt werden, die erforderlich sind, aber nicht mehr als ausreichen, um möglichst klar und genügend redundant die Bedeutung zu motivieren, die durch das Wort bezeichnet werden soll (vgl. René de Saussure 1910: 74; Kalocsay & Waringhien<sup>5</sup> 1985: 385–388). So kann z. B. in Abhängigkeit vom Kontext 'Essen' (als Nahrungsmittel oder Handlung) durch *mang-o* 'ess-NR' bezeichnet werden. Fehlen jedoch Kontextinformationen, so ist für 'konkrete Sache' das Suffix *-af-* (*mang-af-o* 'ess-konkrete.Sache-NR') oder für 'Dauer der Tätigkeit' das Suffix *-ad-* (*mang-ad-o* 'ess-andauernde.Tätigkeit-NR') anzufügen. In Esperanto entstehen zahlreiche Synonyme oder Teilsynonyme durch das parallele Wirken von zwei Möglichkeiten der Wortschatzerweiterung. Zum einen werden komplett (meistens romanische, aber zunehmend auch

englische) Konstruktionen übernommen und in Esperanto als Grundmorpheme eingepaßt: *redaktor-o* 'Redakteur-NR (Redakteur)'. Zum anderen werden vorhandene Wortbildungsmorpheme genutzt: *redakt-ist-o* 'redigier-be-ruflich. Tägler-NR (Redakteur)'.

Volapük kennt **Komposition** und **Derivation**. Derivation überwiegt. Es gibt 56 Präfixe und 85 Suffixe, Wortbildungssuffixe und grammatische Affixe inklusive. Darunter sind 20 nur für die Fachsprache der Medizin und der Chemie vorgesehen (vgl. Schmidt 1933: 17).

Bei **denominalen Komposita** wird in die Kompositionsfuge das Genitivmorphem {a} eingefügt: *vin-a-flad* 'Wein-GEN-Flasche (Weinflasche)', *vol-a-pük* 'Welt-GEN-Sprache (NOM.SG) (Weltsprache)'. Bei denominalen Komposita bezeichnet Volapük die syntagma-internen Beziehungen der unmittelbaren Konstituenten manchmal genauer als Esperanto:

- (14) Esperanto: *hom-am-o* 'Mensch-lieb-NR (Menschenliebe, Liebe des [zum] Menschen)'. Volapük: *men-a-löf* 'Mensch-GEN-Liebe (des Menschen Liebe)', *mel-i-löf* 'Mensch-AKK-Liebe (Liebe zum Menschen)'.

In Interlingua ist das Hauptverfahren der Wortbildung die *Derivation*, die jedoch mehr eine Lehre der Etymologie der romanischen Quellsprachen ist als ein produktives Wortbildungsverfahren. Es werden 124 aktive Wortbildungssuffixe aufgeführt, 44 Präfixe und 80 Suffixe (vgl. Gode & Blair<sup>2</sup> 1971: XLVI–XLVII). Statt nominaler Komposition wie in Esperanto oder Derivation wie in Volapük, bevorzugt Interlingua syntaktische Ausdrücke:

- (15) Esperanto: *bus-staci-o* 'Bus-Station-NR (Busstation); *sun-lum-o* 'Sonne-Licht-NR (Sonnenlicht)'; *akv-o-fort-o* 'Wasser-NR-Kraft-NR (Wasserkraft)'; *akv-o-vapor-o* 'Wasser-NR-Dampf-NR (Wasserdampf)'; *sen-akv-a* 'ohne-Wasser-ADJR (wasserlos)'

Interlingua: *station de omnibus; lumine solar; fortia hydraulic; vapor de aqua; sine aqua.*

Syntaktische Ausdrücke können bei Bedarf auch im Esperanto gebildet werden, und zwar sowohl attributive als auch präpositionalen. Die Äquivalente zu (15) sind:

- (16) *bus-a staci-o* 'Bus-ADJR Station-NR', *staci-o de bus-o* 'Station-NR von Bus-

NR'; *sun-a lum-o* ‘Sonne-ADJR Licht-NR’, *lum-o de sun-o* ‘Licht-NR von Sonne-NR’; *akv-a fort-o* ‘Wasser-ADJR Kraft-NR’, *fort-o de akv-o* ‘Kraft-NR von Wasser-NR’; *akv-a vapor-o* ‘Wasser-ADJR Dampf-NR’, *vapor-o de akv-o* ‘Dampf-NR von Wasser-NR’; *sen akv-o* ‘ohne Wasser-NR’.

Mit Wortbildungsmitteln können in Volapük und Esperanto (jedoch nicht in Interlingua) besonders regelmäßig und systemdurchgängig einige Aktionsarten bezeichnet werden. Dazu folgende Beispiele:

(17) Intensiva:

Volapük (*le-*): *le-dun-ön* ‘INTS-Tat-INF (eifrig tun)’; *le-fäl-ön* ‘INTS-Fall-INF (stürzen)’. Esperanto (-*eg*-): *rid-eg-i* ‘lach-INTS-INF (schallend lachen)’; *krieg-i* ‘schrei-INTS-INF (brüllen)’.

(18) Faktitiva:

Volapük (-*iükön*): *red-iükön* ‘Röte-FAKT.INF (rot machen, röten)’; *fin-iükön* ‘Ende-FAKT.INF (beenden)’. Esperanto (-*ig*-): *ruğ-ig-i* ‘rot-FAKT-INF (röten)’, *pli-grand-ig-i* ‘KMPR-groß-FAKT-INF (vergrößern)’. In Interlingua selten: z. B. *purifica-r* ‘sauber-FAKT-INF (säubern)’, *neutral-isa-r* ‘neutral-FAKT-INF (neutralisieren)’.

(19) Kausativa:

Volapük (*a→i*; *a→ä*): *seat-ön* ‘lieg-INF’ → *seit-ön* ‘lieg:KAUS-INF (legen)’; *lag-ön* ‘häng(intr.)-INF’ → *läg-ön* ‘häng(intr.): KAUS-INF (hängen, tr.)’. Esperanto (*ig*-): *kuş-ig-i* ‘lieg-KAUS-INF (etw./jmd. hin legen)’; *pend-ig-i* ‘häng(intr.)-KAUS-INF ((auf)hängen, trans.)’.

(20) Inchoativa:

Volapük (-*ikön*): *red-ikön* ‘rot-INCH.INF (rot werden)’; *fin-ikön* ‘Ende-INCH.INF (zu Ende gehen)’. Esperanto(-*ig*-): *ruğ-ig-i* ‘rot-INCH-INF (rot werden, erröten)’; *pal-ig-i* ‘bleich-INCH-INF (erbleichen)’.

Volapük und Esperanto können die Aspekt-kategorie **Perfektiv** lexikalisch bezeichnen:

(21) Volapük (*fi*-): Infinitiv *dun-ön* ‘Tat-INF (tun)’, Perfektiv *fi-dun-ön* ‘PFV-Tat-INF (zu Ende tun, erledigen)’; *pen-ön* ‘Schreibstift-INF (schreiben)’, *fi-pen-ön* ‘PFV-Schreibstift-INF (fertig schreiben)’.

(22) Esperanto (*fin*- von *fin-i* ‘Ende-INF (beenden)’): *fin-skrib-i* ‘Ende-schreib-INF (zu Ende schreiben)’, *fin-far-i* ‘Ende-tun-INF (zu Ende tun, erledigen)’.

## 5. Grammatische Kategorien und Paradigmen

Die drei Plansprachen kennen kein **Genus**. Beim **Numerus** wird der Singular durch Nullmorphem und der Plural für alle Kasus durch ein Formativ (Volapük: {-s}, Esperanto: {-j}, Interlingua: {-s} bzw. {-es}) morphologisch bezeichnet:

(23) Volapük: *ob-ik gud-ik buk* ‘ich-ADJR gut-ADJR Buch (mein gutes Buch)’, *ob-ik gud-ik buk-s* ‘ich-ADJR gut-ADJR Buch-PL (meine guten Bücher)’. Esperanto: *mi-a bon-a libr-o* ‘ich-ADJR gut-ADJR Buch-NR’, *mi-a-j bon-a-j libr-o-j* ‘ich-ADJR-PL gut-ADJR-PL Buch-NR-PL’. Interlingua: *mi bon libro* ‘mein gut Buch’, *mi bon libro-s* ‘mein gut Buch-PL’.

In Volapük werden die **Kasus** Genitiv, Dativ und Akkusativ morphologisch bezeichnet. Es kennt außerdem den Vokativ. In Esperanto wird nur der Akkusativ durch das monofunktionale {-n} bezeichnet, das ggf. hinter dem Plural-Marker {-j} folgt: *mi vid-as mi-a-j-n bon-a-j-n amik-o-j-n* ‘ich seh-AKT.PRÄS ich-ADJR-PL-AKK gut-ADJR-PL-AKK Freund-NR-PL-AKK (ich sehe meine guten Freunde)’. Für den Nominativ kann ein Nullmorphem (*la amik-o-o* ‘ART Freund-NR-NOM’) angenommen werden. Interlingua kennt keine morphologisch bezeichneten Kasus, hat aber Varianten zur Bezeichnung der Objektformen für einige Personalpronomen (*io* ‘ich’ – *me* ‘mich’, *ille* ‘er’ – *le* ‘ihn’). Sämtliche in anderen Sprachen markierten Kasus werden in Esperanto und Interlingua durch Präpositionalphrasen wiedergegeben (Genitiv: Esperanto: *de la teknik-o* ‘von ART Technik-NR’, Interlingua: *del technica* ‘von:DEF Technik’; Dativ: Esperanto: *al la teknik-o* ‘zu ART Technik-NR’, Interlingua: *al technica* ‘zu Technik’).

Das Deklinationsparadigma für Substantive in Volapük ist regelmäßig (*fat* ‘Vater’):

	Singular	Plural
Nominativ	<i>fat</i>	<i>fat-s</i>
Genitiv	<i>fat-a</i>	<i>fat-a-s</i>
Dativ	<i>fat-e</i>	<i>fat-e-s</i>
Akkusativ	<i>fat-i</i>	<i>fat-i-s</i>
(Vokativ)	<i>o fat</i>	<i>o fat-s</i>

Tab. 144.3: Deklination von Substantiven in Volapük

		Aktiv	Passiv
Volapük	Präs.	<i>vok-öl</i>	<i>pa-vok-öl</i>
	Prät.	<i>ä-vok-öl</i>	<i>pä-vok-öl</i>
	Fut.	<i>o-vok-öl</i>	<i>po-vok-öl</i>
Esperanto	Präs.	<i>vok-ant-a</i>	<i>vok-at-a</i>
	Prät.	<i>vok-int-a</i>	<i>vok-it-a</i>
	Fut.	<i>vok-ont-a</i>	<i>vok-ot-a</i>
Interlingua	Präs.	<i>voca-n-te</i>	
	Prät.		<i>voca-t-e</i>

Tab. 144.4: Partizipien in Volapük, Esperanto und Interlingua

Nach dem gleichen Paradigma werden in Volapük auch Adjektive (wenn sie nicht unmittelbar hinter dem Substantiv stehen), der Artikel, einige Pronomen und Numeralien dekliniert.

Die **Konjugation** ist in allen drei Plansprachen regelmäßig. Für die Bildung der komplexen Zeiten werden **Partizipien** verwendet. {-öl} dient in Volapük zur allgemeinen Bezeichnung von Partizipien (Infinitiv: Volapük *vok-ön*, Esperanto *vok-i*, Interlingua *voca-r* ‘ruf-INF’):

Volapük kann auch Partizipien vom Perfekt, Plusquamperfekt und Futur II bilden, die hier nicht aufgeführt sind. Interlingua kennt nur die Partizipien Präsens Aktiv und Präteritum Passiv. Esperanto hat in Übereinstimmung mit seinen Tempi ein symmetrisches System. Interlingua übernimmt eine Reihe von Unregelmäßigkeiten aus den romanischen Sprachen, so drei Verbalstämme: auf -a- (*voca-r* ‘ruf-INF’), auf -e- (*batti-r* ‘schlag-INF’) und auf -i- (*puni-r* ‘straf-INF’). Bei e-Stämmen wird im Passivpartizip -e- zu -i- (*batti-te* ‘schlag-PART.PASS.PRÄT’).

Sowohl Volapük als auch Esperanto verfügen über die sprachlichen Mittel, Tempi in verschiedenen Modi sehr differenziert darzustellen. Dazu dienen im Esperanto die Partizipien und das Modalverb *est-i* (‘sein-INF’). Neben dem **Indikativ** kennen Esperanto und Interlingua als Modi nur noch **Imperativ** und **Konditional**. Volapük unterscheidet zusätzlich **Konjunktiv** und **Volitiv**. Volapük geht synthetisch, Esperanto und Interlingua sowohl synthetisch als auch analytisch vor. Die wichtigsten Zeitformen des Verbs im Indikativ zeigt Tab. 144.5.

In Esperanto können in den verschiedenen Zeitstufen auch komplexe Ausdrücke mit den

Futurpartizipien Aktiv (-ont-a) und Passiv (-ot-a) gebildet werden:

- (24) *mi est-as vok-ont-a* ‘ich sein-PRÄS ruf-PART.AKT.FUT-ADJR (ich bin im Begriff zu rufen)’, analog dazu: Präteritum *est-is vok-ont-a*, Futur *est-os vok-ont-a* und Konditional *est-us vok-ont-a*.

Dazu gehören die Passivformen:

- (25) *mi est-as vok-ot-a* ‘ich sein-PRÄS ruf-PART.PASS.FUT-ADJR (man ist im Begriff, mich zu rufen)’, analog dazu: Präteritum *est-is vok-ot-a*, Futur *est-os vok-ot-a* und Konditional *est-us vok-ot-a*.

Auch die Bildung synthetischer Formen durch die Anfügung des finiten Hilfsverbs *est-as* ‘sein-PRÄS’, *est-is* ‘sein-PRÄT’, *est-os* ‘sein-FUT’ an das jeweilige Partizip (-ant-a, -int-a, -ont-a bzw. -at-a, -it-a, -ot-a) ist möglich, kommt aber selten vor (*mi labor-as* ‘ich arbeit-AKT.PRÄS’ → *mi est-as labor-ant-a* → *mi labor-ant-as* ‘ich arbeit-PART.AKT.PRÄS-AKT. PRÄS (ich arbeite)’ als Verlaufsform). Analog dazu gibt es Formen im Passiv: *la leter-o skrib-at-is* ‘ART Brief-NR schreib-PART.PASS. PRÄS-AKT.PRÄT (der Brief wurde geschrieben)’ usw.

Nur in Volapük werden die Stufen der **Komparation** morphologisch bezeichnet: *glet-ik* (‘groß-ADJR’), *glet-ik-um* ‘groß-ADJR-KMPR (größer)’, *glet-ik-ün* ‘groß-ADJR-SUP (am größten)’ (Esperanto: *grand-a* ‘groß-ADJR’, *pli grand-a* ‘mehr groß-ADJR’, *plej grand-a* ‘am meisten groß-ADJR’; Interlingua: *grande*, *plus grande*, *le plus grande*).

Unter Esperantologen umstritten ist die Frage, ob die zusammengesetzten Zeiten im Passiv mit *estis -at-alestis it-a* **Aspektualität** ausdrücken oder nur temporale Bedeutung vermitteln (vgl. La Zamenhofa 1961; Lötzsch 1991).

Die in Esperanto erfolgte regelmäßige morphologische Markierung der Verben, Substantive, Adjektive und der abgeleiteten Adverbien, der grammatischen Kategorien des Verbs und des Akkusativs ermöglichen eine klare Markierung der Satzglieder und damit der Satzstrukturen. Die Satzstruktur *la ~o ~e ~as ~an ~on en la ~o* ist durch semantisch kompatible Wörter leicht auffüllbar:

- (26) *la student-o rapid-e*  
ART Student-NR schnell-ADVR  
*leg-as interes-a-n*  
les-AKT.PRÄS interessant-ADJR-AKK  
*libr-o-n en la bibliotek-o*  
Buch-NR-AKK in ART Bibliothek-NR  
'der Student liest schnell ein interessantes Buch in der Bibliothek'

	Aktiv	Passiv
Präsens:		
Volapük	<i>vok-ob</i>	<i>pa-vok-ob</i>
Esperanto	<i>mi vok-as</i> ( <i>mi est-as vok-ant-a</i> )	<i>mi est-as vok-at-a</i>
Interlingua	<i>io voca</i>	<i>io es (= esse) voca-te</i>
Präteritum:		
Volapük	<i>ä-vok-ob</i>	<i>pä-vok-ob</i>
Esperanto	<i>mi vok-is</i> ( <i>mi est-is vok-ant-a</i> )	<i>mi est-is vok-ata</i>
Interlingua	<i>io voca-va</i>	<i>io esse-va voca-te</i>
Perfekt:		
Volapük	<i>e-vok-ob</i>	<i>pe-vok-ob</i>
Esperanto	<i>mi est-as vok-int-a</i>	<i>mi est-as vok-it-a</i>
Interlingua	<i>io ha (= habe) voca-te</i>	<i>io ha essi-te voca-te</i>
Plusquamperf.:		
Volapük	<i>i-vok-ob</i>	<i>pi-vok-ob</i>
Esperanto	<i>mi est-is vok-int-a</i>	<i>mi est-is vok-it-a</i>
Interlingua	<i>io habe-va voca-te</i>	<i>io habe-va essi-te voca-te</i>
Futur I:		
Volapük	<i>o-vok-ob</i>	<i>po-vok-ob</i>
Esperanto	<i>mi vok-os</i> ( <i>mi es-tos vok-ant-a</i> )	<i>mi estos vokata</i>
Interlingua	<i>io voca-rá</i>	<i>io esse-rá voca-te</i>
Futur II:		
Volapük	<i>u-vok-ob</i>	<i>pu-vok-ob</i>
Esperanto	<i>mi est-os vok-int-a</i>	<i>mi est-os vok-it-a</i>
Interlingua	<i>io habe-rá voca-te</i>	<i>io habe-rá essi-te voca-te</i>

Tab. 144.5: Konjugationsbeispiele (Indikativ) für Volapük, Esperanto und Interlingua

Seine klare morphologische Struktur macht Esperanto geeignet für computerlinguistische Zwecke (vgl. Schubert 1997). Vergleichbare Aussagen lassen sich zu Volapük und Interlingua wegen fehlender Untersuchungen nicht machen.

Typologisch ist Volapük eine lexikalisch apriorisch-aposteriorische Mischsprache (vgl. 1) des agglutinierenden Typs. Interlingua ist lexikalisch rein romanisch und hat wie seine Quellsprachen flektierende Züge. Esperanto ist ebenfalls hochgradig agglutinierend, hat aber auch isolierende Züge. Einige Besonderheiten seiner Struktur ergeben sich aus der Spannung zwischen europäischer Herkunft des Morphemmaterials und seiner Verarbeitung, die asiatische Züge aufweist (vgl. Piron 1980: 33 f.; Nuessel 2000: 55–57).

## 6. Linguistische Bedeutung

Die wissenschaftliche Auseinandersetzung mit Plansprachen kann nicht nur helfen, neue Wege zur Lösung praktischer internationaler

Kommunikationsprobleme zu finden, sondern hat auch heuristische Bedeutung für linguistische Erkenntnisse. Das betrifft sowohl die Reflexion linguistischer Einsichten und Fakten sowie sprachphilosophischer Probleme der jeweiligen Epoche, in der die einzelnen Projekte seit dem Mittelalter entstanden sind, als auch die Möglichkeit, an (wenn auch unterschiedlich lange und intensiv funktionierenden) Plansprachen Theorien zu überprüfen und weiterzuentwickeln.

Erwähnt sei z. B. die Frage der Planbarkeit sprachlicher Strukturen und Prozesse und die Rolle des Bewußten in der Entwicklung von Sprachen, die Bewertung sprachlicher Eigenschaften (was ist „leicht“ oder „schwer“ an einer Sprache) u. a. Quasi unter experimentellen Bedingungen lassen sich Fragen des Sprachwandels an in ihrer Entstehung und Entwicklung genau datierbaren Sprachen verfolgen. Bei vielen Autoren von Plansprachen waren Aspekte der komparativen und konfrontativen Linguistik intuitiv vorhanden, oder es wurden entsprechende Untersu-

chungen durchgeführt. Das schließt auch häufig Reflexionen über die allen Sprachen gemeinsamen Züge (Universalien) ein. Für die Entstehung und Entwicklung der Terminologiewissenschaft z. B. waren plansprachliche Impulse von großer Bedeutung (vgl. Blanke 1998 b: 136). Erfahrungen mit Plansprachen haben sich auf die Entwicklung der Theorien mancher Linguisten ausgewirkt (vgl. Martinet 1991).

Eine qualitativ und quantitativ ständig wachsende interlinguistische und esperantologische Fachliteratur erleichtert die Auswertung der mit Plansprachen gewonnenen theoretischen und praktischen Erkenntnisse und Erfahrungen (vgl. Blanke 2003: 155–192).

## 7. Illustrative Texte

Von Volapük über Esperanto zu Interlingua zeigen die Texte für den mit europäischen Sprachen Vertrauten eine zunehmende Verständlichkeit. Der romanische Charakter der Morpheme nimmt deutlich zu. Der synthetische Charakter von Volapük kontrastiert mit dem analytischen von Interlingua. Esperanto weist sowohl synthetische als auch analytische Züge auf.

### 7.1. Volapük

- (27) *Vol-a-pük bin-on pük*  
 Welt-GEN-Sprache sein-3.SG.N Sprache  
*mek-av-ik*  
 Kunst-Wissenschaft-ADJR  
*sek-o pük*  
 Resultat-ADVR Sprache  
*le-din-o*  
 AUG-Sache-ADVR  
*ned-öf-ik.*  
 Bedürfnis-Eigenschaft-ADJR  
 ‘Die Weltsprache ist (eine) künstliche Sprache (die) folglich (eine) Sprache (ist, die) vor allem benötigt (wird).’ (*Sirküläpenäd lü Volapükaflens Valöpo*. Scunthorpe, 2001, 1, 1)

### 7.2. Esperanto

- (28) *La kanton-o-j est-as*  
 ART Kanton-NR-PL sein-PRÄS  
*suveren-a-j kaj rajt-as*  
 souverän-ADJR-PL und recht-AKT.PRÄS  
*re-form-i si-a-n*  
 wieder-form-INF REFL-ADJR-AKK  
*lern-ej-sistem-o-n konform-e*  
 lern-Ort-System-NR-AKK konform-ADVR

*al lok-a-j bezon-o-j.*  
 zu Ort-ADJR-PL bedürfen-NR-PL

‘Die Kantone sind souverän und haben das Recht, ihr Schulsystem zu reformieren, entsprechend (zu) den örtlichen Bedürfnissen.’ (*Monato*. Antwerpen, 2001, 22.2, 15)

## 7.3. Interlingua

- (29) *Post le trovata de un tumba pro*  
 nach DEF Fund von INDEF Grab für  
*un infante neanderthal, le*  
 INDEF Kind Neanderthal DEF  
*experto-s ha deb-ite*  
 Expert-PL hab müss-PART.PASS.PRÄT  
*revide-r lor theoria-s.*  
 revidier-INF DEF.PL Theorie-PL  
 ‘Nach dem Fund eines Grabes für ein Neanderthaler-Kind, die Experten haben gemußt, ihre Theorien zu revidieren’ (*Panorama: In Interlingua*. Beekbergen, 2000, 13.6,17)

## 8. Zitierte Literatur

Back, Otto (1996), “108. Plansprachen”. In: Goebel, Hans et al. (Hrsg.), *Kontaktlinguistik. Contact Linguistics. Linguistique de contact*. Berlin, New York: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft 12.1), 881–888

Becker, Ulrich (2001), “Interlinguistik im Internet”. In: Fiedler, Sabine & Liu, Haitao (Hrsg.), *Interlinguistikaj studioj. Interlinguistische Studien*. Festlibro por/Festschrift für Detlev Blanke. Dobříčovice: KAVA-PECH, 244–253

Bishop, Brian R. (1983), “Šanĝoj en Volapuko”. *Planlingvistiko* 2.5, 2–8

Blanke, Detlev (1981), *Plansprache und National sprache: Einige Probleme der Wortbildung des Esperanto und des Deutschen in konfrontativer Darstellung*. Berlin: Akad. d. Wiss. (Linguistische Studien 85)

Blanke, Detlev (1985), *Internationale Plansprachen*. Berlin: Akademie-Verlag (Sammlung Sprache 34)

Blanke, Detlev (1997), “The Term ‘Planned Language’”. In: Tonkin (Hrsg.), 1–20

Blanke, Detlev (1998 a), “Interlinguistik und Plansprachen”. *Sitzungsberichte der Leibniz-Sozietät* 21.2, 45–76

Blanke, Detlev (1998 b), “Terminology Science and Planned Languages”. In: Oeser, Erhard & Galinski, Christian (Hrsg.), *Eugen Wüster (1898–1977)*.

- Leben und Werk: Ein österreichischer Pionier der Informationsgesellschaft.* Vienna: TermNet, 133–168
- Blanke, Detlev (2000), “Vom Entwurf zur Sprache”. *Interface – Journal of Applied Linguistics – Tijdschrift voor toegepaste linguistiek* 15.1, 37–89
- Blanke, Detlev (2003), “Interlinguistics and Esperanto studies. Paths to the Scholarly Literature”. *Language Problems & Language Planning* 27.2, 155–192
- Cherpillod, André (1995), *Konciza Gramatiko de Volapuko: Gramat brefik Volapükä*. Courgenat: Autor
- Couturat, Louis (1907), *Étude sur la dérivation en Esperanto*. Coulommiers: Brodard
- Couturat, Louis & Leau, Léopold (³2001), *Histoire de la langue Universelle + Les nouvelles langues internationales*. Hildesheim, New York: Olms [Histoire: ¹1903, Les nouvelles: ¹1907]
- Dr. Esperanto 1887 (= Ludwig L. Zamenhof), *Internationale Sprache: Vollständiges Lehrbuch und Vorrede*. Warschau: Gebethner et Wolff
- Duličenko, Aleksandr D. (1990), *Meždunarodnye vspomogatel'nye jazyki*. Tallinn: Valgus
- Gledhill, Christopher (²2000), *The Grammar of Esperanto: A Corpus-Based Description*. München: Lincom Europa [¹1998]
- Gode, Alexander (1953), “The Case for Interlingua”. *The Scientific Monthly* 77.8, 80–90
- Gode, Alexander (²1971), *Interlingua-English: A dictionary of the International Language*. New York: Ungar [¹1951]
- Gode, Alexander & Blair, Hugh E. (²1971), *Interlingua, a Grammar of the International Language*. New York: Ungar [¹1951]
- Gopsill, Peter (1990), *International Languages: A Matter for Interlingua*. Sheffield: British Interlingua Society
- Haupenthal, Reinhard (1982), “Volapük-Bibliographie”. In: Schleyer, Johann Martin (²1982, Hrsg.), *Volapük. Die Weltsprache*. Hildesheim: Olms [¹1880]
- Janton, Pierre (1993), *Esperanto: Language, Literature, and Community* (ed. by Humphrey Tonkin). New York: State University of New York Press
- Jong, Arie de (1931), *Wörterbuch der Weltsprache: Vödabuk Volapükä*. Leiden: Brill
- Kalocsay, Kálmán & Waringhien, Gaston (⁵1985), *Plena analiza gramatiko de Esperanto*. Rotterdam: Universala Esperanto Asocio [¹1935]
- La Zamenhofa (1961) = *La Zamenhofa Esperanto: Simpozio pri atalita*. La Laguna: Stafeto
- Lapenna, Ivo & Lins, Ulrich & Carlevaro, Tazio (1974), *Esperanto en perspektivo*. Rotterdam: Universala Esperanto Asocio
- Large, Andrew (1985), *The Artificial Language Movement*. Oxford, New York: Blackwell, London: Deutsch
- Lötzsch, Ronald (1991), “Tempus, Passiv und Aspektualität im Esperanto”. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 44, 402–414
- Martinet, André (1991), “Sur quelques questions d’interlinguistique: Une interview de François Lo Jacomo et Detlev Blanke”. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 44, 675–687
- Nuessel, Frank (2000), *The Esperanto Language*. New York, Ottawa, Toronto: LEGAS
- Philippe, Benoît (1991), *Sprachwandel bei einer Plansprache am Beispiel des Esperanto*. Konstanz: Hartung-Gorre
- Piron, Claude (1980), *Esperanto: eine linguistische Standortbestimmung*. A. d. Esperanto übers. v. Günther Becker. Saarbrücken: Saarl. Esperanto-Bund [Original in Esperanto: 1977]
- Sakaguchi, Alicja (1998), *Interlinguistik. Gegenstand, Ziele, Aufgaben, Methoden*. Frankfurt/M. usw.: Peter Lang (Duisburger Arbeiten zur Sprach- und Kulturwissenschaft 36)
- Saussure, René de (1910), *La construction logique des mots en Espéranto*. Genève: Kündig
- Schmidt, Johann (1933), *Lehrbuch der Weltsprache Volapük für Deutschland und die deutschsprachigen Länder*. Frankfurt/M.: Brönnner
- Schubert, Klaus (1989), “Interlinguistics – its Aims, its Achievements, and its Place in Language Science”. In: Schubert (Hrsg.), *Interlinguistics: Aspects of the Science of Planned Languages*. Berlin, New York: Mouton de Gruyter (Trends in Linguistics. Studies and Monographs 42), 7–44
- Schubert, Klaus (1997), “Ausdruckskraft und Regelmäßigkeit: Was Esperanto für automatische Übersetzung geeignet macht”. In: Tonkin (Hrsg.), 116–139
- Stepanova, M[arija]. D. (1971), “Die ‘innere Valenz’ des Wortes und das Problem der linguistischen Wahrscheinlichkeit”. In: Helbig, Gerd (Hrsg.), *Beiträge zur Valenztheorie*. Halle: Niemeyer, 133–142
- Szerdahelyi, István (1983), “Entwicklung des Zeichensystems einer internationalen Sprache: Esperanto”. In: Fodor, István & Hagége, Claude (Hrsg.), *Language Reform: History and Future, Bd. III*. Hamburg: Buske, 277–308

Tonkin, Humphrey (1997, Hrsg.), *Esperanto, Interlinguistics, and Planned Language*. Lanham usw.: Univ. Press of America (Papers on the Center for Research and Documentation on World Language Problems 5)

Tonkin, Humphrey & Fettes, Mark (1996), *Esperanto Studies: An Overview*. Rotterdam: Universala Esperanto-Asocio (Esperanto Document 43A)

*Detlev Blanke, Berlin (Deutschland)*

## XVII. Morphologischer Wandel I: Theoretische Probleme

### Morphological change I: Fundamental issues

#### 145. Fundamental concepts

1. Introduction
2. Processes of language acquisition and language use
3. Morphological change proper
4. Morphologization
5. Morphological change due to extragrammatical forces
6. Morphology and contact-induced change
7. Conclusions
8. Uncommon abbreviations
9. References

##### 1. Introduction

For the purpose of this article **morphological change** is defined as any change that results in a different morphological system of a language.

There are a number of traditional concepts regarding morphological change, such as leveling, rule generalization, and, probably the most widespread among these, analogy. These concepts are not very satisfying. On the one hand, they show a great deal of overlap; on the other hand, these terms do not shed light on the general forces underlying the different types of morphological change since they refer to the effects rather than to the processes of change as such.

In the following discussion of the fundamental concepts regarding morphological change, the existing terminology will therefore not constitute the starting-point. Following Koefoed (1978), the view is taken that a language is never an active force in itself. The causes of change must be sought in processes of language acquisition and language use. Any change is the result of such a process acting or reacting upon a given ‘state of affairs’ in the language in question. By a language we do not only mean the grammar (or the linguistic knowledge of a native speaker), but also the output of that grammar.

In many cases a language may show characteristics, either in its rule system or in the

system’s output, that may be said to ‘trigger’ a morphological change. Examples of such triggers are for instance a rule the conditions of which are little transparent, or a set of output strings that may easily be reanalyzed. As was pointed out above, such characteristics of a language are not considered the cause of a change. Rather, we consider them only a factor leading to change.

An insightful discussion of morphological change, then, must show in which way the different processes of language acquisition and language use are triggered by specific linguistic situations, i. e. in which way the characteristics of a language and the processes of language acquisition and language use interact. The relevant processes of language acquisition and language use are discussed in section 2. As to the configurations that may give rise to morphological change, we will distinguish between (i) morphological change proper, i. e. changes originating in the morphological system and/or its output; (ii) morphologization, i. e. the incorporation of (morpho)phonological and syntactic phenomena into morphology; and (iii) morphological change due to extragrammatical forces. These will be discussed in sections 3, 4 and 5, respectively. In section 6 the effects of language contact on morphological systems will be briefly dealt with, while 7 contains the conclusions.

##### 2. Processes of language acquisition and language use

With regard to morphological change, especially the types subsumed under the label of **analogy**, it has been a long tradition to assign a prominent role to the language learning child. Leveling, overgeneralization, reinterpretation are often exclusively ascribed to the child’s ‘language acquisition device’ or to

some specific heuristic principles that form part of it. The child, having no direct access to the grammar of his parents' language, looks for patterned relations within the set of utterances he hears and may arrive at a different grammar than that of his parents' generation. In this way, language acquisition, and particularly the process of **rule discovery**, may lead to **rule change**.

This picture, with its stereotypical distinction between the language learning child and the language using adult speaker, is too simplistic. No doubt, processes of rule discovery play a role in many types of morphological change (cf. 2.1). However, it is clearly wrong to associate all processes of rule discovery or **rule creation** exclusively with language acquisition by young children. Since these activities may also be rooted in language use (cf. 2.2), rule discovery and/or rule creation may also play a role in the speech of adult speakers.

### 2.1. Language acquisition, particularly rule discovery

The first step in rule discovery is to assign a morphological structure to a given set of words. That is, to observe a formal-semantic relation between these words and another set of words, so that the first can be seen as in some way formed on the basis of the latter. The relation can then be interpreted as the product of a rule of word formation. The second step is to establish the precise nature of that rule and the conditions which determine its domain of application. In his search for morphological relations and word formation rules the learner/user is guided by what has been called the **Humboldtian principle**, that is a preference for one-to-one relations between forms and meanings. Both alternation in form and variation in meaning diminish the transparency of a rule and may trigger a change. As to the domain of application, the language learner has a preference for general, clear-cut conditions. Therefore, the transparency or 'discoverability' of a rule depends on three factors: the analyzability of the derived words, the regularity of its effects and the 'straightforwardness' of its conditions. If a rule is less than optimal in one or more of these respects, language acquisition can lead to a change. The rule in question may not only be 'missed' by the language learner, but the complex forms that derive from it may be reinterpreted, the rule may be altered in such a way that its effects become more regular,

and the conditions delimiting the rule's domain may be adjusted. In other words, rule discovery may lead to regularization. In 3.1 a number of such changes will be discussed.

It is important to see, however, that not every formal-semantic relation between words implies the existence of a morphological 'rule'. That is, not all relationships between words that language users may be aware of, are necessarily translated into rules with which new words can regularly be formed. Some of these relationships, for instance, bear upon categories that are unproductive. Besides, in order to learn a language it is not necessary to discover all the morphological relations between the existing words within the lexicon of that language. One can learn complex words as unstructured entities. As a matter of fact, some complex words must be learned in this way before the learner can start to discover morphological structure. This implies that not all morphological rules have the same function and status with regard to the linguistic knowledge of the speakers. Morphological rules may vary in necessity (from the viewpoint of language acquisition) and saliency (from the viewpoint of language use). Rules that apply very often, among them those that define the obligatory categories of a language, cannot be 'missed', of course. That is, they must be learned by all language learners in order to achieve an acceptable level of competence. Evidently, this applies to most inflectional rules. In addition, it may also apply to derivational morphology in that at least a passive knowledge is called for of those derivational rules that are frequently used to derive new words. However, at the other end of this scale one finds rules which only pertain to a handful, not too frequently used words, the occurrence of which is by and large restricted to the formal registers. As to the level of competence that a learner must acquire, these rules represent a kind of luxury; knowledge of these rules is not obligatory to obtain 'full' competence of the language involved.

### 2.2. Language use

Within the rather heterogeneous domain of 'language use', three domains of processes or activities should be distinguished:

- the psycho-physiological mechanisms of speech production and perception;
- the expression of thoughts and intentions by individual speakers by means of the

- conventional language system that they have acquired;
- (c) the establishing of linguistic conventions, essentially a process of negotiation about social contracts.

#### 2.2.1. Speech production and Speech perception

In general, the mechanisms of speech production and perception (domain a) do not directly affect the morphological system of a language. Indirectly, however, they often have repercussions on the morphological system of a language, since these mechanisms may affect the phonological shape of words. This will be briefly discussed in 5.

#### 2.2.2. ‘Looking for words’

Words, their mutual relations and their internal structure, are directly involved when people try to express their thoughts in the language they know (domain b). In expressing his thoughts, in ‘looking for words’ for what he wants to say, a speaker has essentially two sources at his disposal. First, the set of conventional names for concepts and situations (what Saussure 1916 called the *tréSOR* of the language community) and, second, the regular morphological system of his language, which by means of productive rules provides for new, non-conventional expressions. When a speaker, in expressing a concept, applies a productive rule of his language, this can be considered a manifestation of ‘rule-governed creativity’ (Chomsky 1964). The speaker does not exceed the existing possibilities of the conventional language system; his (knowledge of the) language provides for the expressions he needs, the language remains intact. However, as will be pointed out below, in looking for words a speaker may also exceed the limits of the regular system. That is, he may also turn to ‘rule-changing creativity’ (Chomsky 1964).

With regard to the existing complex words, it is important to see that within the activity to express one’s thoughts two processes are at work. The first process leads to the loss of word structure due to conventionalization and lexicalization. This will be discussed in 2.2.3. The second process leads to the systematization of the lexical stock (i. e. the set of existing complex words).

In formulating their thoughts speakers not only use their language, but constantly ‘re-examine’ it. As was already clearly observed by Von Humboldt, the lexical stock of lan-

guages should not be equated with “eine fertig daliegende Masse”, since it is “ein fortgehendes Erzeugniss und Wiedererzeugniss des wortbildenden Vermögens” (Von Humboldt 1836: 109–110). It is this activity of reflection on the set of existing expressions (the language as the speaker’s *repertoire* of linguistic expressions) that lies at the root of different types of reinterpretation (among which back formation, restructuring, folk-etymology, etc.) and adaptation. What these processes have in common is that they involve the systematization of the lexical stock.

The inspection of the lexical stock of a language by its speakers does not only lead to systematization of the existing complex words, but this activity may also underlie ‘creative coinage’. Words formed in this way exceed the limits of the regular system, meaning that they represent instances of rule-changing creativity. That is, the activity of expressing one’s thoughts and intentions does not only have repercussions on the existing words, but it may also lead to the formation of new words. Precisely because of the fact that the lexical stock of a language is not a given morphologically structured whole, the inspection of the lexical stock may result in all sorts of discoveries, which in their turn may underlie all kinds of ‘creative’ coinings. Put differently, the lexicon of a language also contains the germ of all kinds of creative coinage. Thanks to this inspection of the lexical stock, unproductive relations may be revitalized, new relations between words may be discovered, a relation between not more than two existing words may be extrapolated to other words, a rule may be ‘inverted’ so that back formations are produced, and so on.

One type of situation in which this activity is called for is the one in which the speaker’s knowledge of his language does not provide a conventional expression for the concept he wants to refer to. This situation is typical of children and that is why manifestations of creative coinage are more frequent in the speech of children than in that of adults. However, it is certainly not so that children have a monopoly on rule-changing creativity. It may also be the case that there is a referential gap in relation to the concept in question, i. e. the language does not provide conventional means to express that concept. In that case, adult speakers may also take recourse to creative coinage. Creative coinage, finally, can also be invoked because of socio-stylistic

reasons. Evidently, the desire to be original can also be the motive to coin new expressions instead of using ready-made conventional names. Like in the preceding case, creative coining because of socio-stylistic reasons presupposes relatively high level of competence, meaning that it is particularly common in the language of adults.

### 2.2.3. The establishing of linguistic conventions

As was pointed out in 2.2.2, looking for and using usual expressions (expressions that one has heard before and has remembered) may lead to **lexicalization**, through an intermediate step which we call '**conventionalization**' (domain c). Conventionalization implies that a linguistic expression, although regularly formed by means of productive rules, acquires a kind of 'currency' within the usage of the community. More and more speakers use the expression as a ready-made form, instead of coining it (or another new word) to express the concept. The essential moment in this process of conventionalization is that, in a growing number of speech situations, the expression functions as a whole. That is, its internal structure, although for most speakers still recognizable, is no longer essential for its referential function. In this respect, the regular complex expression already functions in the same way as an unanalyzable simplex word. Once conventionalized, the expression becomes susceptible to processes affecting its systematic properties. Sound changes may obscure its internal structure and the relation with the original base, while shifts in meaning may obscure its semantic transparency (of these, the sound changes originate in domain a, the shifts in meaning in extralinguistic social reality). The result will be a lexicalized word with unpredictable phonological and/or semantic properties, which has to be learned separately. Eventually, lexicalized words may lose their internal structure completely. Crucial to this development is that it is the process of conventionalization that makes it possible that processes of speech production and perception do their 'destructive' work with regard to morphological structure.

The establishing of linguistic conventions, however, is not only relevant to the conventionalization of individual complex words, but it is also relevant to the morphological system. The fact is, that morphological rules may have a specific social status as well. It is not uncommon, for instance, that a derivational

rule is only used in a specific register of the language (cf. Art. 33). Besides, whether or not a morphological rule is considered productive is not a question of its systematic properties alone, but also of its conventional status. Changes in **productivity** or in the stylistic properties of morphological categories or rules must therefore also be considered morphological changes. These changes lead to different morphological system as well.

### 2.2.4. Summary

In the preceding, two opposite processes were distinguished within the central activity to express one's thoughts and intentions. First, the inspection of the lexical stock may result in the systematization of existing words. In addition, the same activity may underlie creative coining, too. Second, the use of complex words may also result in conventionalization. This process entails that the internal structure of the words in question is no longer essential for their referential functioning. Subsequently, this may lead to lexicalization: the conventionalized words develop unpredictable semantic and formal properties.

## 3. Morphological change proper

Under this heading those types of morphological change are discussed which originate in a language's morphology and/or the words it has produced. That is, in all instances of morphological change that will be treated in this section the starting point, i. e. the configuration that triggers the change, is to be situated either in the morphological rule system or in its output, i. e. the set of existing ('conventional') complex words.

### 3.1. Changes triggered by the rule system: regularization

This class of morphological changes comprises those developments which somehow or other involve the 'streamlining' of the morphological rule system. As a general term for these changes we propose regularization. In many cases, regularization of a rule implies **generalization**, i. e. the extension of its domain of application. This, however, is not always the case since regularization may be more dramatic than the mere streamlining of the morphological rule system (cf. below).

The most obvious cases of **regularization** are those in which a given rule becomes less restricted (more general) or more transpar-

ent. Note that rule generalization may both relate to the elimination of accidental exceptions and to the elimination of systematic 'gaps' in the system. Generalization, finally, may also come into effect in the delimitation of rival categories.

An example of the elimination of an accidental exception is the coming into existence of regular preterites such as *scheerde* next to irregular ('strong') *schoor* in Dutch (cf. *scheren* 'to shave'). Irregular forms of strong verbs, so to say, run the risk of being ousted by their regular ('weak') counterparts. In many cases, these regular forms are not in common use, of course, but they may be formed any moment by the regular application of the rule of preterite formation.

The elimination of a systematic gap can be illustrated by means of the female personal names in *-e* in Dutch. When these forms became a more or less regular pattern in the present-day language, they refrained from pluralization. In the course of time, however, they gradually came under the reach of the plural system, resulting in plurals such as *studente-s* 'female students'. (Note, however, that this process has not yet been completed, meaning that there are still quite some female personal names in *-e* which do not readily underlie a plural form.)

In the case of rival categories – particularly when more than one is productive – there can be question of gradual changes in the division of labour between the categories involved, leading to a greater transparency of the system. An example of such a shift is the pluralization of simplex words ending in *-ing* in modern Dutch, examples being *haring* 'herring', *ketting* 'chain', and *koning* 'king'. In standard Dutch these words pluralize in *-en*, irrespective of the fact that the syllable in *-ing* does never bear main stress. In all kinds of non-prestigious varieties of Dutch (dialects, child language), and also in Afrikaans, however, these words have a plural in *-s*, which is completely in line with the general pattern of Dutch (and Afrikaans) plural formation, according to which words with non-final main stress generally have *-s* plurals.

Another type of regularization involves the streamlining of the semantic properties of morphological rules. An interesting example of the latter development is presented by Dutch deverbal adjectives in *-baar* '-able', which represent a relatively recent pattern in the language. During the period of their in-

roduction, adjectives in *-baar* were little transparent as far as their meaning was concerned. Specifically, they could have both active and passive readings. In the nineteenth century, however, the passive meaning became more and more dominant. In contemporary Dutch the productive rule invariably produces words with a passive meaning (cf. e.g. *converteerbaar* 'able to be converted'). It is evident that the process of 'semantic leveling' that the category in *-baar* has undergone had consequences for the way in which coining in *-baar* is conditioned in the present-day language. In modern Dutch, adjectives in *-baar* can only be regularly derived from transitive verbs (or, more precisely, from verbs with a transitive reading), meaning that the rule's domain has become more restricted.

The above examples, particularly the latter one, already made clear that regularization is a phenomenon which is more comprehensive than the mere elimination of exceptions to a given rule. That is, **regularization** is more comprehensive than **rule generalization**. In the case of the semantic streamlining of the adjectives in *-baar*, there was no question of the elimination of gaps (neither accidental nor systematic) in the rule's morphological domain. The following example, although of a completely different nature, illustrates the same fact: regularization is a much wider phenomenon than the mere elimination of gaps in the input of a rule. In Van Marle (1978) attention was drawn to the regularization of formations expressing 'to be able/to be unable' in the eastern dialects of Turkish and in closely related Azeri (or Azerbaijani). In modern Turkish the notion 'to be able' is expressed by means of compounding: the verb *bil* 'to know' is added to the required verb stem which as such is extended by the element *-el-a*, e.g. *gel-ir-im* 'come-PRES-1.SG (I come)' versus *gel-e-bil-ir-im* 'come-ø-know-PRES-1.SG (I can come)'. A remarkable feature of modern Turkish is, that the negative counterparts of the above verb forms – i.e. the forms expressing 'to be unable' – are not formed in a parallel way, as is the case in other negative verbs. As can be inferred from the following example, the so-called impotential verb does not involve compounding at all but suffixation instead. The impotential verb is formed by means of adding the suffix *-emel-ama* (to the verb stem: e.g. *gel-eme-di-m* 'come-IMPOT-PAST-1.SG (I could not come)' versus *gel-di-m* 'come-PAST-1.SG (I

have come)'. As a result, in the impotential verb the element *bil-* 'to know' is lacking. In closely related Azeri/Azerbaijani and in the adjacent Turkish dialects, this anomaly in the Turkish verb has been regularized (Cafe-roğlu & Doerfer 1959: 304). In these varieties the impotential verb is no longer characterized by a unique pattern – i. e. a pattern completely independent of the potential verb – but it is formed in a way that is completely parallel to the potential verb, e. g. impotential *gäl-ä-bil-mä-di-m* 'come-ø-can-NEG-1.SG (I have not been able to come)' which, except for the negative suffix *-mä*, is similar to potential *gäl-ä-bil-di-m* 'come-ø-can-PAST-1.SG (I have been able to come)'. Crucial to this regularized pattern is, that the impotential verb, like the potential verb, involves the compounding of *bil-* 'to know' to the verb stem.

In conclusion, regularization involves the replacement of an idiosyncrasy in the system by a form or pattern which directly conforms to the system. This process may vary from the straightforward elimination of accidental exceptions to much more dramatic developments leading to fundamental changes in the system (see also Koch 1995 for other types of regularization with far-reaching consequences).

### 3.2. Changes triggered by characteristics of existing words: systematization

Many morphological changes have their origin in the constant, never ending inspection of lexical stock from the part of speakers of the language (cf. 2). This activity of the speakers may manifest itself in essentially two ways: reinterpretation and adaptation. The common characteristic of both processes is that they both involve the systematization of the lexical stock (however limited this may be). Eventually, reinterpretation may give rise to rule creation.

#### 3.2.1. Reinterpretation

**Reinterpretation** involves the assignment of a structure to existing words which involves a change with respect to the original situation. There are many different types of reinterpretation. First, we have to do with at least three possibilities: (i) a simplex word is assigned a morphological structure, (ii) a complex word is considered simplex, and (iii) a complex word is assigned a structure that differs from its original structure. Second, the process of reinterpretation may be governed

by all kinds of additional factors which may bear upon phonology, semantics and 'use'. Consequently, the notion of reinterpretation bears upon a rather heterogeneous set of phenomena, ranging from folk etymology to affix telescoping.

A well-known instantiation of the first type of reinterpretation mentioned above, i. e. a simplex word is considered complex, is the phenomenon of **back formation** (Rückbildung). In Bloomfield (1933: 412) this phenomenon is illustrated with the example of the reinterpretation of Middle English *cheris* 'cherry' (from Old French *cherise*). The final *s* of *cheris* was associated with the formally identical plural ending, which resulted in the coming into existence of a new singular form *cherry*. In this new system, the form *cherries* – which was the original singular – functions as plural form. Similar examples are not hard to come by. In Dutch, singular forms ending in *-en*, so to say, run the risk of being reinterpreted as plural forms. The fact is, that the most frequently used plural ending in Dutch reads *-en*. In this way, *cyclamen* 'id.' became *cyclaam*, *varken* 'pig' became *vark* (in child language, also in Afrikaans), and *molen* 'mill' became *mool* (in American Dutch).

The second type of reinterpretation distinguished above, involves the loss of internal structure of a complex word. Most instances of complex words losing their internal structure involve conventionalization/lexicalization rather than reinterpretation (see 2 and 5). However, reinterpretation, too, may result in loss of structure. An example is formed by words such as *biels* '(railroad) sleeper/tie' or *rails* 'rail' in modern Dutch. For many speakers of the present-day language, these words are simplex. Diachronically, however, these words are plurals, be it plurals which did not join in with the general characteristics of Dutch plural formation. According to this system, monosyllabic words that are consonant-final take the plural ending *-en*, which means that regular plural forms are never monosyllabic. That is, in terms of the living structure of Dutch, the plurals *biels* and *rails* looked like singulars, and that is how they were reinterpreted.

The third type of reinterpretation, i. e. the type which leads to a change in the original structure of complex words, can best be illustrated with the help of the phenomenon that is known as affix telescoping. In Paul (1920: 245) it was already pointed out that

derivatives of the second degree may develop into derivatives of the first degree, in consequence of the fact that the original derivative of the second degree is directly related to the simplex base. This process results into the amalgamation of the two suffixes that are involved (Van Marle 1984; 1990 b; Haspelmath 1995). Evidently, the reinterpreted words may develop into a new derivational pattern, on the basis of which new words may be directly formed by means of the amalgamated suffix. This may be illustrated by means of Dutch female personal names in *-er-es*, e.g. *dans-er-es* 'female dancer'. Words of this type are generally looked upon as derivatives in *-es* which have neutral personal names in *-er* (*dans-er* 'dancer') as their starting point. However, the possibility should not be excluded that female personal names of the *dans-er-es* type are considered to be directly derived from the verb *dans-en* 'to dance', the result of which is the amalgamation of *-er* and *-es* into *-eres*. In this way, the formations in *-eres* become identical to their productive rivals in *-ster*, by means of which deverbal female personal names are coined. Cf. *roei-ster* 'female rower' which has the verb *roei-en* 'to row' as its base. That the formations in *-er-es* are really reinterpreted as formations in *-eres*, is proven by the fact that we find words in *-eres* for which there is no ready base in *-er* at hand. Examples are *post-eres/postbod-eres* 'female postman' and *strip-teas-eres* 'female stripper' which can best be looked upon as derivatives in *-eres* which have *post/postbode* 'postman' and *striptease* 'id.' as their starting point.

In the example just given, the reinterpretation of a derivative of the second degree as a derivative of the first degree was, to a certain extent, determined by the morphological system. In Van Marle & Koefoed (1980) instances of this type of reinterpretation were characterized as 'short cuts' in the morphological system. A derivative of the second degree is directly related to the simplex base, which results in the coming into existence of a derivative of the first degree.

Complex words may also fall victim to processes of reinterpretation of a completely different kind. They may also be reinterpreted on the basis of their 'phonological' properties, i.e. on the basis of the phoneme sequences of which they consist. This type of reinterpretation can best be characterized as a process of **resegmentation**. The prime force behind this process does not involve the

form-meaning characteristics of the words in question, but the classification (and concomitant segmentation) of these words on the basis of their shared phonological properties. Consider the following personal names in Dutch: *leugen-aar* 'liar', *molen-aar* 'miller' and *Nijmegen-aar* 'inhabitant of Nijmegen'. In Dutch neutral personal names take *-aar* when the base ends (among other things) in *en*. This pattern has given rise to a new suffix *-enaar*, which we meet in e.g. *bult-enaar* 'hunch back' and *Delft-enaar* 'inhabitant of Delft', which are based on *bult* 'hunch' and *Delft* respectively. The coming into existence of the new suffix *-enaar* is interesting in that it makes clear that complex words may be classified and segmented on the basis of their phonological properties while their morphological properties are glossed over. A salient characteristic of Dutch personal names, apparently, is that they often share the final sequence *enaar*. Originally, this 'surface characteristic' of the words in question had no morphological status, *-aar* being the suffix and *en* only being the final part of the simplex stem (without any meaning whatsoever). However, this phonological property of the personal names was assigned morphological importance, meaning that *enaar* became a suffix. This is evidenced by the fact that new words were coined by means of 'resegmented' *-enaar*.

The rise of *-enaar* in Dutch is also interesting in that it makes clear that this phenomenon of resegmentation need not be governed by syllable boundaries, as is sometimes suggested in the literature. The syllable boundary, to be sure, never lies before the schwa of *-enaar*, the syllable boundary always lies before the *n*, i.e. 'within' the suffix. That is, as far as prosody is concerned, *-enaar* is no unit. This is not to say that syllable structure is never relevant. A classic example of resegmentation which is generally considered to be governed by syllable structure, is the coming into existence of the suffix *-ling*, which used to be *-ing*. In formations such as *Karol-ing* the syllable boundary lies before or 'within' the *l* preceding *-ing*. It has often been suggested that this fact may have contributed to the coming into existence of the suffix *-ling*. Evidently, in the case of *-ling* the position of the syllable boundary may have played a part. However, Dutch *-enaar* makes clear that syllable structure is no necessary factor in the phonologically governed resegmentation of complex words, in that the initial

schwa of *-enaar* exceptionlessly forms part of another syllable than remaining *naar*. Still, the phonologically determined resegmentation of words such as *molenaar* ‘miller’ has resulted in a new suffix *-enaar*.

To conclude, Dutch *-enaar* illustrates that the phonological classification and segmentation of complex words may be governed by phonological similarity alone. As was pointed out above, it may well be that there are instances of phonologically determined resegmentation in which syllable structure plays a part, but this is certainly no general condition for this process to take place.

### 3.2.2. Adaptation

**Adaptation** is the effect of paradigmatic pressure. Individual words may undergo a change which makes them more ‘fit’ into the morphological patterns of the language as a whole. Like reinterpretation, morphological adaptation may adopt quite different shapes. The common trait of adaptive processes is that they are primarily triggered by the surface forms and that they cannot be linked to the underlying rule system in a direct way. (See 3.2.1 where the relevance of ‘surface characteristics’ came up for discussion as well.)

A well-known type of adaptation is **hypercharacterization** (Van Marle 1978). This kind of morphological change represents the formal adaptation of a given word to the regular patterns that it is semantically related to, but not formally. Consider e.g. an irregular (‘strong’) preterite such as *liep* (cf. *lopen* ‘to walk’). This irregular preterite lacks the formal characteristics of regular preterites which take *-te* or *-de* (depending on the voice quality of the stem-final segments). The adaptive change that such forms may undergo, involves the ‘superfluous’ addition of *-del-te* to the irregular preterites, resulting in ‘hypercharacterized’ forms such as *liep-te*. The latter forms are sometimes referred to as ‘double preterites’, but this characterization is somewhat misleading in that these adapted preterites are normal preterites as far as their meaning is concerned. Hypercharacterization is not only prominent in inflection – cf. also ‘double plurals’ such as Dutch *musea* ‘museums’ besides ‘correct’ *musea* – but also in derivation. Throughout the history of Dutch, personal names which are not formed by means of *-er* (*-der*), have been assigned this suffix superfluously. Cf. e.g. *dragon-der* ‘dragoon’ and *medicijn-er* ‘doctor’ (from French

*dragon* and *médecin* respectively) which represent older examples, while *aio-er*, *uhd-er* and *postdoc-er* for *aio* (an acronym based on *assistent in opleiding* ‘research assistant’), *uhd* (an acronym based on *universitair hoofddocent* ‘university lecturer’) and *post-doc* ‘idem’ represent recent adaptations.

Crucial to **morphological adaptation** is that it cannot be explained in terms of the rule system proper. Morphological adaptation can only be understood in terms of the ‘output’ of this system. That is, regular preterites are formed by means of *-del-te*. Preterites which are not formed in this way, are assigned the formal characteristics of their regular counterparts all the same. Morphological adaptation, then, is a process that is triggered by the canonical forms that the rule system produces, but not by the rule system as such.

### 3.2.3. Summary

In general, systematization (i.e. reinterpretation and adaption) contribute to the coherence of the lexical stock of a language. In addition, reinterpretation may result in new affixes which may become the formal exponent of new productive patterns (Dutch *-enaar*). In that way reinterpretation may – indirectly – contribute to the ‘lexical enrichment’ of a language.

## 3.3. Changes triggered by characteristics of inflectional paradigms

As we have seen in 3.2, morphological change need by no means be triggered by the rule system, since in many cases morphological change is rooted in properties of the output. Particularly morphological adaptation makes this clear (see 3.2.2). Another phenomenon which illustrates this point, is **paradigm leveling** (which, in contrast to morphological adaptation, is well-known and frequently discussed). The forms constituting an inflectional paradigm either tend to give up formal differences that do not correspond with differences in meaning (leveling), or tend to generalize them throughout the paradigm (extension). Well-known, for instance, is the elimination of the vowel alternation between singular and plural in the preterite of Dutch ‘strong’ verbs such as *binden* ‘to bind’. Originally, the singular had *a* as stem vowel (*bant*) vis-a-vis the *o* in the plural (*bonden*). This difference between singular and plural has been eliminated (leveled), resulting in modern Dutch *bont* (spelled *bond*) *-bonden*.

Crucial to changes of this kind is that they, like the instances of adaptation (cf. 3.2.2), bear upon the surface characteristics of the forms in question. However systematic the above discussed vowel alternation may have been, the difference at the surface between the singular and the plural constituted the trigger for the change in question. This change involves the elimination of the meaningless difference in form between the singular and the plural. Consequently, paradigm leveling represents an adaptive process: the singular forms (with *a*) have adapted their form to the plural forms (with *o*), a development which results in new singulars (with *o*).

#### 4. Morphologization

Generally speaking, morphologization involves the incorporation of originally (morpho-)phonological or syntactic phenomena into morphology. On the whole, the incorporation of extra-morphological phenomena into the morphological system involves complex changes in which all kinds of factors interact. Morphologization, finally, forms part of the more comprehensive phenomenon of grammaticalization.

##### 4.1. The incorporation of (morpho)phonological phenomena

The classic example of the morphologization of a non-morphological phenomenon is the occurrence of umlaut in cases such as *Bruder* ‘brother’ versus *Brüder* ‘brothers’ or *Vater* ‘father’ versus *Väter* ‘fathers’. As is well-known, the history of **umlaut** represents a development from a phonetic to a phonological, and from a phonological to a morphophonological process. In present-day German, umlaut is predominantly morphophonological in nature, since its occurrence is determined by morphology. In *Baum* ‘tree’ – *Bäum-e* ‘trees’ or *Mann* ‘man’ – *Männ-er* ‘men’, for instance, umlaut is concomitant to pluralization by means of *-e* and *-er*, respectively. However, in the case of *Bruder* ‘brother’ – *Brüder* ‘brothers’ or *Vater* ‘father’ – *Väter* ‘fathers’, things are different, in that in these cases umlaut is the only bearer of the notion ‘plural’. That is, in these cases umlaut is no longer concomitant to pluralization. In the words in question umlaut has become plural marker itself, meaning that it has developed into a purely morphological phenomenon.

The **morphologization** of originally morphophonological phenomena is not particu-

larly rare. Morphonological phenomena are concomitant, i.e. they are associated with a particular morphological process. If the overt marker of that process – i.e. the suffix – disappears, the concomitant phenomenon may take over the function of overt morphological marker. This also happened to plural formation in certain varieties of Spanish, particularly in the dialect of Granada. In certain nouns the notion of plurality is exclusively expressed by vowel laxing. Cf. e.g. [kaβéθa] ‘head’ versus [kɑβéθɑ] ‘heads’ or [sélvɑ] ‘forest’ versus [sélvɑ] ‘forests’ (all vowels that are marked are lax). In other nouns the notion of plural is expressed by a combination of vowel laxing and aspiration. Cf. e.g. [lóʃo] ‘wolf’ versus [lóʃøʰ] ‘wolves’ or [píso] ‘floor’ versus [píʃøʰ] ‘floors’ (all the above data stem from Alonso et al. (1950) and are cited in Hooper (1976: 36)). Like in German, in this variety of Spanish the plural ending (*-s*) has disappeared. In Spanish this happened via a process of aspiration, which in its turn functioned as the phonetic trigger of vowel laxing. Exactly like in German, the phonological phenomena which were originally concomitant to pluralization have developed into the markers of plurality when the suffix disappeared. In the case of Spanish this implies that in the words in question vowel laxing and aspiration have lost their concomitant morphonological character and have developed into true morphological phenomena.

Morphologization of originally non-morphological phenomena may come into existence in another way as well. Consider the morphologization of tone in the dialects spoken in the Dutch and Belgian Provinces of Limburg. As is pointed out in Goossens (1964), apocope of final schwa in bisyllabic words may lead to a tone shift in the first syllable. Specifically, the original dragging tones is replaced by a falling tone. In the dialects in question, this phenomenon cannot only be observed in certain plural forms, but also in adjectival declension and in a few gerunds. In some nouns a difference in tone is the only marker of plurality, e.g. *ste:\_n* ‘stone’ versus *ste:\_n* ‘stones’ where the difference in tone is the sole marker of plurality (\*-subscript indicates a dragging tone, whereas \*-superscript indicates a falling tone). The same difference can be found in the verb system. Verbs like *du:\_n* ‘to do’ and *zi:\_n* ‘to see’ have a dragging tone in the infinitive, whereas they have a falling tone in the gerund (the form used after the preposi-

tion *te*), cf. *tə du<sup>\*</sup>n* and *tə zi:<sup>\*</sup>n* (in all cases Goossens' notation has been used). Originally, these gerunds ended in *-ə*. When this ending got lost, there was a tone shift, meaning that the new monosyllabic gerunds got a falling tone. Goossens suggests, then, that in these examples it is not the case that an originally concomitant phenomenon has taken over the primary function, but that an originally non-morphological phenomenon is invoked to signal a morphological distinction, thus becoming a morphological phenomenon itself.

#### 4.2. From word to affix

It is not only (mor)phonological phenomena which can be incorporated into morphology. Syntactic phenomena may be incorporated into morphology as well. This latter development typically results in the rise of new affixes (as was pointed out in 3.2.1, reinterpretation may also result in new affixes). The prototypical form that this development adopts is that words lose their individual status and turn into 'affix-like' elements and eventually become true affixes. However, this is not the only way in which morphologization may give rise to new affixes. Another way in which new suffixes may arise is the gradual transition of the second constituent of compounds into suffixes. This development, too, may be considered the result of morphologization.

An interesting example of this gradual shift of a word into an affix is provided by the Turkish postposition *ile* 'with' which has developed into a suffix (but which still functions as a postposition as well). This change in status is reflected by the fact that the new suffix *-le* (after consonants)/*-yle* (after vowels) is, like all regular suffixes, subject to vowel harmony (but see below), whereas postpositions are not. So, in modern Turkish we have both the original construction *vapur ile* 'boat with<sub>1</sub>' (by boat) with bisyllabic *ile* and the new suffixed variant *vapur-la* 'boat-with<sub>2</sub>' (by boat) (with reduced, i.e. monosyllabic *-la*). The fact that developments like these are gradual by definition is evidenced by the fact that there are cases where the original postposition appears already as an affix, but where it is 'not yet' subject to vowel harmony. (The affix-status of *-yle/-le* is clear from its condensed form: it has lost its initial vowel, i.e. *i*, meaning that it has lost its bisyllabic character.) This is the case when it is used after a third person possessive suffix,

e.g. *kari-si-yle* 'wife-poss:3.SG-with<sub>1</sub> (with his wife)' or *omuz-u-yle* 'shoulder-poss:3.SG-with<sub>1</sub> (with his shoulder)'. Note, however, that the latter phenomenon – after a third person possessive suffix, *-yle* is invariable – is a characteristic of, as Lewis (1967: 87) calls it, 'educated pronunciation' which is neglected by many speakers of modern Turkish. In the language of these speakers, *yle* is no longer invariable after the third person possessive suffix as is evidenced by e.g. *kari-si-yla* and *omuz-u-yla* where it has become subject to vowel harmony.

Evidently, the language of the latter speakers reflects a stage that can best be considered one step further into the direction of 'true suffix' that the postposition *ile* has taken. The development of *ile* not only illustrates the formal change from word to affix, but also the incorporation of syntactic phenomena into the morphological system: a function word gradually develops into a case suffix (5.1).

The change from word to affix is also prominent in a completely different context, viz. in compounds. The second constituent of compounds may develop into a suffix-like element fairly easily. This phenomenon underlies the fact that in many languages suffixes can be found that have word-like properties. In Dutch there is a whole series of words that end in *-boer*. Cf. e.g. *groente-boer* 'greengrocer', *melk-boer* 'milkman', *patat-boer* 'seller of French fries', *vis-boer* 'fishmonger', *vodden-boer* 'old-clothes-man', etc. This element is formally identical to the word *boer* 'farmer'. However, in many of the compounds with *boer* as a second constituent, this element cannot simply be equated with the independently occurring and formally identical word. That is, in these cases there is no direct semantic link between the second constituent *-boer* in compounds and the word *boer* 'farmer'. What has happened, is that the word *boer* in compounds such as *groenteboer* and *melkboer* has developed into a suffix-like element, meaning 'seller of X'. This 'seller of' interpretation has gained a certain popularity as far as the coining of new words is concerned. In these newly coined words (e.g. *patat-boer* 'seller of French fries', *sigarenboer* 'seller of cigars', etc.) there need not be question of goods produced by farmers any longer, which means that the link between the suffix-like element *-boer* in the compounds in question and the word *boer* 'farmer' has become weaker and weaker.

Both in the case of Turkish *ile vis-a-vis -ylel-le* and in the case of Dutch *boer vis-a-vis -boer*, the original word and the newly arisen affix still form part of the language. This need not be the case, however. Not infrequently, the original word has disappeared from the language. The Dutch suffix *-loos*, for example, has no doubt developed from the adjective *loos* ‘deprived of’. However, this word no longer forms part of modern Dutch, its sole remnant (apart from some frozen expressions) is the word-like suffix *-loos*. Complex adjectives in *-loos* still display a compound-like character. This is, among other things, evidenced by its phonological characteristics: *-loos* is still a phonological word. As far as grammar is concerned, however, *-loos* is a suffix, since it does not occur independently. In other cases, the affix has undergone such drastic changes that the link with the word it has developed from, is no longer clear. In Dutch, many prefixes have also originated from independent words. This is for instance the case with the prefixes *be-* and *ver-*, both with schwa. However, for the speakers of present-day Dutch, these prefixes are no longer related to the words they have developed from, viz. *bi* and *for* respectively (in the modern language *bij* ‘with’ and *voor* ‘for’).

Morphologization, finally, is a very important process – particularly when it is followed by (or interacts with) regularization – because of its potential drastic character. **Morphologization** is the force underlying dramatic developments such as the development of postpositions into case endings (see above) or the development of auxiliaries into tense markers. Particularly in the case that such changes lead to large-scale regularization, they may result in morphological systems that are fundamentally different from those that preceded them.

It has been suggested that all affixes stem from words. The correctness of this far-reaching hypothesis concerning the origin of affixes is doubtful, however (see also 3.2.1). What has been overlooked, among other things, is that phonological phenomena may be assigned a morphological interpretation as well (also 4.1). One way in which this incorporation of purely phonological (i. e. meaningless) elements into morphology may come about, is the development of ‘phonological chunks’ into suffixes. This phenomenon is known as **secretion** (Jespersen 1922: 384–386). A particularly interesting example of

this development is presented in Jespersen (1939). In that study it is pointed out that the variation in English between forms with and without final *n* sometimes resulted in the addition of *n/en* to words, without changing their meaning. In this way, Jespersen points out, forms such as *hidden* (next to *hid*) and *olden* (next to *old*) came into being. Verbs, too, were candidates for extension by means of *en*. This extension is particularly interesting in the case of verbs derived from adjectives. Originally, these verbs were formally identical to their adjectival base, i. e. they were formed by means of conversion. In this stage, there was no formal difference between the adjective *hard* and the verb *hard* which is formed on the basis of the adjective. These deadjectival verbs could undergo *en* extension as well and in this way verbs such as *darken*, *deepen*, *harden*, *ripen* and *sharpen* came into being. Originally, the element *en* in these latter verbs was not linked to any grammatical and/or semantic value, meaning that it is completely comparable to *en* in e. g. *often* (from older *oft*) (Jespersen 1939).

However this may be, the extended deadjectival verbs became subject to reanalysis. These verbs were no longer regarded as extended verbs, but they were considered to be directly derived from the adjectival base by means of the suffix (!) *-en*. This new status of the verbs in *-en* is evidenced by the fact that the newly established derivational pattern led to new coinings. An example is the verb *biggen*. There has never been a verb *big*, which means that the verb *biggen* cannot be conceived of as an extended version of it. Put differently, the verb *biggen* can only be considered to be directly derived from the adjective *big*, meaning that the element *en* has developed into a derivational suffix *-en*. The same holds for verbs such as *cheepen*, *coarsen*, *laten*, *louden*, and *ruden*. These verbs make clear that it is not necessarily the case that all ‘bound forms’ (i. e. affixes) stem from originally free forms. As said, according to Jespersen (1939), the suffix *-en* in the de-adjectival verbs in question does not stem from a free form: *en* has never been a word. In older English *en* was only the phonological extension of the original verb. This makes clear that purely phonological entities may also be incorporated into morphology, in this case in consequence of the reinterpretation of the verbs they are part of. The net result of this reinterpretation is that a purely phonological sequence develops into a suffix.

## 5. Morphological changes due to extragrammatical forces

In this section two distinct sets of phenomena will be dealt with. In both cases the phenomena in question relate to different aspects of language use (see 2.2). First of all, changes will be discussed that relate to the indirect repercussions of speech production and speech perception on the morphological system. Subsequently, it is the effects of conventionalization that will be discussed. Clearly, these are the effects of, respectively, the first and the third domain of language use that were discerned in 2.2. (The relevance of the second domain of language use ('looking for words') for morphological change was discussed in section 2.2.2, where it was associated with the never-ending inspection of the lexical stock by the speakers of a language (also 3.2)).

### 5.1. The effects of speech production and speech perception

As was pointed out in section 2.2, the processes of speech production and speech perception particularly affect the phonological form of words. This means that they affect morphology only indirectly. In general, it is particularly the final parts of words that fall victim to these processes, the net result of which is the weakening and eventual loss of word-final elements. Since these latter elements are often (part of) suffixes, the net result is that these phonological forces are particularly devastating with respect to suffixes. In the case of complex words it is not only the word form that may be affected by these forces, but the grammatical information associated with the affected suffixes may get lost as well. Particularly in the case that suffixes do not bear stress, they easily fall victim to phonological weakening and loss. It has been suggested that the phonological factors in question may become operative more easily when the semantic content of the endings is opaque. Case systems, for instance, are considered to resist phonological weakening better when each case has a transparent semantic content and, consequently, has a neatly defined position within the case system as a whole. However this may be, even cases whose semantic characteristics are clear, may fall victim to phonological erosion when they do not bear stress. One of the reasons why the case system of a language like Turkish is so strong, is that the case endings in Turkish

are stressed. As a matter of fact, the Turkish case system is even expanding: the above discussed suffix *-ylel-le* (which stems from the postposition *ile*, see 4.2) clearly develops into the direction of a new case suffix (viz. that of instrumental/comitative).

### 5.2. The effects of conventionalization

As will be clear by now, the processes of conventionalization and lexicalization first of all affect individual words. Once formed, complex words may become part and parcel of the conventional lexical stock, which means that their internal morphological properties are no longer relevant to their functioning. As was pointed out in section 2, as to their functioning conventionalized complex words are similar to simplex words. In consequence of the 'irrelevance' of their structural properties, conventionalized complex words are an easy target for phonological and semantic forces affecting and obscuring their internal structure. However, the conventionalization of complex words may in the long run also affect the morphological system. Morphological processes may lose their productivity when their semantic properties get obscured. This may be the case when too many of the words that they have produced have become semantically opaque. This happened for instance with the Dutch deverbal adjectives in *-lijk* (e.g. *wenselijk* 'desirable', cf. *wens-en* 'to desire'). In Van Marle (1988) it is suggested that this once very productive pattern has lost its productivity in consequence of the semantic diversity that the words in *-lijk* exhibited.

Another type of change which should be discussed in this section on conventionalization, are changes in the register-dependent character of morphological processes. As was pointed out in Art. 33, many morphological processes are conventionally linked to a specific (set of) register(s). The category of deverbal adjectives in *-baar* '-able', for instance, is a typical representative of a process with a clear High-character (Van Marle 1990 a), and so is the category of deverbal abstract nouns in *-ing* (Van Haeringen 1971). In Art. 33 it was pointed out, too, that categories with a High-character tend to have morphological domains which are clearer defined (and, not infrequently, more restricted) than categories with a Low-character. This implies that changes in the register-dependent character of morphological processes may have consequences for the range of words which may

potentially serve as input for the process in question. In Van Marle (1990a) it is pointed out that this is indeed the case. The category in *-baar*, for instance, which is the successor of the category in *-lijk*, has a much more restricted domain than its predecessor. In contrast to the category in *-lijk*, the category in *-baar* is a typical exponent of the High-registers. In Afrikaans, the daughter language of Dutch which ‘copied’ Dutch derivational morphology during its standardization, the category in *-baar* lost its High-character, the consequence of which is that in Afrikaans deverbal adjectives in *-baar* can be formed much more freely than in Dutch (Van Marle 1996). Similarly, there are dialects in Dutch where coining in *-ing* is remarkably popular. In these dialects, and different from the standard language, the category in *-ing* has no clear High-character, a fact which directly fits in with the observation that in general there is no strong trend in dialects to distinguish between distinct registers. Not surprisingly, in these dialects coining in *-ing* runs much more smoothly than in the standard language.

A quite uncommon type of morphological change which also belongs to this category, finally, is the development of a morphological process into the direction of a register marker. This is what happened to Afrikaans diminutive formation. Apart from normal diminutives, Afrikaans has a whole series of a-typical **diminutives**, the occurrence of which is restricted to ‘affective usage’ (see Art. 33). Interesting about these a-typical diminutives is, that they may be based on pronouns, prepositions and verbs and that the attachment of the diminutive suffix is not associated with a change in word class. Specifically, if an a-typical diminutive is formed on the basis of a pronoun or verb, the resulting word is still a pronoun or verb. Cf. *hom* ‘him’ versus *hompie*, where the latter diminutivized form is still a pronoun. Typical of these a-typical diminutives in Afrikaans is, that they have lost their referential function, in consequence of which the development underlying their coming into existence can best be considered an instance of ‘demorphologization’.

## 6. Morphology and contact-induced change

In our above discussion of morphological change we have concentrated on monolingual situations. All forces discussed operate

language-internally, i.e. in a situation in which a given language is transferred from one generation to another without there being other languages spoken in the same community at the same time. That is, contact-induced change was left out of consideration completely. Evidently, it is beyond the scope of this article to discuss the effect of language contact on morphological systems in detail. Language contact may adopt various shapes, and the effects on the morphology of a language may vary accordingly. In this connection we confine ourselves to some brief remarks.

### 6.1. Borrowing

In 4 we discussed the fact that morphological systems may expand due to the incorporation of phonological and syntactic phenomena. Another well-known type of morphological expansion involves language contact, i.e. the incorporation of loan affixes (cf. also Art. 152). Note, however, that as a rule this type of expansion is – as far as morphology is concerned – only an indirect effect of the bilingual situation in question. More often than not, these so-called loan affixes have not been directly borrowed at all. In most cases it is not the affixes that are borrowed but the complex words that these affixes are part of. Put differently, in most cases the **borrowing** of words is a predominantly (or even exclusively) lexical process which results in an expansion of the lexicon in the first place. However, since the lexical stock – native or non-native – is subjected to a constant process of inspection from the part of the speakers of the language (see 2), the affixes figuring in complex words may be recognized and subsequently ‘isolated’ from the borrowed words in question. The implication of this is that in many cases the occurrence of loan affixes is no directly contact-induced phenomenon and neither is their ‘recognition’. What is contact-induced, is the lexical process of borrowing (lexical expansion).

In cases of long-term cultural contact, borrowing may result in a lexical stock and a morphological system in which two **strata** can be distinguished, a native and a nonnative one. This is the synchronic reflex of the contact-induced process of borrowing.

### 6.2. Morphological erosion

In 5.1 it was discussed that morphological systems may experience erosion, since the processes of speech production and speech

perception often affect the phonological form of words. As a rule, the erosion of morphological systems due to performance factors represents a process with a gradual character. However, morphological erosion may also have a contact-induced character, since certain types of language contact may lead to morphological erosion as well. This loss of morphological distinctions may affect both the (original) first language of the speakers due to language shift and language loss, and the (original) second language due to imperfect learning. Note that the decay of a morphological system due to language contact is much more abrupt than the gradual erosion of a morphological system due to the effects of speech production and speech perception. In American Dutch, for instance, both the plural suffixes of nouns and the verb endings are spontaneously dropped, simply because of the fact that the present-day speakers of American Dutch do no longer have full command of their language (Smits 1996).

Evidently, the effects of language contact on morphological systems may be quite dramatic, depending among other things on the intensity of the contact situation. Morphology, particularly inflection, appears to be highly vulnerable to language contact, since one of the most radical contact-induced developments typically affects this part of a language. In cases of intense contact, (the greater part of) morphological systems may fall victim to large-scale erosion, meaning that large parts of the original system are lost more or less completely.

## 7. Conclusions

In the preceding sections an attempt has been made to arrive at a classification of morphological changes based on the idea that any morphological change is the result of a certain type of process acting upon a specific domain or component of a language. By way of conclusion the various types of morphological change discussed so far will be briefly characterized in terms of the language domains in question and the processes of language acquisition and use that are involved.

The first two types of morphological change that were discussed are regularization and systematization. These two types of change originate in the morphological system itself, i. e. in the set of existing words, in the

relations between them, or in the rules by means of which they are formed.

Regularization bears upon those changes that are triggered by idiosyncrasies in the rule system. As its name indicates, this process is directed towards more regular and transparent morphological rules. This may result in the well-known phenomenon of **rule generalization**, i. e. the increase of a rule's domain by the elimination of its exceptions. On the other hand, the effect may be the restriction of a rule's domain as well, as a consequence of the fact that its conditions may also become more strictly delimited.

Systematization comes in two types: **reinterpretation** and **adaptation**. Reinterpretation comprises a number of changes in the morphological analysis of existing words. The factor at work is the speaker's tendency to look for transparent and clearly recognizable structure and segmentation of the words of his or her language. Well-known subtypes of reinterpretation are **back formation**, restructuring, **folk etymology**, and the isolation of affixes out of complex loan words. (In the latter case this may lead to the coming into existence of a new productive morphological rule.) **Adaptation** refers to the morphological adjustments of existing words, due to the paradigmatic pressure of semantically similar but formally different words. The best known subtypes are **leveling** and **hypercharacterization**. In the first case, a word (or class of words) loses its (their) characteristics which makes it (them) deviant from semantically similar words, in the second case a word (or class of words) acquires the formal characteristics of the category whose meaning it shares.

As said, both regularization and systematization have their roots in the morphological system. In terms of the processes involved, they are different, of course. Regularization is primarily associated with language acquisition (particularly with rule discovery), whereas systematization is primarily considered to be the outcome of language use (viz. the activity of 'looking for words'). The view that systematization is a manifestation of language use is further corroborated by the fact that the processes of reinterpretation and adaptation presuppose that the relevant systematic properties of the language have already been acquired. On a deeper level, however, regularization and systematization are

related processes. Both processes are the result of essentially the same activity: the ongoing search by language learners/language speakers for structure, patterns and regularity within the lexical stock and its underlying system. In addition, both processes tend to have a 'streamlining effect' on the domain they bear upon: regularization involves the streamlining of the rule system, whereas systematization involves the streamlining of the lexical stock.

The second group of changes that were discussed can be subsumed under the heading of morphologization. These changes do not originate in the morphological system of a language itself; rather they arise as reactions upon originally (mor)phonological or syntactic configurations in the output of a language (which, in their turn, may have come into existence due to the efficacy of performance factors). The essential process is the same as in the first group of morphological changes: the search for transparent morphological patterns. A common type of morphologization is the incorporation of (mor)phonological distinctions into morphology, which involves, among other things, the reinterpretation of a meaningless formal characteristic of a given set of words as meaningful. A second type of **morphologization** is the transition of words to affixes. Particularly well-known is the development of function words into affixes. Another change that may be mentioned here is the development of 'part of a compound' into suffix. Evidently, due to the latter developments new affixes come into existence. Note, however, that this does not necessarily imply that all affixes derive from independent words. Phonological chunks in the output may be morphologized as well, which means that new affixes may also have a phonological origin. In general, morphologization results in the growth of morphological systems. As was pointed out above, however, this does not imply that the forces underlying morphologization are fundamentally different from those in which morphological change proper is rooted. On the contrary, both morphological change proper and morphologization are largely rooted in the same source: the speaker's disposition to look for meaningful and transparent word structure. As a matter of fact, this seems to be the most fundamental force in morphology, in that it creates, maintains, extends and recreates morphological systems.

The remaining types of morphological change are due to performance factors. In general, these forces lead to the loss of morphological structure. First, the mechanisms of speech production and speech perception often bring about the loss of morphologically relevant distinctions and finally to deflection. Second, spontaneous use of language in concrete situations easily leads to the conventionalization of complex words, a phenomenon which opens the way to lexicalization. In its turn, lexicalization may result in the loss of productivity of morphological rules, due to the fact that a rule may lose its transparency when too many of the words it has produced have become lexicalized and have become opaque, either formally or semantically.

Finally, a type of change which is also due to performance factors but which does not necessarily lead to the loss of morphological structure, are changes in the conventional status of morphological rules. Since many morphological categories, particularly categories belonging to derivational morphology, are to a certain extent sensitive to stylistic conditions, various social and cultural processes may lead to a change in their conventional status. Such changes may lead both to the increase and the decrease of the morphological potential of these rules, depending on the direction that the change in conventional status has taken.

By way of conclusion, a brief remark on the overall nature of morphological change. More often than not, morphological change displays a little mechanistic nature. This is very clear in the case of systematization, but even in the case of regularization – i.e. in the case of morphological change proper! – morphological change may be remarkably idiosyncratic. This non-mechanistic and often unpredictable character of morphological change is due to two factors. First, many types of morphological change do not react upon characteristics of the morphological system as such, but upon properties of individual words, specific groups of words, or output strings. Second, the main force in morphological change itself is not a mechanistic but rather a creative process, stemming from the speaker's aptness to assign – i.e. to recognize or to create – internal structure to words.

## 8. Uncommon abbreviations

IMPOT impotential

## 9. References

- Alonso, Dámaso & Vicente, Alonso Zamora & Cannellada, María Josefa (1950), "Vocales Andaluzas". *Nueva Revista de Filología Hispánica* 4, 209–230
- Bloomfield, Leonard (1933), *Language*. London: Allen & Unwin
- Caferoğlu, Ahmet & Doerfer, Gerhard (1959), "Das Aserbeidschanische". In: Deny, Jean & Gronbech, Karel & Scheel, Helmuth & Tongan, Zeki Velidi (eds.), *Philologiae Turcicae Fundamenta*. Wiesbaden: Steiner, 280–307
- Chomsky, Noam (1964), *Current Issues in Linguistic Theory*. The Hague etc.: Mouton
- Goossens, Jan (1964), "Een tweede Limburgs gerundium". *Taal en Tongval* 16, 103–106
- Haeringen, Coenraad Bernardus van (1971), "Het achtervoegsel -ing: mogelijkheden en beperkingen". *De Nieuwe Taalgids* 64, 449–468
- Haspelmath, Martin (1995), "The Growth of Affixes in Morphological Reanalysis". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology 1994*. Dordrecht: Kluwer, 1–29
- Hooper, Joan B. (1976), *An Introduction to Natural Generative Phonology*. New York etc.: Academic Press
- Humboldt, Wilhelm von (1836), *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*, ed. by E. Wasmuth. Berlin: Schneider
- Jespersen, Otto (1922), *Language, its Nature, Development & Origin*. London: Allen & Unwin
- Jespersen, Otto (1939), "The History of a Suffix". *Acta Linguistica* 1, 48–56
- Koch, Harold (1995), "The Creation of Morphological Zeroes". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology 1994*. Dordrecht: Kluwer, 31–71
- Koefoed, Geert (1978), "Taalverandering in het licht van taalverwerving en taalgebruik". In: Koefoed, Geert & Marle, Jaap van (eds.), *Aspecten van taalverandering*. Groningen: Wolters-Noordhoff, 11–70
- Lewis, Geoffrey L. (1967), *Turkish Grammar*. Oxford: Oxford University Press
- Marle, Jaap van (1978), "Veranderingen in woordstructuur". In: Koefoed, Geert & Marle, Jaap van (eds.), *Aspecten van taalverandering*. Groningen: Wolters-Noordhoff, 127–176
- Marle, Jaap van (1984), "Morfologische veranderingen in breder perspectief". *Tijdschrift voor Nederlandse Taal- en Letterkunde* 100, 131–153
- Marle, Jaap van (1988), "On the Role of Semantics in Productivity Change". In: Booij, Geert & Marle, Jaap van (eds.), *Yearbook of Morphology 1988*. Dordrecht: Foris, 139–154
- Marle, Jaap van (1990a), "A Case of Morphological Elaboration: the History of Dutch -baar". *Folia Linguistica Historica* 9, 213–234
- Marle, Jaap van (1990b), "Rule-creating Creativity: Analogy as a Synchronic Morphological Process". In: Dressler, Wolfgang U. & Luschützky, Hans C. & Pfeiffer, Oskar E. & Rennison, John R. (eds.), *Contemporary Morphology*. Berlin, New York: Mouton de Gruyter, 267–273
- Marle, Jaap van (1996), "On the Interplay of Inherited and Non-inherited Features in Afrikaans Derivational Morphology". In: Nielsen, Hans F. & Schøsler, Lene (eds.), *The Origins and Development of Emigrant Languages*. Odense: Odense University Press, 103–115 (RASK Supplement Vol. 6, NO-WELE Supplement Vol. 17)
- Marle, Jaap van & Koefoed, Geert (1980), "Humboldtiaanse taalveranderingen, morfologie en de creativiteit van taal". *Spektator* 10, 111–147
- Paul, Hermann (51920), *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer [81970]
- Saussure, Ferdinand de (1916), *Cours de linguistique générale*. Paris: Payot
- Smits, Caroline (1996), *Disintegration of Inflection: The Case of Iowa Dutch*. The Hague: HAG (HIL Dissertations 22)
- Geert Koefoed, Utrecht (The Netherlands)  
Jaap van Marle, Heerlen (The Netherlands)

## 146. Grammaticalization: from syntax to morphology

1. Introduction
2. An overview of grammaticalization
3. The process of grammaticalization
4. Uncommon Abbreviations
5. References

### 1. Introduction

#### 1.1. History of grammaticalization studies

That bound morphological formatives often have their ultimate origin in independent lexical items has been a commonplace observation since the early 19th century (cf. Bopp 1816; Humboldt 1825). The phenomenon was not of primary interest to historical linguists more concerned with morphological constructions which could be reconstructed for a proto-language than with secondary developments, but its importance is explicitly recognized in discussions of general principles of diachronic linguistics such as Whitney (1875), Paul (¹1920 [¹1880]), and von der Gabelentz (1891).

Meillet (1912) first applied the term **grammaticalization** to the process by which lexical items enter into the grammatical system (“le passage d'un mot autonome au rôle d'élément grammatical”), a process whose endpoint is the development of new morphological constructions. Already in this discussion Meillet anticipates one of the most significant facts about the process, viz. that it is a gradual process rather than a sudden categorial shift. He distinguishes four degrees of grammaticalization of the French copula, from its lexical use in equational sentences (*je suis celui qui suis*) to its use as a tense auxiliary (*je suis parti*), and points out the ambiguous category of French *faire*, which is both a lexical verb ‘do, make’ and a causative:

“laissez peut être un mot principal, dans *laissez cela* par exemple; mais ici [in *laissez venir à moi les petits enfants*] *laissez venir* forme un ensemble, où *laissez* est, en quelque mesure, un auxiliaire.” (Meillet 1912 [1926: 134])

Since Meillet's recognition and naming of grammaticalization as a distinct phenomenon worthy of study, the topic has attracted the attention of a few scholars, notably Kurylowicz, who defined it in similar terms:

“Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more

grammatical status, e.g. from a derivative formant to an inflectional one.” (Kurylowicz 1965: 69)

However, the phenomena of grammaticalization have not been of central interest to most scholars of historical or synchronic linguistics, and it is only since about 1970 that it has begun to be systematically studied; it became the subject of sustained cross-linguistic investigation by the scholarly community only in the 1980's. Important contemporary works and collections include Givón (1979), C. Lehmann (1986; 1995), Heine & Reh (1984), Heine et al. (1991), Heine (1993; 1997), Hopper & Traugott (1993), Traugott & Heine (1991, eds.), Rissanen et al. (1997, eds.), Ramat & Hopper (1998, eds.), and Fischer et al. (2000, eds.).

The word *grammaticalization* (*grammaticalization* and *grammatization* are also used in the same sense) implies a process of becoming “grammatical”. The reference can be taken as being to lexical morphemes becoming grammatical ones, or, more broadly, to any linguistic construct (a morpheme, syntactic construction, or discourse pattern) becoming part of the grammatical system of a language. Recently scholars have begun to use the term to refer to shifts from more pragmatic to more grammatical function of syntactic constructions, e.g. the development of subject from topic constructions (see e.g. Givón 1979; 1989).

Grammaticalization involves changes in each of the three basic areas of linguistic structure: semantics/pragmatics, morphology/syntax, and phonology. The shift of a lexical form to a grammatical function involves, first, some shift in its semantic and/or pragmatic function. This is the necessary precondition for a shift in syntactic category, a reanalysis of the syntactic construction. These two shifts in turn result in destressing of the grammaticalized morpheme, resulting in phonological reduction and cliticization, which in turn can lead to morphologization. The process is essentially unidirectional (Haspelmath 1999); the development of lexical from grammatical forms, while attested, is rare. Two fundamental questions outline the topic of grammaticalization: when, how, and why does a lexical form grammaticalize, and what specific types of grammatical formative develop from what specific types of lexical item?

### 1.2. Grammatical and lexical meaning

The phenomenon of grammaticalization has important implications for the traditional notion of “lexical” vs. “grammatical” morphemes. On the one hand, the traditional conception of grammaticalization as the passage of a form from the first to the second of these categories relies crucially on this distinction. On the other, the actual phenomena which we discover in studying cases of grammaticalization show that, like many other dichotomous categorizations in linguistics, this distinction is in fact a gradient rather than a clear split, and that forms must be thought of as more or less grammatical rather than simply as grammatical or lexical.

While it is easy to find unambiguously lexical and unambiguously grammatical morphemes, it is notoriously difficult to draw a clear dividing line. Thus prepositions and subordinating conjunctions are more lexical than case inflections, but less lexical, and more likely to have syntactic as opposed to lexical value, than ordinary nouns and verbs. Grammaticalization research shows that historical change is almost always from more lexical to more grammatical status. All such change is grammaticalization, the shift of a relational noun or serialized verb into an adposition just as much as the shift of a adposition to a case desinence.

The gradual nature of grammaticalization also calls into question the common conception of purely grammatical meaning. There is a strong correlation between the semantics of lexical items and their potential for grammaticalization – for example, future tense constructions develop from verbs originally meaning ‘want’ or ‘go’. If the meanings of grammatical forms derive by regular processes from lexical meaning, then grammaticalization provides a key to our understanding of the notion of grammatical meaning (Bybee 1988; Sweetser 1988; Heine et al. 1991; Heine 1993). The fact that not only do grammatical morphemes develop from lexical morphemes, but that specific types of grammatical morpheme regularly develop from lexical forms with particular meanings, suggests that grammatical meaning must to some degree have the same sort of semantic content as lexical meaning, rather than being purely structural.

### 1.3. Theoretical significance of grammaticalization studies

Already in 1912 Meillet points out that grammaticalization, though equally as important as analogy in the development of new gram-

matical forms, had received much less attention in historical linguistics. From a modern perspective, informed by knowledge of a wider range of languages, we can assert that grammaticalization is in fact much more important than analogy; nevertheless until recently it has continued to receive less attention than it merits in historical and general linguistics.

One reason for this neglect is that the facts of grammaticalization pose a serious challenge to a fundamental aspect of structuralist synchronic analysis, an aspect which remains fundamental to omit much contemporary work in the generative paradigm. A structuralist or generative model consists of an inventory of syntactic categories and a set of rules for combining them into larger structures. The initial problem posed by examples such as Meillet’s is that in such data we seem to find one and the same morpheme as a member of two or more categories (e.g. in Meillet’s example as both copula and tense auxiliary). On further examination of such data the problem worsens, as it often is difficult or impossible, at least on any non-ad hoc basis, to assign some uses of an etymon to one or another category.

Grammaticalization sensu strictu involves the shift of morphemes from one form class to another, and often involves the innovation of a new form class. Thus, in grammaticalization we find evidence bearing on the nature of morphosyntactic categories and their place in the organization of grammar. We regularly find cases in which grammaticalizing forms occupy intermediate categorial status. For example, in modern English we find a range of more and less verb-like characteristics among grammaticalizing forms such as *used to*, *want to*, *ought to*, etc. (Bolinger 1980). Such data call into serious question the adequacy of any model of linguistic structure which takes the notion of discrete, clearly defined morphosyntactic categories as a theoretical given.

Another traditional distinction which requires reevaluation in the light of modern studies of grammaticalization is the opposition of **synchronic** and **diachronic analysis**. Some scholars (especially Hopper 1987; 1991; Givón 1989), are now suggesting that our traditional notion of a static synchronic “state” of a language in which every morpheme and construction can be unambiguously assigned a categorial place in the grammar is not only an idealization, but an unre-

alistic one, and that viewing grammar in terms of the fluid categories and indeterminately syntacticized constructions which are the stuff of the diachronic study of grammaticalization provides a more adequate basis for understanding the “synchronic” structure as well as the diachronic changes in language.

## 2. An overview of grammaticalization

### 2.1. Grammaticalization exemplified

The study of grammaticalization reveals recurrent patterns for the origin of particular grammatical structures. For example, causative morphemes regularly develop from serialized or complement-taking verbs with meanings like ‘make’, ‘give’, or ‘send (on an errand)’. Dative case markers originate in ‘give’ verbs or in locative/allative markers, and locative and allative markers derive diachronically from relational nouns or from verbs of location, position, and motion. One important aspect of ongoing research on grammaticalization is the cataloguing of typical sources for various grammatical constructions, which show common patterns throughout the world. As a typical example, we may consider the origin of grammatical tense categories. While Bopp is still criticized for his promiscuous identification of Sanskrit conjugational endings as grammaticalized copulas, the mechanism which he suggested as the origin of verbal desinences in Sanskrit is one which is widely attested in languages of the world (e.g. Givón 1971; Haas 1977; Heine & Reh 1984). A well-known example is the development of the modern French, Spanish, and Italian synthetic future conjugations from fusion of auxiliary *habere* ‘have’ with the infinitive, so that e.g. Fr. *chanterai* ‘I will sing’ reflects a Vulgar Latin morphologization of an earlier infinitive + auxiliary construction *cantare habeo* (cf. Benveniste 1968; Hopper & Traugott 1993). While in the world’s languages we can find many tense affixes for which no etymology is recoverable from available data, and there are occasional examples of tense morphology with other origins (e.g. in morphologized adverbs; cf. Art. 110), in the vast majority of cases, if the origin of morphological tense markers can be traced, they will be found to originate in auxiliary verb constructions of some sort.

Auxiliary verbs, in turn, represent grammaticalizations of originally biclausal syntac-

tic constructions, in which the potential auxiliary is a finite verb with a complement clause, or the highest verb in a **serialization** construction (cf. Art. 78). We can illustrate the sequence with a grammaticalization series from modern Central Tibetan (DeLancey 1991; 1997). Tibetan, like many verb-final languages, makes abundant use of a **clause-chaining** structure in which only the last of a sequence of clauses has tense/aspect/evidentiality marking. Preceding clauses are marked with a subordinator, here glossed ‘NF’ for “non-final”, which functions only to mark them as non-final clauses in a chain. It is common for two chained clauses to share all arguments, with the result that their verbs occur in sequence, separated only by the ‘NF’ morpheme:

- (1) *khos kha=lags zas-byas phyin-song*  
he:ERG meal ate-NF went-PF  
'He ate and left.'

There is a small set of verbs which can occur as the second member of such a sequence without intervening ‘NF’ marking. These thus constitute a distinct syntactic subcategory, which we can categorize as auxiliaries. The difference between the two constructions can be illustrated using the same etymon, *bzhag*, which as a lexical verb means ‘put’, and as a grammaticalized auxiliary forms a perfect construction:

- (2) *kho phyin-byas bzhag-pa\_red*  
he went-NF put-PF  
'He went and put it there.'
- (3) *kho phyin bzhag-pa\_red*  
he went put-PF  
'He has gone.'

In the auxiliary construction (3), *bzhag* is still clearly the main verb; it carries full lexical tone, and takes tense/aspect marking. The same etymon occurs also in an even more grammaticalized construction, as a suffix marking inferential perfect:

- (4) *kho phyin-zhag*  
he went-INF.PF  
'He has left (I infer).'

Here the lexical verb *phyin* is the main verb. *Zhag* is unstressed and phonologically reduced, and represents still another morphosyntactic category, that of tense/aspect/evidentiality suffix.

## 2.2. Stages of grammaticalization

Discussion of processes of historical change traditionally present a series of “stages” leading from an initial to a final state, with the implication that these stages are themselves distinct states of linguistic structure. This can be a useful idealization, as long as it is recognized as an idealization of what is in fact a continuous phenomenon (cf. C. Lehmann 1985, which develops a scale for assessing the degree of grammaticalization of a morpheme). But the overall process of grammaticalization is perhaps better conceptualized in terms of a number of processes, some of which facilitate, precondition, or promote others, but which do not necessarily proceed strictly serially.

The starting point of the process is a productive syntactic construction: noun phrase with genitive dependent, matrix with complement clause, conjoined or chained clauses, etc. The precondition for grammaticalization is that there be some lexeme or lexemes which occur frequently in this construction for some semantic/pragmatic reason: potential relational nouns like ‘top’, ‘face’, ‘back’, regularly used to provide further locational specification in noun phrases used as locatives, phasal or other complement-taking verbs like ‘finish’ or ‘want’, semantically non-specific transitive verbs like ‘use’ or ‘hold’ conjoined or serialized with more specific verbs. This usually involves a lexeme with a very general meaning, which can therefore be used in a wide range of contexts. Thus a verb meaning ‘finish’ is much more likely to metamorphose into an aspectual marker, and thus potentially to enter the morphological system, than one meaning ‘buy’ or ‘repair’, which are initially usable in a much smaller range of semantic/pragmatic contexts.

This situation, in which a particular construction – a productive syntactic structure with a specific lexeme in a specific slot – is a useful and regularly-used locution in the language, i.e. where speakers regularly refer to ‘the face of [noun phrase]’, ‘finish [verb phrase]’, etc., is the initial point of grammaticalization. We can refer to this situation as “functional specialization” of the construction. The next step is for the lexeme to undergo a certain amount of **semantic bleaching**, to use Givón’s term, or, put another way, for the locution to be used in an extended set of contexts, including some in which the literal meaning of the lexeme is not applicable:

“... without an external change of its exponent a category may undergo important internal (functional) changes due simply to an extension of a limitation of its range. The logical principle of the mutual relation of range and content has to be applied in such a case: the *increase of the range* of a given category entails the impoverishment of its content, and vice versa.” (Kuryłowicz 1965: 57f.)

This can often be observed even in forms which have not undergone any formal grammaticalization. For example, *finish* in English is often used to mean simply ‘stop’, so that one can say *I’ve finished writing for today*, even when the project on which the speaker is working is far from completed. A more grammaticalized example is the Tibetan verb *sdad* ‘sit’, which is currently developing into a progressive auxiliary. It can now occur in sentences such as (5):

- (5) *kho rgyugs=shar\_slod(-ni) sdad-zhag*  
 he run(-NF) sit-PF  
 ‘He was running/kept running/was always running.’

The semantic incompatibility of ‘run’ and ‘sit’ is such that *sdad* cannot possibly be interpreted with its original lexical sense here. The structure thus cannot be a simple case of verb serialization, but represents an early stage of grammaticalization.

The next stage of grammaticalization occurs when this lexeme begins to “decategorialize” (Heine et al. 1991; for further examples see Matisoff 1991), i.e. to lose the morpho-syntactic characteristics of its original category. (This is a broad notion which includes more specific concepts such as C. Lehmann’s (1985: 307) “morphological degeneration ... the loss of ability to inflect”). For example, in the English construction *on top of* [noun phrase], *top*, while clearly a noun in origin, is un-nounlike in several respects. It lacks an article, and it cannot pluralize: we can say *on top of all the houses*, with *top* as a relational noun, or *on the tops of all the houses*, with *top* as an ordinary noun, but we cannot pluralize the relational noun: \**on tops of all the houses*.

Note that *top* in its relational noun use is already semantically bleached, in that *the top of* [noun phrase] necessarily refers to a specific part of the object, while *on top of* [noun phrase] simply refers to whatever side of it is uppermost at the moment. (E.g. if a refrigerator is lying on its back, something resting on the door, which is the uppermost surface, is *on top of the refrigerator*, but is not *on the*

*top of the refrigerator*). This is typical, and apparently universal; it is easy to find lexemes, like English *finish*, which have undergone some semantic bleaching without any morphosyntactic decategorialization, but decategorialization does not occur without some prior functional shift, toward either grammaticalization or lexicalization.

From this point we have a continuous process of further decategorialization and phonological reduction. One possible outcome is **recategorialization** (cp. C. Lehmann's 1985 "paradigmaticization"). There may already exist a morphosyntactic category into which the grammaticalizing form will fit and which it can enter. For example, the English preposition *atop* represents one endpoint of the grammaticalization of a noun 'top'. Or, if similar changes are occurring involving several functionally related morphemes, they may become a new paradigmatic category – a famous example is the English modals.

The final stages of grammaticalization are **cliticization** and **morphologization**, in which a grammaticalized form becomes increasingly bound phonologically and syntactically to a lexical head. While outlines of grammaticalization sometimes present this as a necessary final step, it is not; grammaticalized forms may remain as auxiliaries or "particles", never becoming bound. The exact conditions which determine whether and to what extent morphologization will occur are as yet not well-understood, though they appear to have to do with stress patterns and with word-order typology – morphologization appears to be more common in verb-final than in verb-medial languages, for example. (For some discussion see W. Lehmann 1973; Donegan & Stampe 1983.)

### 2.3. The cycle

The best-known version of the idea of discrete ordered stages in historical change is the hypothesis of the "**morphology-syntax cycle**". The traditional version describes an essentially mechanical sequence. Morphology is by its nature subject to phonological erosion, which over time reduces the distinctiveness of affixes to the point where essential distinctions are lost. In the face of this languages resort to grammaticalization as a therapeutic measure, creating new periphrastic grammatical constructions to replace lost morphological forms. Since grammatical for-

matives, whether free or bound, do not normally carry accent, newly-grammaticalized forms are now subject to phonological reduction, and over time cliticize, and eventually develop into new morphological constructions. These now are by nature subject to phonological erosion, and over time the new formation loses distinctiveness, and a new cycle takes place.

It has sometimes been suggested that this is a typological cycle, i.e. that every language passes through successive analytic and synthetic stages (cf. Hodge 1970). This is certainly open to doubt, and in any case would be impossible to document for most languages and families. A more plausible claim is that the cycle operates at the level not of whole-language typology, but of the instantiation of particular functions, so that, for example, tense marking, or even more specifically the expression of a particular tense category, in a language will alternate between periphrastic and morphological encoding. Examples at this level are easier to find; consider for example the contemporary competition of the French morphological future – itself the end-point of grammaticalization of an earlier periphrastic construction with 'have' – with a new periphrastic future in *aller* 'go': *je chanterai* vs. *je vais chanter* 'I will sing'.

While there is no doubt that the cycle is a descriptively useful schema for interpreting historical change, the traditional explanatory account which accompanies it is empirically inadequate. In the traditional description, the rise of new periphrastic constructions is motivated by, and thus follows, the loss of older morphological ones:

"Das, was man Aufbau nennt, kommt ja, wie wir gesehen haben, nur durch einen Verfall zu Stande, und das, was man Verfall nennt, ist nur die weitere Fortsetzung dieses Prozesses. Aufgebaut wird nur mit Hilfe der Syntax. Ein solcher Aufbau kann in jeder Periode stattfinden, und Neuaufgebautes tritt immer als Ersatz ein da, wo der Verfall ein gewisses Mass überschritten hat." (Paul 1920: 351)

Such interpretations of diachronic development effectively assume an information-theoretic functional model of language in which there is a fixed set of functions which linguistic structure must be able to handle, and some inherent pressure to avoid redundancy. The simple picture of the cycle is one

in which a language at any given synchronic stage has one mechanism to carry out each essential function, and as this mechanism wears out it must be replaced by a new one so that the function will not be lost. Thus, for example, Benveniste (1968), in discussing the innovation of periphrastic grammatical constructions which constitutes the first step of what we call grammaticalization, labels it **conservative mutation**, explicitly assuming that **periphrastic constructions** develop to replace earlier morphological constructions in the same function.

However, there is abundant evidence that this model is too simple. It is not the case that a language will have only one functioning means of expressing a particular meaning at a given stage, or that serious phonological erosion of one construction is a necessary condition for the rise of another (DeLancey 1985; Hopper 1991; Hopper & Traugott 1993). Consider, as a simple example, the gradual replacement of the French inflected past and future verb forms by periphrastic constructions with *avoir* and *aller*. There is clearly a stage in this development (i.e. Modern French) where the inflectional and periphrastic constructions co-exist, where the new construction is already grammaticalized while the older one remains functional. Arguably we could claim that the traditional account has the direction of causation backwards, that it is in fact the development of the new construction which leads to the loss of the old, rather than the decay of the older one leading to the development of a new replacement (Bybee 1985; DeLancey 1985).

#### 2.4. Sources and pathways

Grammaticalization processes are not random; particular types of grammatical formative tend to develop from specific lexical sources. Recent studies have made considerable progress in developing a catalogue of such pathways of development (Traugott 1978; 1988; Ultan 1978; Givón 1979; Heine & Reh 1984; Bybee 1988; Bybee et al. 1994; Bybee & Dahl 1989; Heine et al. 1991; various papers in Traugott & Heine 1991, eds.).

The best studied category of nominal morphology is **case** inflection. Since the functions of case inflection and adpositional marking show considerable overlap – indeed in many languages case marking is accomplished by adpositions rather than inflections – we should expect to find a diachronic connection between the two. Indeed, case inflection does

ordinarily develop from cliticization and subsequent morphologization of postpositions (Kahr 1975); examples of morphologization of prepositions appear to be rare. Postpositions and prepositions both develop either from **serial verbs** or from **relational noun** constructions. Other categories of nominal inflection have not been as well-studied, but see for example Greenberg (1978) on the origin of gender marking.

Verbs, of course, show a much wider range of common inflectional categories than nouns (cf. Art. 72), but most of these come from the same proximate source, auxiliary verbs, which are the typical source of tense/aspect and modality marking, deictic specification, and causative, benefactive, and other “applicative” constructions. The other common morphological category in verbs, person and number agreement, arises most commonly from the morphologization of unstressed, non-contrastive pronouns.

**Auxiliary verbs** develop from lexical verbs through two types of construction: either complementation or clause-chaining. Since European languages are not prone to clause-chaining, many of the best-known examples of the development of auxiliaries and hence verbal inflection arose from complement constructions; for example, the French synthetic future represents reanalysis and subsequent morphologization of the infinitive form as a complement of the verb *habere* ‘have’. The path from **clause-chaining** through **verb serialization** to auxiliarization is also widely attested. For example, the Lhasa Tibetan perfect/perfective paradigm includes, among other forms, an inferential perfect *-zhag l-ša/* and a volitional perfective *-pa yin l-payii/*, reduced in running speech to *l-pii/*. The first of these developed from a serial verb construction with the verb *bzhag* ‘put’, originally a grammaticalization of clause-chaining constructions along the lines of (6):

- (6) *sha btsabs-nas bzhag*  
meat chop-NF put  
'chop the meat and put it (aside)'

The second developed from a nominalized clause (*-pa* is a nominalizer) as the complement to a higher copula (*yin*), so that the modern perfective construction *V-pii* reflects an older structure *\*[S-pa] yin*. Special note should be made of this second pattern, which exemplifies a cross-linguistically very com-

mon phenomenon of tense/aspect and/or evidentiality marking through the grammaticalization of copulas.

### 3. The process of grammaticalization

#### 3.1. Functional aspects of grammaticalization

We have referred to the notion of “functional specialization” as prerequisite to grammaticalization. The essential precondition for grammaticalization is that a lexical form have some special functional status which distinguishes it from other members of its syntactic category. We can describe three paths to grammaticalization: semantic specialization through metaphor, reanalysis through pragmatic inference, and what we may call **referent conflation**. It must be emphasized that these are not presented here as competing explanations; all are clearly attested as occurring. Moreover, they are not mutually exclusive, and particular cases of grammaticalization may be best explained in terms of some combination of these processes.

Many examples of shift from a lexical to a more grammatical meaning for a morpheme involve **metaphorical extension**. One example that has been well-explored is the “body-part model” for the development of spatial expressions. In a wide range of languages some or most of the relational nouns which express spatial relations such as ‘front’ and ‘back’, and eventually adpositions derived from these relational nouns, are drawn from body-part vocabulary (see e.g. Friedrich 1969; Brugman 1983; Heine et al. 1991). Among the commonest such semantic transfers are expressions for ‘on’ derived from nouns meaning ‘head’, for ‘behind’ from ‘back’, for ‘in front of’ from ‘face’, and for ‘inside’ from ‘belly’.

Another well-explored metaphorical field is the wide range of “localist” phenomena in which spatial expressions are used with temporal, logical or other meaning (Anderson 1973; Traugott 1978). The most prevalent example of this is the use of forms meaning ‘at’, ‘from’, and ‘to’ to express temporal as well as spatial relations, which is so universal as to leave open to question in what sense it is appropriately considered to be metaphorical. Another is the common development of perfect or past tense constructions from con-

structions with ‘come’, and of futures from ‘go’ (Traugott 1978). A well-explored type of abstraction from one cognitive domain to another is the development of grammatical from spatial (or “local”) case forms (Anderson 1971; DeLancey 1981); the most widely attested development is the origin of dative case markers from locative or allative adpositions (e.g. English *to*, French *à*, and their cognates in other Germanic and Romance languages).

Another line of explanation for **semantic shift**, sometimes referred to as metonymic (Heine et al. 1991), finds the initial shift in the transfer of the primary meaning of the construction from one to another aspect of the situation to which it applies. For example, we often find perfective constructions developing from verbs meaning ‘go’, ultimately deriving from serial verb constructions meaning ‘did X and then went’; the obvious inference that the doing was completed before the doer left ceases to be merely an inference and becomes the primary semantic content of the construction.

As we will note below, grammaticalization typically involves the flattening of a syntactic structure – a biclausal construction is reanalyzed as a single clause, or a noun phrase with another noun phrase embedded within it is reinterpreted as a single adpositional phrase. This **syntactic reanalysis** is driven by a semantic reinterpretation in which two conceptually distinct referents are reinterpreted as one. For example, in an ordinary genitive construction like *the leg of the table*, each lexical noun refers separately; although the leg is part of the table, the noun phrase makes distinct reference to the table and its leg. An adpositional construction such as *on the table*, in contrast, refers only once. Thus in the development of an adposition from a relational noun, e.g. *atop* from \**on the top of*, we see the conflation of two referents into one.

#### 3.2. Syntactic aspects of grammaticalization

Functional specialization is an ubiquitous phenomenon; lexemes with certain semantic characteristics tend to be almost automatically specialized in this sense to some degree. While these functional factors are the engine which drives the process, most linguists do not recognize grammaticalization per se until actual changes of grammatical structure have occurred (but cf. Hopper 1987; Givón 1989

for suggestions that this distinction is to some extent artificial). The critical patterns of structural change relevant to the development of morphology are concerned with changes in constituent structure, particularly head-dependent relations, and changes in the status of the morphosyntactic boundaries setting off the grammaticalizing morpheme.

Consider again the three stages of grammaticalization illustrated by the Lhasa Tibetan examples (2)–(4) (see 2.1). In (2), we can identify *bzhag* as both the syntactic head and the semantic/pragmatic center of the construction. Syntactically, it is the finite verb, and in the rightmost position which in Tibetan is characteristic of heads. Semantically, the ‘putting’ is the primary foregrounded information; the informational contribution of ‘went’ is essentially adverbial. In (3) the syntactic and functional analyses are no longer congruent: while *bzhag* remains the syntactic head, as noted previously, *phyin* is the lexical verb and the information focus. By the stage illustrated in (4), the shift of informational focus is complete, and by many analyses the head-dependent relations have also shifted. Some modern interpretations of the head-dependent relation would identify *zhag* as still the head of its construction; in any framework with this feature the description of the changes involved in grammaticalization is simpler, since it does not have to involve reanalysis of the head-dependent relations in the construction.

Accompanying each of these shifts is a downgrading of the boundary between the two verbs. In (2), the two verbs are in two distinct clauses, and a clausal boundary separates them. In (3) there is only one clause, but the lexical verb *phyin* and the auxiliary *bzhag* remain separate words. In (4) -*zhag* is a suffix, with a morpheme boundary corresponding to the word boundary in (3) and the clause boundary in (2). The mechanics of this sort of boundary downgrading and erasure are discussed at length by Langacker (1977).

### 3.3. Grammaticalization and lexicalization

A lexical morpheme which loses its autonomy may follow various career paths, which we can broadly categorize as grammaticalization and **lexicalization**. (I use “lexicalization” here to refer to the process by which originally independent lexemes become parts of

new lexical items; the term has several other uses, including one, essentially the converse of grammaticalization, in which it refers to an inflected form leaving its paradigm and becoming a distinct lexeme; the best-known category of example is the development of adverbials from oblique case-forms of nouns; cf. Art. 150). The difference is in whether the output of the process is a new lexeme or a new construction. For example, consider the Early New English and now dialectal progressive prefix *a-* in English, as in:

- (7) *I'm a-wanting for to go.*

This represents grammaticalization of an earlier preposition. Exactly the same process of phonological reduction and cliticization of a preposition is the source of the *a-* in *atop*. In this case, the end product of the process is, not a new construction, but a single lexical item, and thus an example of lexicalization. (As discussed above, the shift of *top* out of the open class of nouns, and the creation of a new member of the closed class of prepositions, are more profitably discussed as grammaticalization.)

Obviously this distinction depends on the distinction between lexical and grammatical morphemes, and thus like that distinction must be viewed as a continuum rather than a dichotomy. For example, the same reduction of a preposition discussed in the preceding paragraph also produced the odd family of *a-* adjectives (*awake*, *asleep*, *alone*, etc.). In terms of productivity this development lies somewhere between that of *atop* and the entirely productive progressive prefix. Here the process created a new subcategory; the resulting forms differ from most adjectives in that they cannot occur in prenominal position (\**an alone man*). But as a new category this has no other repercussions in the grammar, since there is no new grammatical construction associated with it, and it is not particularly useful to consider this as an example of grammaticalization.

Occupying the middle ground between lexicalization and pure grammaticalization is the development of derivational morphology, which as Paul points out is not systematically distinguishable from the origin of inflection:

“Auf die gleiche Weise wie die Ableitungssuffixe entstehen Flexionssuffixe. Zwischen beiden gibt es ja überhaupt keine scharfe Grenze.” (Paul 1920: 349)

For the purposes of grammaticalization theory the most useful criterion for identifying grammaticalization is the degree to which the output of the process is a new productive construction, i.e. a new element of grammatical structure, as opposed to simply a new set of (one or more) lexical forms.

#### 4. Uncommon abbreviations

NF	non-final clause marker
INFR.PF	inferential perfect

#### 5. References

- Anderson, John (1971), *The Grammar of Case: Towards a Localistic Theory*. Cambridge: Cambridge University Press
- Anderson, John (1973), *An Essay Concerning Aspect*. The Hague: Mouton
- Benveniste, Emile (1968), "Mutations of Linguistic Categories". In: Lehmann, Winfred, & Malkiel, Yakov (eds.), *Directions for Historical Linguistics: a Symposium*. Austin: University of Texas Press, 83–94
- Bolinger, Dwight (1980), "Wanna and the Gradience of Auxiliaries". In: Bretschneider, Gunter & Lehmann, Christian (eds.), *Wege zur Universalienforschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*. Tübingen: Gunter Narr, 292–299
- Bopp, Franz (1816), *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprachen*. Frankfurt: Andreäische
- Brugman, Claudia (1983), "The Use of Body-Part Terms as Locatives in Chalcatongo Mixtec". In: *Report #4 of the Survey of Californian and Other Indian Languages*. Berkeley: University of California, 235–290
- Bybee, Joan (1985), *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam, Philadelphia: Benjamins
- Bybee, Joan (1988), "Semantic Substance vs. Contrast in the Development of Grammatical Meaning". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 14, 247–264
- Bybee, Joan & Dahl, Östen (1989), "The Creation of Tense and Aspect Systems in the Languages of the World". *Studies in Language* 13, 51–103
- Bybee, Joan & Perkins, Revere & Pagliuca, William (1994), *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press
- DeLancey, Scott (1981), "An Interpretation of Split Ergativity and Related Patterns". *Language* 57, 626–57
- DeLancey, Scott (1985), "The Analysis-synthesis Cycle in Tibeto-Burman: A Case Study in Motivated Change". In: Haiman, John (ed.), *Iconicity in Syntax*. Amsterdam: Benjamins, 367–89
- DeLancey, Scott (1991), "The Origins of Verb Serialization in Modern Tibetan". *Studies in Language* 15, 1–23
- DeLancey, Scott (1997), "Grammaticalization and the Gradience of Categories: Relator Nouns and Postpositions in Tibetan and Burmese". In: Bybee, Joan L. & Haiman, John & Thompson, Sandra A. (eds.), *Essays on Language Function and Language Type*. Amsterdam: Benjamins, 51–69
- Donegan, Patricia, & Stampe, David (1983), "Rhythm and the Holistic Organization of Language Structure". *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago: Chicago Linguistic Society, 337–353
- Fischer, Olga & Rosenbach, Anette & Stein, Dieter (2000, eds.), *Pathways of Change: Grammaticalization in English*. Amsterdam: Benjamins
- Friedrich, Paul (1969), *On the Meaning of the Tarascan Suffixes of Space*. (Memoir 23 of the International Journal of Linguistics; supplement to the *IJAL* 35.4)
- Gabelentz, Georg von der (1891), *Die Sprachwissenschaft: Ihre Aufgaben, Methoden, und bisherigen Ergebnisse*. Leipzig: Weigel
- Givón, Talmy (1971), "On the Verbal Origin of the Bantu Verb Suffixes". *Studies in African Linguistics* 2.2, 145–163
- Givón, Talmy (1979), *On Understanding Grammar*. New York, London: Academic Press
- Givón, Talmy (1989), *Mind, Code, and Context: Essays in Pragmatics*. Hillsdale/NJ: Erlbaum
- Greenberg, Joseph (1978), "How Does a Language Acquire Gender Markers?". In: Greenberg et al. (eds.), *Vol. III*, 47–82
- Greenberg, Joseph & Ferguson, Charles & Moravcsik, Edith (1978, eds.), *Universals of Human Language, Vol. 1–4*. Stanford: Stanford University Press
- Haas, Mary R. (1977), "From Auxiliary Verb Phrase to Inflectional Suffix". In: Li (ed.), 525–537
- Haspelmath, Martin (1999), "Why is Grammaticalization Irreversible?". *Linguistics* 37.6, 1043–1068

- Heine, Bernd (1993), *Auxiliaries: Cognitive Forces and Grammaticalization*. New York: Oxford University Press
- Heine, Bernd (1997), *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press
- Heine, Bernd & Claudi, Ulrike & Hünnemeyer, Friederike (1991), *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press
- Heine, Bernd & Reh, Mechthild (1984), *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske
- Hopper, Paul (1987), "Emergent Grammar". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 13, 139–157
- Hopper, Paul (1991), "On some Principles of Grammaticalization". In: Traugott & Heine (eds.), *Vol. I*, 17–35
- Hopper, Paul, & Traugott, Elizabeth (1993), *Grammaticalization*. Cambridge: Cambridge University Press
- Humboldt, Wilhelm von (1825), "Über das Entstehen der grammatischen Formen und ihren Einfluß auf die Ideenentwicklung". *Abhandlungen der Königlichen Akademie der Wissenschaft zu Berlin*, 401–430
- Kahr, Joan (1975), "The Renewal of Case Morphology: Sources and Constraints". *Stanford Working Papers in Linguistic Universals* 20, 107–151
- Kuryłowicz, Jerzy (1965), "The Evolution of Grammatical Categories". *Diogenes* 51, 55–71 [reprinted in: Kuryłowicz, Jerzy (1975), *Esquisses linguistiques II*. München: Fink]
- Langacker, Ronald (1977), "Syntactic Reanalysis". In: Li (ed.), 57–139
- Lehmann, Christian (1985), "Grammaticalization: Synchronic Variation and Diachronic Change". *Lingua e Stile* 20.3, 303–318
- Lehmann, Christian (1986), "Grammaticalization and Linguistic Typology". *General Linguistics* 26.1, 3–23
- Lehmann, Christian (1995), *Thoughts on Grammaticalization*. Munich: Lincom Europa (LINCOM Studies in Theoretical Linguistics 1)
- Lehmann, Winfred P. (1973), "A Structural Principle of Language and its Implications". *Language* 49, 47–66
- Li, Charles (1977, ed.), *Mechanisms of Syntactic Change*. Austin, London: University of Texas Press
- Matisoff, James A. (1991), "Areal and Universal Dimensions of Grammaticalization in Lahu". In: Traugott & Heine (eds.), *Vol. II*, 383–453
- Meillet, Antoine (1912), "L'évolution des formes grammaticales". *Scientia (Rivista de scienza)* 12.26 [reprinted in: Meillet, Antoine (1926): *Linguistique historique et linguistique générale*, 130–148]
- Paul, Hermann (1920), *Prinzipien der Sprachgeschichte*. Halle: Niemeyer [1880]
- Ramat, Anna, & Paul Hopper (1998, eds.), *The Limits of Grammaticalization*. Amsterdam: Benjamins (Typological Studies in Language 37)
- Rissanen, Matti & Kytö, Merja & Kirsi Heikkonen (1997, eds.), *Grammaticalization at Work*. Berlin: Mouton de Gruyter
- Sweetser, Eve (1988), "Grammaticalization and Semantic Bleaching". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 14, 389–405
- Traugott, Elizabeth (1978), "On the Expression of Spatio-Temporal Relations in Language". In: Greenberg et al. (eds.), *Vol. III*, 369–400
- Traugott, Elizabeth (1988), "Pragmatic Strengthening and Grammaticalization". *Berkeley Linguistics Society: Proceedings of the Annual Meeting* 14, 406–416
- Traugott, Elizabeth, & Heine, Bernd (1991, eds.), *Approaches to Grammaticalization, Vol. 1–2*. Amsterdam: Benjamins (Typological Studies in Language 19)
- Ultan, Russell (1978), "The Nature of Future Tenses". In: Greenberg et al. (eds.), *Vol. III*, 83–123

*Scott DeLancey, Eugene (U.S.A.)*

## 147. Morphologisierung: von der Phonologie zur Morphologie

1. Was ist Morphologisierung?
2. Zur Geschichte des Konzepts
3. Parameter der Morphologisierung
4. Zitierte Literatur

### 1. Was ist Morphologisierung?

Von Morphologisierung spricht man im allgemeinen, wenn im Laufe der Sprachgeschichte nichtmorphologische grammatische Erscheinungen zu morphologischen Erscheinungen werden. Will man den Begriff der Morphologisierung etwas genauer umschreiben, so geschieht das am zweckmäßigsten, indem man von der Bestimmung der Morphologie, genauer gesagt der Flexionsmorphologie, ausgeht. Die **Flexionsmorphologie** ist im Rahmen des Sprachsystems dadurch gekennzeichnet, daß sie erstens grammatische Kategorien formal durch Marker symbolisiert, und daß sie zweitens innerhalb der Grenzen des Wortes operiert. Eine Morphologisierung findet immer dann statt, wenn eine grammatische Erscheinung, die bisher nur eine dieser beiden für die Flexionsmorphologie konstitutiven Eigenschaften zukam, infolge von Sprachwandel auch die jeweils zweite dieser Eigenschaften herausbildet. Damit ergeben sich grundsätzlich zwei unterschiedliche Möglichkeiten der Morphologisierung. Zum einen können syntaktisch-analytische Marker (Marker mit Wortcharakter), die ja bereits grammatische Kategorien symbolisieren, durch Tilgung von Wortgrenzen zu morphologischen Markern werden. Zum anderen können phonologische Alternationen, die innerhalb des Wortes operieren und die Formen von Flexionsparadigmen betreffen, durch Ersetzung ihrer phonologischen Bedingungen durch morphologische zu morphologischen Markern werden. Allen Morphologisierungen ist gemeinsam, daß sich im Wort, der Domäne der Morphologie, neue Zeichenverhältnisse herausbilden, neue Relationen zwischen Zeicheninhalten, also grammatischen Kategorien, und Zeichenformen, also grammatischen Markern. Dennoch unterscheiden sich die beiden Typen von Morphologisierung hinsichtlich ihres Status und ihrer Ergebnisse beträchtlich. Im ersten Fall ist die Morphologisierung nur eine Etappe eines umfassenden Grammatikalisierungsprozesses; sie führt zur Herausbildung von additiven Markern (Markern mit Morphemcharakter)

und ist entsprechend in den Zusammenhang der Grammatikalisierung einzuordnen (vgl. Art. 145). Im zweiten Fall hingegen stellt die Morphologisierung selbst einen eigenständigen, meist relativ komplexen Wandelprozeß dar; sie führt typischerweise (wenn auch nicht ausschließlich; vgl. 3.3) zur Herausbildung von modifikatorischen Markern (Markern mit Alternationscharakter). Die Morphologisierung phonologischer Alternationen (im weiteren kurz Morphologisierung) konstituiert einen spezifischen Bereich der historischen Morphologie.

Da phonologischen und morphologischen Alternationen jeweils entsprechende Regeln zugrundeliegen, läßt sich auch die Morphologisierung am angemessensten bezogen auf den Regelbegriff definieren:

**Morphologisierung** liegt genau dann vor, wenn an die Stelle einer Regel/mehrerer Regeln, die in einem phonologischen Kontext eine formale Operation vornimmt/vornehmen, eine Regel tritt/mehrere Regeln treten, die die entsprechende Operation oder deren Inversion in einem morphologischen Kontext vornimmt/vornehmen.

Es findet also, kurz gesagt, ein Übergang von phonologischen Regeln zu morphologischen Regeln statt, wobei dieser Übergang kein unmittelbarer sein muß, sondern auch vermittelt über morphonologische Regeln erfolgen kann (vgl. 3.2). Eine notwendige Voraussetzung für die Morphologisierung einer phonologischen Regel ist, daß diese formale Alternationen innerhalb von Flexionsparadigmen bewirkt. Der resultierende morphologische Kontext besteht aus grammatischen Kategorienmerkmalen (wie Numerus-, Kasus-, Tempusmerkmalen); in Sprachen mit Flexionsklassen treten auch entsprechende Klassenmerkmale auf. Bei Morphologisierungen gibt es nicht notwendigerweise ein Eins-zu-eins-Verhältnis zwischen einer phonologischen und einer morphologischen Regel. Das zeigen z. B. die verschiedenen Morphologisierungen des Umlauts im Deutschen; im Laufe der Sprachgeschichte wird eine einheitliche phonologische Umlautregel durch eine ganze Anzahl von morphologischen Umlautregeln ersetzt. Ein Beispiel dafür, daß eine resultierende morphologische Regel die inverse Operation der ursprünglichen phonologischen Regel vornimmt, ist der morphologische ‘Rückumlaut’ der deutschen Verben

des Typs *brennen* – Präteritum *brannte* mit Umkehrung der ursprünglichen Ableitungsrichtung von ‘a → e’ zu ‘e → a’ (vgl. 3.4).

Mit der Morphologisierung von phonologischen Regeln ändert sich deren Status grundsätzlich. Eine Regel, deren (vollständig oder doch partiell erhaltene) Funktion darin bestand, eine Klasse von Lautfolgen an den phonologischen Kontext anzupassen, erhält die neue Funktion der formalen Symbolisierung morphologischer Kategorien, wie sie allen morphologischen Regeln zukommt. Damit wird die phonetische Motivierung der Regel durch eine semiotische Motivierung ersetzt; aus lautlichen Alternationen werden sprachliche Zeichen; vgl. dazu z. B. urnordisch *londa* ‘Land’ – Nom./Akk. Pl. *lond* mit Assimilation [a] > [ɔ] an das [u] der Folgesilbe und neuisländisch *land* – Nom./Akk. Pl. *lönd* mit morphologischer *a/ö*-Alternation als alleinigem Numerus-Kasus-Marker.

## 2. Zur Geschichte des Konzepts

Bereits die traditionelle Sprachwissenschaft der zweiten Hälfte des 19. Jahrhunderts stößt bei ihren sprachhistorischen Untersuchungen auf die Problematik der Morphologisierung, ohne daß freilich der Terminus auftritt. So stellt Paul (1880) fest, daß es normalerweise kein “lautgesetz” gibt, “das nicht, sobald es einmal in einer anzahl von fällen das etymologisch eng zusammenhängende lautlich differenziert hat, auch eine reaction gegen diese differenzierung hervorriefe” (Paul 1880: 104). Paul konstatiert, daß eine solche Entwicklung jedoch nicht eintritt, wenn die formale Differenzierung mit einem Funktionsunterschied, d. h. einem morphologischen Unterschied, zusammenfällt. Dieses Zusammenfallen kann dann nämlich “die ursache zu dauernder bewahrung eines lautlichen unterschiedes sein, und dies vor allem, wenn er zugleich [...] durch die formale analogie widerstandsfähig gemacht wird” (Paul 1880: 114), d. h. wenn die Alternation morphologisch systematisiert wird:

“Bei dem zusammentreffen dieser beider umstände kann sich die vorstellung von dem lautlichen unterschiede so fest mit der von dem functionsunterschied verbinden, dass dem sprachgefühl beides unzertrennbar erscheint. Auf diese weise wird allmälig der zufällig entstandene bedeutungslose unterschied zu einem bedeutungsvollen. Er wird es um so mehr, je weniger die bedeutungsverschiedenheit durch sonstige unterschiede in der lautgestaltung deutlich gekennzeichnet ist. So vermag sich die sprache einen ersatz zu schaffen für den in folge

des lautlichen verfalls eintretenden verlust der charakteristischen merkmale des functionsunterschiedes.” (Paul 1880: 114; Hervorhebung im Original)

Um die theoretische Einordnung der Morphologisierung haben sich dann bereits zur Zeit der Junggrammatiker vor allem namhafte Vertreter des frühen russischen Strukturalismus verdient gemacht. Baudouin de Courtenay, der offenbar auch den Terminus “Morphologisierung” prägt, unterscheidet in seiner “Theorie phonetischer Alternationen” (Baudouin de Courtenay 1895) u. a. drei für den Morphologisierungszusammenhang einschlägige Typen von Alternationen, nämlich (i) “neophonetische”, d. h. phonologisch reguläre Alternationen, (ii) “traditionelle”, d. h. phonologisch und morphologisch irreguläre Alternationen und (iii) “psychophonetische”, d. h. phonologisch irreguläre, aber morphologisch reguläre Alternationen (Baudouin de Courtenay 1895: 110). Die Morphologisierung wird bei Baudouin in umfassendere grammatische Zusammenhänge gestellt. Für ihn hat die Sprache eine phonetische, eine semasiologische (semantische) und eine morphologische Seite, wobei sich letztere auf die grammatische Struktur der Sprache bezieht. Dabei verbinden sich phonetische Vorstellungen mit semasiologischen oder morphologischen; Lautsegmente oder auch phonetische Merkmale werden semasiologisiert oder morphologisiert (Baudouin de Courtenay 1908: 10f.). Semasiologisierung und Morphologisierung haben jedoch einen recht unterschiedlichen Status. Alle (distinktiven) Unterschiede werden generell im Rahmen von Morphemen semasiologisiert, d. h. zur Unterscheidung der Morpheme als Bedeutungsträger genutzt.

“Demgegenüber ist die Morphologisierung bestimmter Unterschiede zwischen artikulatorisch-auditiven Elementen in der Geschichte der Sprachen eine Übergangsscheinung. Sie wird gewöhnlich durch bestimmte historisch-phonetische Prozesse hervorgerufen, die zur Auflösung eines Phonems in zwei oder mehr führen; und diese neu entstandenen Abwandlungen eines ehemals einzigen Phonems [...] können psychophonetische Alternationen werden, die mit einem festen Unterschied von Formen verknüpft sind.” (Baudouin de Courtenay 1922: 67; deutsche Übersetzung nach Mugdan 1984: 75)

Die Morphologisierung erfolgt also in zwei Schritten: “Neophonetische” Alternationen verlieren zunächst durch Lautwandel ihre phonologische Motivierung und werden zu “traditionellen” Alternationen (i > ii). Wenn

diese dann morphologisch systematisiert werden, bilden sich “psychophonetische” Alternationen heraus. Es tritt Morphologisierung ein (ii > iii). Wenn das nicht geschieht, bleiben sie als rein traditionelle, irreguläre Alternationen erhalten (für Details vgl. Mugdan 1984: 73 ff.).

Auf Baudouin aufbauend legt dann sein Schüler Kruszewski (der auch Baudouins erst später publizierte Auffassungen zur Morphologisierung kennt) ein geradezu erstaunlich modernes Konzept der unterschiedlichen Typen von “Lautabwechslungen”, d. h. von Alternationen, und der sprachhistorischen Übergänge zwischen ihnen vor (Kruszewski 1881), das in seinen Grundzügen noch heute gilt. Er unterscheidet als erster zwischen (i) phonetisch bedingten Alternationen, (ii) morphphonologisch bedingten Alternationen und (iii) morphologisch bedingten Alternationen. Es gibt einen sprachhistorischen Entwicklungsprozeß von (i) zu (iii), wobei der Wandel von (i) zu (ii) “den Anfang der Verbindung von phonetischen Erscheinungen mit morphologischen Kategorien” darstellt (Kruszewski 1881: 23). Da die morphonologisch bedingten Alternationen eine Art von Umbauperiode zwischen (i) und (iii) markieren, stellen sie sich zunächst als scheinbar unregelmäßig dar; sie tendieren aber quasi gesetzmäßig zu einer neuen Regularität. Für die weitere Entwicklung gibt es zwei Möglichkeiten. Entweder erfolgt ein Wandel von (ii) zu (iii) in Form einer “Differenzierung”, d. h. einer Funktionalisierung der Alternation, wie beim deutschen Umlaut als Pluralmarker in Fällen wie *Loch – Löcher*; die Alternation ist dann morphologisiert. Oder es findet eine morphologische “Assimilation”, ein Ausgleich innerhalb des Paradigmas, statt, wie im Konsonantismus bestimmter starker Verben des Deutschen, vgl. ahd. *lesan – las – lärum* > nhd. *lesen – las – lasen*. In diesem Fall verschwindet die Alternation aus der Grammatik (Kruszewski 1881: 24 ff.; vgl. dazu Klausenburger 1994: 2562).

Nach einer langen Zeit, in der die Morphologisierung in der theoretischen Linguistik faktisch keine Rolle spielte, wird die Problematik in den siebziger Jahren wieder aufgegriffen. Ausgangspunkt dafür ist die zusammenfassende Darlegung der generativen phonologischen Standardtheorie durch Chomsky & Halle (1968). Diese Theorie betrachtet sämtliche Alternationen von Lauteinheiten ungeachtet ihrer jeweiligen Bedingtheit und Funktion einheitlich als phonologi-

sche Regeln. In den dadurch hervorgerufenen Diskussionen wird auch die Frage nach den zu unterscheidenden Typen von Alternationsregeln und den Kriterien ihrer Abgrenzung voneinander thematisiert. Damit wird dann auch wieder die Problematik der Genese und Entwicklung von Alternationsregeln theoretisch relevant. So entsteht eine ganze Anzahl von Arbeiten, die sich unter unterschiedlichen Gesichtspunkten mit der Morphologisierung auseinandersetzen. Dabei steht meist die Frage nach den Triebkräften von Morphologisierungsprozessen im Mittelpunkt.

Unmittelbaren Bezug auf die Schwachstellen der generativen Phonologie nimmt Skousen (1975). Bei der Behandlung der finnischen Konsonantengradation kommt er zu dem Ergebnis, daß diese durch (vermeintliche) phonologische Regeln und die entsprechenden abstrakten Repräsentationen nicht in auch nur einigermaßen angemessener Weise erfaßt werden kann, denn ihre Domäne ist im modernen Finnischen durch nichtphonologische, nämlich morphologische Bedingungen eingeschränkt bzw. ausgeweitet. Die ursprüngliche phonologische Regel ist im Laufe der Sprachgeschichte zu einer morphologischen Alternation geworden. “The motivating force” für die Morphologisierung von Regeln ist “a desire for surface regularity”, das Bestreben nach morphologischer Regularität (Skousen 1975: 125). Hooper (1976) hebt die semiotische Seite von Morphologisierungsprozessen hervor: “After a phonetic alternation appears in the language, the tendency is for this alternation to walk its way up towards the meaning end of the grammar, moving from a purely phonetic function to a semantic function” (Hooper 1976: 86). Hier setzt dann auch Dressler (1977) an, der Morphologisierungen generell damit erklärt, daß “morphologische Funktionen semiotisch wichtiger sind als phonologische” und demgemäß in entsprechenden Fällen “die morphologische Funktion [...] immer wichtiger, die phonologische Funktion hingegen immer unwichtiger wird” (Dressler 1977: 24). Klausenburger (1979) macht auf das Faktum aufmerksam, daß phonologische Regeln nur dann morphologisiert werden können, wenn sie über ein “preexisting morphological conditioning” verfügen (Klausenburger 1979: 32), d. h. wenn die Alternationen von Anfang an (zufälligerweise) an morphologische Kategorien gebunden sind. Alle diese Arbeiten haben gemeinsam, daß Morphonologisierungen und Morphologisierun-

gen ausschließlich als durch morphologische Triebkräfte bedingt angesehen werden.

Einen gegensätzlichen Standpunkt vertritt Robinson (1975) bei einer historischen Analyse des deutschen Umlauts. Er setzt sich (ähnlich wie Skousen 1975 hinsichtlich der finnischen Gradation) mit der abstrakt-generativen Behandlung des neuhochdeutschen Umlauts als phonologischer Regel auseinander und wertet diesen (Wurzel 1970 folgend) als morphologisch bedingt. Die Morphologisierung der phonologischen Umlautregel findet seiner Meinung nach einheitlich zu dem Zeitpunkt statt, zu dem die phonologischen Umlauffaktoren durch phonologische Tilgungen und Neutralisierungen abgebaut worden sind (Robinson 1975: 4f.). Die Morphologisierung des deutschen Umlauts ist damit ausschließlich phonologisch bedingt, was im übrigen der gängigen germanistischen Lehrmeinung entspricht (vgl. jedoch dazu 3.1). Des weiteren wird angenommen, daß im Fall des deutschen Umlauts die Morphologisierung ohne morphonologische Zwischenstufe erfolgt.

In Wurzel (1980) wird dann anhand von einschlägigen Beispielfällen gezeigt, daß sowohl morphologisch als auch phonologisch ausgelöste Morphologisierungen anzunehmen sind. Dabei wird auch die oft sehr große innere Komplexität von scheinbar einfachen Morphologisierungsprozessen im Spannungsfeld von Phonologie und Morphologie herausgearbeitet. Schließlich wird eine Klassifizierung von Morphologisierungen nach den Parametern der Bedingtheit, des Typs der Ausgangsregel, der Anzahl der Schritte im Gesamtprozeß, der Anzahl der Ausgangsregeln und der resultierenden Regeln sowie des Auftretens bzw. Nichtauftretens von Regelinversion vorgeschlagen (Wurzel 1980: 457 ff.).

Die Beiträge zur Morphologisierung aus den achtziger und neunziger Jahren thematisieren speziell die Unidirektionalität von Morphologisierungsprozessen und deren Erklärbarkeit. Dressler (1985) ordnet die Morphologisierung in einen semiotischen Theorierahmen ein: Wörter sind primäre sprachliche Zeichen; Morpheme und morphologische Regeln sind Zeichen von Zeichen (den Wörtern), d. h. damit sekundäre Zeichen; und Phoneme sind Zeichen von Zeichen (den Morphemen) von Zeichen (den Wörtern), d. h. tertiäre Zeichen. Damit hat das Lexikon semiotische und epistemologische Priorität gegenüber der Morphologie und die Morphologie gegenüber der Phonologie (Dressler

1985: 283). Das bedeutet, daß auch morphologische Indexikalität höher zu bewerten ist als phonologische Indexikalität. Wenn man weiter berücksichtigt, daß (grammatisch bedingter) Sprachwandel zu effizienteren grammatischen Strukturen von Subsystemen führt, so ergibt sich daraus die Tendenz des Übergangs von phonologischen Regeln zu morphonologischen Regeln und weiter zu morphologischen Regeln. Die umgekehrte Entwicklungsrichtung ist dagegen nicht möglich (Dressler 1985: 311).

Diese auf Kruszewskis Konzept zurückgehende Gerichtetheit der Übergänge zwischen den drei Regeltypen kann heute weitgehend als allgemein akzeptiert gelten. Anders argumentieren jedoch Morin et al. (1990), daß im Fall der Herausbildung des französischen ‘o-Tensing’ im Auslaut die eingetretene Sprachveränderung überhaupt nur dann erklärbar ist, wenn man annimmt, daß morphonologische Regeln ‘rephonologisiert’ werden können, was bedeuten würde, daß die Unidirektionalität unhaltbar wäre. Klausenburger (1994) arbeitet jedoch heraus, daß in diesem Fall nicht eine (ohnehin fragwürdige) morphonologische Regel, sondern ein phonetischer bzw. phonotaktischer Sachverhalt der französischen Aussprache phonologisch generalisiert wird (Klausenburger 1994: 2563 f.). Um zu ermitteln, ob ein ‘Weg zurück’ unter entsprechenden Bedingungen möglich ist, überprüft er Fälle aus der romanischen Sprachgeschichte, in denen die Sprecher in ihrer Sprache gegebene Alternationen prinzipiell entweder phonologisch oder morphologisch werten könnten. Der weitere Verlauf der Entwicklung erweist in allen Fällen, daß jeweils die morphologischen Bedingungen für die Alternation gewählt wurden. Das gleiche Ergebnis zeigen auch Experimente, in denen sowohl phonologisch als auch morphologisch interpretierbare Vokalalternationen in spanischen Nonsense-Verben von nativen Sprechern eindeutig an morphologischen Kategorien festgemacht wurden (Bybee 1985: 66 ff.). Der Weg von einer morphonologischen Regel zu einer phonologischen Regel, der der semiotischen Priorität der Morphologie gegenüber der Phonologie widerspricht, ist aller vorliegenden Evidenz nach ausgeschlossen.

### 3. Parameter der Morphologisierung

Alle Morphologisierungen sind (*per definitio nem*) durch den Übergang von phonologischen zu morphologischen Regeln gekenn-

zeichnet, doch darüber hinaus haben sie kaum durchgängige Gemeinsamkeiten. Die vorkommenden Morphologisierungen unterscheiden sich in einer ganzen Reihe von relevanten Eigenschaften, die zudem auch noch jeweils weitgehend unabhängig voneinander sind. Eine wirklich erschöpfende Typologisierung von Morphologisierungen stellt sich unter diesen Bedingungen als sehr kompliziert dar und wäre nur im Rahmen einer umfassenden Darstellung anhand einer großen Anzahl von Beispielen möglich. Hier können demgegenüber nur einige wenige einschlägige Beispiele vorgestellt und aufgrund der wichtigsten **Parameter der Morphologisierung** zueinander in Bezug gesetzt werden. Diese Parameter sollen sein: die Initiierung der Morphologisierung, der Weg von der phonologischen zur morphologischen Regel, der Typ der resultierenden morphologischen Regel, die Ableitungsrichtung der morphologischen Regel und der Status des neuen Kategorienmarkers.

### 3.1. Die Initiierung von Morphologisierungen: phonologisch versus morphologisch

Jede Morphologisierung wird durch einen Sprachwandel ausgelöst. Das kann entweder ein phonologischer oder ein morphologischer Wandel sein, womit sich ein erster Klassifizierungsparameter für Morphologisierungen ergibt. Vgl. dazu zunächst die folgenden Beispiele.

**Beispiel 1:** Im Althochdeutschen zeigen starke Verben mit einem hinteren Stammvokal im Infinitiv wie *graban* ‘graben’ bedingt durch die phonologische Umlautregel im Präsens Indikativ entsprechende Alternationen: 1. Sg. *grabu*, 2. Sg. *grebis*, 3. Sg. *grebit*; 1. Pl. *graben*, 2. Pl. *grabet*, 3. Pl. *grabent*. Vor einem [i] der Folgesilbe erscheint der umgelautete Vokal, sonst der unumgelautete. Beim Übergang zum Mittelhochdeutschen tritt ein phonologischer Wandel ein, durch den die unbetonten Endsilbenvokale zu [ə] (phonologisch /e/) neutralisiert werden; vgl. Infinitiv *graben*; 1. Sg. *grabe*, 2. Sg. *grebest*, 3. Sg. *grebet*; 1. Pl. *graben*, 2. Pl. *grabet*, 3. Pl. *grabent*. Der Umlaut hat seine phonologische Bedingtheit verloren (vgl. 3. Sg. *grebet* vs. 2. Pl. *grabet*). Er tritt in den Präsensparadigmen der starken Verben jetzt im morphologischen Kontext ‘2./3. Sg. Präs. Ind.’ auf und ist damit zum Kategorienmarker geworden. An die Stelle der phonologischen Umlautregel ist bei

diesen Verben eine morphologische Umlautregel getreten, die noch heute gilt; vgl. *ich grabe*, *du gräbst* usw.

**Beispiel 2:** Einen parallelen Ausgangspunkt (wenn auch einen anderen Verlauf) hat eine Morphologisierung im Nordischen. Hier alternieren im späteren Urnordischen entsprechend einer phonologischen *i*-Umlautregel nichtumgelautete und umgelautete Flexionsformen im Präsens Indikativ der langsilbigen starken Verben auf gleiche Weise wie im Althochdeutschen; vgl. Infinitiv *blōta(n)* ‘opfern’; 1. Sg. *blōtu*, 2. Sg. *blōtiR*, 3. Sg. *blōtiR*; 1. Pl. *blōtum* usw. Die Morphologisierung beginnt in diesem Fall mit einem morphologischen Wandel, der Systematisierung des Verhältnisses zwischen Funktion und Form im Paradigma. Die 1. Person Singular wird im Vokalismus an die beiden anderen Singularformen angeglichen; es ergibt sich das Teilparadigma: Inf. *blōta(n)*; 1. Sg. *blōtu*, 2. Sg. *blōtiR*, 3. Sg. *blōtiR*; 1. Pl. *blōtum*, 2. Pl. *blōteð*, 3. Pl. *blōta(n)*. Die Umlautregel gilt weiterhin für die Vokale vor [i], aber zusätzlich dazu morphologisch bedingt für den Vokal der 1. Person Singular. Damit erfaßt die Alternation durch Ausdehnung ihrer Domäne außer den *i*-Instanzen anderer Paradigmen jetzt auch den gesamten Singular Präsens Indikativ der Verben des Typs *blōta(n)*. Aufgrund von phonologischen Tilgungen verschwindet dann später das (ursprünglich) umlautbewirkende [i]; vgl. das resultierende altsländische Paradigma: Inf. *blōta*; 1. Sg. *blōt*, 2. Sg. *blōtr*, 3. Sg. *blōr*; 1. Pl. *blōtum* usw. Die Alternation ist für diese Verben zu einer morphologischen Regel geworden. Dieser schließen sich dann auch die kurzsilbigen starken Verben an, vgl. *taka* ‘nehmen’, *tek*, *tekr* usw., wodurch die eine morphologische Regel belastende phonologische Einschränkung der Langsilbigkeit verschwindet.

**Beispiel 3:** Im frühen Althochdeutschen zeigen die Maskulina der *i*-Deklination wie *gast* ‘Gast’ die folgenden Flexionsformen: Sg. Nom./Akk. *gast*, Gen. *gastes*, Dat. *gaste*, Instr. *gestiu*; Pl. Nom./Akk. *gesti*, Gen. *gestio*, Dat. *gestim*. Die Alternation ist bedingt durch die erwähnte phonologische Umlautregel. Im späteren Althochdeutschen tritt dann in den entsprechenden Paradigmen ein ‘analogischer’ Ausgleich ein, durch den bei weiterhin vorhandenem [i] die Instrumental Singular Form *gestiu* durch *gastiū* ersetzt wird. Die Domäne der phonologischen Umlautregel wird morphologisch eingeschränkt; der

Umlaut tritt nur noch, und zwar durchgängig, im Plural auf. Beim Übergang zum Mittelhochdeutschen werden auch hier die Vokale der Flexionsendungen zu [ə] reduziert. Des weiteren verschwindet die Kategorie des Instrumentals aus dem Flexionssystem. Das Substantiv *gast* flektiert dann in folgender Weise: Sg. Nom./Akk. *gast*, Gen. *gastes*, Dat. *gaste*; Pl. Nom./Akk./Gen. *geste*, Dat. *gesten*. Die phonologische Umlautregel ist durch eine morphologische Regel ersetzt worden, die im Kontext der grammatischen Kategorie Plural den Vokal verändert, wenn das Wort zur Umlautklasse gehört. Diese Klasse ist durch ein Flexionsmerkmal spezifiziert, das diese von der Klasse der sonst gleich flektierenden nichtumlautenden starken Maskulina des Typs *tac* ‚Tag‘ unterscheidet; vgl. *gast* – Nom. Pl. *geste* vs. *tac* – Nom. Pl. *tage*.

Im ersten Beispiel wird die Morphologisierung also durch einen phonologischen Wandel ausgelöst, in den beiden anderen Beispielen durch einen morphologischen Wandel. Das bedeutet jedoch nicht, daß diese Morphologisierungen jeweils ausschließlich phonologisch bzw. ausschließlich morphologisch bedingt sind. In allen drei betrachteten Fällen wirken phonologische und morphologische Bedingungen zusammen. Im Beispiel 1 folgt dem phonologischen Abbau des umlautbewirkenden [i] eine morphologische Reanalyse, indem das Auftreten des Umlautvokals mit den entsprechenden grammatischen Kategorien verbunden wird. Und in den Beispielen 2 und 3 werden nach dem jeweils auslösenden morphologischen Wandel die umlautbewirkenden Vokale durch phonologische Reduzierungen bzw. Tilgungen abgebaut. Da weder eine phonologisch ausgelöste Morphologisierung ohne morphologische Reanalyse, noch eine morphologisch ausgelöste Morphologisierung ohne phonologische Tilgung des ursprünglichen phonologischen Kontextes vorstellbar ist, ergibt sich die Schlußfolgerung, daß alle Morphologisierungen zugleich phonologisch und morphologisch bedingt sind. Deshalb soll hier nicht (wie häufig) zwischen phonologisch und morphologisch bedingten, sondern genauer zwischen **phonologisch** und **morphologisch initiierten Morphologisierungen** unterschieden werden.

Phonologisch initiierte Morphologisierungen kommen morphologisch gesehen zufällig zustande; die phonologischen Bedingungen der Alternation verschwinden und werden notgedrungen durch morphologische Bedin-

gungen, d. h. grammatische Kategorien, ersetzt. Demgegenüber werden morphologisch initiierte Morphologisierungen von den Sprechern ‚gezielt‘ in die Wege geleitet. Eine Alternation wird trotz ihrer fortbestehenden phonologischen Bedingtheit von den Sprechern als zusätzlicher Marker derjenigen Kategorie oder Kategorien reanalysiert, in der oder denen sie typischerweise auftritt. Das morphologisch unsystematische Nichtauftreten der Alternation wie im Fall des urnordischen 1. Sg. Präs. Ind. *blōtu* bzw. ihr unsystematisches Auftreten wie im Fall des althochdeutschen Instr. Sg. *gestiu* wird durch Ausdehnung bzw. Einschränkung der Domäne der Regel beseitigt. Bei den morphologisch initiierten Morphologisierungen zeigt sich damit deutlich die (von verschiedenen Linguisten konstatierte) semiotische Priorität der Morphologie gegenüber der Phonologie. Zu ergänzen ist, daß beide Typen der Morphologisierung von phonologischen Regeln die im System vorhandenen Flexionskategorien voraussetzen. Durch sie entstehen neue Marker für bereits gegebene Kategorien, aber (anders als bei der Morphologisierung syntaktischer Marker) niemals neue Flexionskategorien.

### 3.2. Der Weg von der phonologischen zur morphologischen Regel: direkt vs. indirekt

Im Rahmen von Morphologisierungsprozessen sind (wie gesagt) drei Klassen von Regeln zu unterscheiden. **Phonologische Regeln** nehmen formale Operationen im Kontext phonologischer Einheiten wie Merkmale, Segmente und Silben vor. Sie enthalten grundsätzlich keine grammatischen Kategorienmerkmale wie ‚Plural‘, ‚Genitiv‘, ‚Präsens‘ usw. und keine Flexionsklassenmerkmale. **Morphologische Regeln** vollziehen demgegenüber solche Operationen im Kontext von grammatischen Kategorien, in Sprachen mit Flexionsklassen bezogen auf diese. **Morphologische Regeln** nehmen formale Operationen in einem Kontext vor, der sowohl phonologische Einheiten als auch grammatische Kategorien enthält. Ein Blick auf die Beispiele zeigt, daß im Prozeß der Morphologisierung der Übergang von phonologischen zu morphologischen Regeln entweder direkt oder indirekt, d. h. über eine morphonologische Regel, erfolgen kann.

Im Beispiel 1 vollzieht sich die Morphologisierung direkt. Aufgrund des phonologischen Abbaus sämtlicher umlautbewirkender [i]-Instanzen im Paradigma werden diese voll-

ständig durch grammatische Merkmale ersetzt. Anders im Beispiel 3. Hier wird der im Prinzip weiter funktionierende phonologische Umlaut für die Instrumental Singular Formen von Substantiven des Typs *gast* blockiert. Der althochdeutsche Umlaut erfährt darüber hinaus noch eine morphologisch bedingte Blockierung, d. h. eine weitere Morphonologisierung.

Beispiel 4: Im älteren Althochdeutschen zeigen die schwachen Maskulina wie z. B. *hano* ‘Hahn’ die folgenden Flexionsformen: Sg. Nom. *hano*, Gen./Dat. *henin*, Akk. *hanun*; Pl. Nom./Akk. *hanun*, Gen. *hanōnō*, Dat. *hanōm*. Vor folgendem [i] tritt also ‘lautgesetzlich’ Umlaut auf. Dieser Umlaut im Genitiv und Dativ Singular wird schon sehr früh durch einen morphologischen Wandel beseitigt; *henin* wird durch *hanin* ersetzt, wodurch im Paradigma jetzt durchgängig der unumgelautete Vokal erscheint. Beim Übergang zum Mittelhochdeutschen wird auch das [i] des Markers *-in* phonologisch zu [ə] reduziert. Die Flexionsformen des Paradigmas lauten jetzt Sg. Nom. *hane*, Gen./Dat./Akk. *hanen*; Pl. *hanen*.

Die gemeinsame Zwischenstufe dieser Entwicklungen, die spätalthochdeutsche Umlautregel, lässt sich damit wie folgt skizzieren:

$$V \rightarrow [-\text{hinten}] / \left\{ \begin{array}{ll} \text{Mask} & (\text{a}) \\ \sim i\text{-Dekl} \\ -\text{Pl} \\ \sim [n\text{-Dekl}] & (\text{b}) \\ \_\_ K_1 i & (\text{c}) \end{array} \right.$$

Abb. 147.1: Spätalthochdeutsche Umlautregel

Diese Regel ist eine morphonologische Regel; ihre Anwendung ist partiell (noch) phonologisch und partiell (schon) morphologisch bedingt. Die morphologischen und phonologischen Bedingungen der Alternation lassen sich hier nicht unabhängig voneinander formulieren. Die Teilregeln sind extrinsisch geordnet, denn die blockierenden morphologischen Teilregeln (a) und (b) müssen der phonologischen Teilregel (c) vorausgehen. Im Mittelhochdeutschen ist dann der phonologische Kontext der Regel abgebaut. Die Teilregel (a), die den Vokalwechsel im Paradigma *gast* systematisiert, findet ihre Entsprechung in einer morphologischen Regel, die dann den Vokal im inversen (positiven) Kategorienkontext ‘*i*-Dekl + Pl’ umlautet. Die ursprüngliche phonologische Regel ist auf indi-

rektem Wege morphologisiert worden. Die Teilregel (b), die den Wechsel im gesamten Paradigma *hano* blockiert, verschwindet aus der Grammatik, da sie keine Kategorienmerkmale enthält. In solchen Fällen tritt ein morphologisch bedingter Ausgleich im Paradigma (eine ‘Assimilation’ in Kruszewskis Sinne), aber keine Morphonologisierung einer Alternation ein. Nicht jede Morphonologisierung führt also zur Morphonologisierung.

Auch im Beispiel 2 erfolgt die Morphonologisierung über eine morphonologische Zwischenstufe. Die zunächst nur vor einem [i] der Folgesilbe eintretende Alternation wird auf die Umgebung ‘Stark, – Prät, – Konj, – Pl’ ausgedehnt; sie folgt wiederum teils (schon) morphologischen, teils (noch) phonologischen Bedingungen. Die morphologische und die phonologische Teilregel sind zwar nicht untrennbar miteinander verzahnt wie im vorangehenden Beispiel (da die morphologische Teilregel einen positiven Kontext hat, ließen sich die beiden Teilregeln mit etwas größerem Aufwand auch als selbständige Regeln formulieren und zeigen keine extrinsische Ordnung), doch sie nehmen die gleiche formale Operation vor und haben sich überschneidende Anwendungsbereiche. Mit der Tilgung der [i]-Instanzen im Paradigma verschwindet die phonologische Teilregel, übrig bleibt eine morphologische Regel.

Die diskutierten (wie auch viele weitere) Beispiele scheinen dafür zu sprechen, daß phonologisch initiierte Morphonologisierungen direkt zu morphologischen Regeln führen (Beispiel 1), während morphologisch initiierte (Beispiele 2 und 3) den Weg über morphonologische Regeln nehmen. Doch vgl. dazu den folgenden Fall:

Beispiel 5: Bedingt durch die phonologische *u*-Umlautregel tritt im späteren Urnordischen u. a. auch in den Paradigmen der maskulinen und neutralen *a*-Substantive eine Alternation zwischen [a] und [ɔ] (ø) auf. Vor einem [u] im Flexiv erscheint im Stamm [ɔ], sonst [a]. Vgl. z. B. Maskulinum *armaR* ‘Arm’; Sg. Nom. *armaR*, Gen. *armas*, Dat. *armē*, Akk. *arma*; Pl. Nom. *armōR*, Gen. *armō*, Dat. *ormumR*, Akk. *arma(n)* und Neutr. *landa* ‘Land’; Sg. Nom./Akk. *landa*, Gen. *landas*, Dat. *landē*; Pl. Nom./Akk. *londu*, Gen. *landō*, Dat. *londumR*. In der Entwicklung zum Altländischen werden die Flexionsendungen aufgrund von eindeutig phonologischen Veränderungen getilgt bzw. reduziert, so daß sich die Formen Sg. Nom.

*armr*, Gen. *arms*, Dat. *armi*, Akk. *arm*; Pl. Nom. *armor*, Gen. *arma*, Dat. *ormum*, Akk. *arma* und Sg. Nom./Akk. *land*, Gen. *landas*, Dat. *landi*; Pl. Nom./Akk. *lond*, Gen. *lando*, Dat. *londum* ergeben. Während der Umlaut im Dativ Plural bei erhaltenem [u] weiterhin phonologisch bedingt ist, wird er (notgedrungen) im Nominativ und Akkusativ Plural der Neutra mit dem entsprechenden morphologischen Kategorienkontext assoziiert (ähnlich übrigens auch in anderen Paradigmen). Der *u*-Umlaut ist im Altländischen damit zu einer morphonologischen Regel geworden. Die weitere Entwicklung ist dadurch gekennzeichnet, daß zum Neuisländischen hin durch phonologische Epenthese neue [u]-Instanzen in den Flexiven entstehen, vgl. altländisch *armr* > neuisländisch *armur*, wobei (erwartungsgemäß) kein Umlaut auftritt. Es erfolgt weiterhin der Übergang von [u] zu [y] (orthographisch weiter *u*) und von [ø] (ø) zu [ø] (ö). Umlaut erscheint also in den Paradigmen nicht mehr im phonologischen Kontext 'vor [y]', sondern im morphologischen Kontext 'vor dem Dat. Pl. Flexiv -um [ym]'; vgl. Dativ Plural *örnum* und *löndum*, aber Nominativ Singular *armur*. Daß hier tatsächlich der Kategorienkontext entscheidend ist, zeigen auch Neubildungen wie *farum* 'darüber' (aus *far* 'dort' und *um* 'um, über'), die ebenfalls umlautlos bleiben. Damit ist auch die phonologische Teilregel der bisherigen morphonologischen Regel morphologisiert worden. Im Laufe des Morphologisierungsprozesses sind an die Stelle der ursprünglichen phonologischen *u*-Umlautregel u. a. (die Regel unterliegt auch weiteren Morphologisierungen) zwei für die betrachteten Paradigmen geltende morphologische Regeln getreten, eine für den Nominativ/Akkusativ Singular Neutr. vgl. *lond*, und eine für den Dativ Plural beider Genera, vgl. *örnum*, *löndum* (für Details vgl. Árnason 1985).

Hier führt eine phonologisch initiierte Morphologisierung anders als im Beispiel 1 zunächst zu einer morphonologischen Regel. Dieser Unterschied ergibt sich daraus, daß durch den auslösenden phonologischen Wandel die umlautbewirkenden [u]-Instanzen nicht vollständig, sondern nur partiell getilgt werden. Es läßt sich somit generalisieren: Phonologisch initiierte Morphologisierungen, bei denen im ersten Schritt sämtliche bedingenden phonologischen Kontexte abgebaut werden, führen direkt zu morphologischen Regeln, alle anderen Morphologisierungen dagegen über morphonologische Regeln.

Beispiel 6: Diese bedingt in den Paradigmen der *ir*-Stämme des Typs *hrind* 'Rind' Alternationen zwischen [ir] und Ø der folgenden Art: Sg. Nom./Akk. *hrind*, Gen. *hrind-ir-as*, Dat. *hrind-ir-a*, Instr. *hrind-ir-u*; Pl. Nom./Akk. *hrind-ir-u*, Gen. *hrind-ir-o*, Dat. *hrind-ir-um*. Später tritt ein phonologischer Wandel ein, durch den ein auslautendes [u] unter bestimmten Bedingungen beseitigt wird. Die frühalthochdeutsche Entsprechung des Paradigmas sieht dann so aus: Sg. Nom./Akk. *hrind*, Gen. *hrind-ir-es*, Dat. *hrind-ir-e*, Instr. *hrind-ir-u* (hier ist die Erhaltung des [u] durch Ausgleich bedingt); Pl. Nom./Akk. *hrind-ir*, Gen. *hrind-ir-o*, Dat. *hrind-ir-um*. Es ergeben sich also Instanzen von [ir] im Auslaut; das Nichtvorkommen bzw. Vorkommen von [ir] im Paradigma ist nicht mehr phonologisch faßbar und wird statt dessen an die grammatischen Kategorien gebunden. Da das Element [ir] nicht in der Nominativ Singular Form des Paradigmas, wohl aber in den meisten abgeleiteten Formen auftritt, wird es als ein Flexiv interpretiert, das in den Kategorien des Genitivs, Dativs und Instrumentals Singular und des Plurals durch eine morphologische Regel eingeführt wird. Die phonologische Regel ist durch eine morphologische Regel ersetzt worden. In einem weiteren Schritt wird [ir] im Verlauf des Althochdeutschen dann aufgrund seiner Verteilung im Paradigma (sowie weiterer Faktoren; vgl. Wurzel 1980: 445 ff.; 1992: 282 ff., 291 f.) als Pluralmarker reanalyisiert und entsprechend in seinen Singularvorkommen beseitigt, vgl. 'normalahd.' Sg. Nom./Akk. *hrind*, Gen. *hrind-es*, Dat. *hrind-e*, Instr. *hrind-u*; Pl. Nom./Akk. *hrind-ir*, Gen. *hrind-ir-o*, Dat. *hrind-ir-um*. Die unsystematische morphologische Regel wird so in eine Pluralregel umgewandelt, die als *er*-Pluralregel bekanntlich noch heute existiert; vgl. *Rind – Rinder*.

Hier ist das Ergebnis der Morphologisierung einer phonologischen Regel keine **modifikatorische**, sondern eine **additive morphologische Regel**. Für diese Entwicklung gibt es zwei Gründe: Erstens ist die Ausgangsregel eine phonologische Tilgungsregel und keine Alternationsregel, und zweitens operiert sie am Wortende. Daraus läßt sich verallgemeinernd der Schluß ziehen, daß die Morphologisierung von phonologischen Regeln, die am Wortende phonologische Substanz tilgen oder einführen, zu additiven morphologischen Regeln führt. Das gilt in dieser Form natürlich nur für Suffixsprachen; es ist jedoch

zu erwarten, daß das gleiche auf am Wortanfang operierende Tilgungs- bzw. Epentheseregeln in Präfixsprachen zutrifft.

### 3.3. Die Ableitungsrichtung der resultierenden morphologischen Regel: Beibehaltung vs. Inversion

Bei Morphologisierungen von phonologischen Regeln kann die Ableitungsrichtung der Regeln beibehalten oder umgekehrt werden. In den Beispielfällen 1, 2, 3 und 5 bleibt die Ableitungsrichtung auch nach der Morphologisierung die gleiche. Unumgelautete Vokale werden in bestimmten Kontexten in umgelautete Vokale überführt. Anders im Beispiel 6. Hier wird bei der Morphologisierung eine phonologische Regel mit der Operation ‘ir → Ø’ durch eine morphologische Regel mit der inversen Operation ‘Ø → ir’ ersetzt. Auch über die Beispiele hinaus ist die **Beibehaltung der Ableitungsrichtung** weitaus häufiger als ihre **Inversion**, was darauf hindeutet, daß die Inversion spezifische Bedingungen voraussetzt, was durch einen weiteren Fall von Regelinversion, der Morphologisierung des sogenannten Rückumlauts im Deutschen, bestätigt wird:

Beispiel 7: In einer Teilkasse der schwachen Verben stehen sich im frühen Althochdeutschen bedingt durch die phonologische Umlautregel umgelautete Infinitiv- und Präsensformen einerseits und unumgelautete Präteritalformen andererseits gegenüber. Vgl. z. B. Infinitiv *brenjan* ‘brennen’, 1. Sg. Präs. Ind. *brennu*, 2. Sg. Präs. Ind. *brenis* usw. – 1. Sg. Prät. Ind. *branta*, 2. Sg. Prät. Ind. *brantōs* usw. und Infinitiv *stelljan* ‘stellen’, 1. Sg. Präs. Ind. *stellju*, 2. Sg. Präs. Ind. *stelis* usw. – 1. Sg. Prät. Ind. *stalta*, 2. Sg. Präs. Ind. *staltōs* usw. (Die Geminaten sind dabei durch das folgende unsilbische [i] hervorgerufen.) In der weiteren Entwicklung wird durch phonologischen Wandel das [i] getilgt, wobei die Infinitivendung in [en] übergeht. Es ergeben sich die Formen *brennen*, *brennu*, *brenis* – *branta*, *brantōs* bzw. *stellen*, *stellu*, *stelis* – *stalta*, *staltōs*. Das Auftreten des Umlautvokals [e] im Infinitiv und in den meisten Präsensformen (allen außer der 2./3. Person Singular) ist nicht mehr phonologisch faßbar. Da bei diesen Verben im Infinitiv und im Präsens durchgängig umgelautete Vokale und im Präteritum durchgängig unumgelautete Vokale erscheinen, wird der Vokalwechsel (als eine Art ‘neuer Ablaut’) mit diesen Kategorien assoziiert und damit morphologisiert.

Dabei wird als Ausgangsvokal nicht das umgelautete [a] des Präteritums, sondern das umgelautete [e] des Infinitivs und des Präsens gewählt; die Ableitungsrichtung der Regel wird also von ‘a → e’ zu ‘e → a’ umgekehrt. Daß die Annahme der Inversion korrekt ist, zeigt der spätere weitgehende Abbau der Vokalalternation, der stets zugunsten des Präsensvokals erfolgt; vgl. *stellen* – *stellte*.

Weshalb tritt nun in den beiden letzten Beispielen eine Inversion der Ableitungsrichtung ein? Die anderen Morphologisierungen betrafen jeweils phonologische Regeln, die innerhalb des Paradigmas bestimmte abgeleitete Flexionsformen, aber nicht die Grundform erfaßten; vgl. z. B. ahd. *graban* – 2. Sg. Präs. Ind. *grebis* und urnord. *landa* – Nom. Pl. *londu*. Die formale Ableitungsrichtung der Regeln entspricht der grammatischen Ableitung der Kategorien in den Paradigmen. In solchen Fällen gibt es keinen Grund zur Inversion; die Ableitungsrichtung der Regel bleibt auch nach der Morphologisierung erhalten. Dagegen sind die beiden hier relevanten Beispiele dadurch charakterisiert, daß jeweils eine phonologische Regel morphologisiert wird, die gerade die Grundform des Paradigmas, aber nicht bestimmte abgeleitete Flexionsformen erfaßt; vgl. *hrind* mit *ir*-Tilgung, aber Nom. Pl. *hrindiru* ohne *ir*-Tilgung sowie Infinitiv *brenjan* mit Umlaut, aber Präteritum *branta* ohne Umlaut. Die formale Ableitungsrichtung der Regel ist der grammatischen Ableitung der Kategorien in den Paradigmen gegenläufig. In solchen Fällen wird bei Verlust der phonologischen Bedingungen (zumindest normalerweise) die Ableitungsrichtung der Regel umgekehrt, so daß dann die in der Grundform auftretende Alternative die Eingabe der neuen morphologischen Regel darstellt; vgl. Nom. Sg. *hrind* ohne *ir*-Einführung, aber Nom. Pl. *hrindir* mit *ir*-Einführung sowie Infinitiv *brennen* ohne Vokalwechsel, aber Präteritum *branta* mit Vokalwechsel ‘e → a’. Damit spezifiziert die morphologisierte Regel dann – wie in den Fällen ohne Inversion – einen Marker für eine abgeleitete Kategorie, in den Beispielen den Pluralmarker [ir] bzw. den Präteritalmarker ‘Vokalwechsel’. Daß Inversionen bei Morphologisierungen relativ seltener vorkommen als die Beibehaltung der Ableitungsrichtung resultiert einfach daraus, daß phonologische Alternationen typischerweise durch Segmente in Flexiven hervorgerufen werden, und daß in abgeleiteten (markierten)

Kategorien eher Flexive auftreten als in (unmarkierten) Basiskategorien (Greenberg 1963: 74 f.). Man vgl. dazu noch einmal die [i]- und [u]-haltigen Flexive des Germanischen.

### 3.4. Der Status des neuen Kategorienmarkers: Hauptmarker vs. Nebenmarker

Die Kategorienmarker der Flexionsformen haben nicht alle den gleichen Status. Ein Marker kann die einzige Symbolisierung (*Hund – Hund-e*, *Vater – Väter*), die primäre Symbolisierung (*Wolf – Wölf-e*) oder eine sekundäre Symbolisierung (*Wolf – Wölf-e*) einer Kategorie am Wort sein. In den ersten beiden Fällen liegt ein Hauptmarker, im letzten Fall ein Nebenmarker der entsprechenden Kategorie vor. Welcher Marker bei Mehrfachsymbolisierung den Status des Hauptmarkers hat, ergibt sich aus der Systematik des Sprachsystems. Wenn, was für Sprachen wie das Deutsche typisch ist, eine Kategorie durch einen additiven und einen modifikatorischen Marker symbolisiert wird, so ist im allgemeinen der additive Marker der Hauptmarker, weil dieser in allen Wörtern der betreffenden Flexionsklasse vorhanden ist, während der modifikatorische Marker nur eingeschränkt auftritt. Man denke etwa an die Substantive der deutschen *er*-Pluralklasse wie *Kalb – Kälber* und *Rind – Rinder*, wo das Suffix bei allen Wörtern, der Umlaut (bedingt durch seine phonologische Genese) aber nur bei Wörtern mit hinterem Stammvokal erscheint. Doch durch Morphologisierungen von phonologischen Regeln können sowohl **Hauptmarker** als auch **Nebenmarker** entstehen, wie die Beispiele im einzelnen erweisen.

Morphologisch initiierte Morphologisierungen wie die Fälle 2 und 3 sind immer dem Bestreben der Sprecher geschuldet, formale Alternationen an die grammatischen Kategorien anzupassen. Im Beispiel 2 wird die 1. Person Singular an die anderen Singularformen angeglichen, vgl. 1. Sg. *blōtu* > *blōtu* wie 2./3. Sg. *blōtiR*; im Beispiel 3 der Instrumental Singular an die übrigen Singularformen, vgl. Instr. Sg. *gestiu* > *gastiū* wie Nom./Akk. Sg. *gast* usw. Im ersten Fall resultiert daraus ein Präsens Singularmarker, im zweiten ein Pluralmarker. Da in beiden Fällen vorher keine einheitlichen Marker für die betreffenden Kategorien existierten, erhalten die neuen Marker jeweils den Status von Hauptmarkern. Hingegen ist die Herausbildung

von Nebenmarkern durch morphologisch initiierte Morphologisierungen nur schwer vorstellbar. Dazu müßte bereits ein einheitlicher Marker für die betreffende Kategorie vorhanden sein. Ein solcher einheitlicher Marker würde sich (zumindest in Fällen, die den hier diskutierten Beispielen vergleichbar sind) aber dann auch in einheitlicher Weise auf die Form des Stammes auswirken. Alternationen zwischen den verschiedenen Formen der Kategorie (etwa den einzelnen Kasusformen des Plurals) und damit morphologisch initiierte Morphologisierungen (im Sinne von Ausgleichen) sind unter solchen Bedingungen überhaupt nicht möglich.

Unter den phonologisch initiierten Morphologisierungen sind zwei Gruppen von Fällen zu unterscheiden. Bei der Morphologisierung der Alternationen ahd. 1. Sg. *grabu* – 2. Sg. *grebis* > mhd. 1. Sg. *grabe* – 2. Sg. *grebest* (s. Beispiel 3) und frühahd. Inf. *brenjan* – Prät. Ind. 1. Sg. *branta* > ahd. Inf. *brennen* – Prät. Ind. 1. Sg. *branta* (Beispiel 7) bleiben trotz der phonologischen Reduktionen die alten additiven Marker als Hauptmarker erhalten. Die neuen modifikatorischen Marker erhalten folglich den Status von Nebenmarkern. Davon abweichend verschwinden bei der Entwicklung von vorahd. Nom./Akk. Sg. *hrinda* – Nom./Akk. Pl. *hrindiru* > ahd. Nom./Akk. Sg. *hrind* – Nom./Akk. Pl. *hrindir* die ursprünglichen Marker, so daß das Element [ir] zum alleinigen Pluralmarker und damit zum Hauptmarker wird. Interessant ist in diesem Zusammenhang das Beispiel 5, die Morphologisierung des *u*-Umlauts in Paradigmen des Typs *urnord*. Nom./Akk. Sg. *landa* – Nom./Akk. Pl. *londū*, Dat. Pl. *londum* > nisl. Nom./Akk. Sg. *land* – Nom./Akk. Pl. *lönd*, Dat. Pl. *löndum*. Hier wird aufgrund der phonologischen Reduzierung der alte Marker des Nominativs und Akkusativs Plural beseitigt; der *u*-Umlaut wird zum alleinigen Marker, d. h. zum Hauptmarker. Dagegen bleibt im Dativ Plural der alte Marker erhalten, der neue Marker '*u*-Umlaut' ist entsprechend der Nebenmarker.

Es sei noch darauf verwiesen, daß durch Morphologisierung entstandene Marker im Laufe der weiteren Entwicklung unter entsprechenden Bedingungen durchaus Veränderungen in ihrem Status erfahren können, wie das Beispiel 3 anschaulich zeigt: Mit dem Übergang von Wörtern des Typs *gast* zur mittelhochdeutschen Flexion Sg. Nom./Akk. *gast*, Gen. *gastes*, Dat. *gaste*; Pl. Nom./Akk./

	Initiierung	Weg phon. Regel > morph. Regel	resultierender Regeltyp	Ableitungs- richtung	Markerstatus
B. 1: mhd. <i>grebest</i>	phonologisch	direkt	modifikatorisch	Beibehaltung	Nebenmarker
B. 2: urnord. <i>blōtu</i>	morphologisch	indirekt	modifikatorisch	Beibehaltung	Hauptmarker
B. 3: ahd. <i>gastiū</i>	morphologisch	indirekt	modifikatorisch	Beibehaltung	Hauptmarker
B. 4: ahd. <i>hanin</i>	(morphologisch)	(indirekt)	–	–	–
B. 5: isld. <i>lönd; löndum</i>	phonologisch	indirekt	modifikatorisch	Beibehaltung	Hauptmarker; Nebenmarker
B. 6: ahd. <i>hrind; hrindir</i>	phonologisch	direkt	additiv	Inversion	Hauptmarker
B. 7: ahd. <i>brennen; branta</i>	phonologisch	direkt	modifikatorisch	Inversion	Nebenmarker

Tab. 147.1: Übersicht

Gen. *geste*, Dat. *gesten* entsteht auf phonologischem Weg ein neuer, einheitlicher additiver Pluralmarker *-e* [ə] (vgl. Dat. Pl. *gest-e-n*) mit dem Status des Hauptmarkers; der Umlaut wird so zum Nebenmarker. Doch noch während des Mittelhochdeutschen wird der Umlaut u. a. auf Maskulina übertragen, die über keinen Pluralmarker am Wort (mehr) verfügen, vgl. *vater – veter, nagel – negel*. In diesen Wörtern ist der Umlaut damit wieder alleiniger Marker, also Hauptmarker, für den Plural; bei den übrigen Substantiven bleibt er Nebenmarker.

### 3.5. Zusammenfassung

Abschließend soll in Tab. 147.1 noch kurz zusammenfassend dargestellt werden, wie sich die diskutierten Fälle von Morphologisierungen hinsichtlich der einzelnen Parameter verhalten. Auf diese Weise ergibt sich eine Kreuzklassifikation von Morphologisierungssprozessen, die zwar bei weitem kein vollständiges, aber doch wohl ein recht instruktives Bild des Gesamtphänomens der Morphologisierung von phonologischen Regeln und damit des Weges von der Phonologie zur Morphologie bietet.

## 4. Zitierte Literatur

Árnason, Kristján (1985), “Morphology, Phonology and *u*-umlaut in Modern Icelandic”. In: Gußmann, Edmund (Hrsg.), *Phono-Morphology: Studies in the Interaction of Phonology and Morphology*. Lublin: Catholic University, 9–22

Baudouin de Courtenay, Jan (1895), *Versuch einer Theorie phonetischer Alternationen: Ein Capitel aus der Psychophonetik*. Straßburg: Trübner [Im Faksimile abgedruckt in: Jan Baudouin de Courtenay (1984), *Ausgewählte Werke in deutscher Sprache* (hrsg. von J. Mugdan), München: Fink]

Baudouin de Courtenay, Jan (1908), “O związku wyobrażeń fonetycznych z wyobrażenami morfologicznemi, syntaktycznemi i semazjologicznemi”. In: *Sprawozdania z podziedzeń Towarzystwa Naukowego Warszawskiego*, r. 1, Wydział językoznawstwa i literatury (z. 5/5), 9–28

Baudouin de Courtenay, Jan (1922), *Zarys historji jęziku polskiego*. Warszawa: Polska Składnica Pomocy Szkolnych

Bybee, Joan (1985), *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam: Benjamins

Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York, Evanstone: Harper and Row

Dressler, Wolfgang Ulrich (1977), *Grundfragen der Morphonologie*. Wien: Verlag der Österreichischen Akademie der Wissenschaften

Dressler, Wolfgang Ulrich (1985), *Morphonology: The Dynamics of Derivation*. Ann Arbor: Karoma

Greenberg, Joseph H. (1963), “Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements”. In: Greenberg, Joseph H. (Hrsg.), *Universals of Language*. Cambridge/MA: MIT Press, 59–90

Hooper, Joan B. (1976), *An Introduction to Natural Generative Phonology*. New York: Academic Press

Klausenburger, Jürgen (1979), *Morphologization: Studies in Latin and Romance Morphophonology*. Tübingen: Niemeyer

Klausenburger, Jürgen (1994), “Morphologization”. In: Asher, R. E. & Simpson, J. M. Y. (Hrsg.),

- The Encyclopedia of Language and Linguistics*, Bd. V. Oxford, usw.: Pergamon Press, 2562–2567
- Kruszewski, N[ikolaj] (1881), *Ueber die Lautabwechslung*. Kasan: Universitätsbuchdruckerei
- Morin, Yves-Charles & Langlois, Marie-Claude & Varin, Marie-Eve (1990), “Tensing of Word-final [ɔ] to [o] in French: The Phonologization of a Morphophonological Rule”. *Romance Philology* 43, 507–528
- Mugdan, Joachim (1984), *Jan Baudouin de Courteau (1845–1929): Leben und Werk*. München: Fink
- Paul, Hermann (1880), *Principien der Sprachgeschichte*. Halle: Niemeyer
- Robinson, Orrin W. (1975), “Abstract Phonology and the History of Umlaut”. *Lingua* 37, 1–29

Skousen, Royal (1975), *Substantive Evidence in Phonology*. The Hague: Mouton

Wurzel, Wolfgang Ullrich (1970), *Studien zur deutschen Lautstruktur*. Berlin: Akademie-Verlag (Studia grammatica 8)

Wurzel, Wolfgang Ullrich (1980), “Ways of Morphologizing Phonological Rules”. In: Fisiak, Jacek (Hrsg.), *Historical Morphology*. The Hague: Mouton, 443–462

Wurzel, Wolfgang Ullrich (1992), “Morphologische Reanalysen in der Geschichte der deutschen Substantivflexion”. In: *Folia Linguistica Historica* 8.1–2, 279–307

Wolfgang Ullrich Wurzel †, Berlin  
(Deutschland)

## 148. Analogical change

1. Introduction
2. Definition and exemplification
3. Previous analyses
4. Optimality Theory
5. Conclusion
6. References

### 1. Introduction

Analogical change is a topic which has received relatively little attention in recent handbooks on phonology, morphology, or historical linguistics. The reason for this may well be that it touches on so many different aspects of our linguistic capacity and on different components of linguistic theory that it is not clear where it is most properly treated. This also makes it one of the most interesting topics of investigation, and testing grounds for particular linguistic theories, especially those which seek to unite such different aspects in a unified model.

**Analogical change** is a specific type of **sound change**, simply because the sounds of a word change in the process. However, it has been assumed for centuries that the motivation for analogical changes is to be found not so much in the phonetic/phonological environment of the sound(s) concerned, but rather that other words in the language to which the changing word is morphologically related, play a role. This is where morphol-

ogy and the organisation of the lexicon come in. Finally, all processes of change must be examined in the light of theories of acquisition and variation, which have now developed to such an extent that we may test them against new sets of data. The sometimes erratic process of analogical change is then test material just like more regular sound changes.

### 2. Definition and exemplification

The concept of **analogy** has been used in various branches of science for ages. In linguistics, the Alexandrian grammarians used it to establish grammatical correctness in Homer. More recently, analogy has been used as a shorthand term for the historical phenomenon of analogical change, and, loosely, as a way of indicating the explanation for such historical language changes (cf. Lehmann 1995).

Let us first give a simple example (cited in Joseph 1998; cf. also Meillet & Vendryès 1924: 77; Schindler 1974: 3, 6). In Early Latin, the word for ‘small’ was *parwos* ‘small. NOM.SG’, with genitive *parwī* (see 1a). A process then became active in the language which deleted the rounded semivowel *w* before round vowels (1b).

- (1) (a) Early Latin  
*parwos* ‘small.NOM.SG’  
*parwī* ‘small.GEN.SG’  
(b)  $w \rightarrow \emptyset / \_ [V, +round]$   
(c) Intermediate: *paru:s*, *parvi*  
(d) Classical Latin: *parvus*, *parvī*

As a result, an alternation between *w* and  $\emptyset$  was introduced into the paradigm of the word for ‘small’ (1c). By the time of Classical Latin, this alternation was destroyed again, or levelled out, by restoring the original consonant sound in the nominative, resulting in *parvus*, *parvī* (1d). The latter change is usually referred to as paradigmatic levelling, as its motivation is generally felt to be a restoration of uniformity in the paradigm of this word.

### 3. Previous analyses

The assumption that the cause for analogical levelling is connected to regularity in morphological paradigms dates back at least to the Neogrammarians (cf. Art. 10). Hermann Paul (1920: 205), for instance, concluded that paradigmatic pressure (“Festigkeit des Zusammenhangs der etymologischen Gruppen”), i.e. the pressure to keep a paradigm as uniform as possible, played an important role in this type of change. Note that paradigmatic pressure usually operates in retrospect: it restores regularity after regular phonetic/phonological changes have created alternations in a paradigm: it does not prevent such irregularities from occurring (cf. Art. 32).

For the Neogrammarians, as historical linguists, the linguistic structure of words had very much still a historical outlook: for them a word was defined as its history, and not as its role and function in the synchronic system. This is the reason why the Neogrammarians could not yet gain a full insight into its mode of occurrence. To attain this, comprehension of the Saussurean concept of language as a synchronic system was an indispensable precondition (see Jankowsky 1972: 137). Saussure would also develop the so-called proportional solution to analogical changes, already mentioned by Paul (e.g. Paul 1920: 117), for instance to shed light on another case of analogical levelling, the well known (partial) levelling of the [s] ~ [r] alternation caused by **rhotacism** in Latin (see Kiparsky 1982 for further discussion, among many other sources):

- (2) *ōrātōrem* : *ōrātor* = *honōrem* : x  
x = honor

Saussure also notes the capricious character of analogy. In German, for instance, the plurals in (3a) are regular i.e. etymological Umlaut formations. The plurals in (3b) have become umlauted analogically. However, the plurals in (3c) have resisted analogy for some reason or another. It is therefore impossible to say in advance how far the imitation of a particular model will extend, or which patterns are destined to provoke it (Saussure 1983: 222).

- (3) (a) *Gast/Gäste* ‘guest.SG/PL’  
*Balg/Bälge* ‘skin.SG/PL’  
(b) *Kranz/Kränze* ‘wreath.SG/PL’  
(from earlier plural *Kranza*)  
*Hals/Hälse* ‘neck.SG/PL’  
(from earlier plural *Halsa*)  
(c) *Tag/Tage* ‘day.SG/PL’  
(not \**Täge*)  
*Salz/Salze* ‘salt.SG/PL’  
(not \**Sälze*)

Proportional solutions as explanations of analogical changes are insightfully intuitive and easy to work with. Still, this line of thinking about analogical change came under attack in the 1960s, under influence of the generative paradigm (Chomsky & Halle 1968), which of course had to redraw its own position with regard to language change in general and analogy in particular. One of the critics of the proportional model as an explanatory device was King (1969). One question which must be asked with respect to proportional analogies like that in (2) is just how similar the inputs have to be to give rise to a possible analogical proportion. In this case, it makes intuitive sense to “compare” *ōrātōrem* and *honōrem*, but the question remains if satisfactory conditions on possible inputs can be formulated. Hence, such proportions are useless as an explanatory device, i.e. a device which will predict what analogical changes are possible and what are impossible changes. A second piece of criticism was that in some cases different proportions must be drawn up for different words, which seem to be undergoing the same levelling process, as in the case of the spread of the plural *s* in English: a different proportion must be drawn up for words ending in a vowel, an obstruent or any other consonant. Much effort has therefore gone into introducing more

structure into analogical proportions, so that its effects are better understood, primarily already by Kuryłowicz (1949) and Mańczak (1958). These efforts show that, without morphological analysis, analogical proportions are useless as predictive devices.

In Chomsky & Halle (1968), grammars were conceived of as lexicons complemented by lists of rules. Language change could therefore be conceived of as changes in lexical entries or in changes of the rules of the language. One important idea of historical generative linguistics was that languages become simpler formally. One way of simplifying grammars is to simplify the rules themselves, e.g. in terms of the number of distinctive features affected by the rule. Another way was to lose rules completely from the grammar. Finally, there was much discussion on the question whether some orderings of rules were simpler than others. Kiparsky (1972) argued that grammars evolved into simplicity and that as a result analogical effects were observed. For instance, the levelling which took place after the rule in (1b) had created an alternation in the paradigm, can be simply formalised as the loss at a later stage of that same rule. Similar effects can be obtained by rule reorderings of different types (see Kiparsky 1972).

In more complicated cases of analogical levelling, however, grammar simplification could not be achieved. Take the Latin rhotacism case briefly mentioned above. Because the change from the earlier underlying form /hono:s/ to the later /honor/ was only partial, it could not be expressed as being motivated by the grammar at all, as Kiparsky (1982: 100) notes, although intuitively the change to a uniform paradigm represents a simplification. If the grammar must be complicated to describe processes of analogical levelling, then we must surely be on the wrong track. In response to this, Kiparsky (1971) proposed that **functional constraints** should also help to establish whether a grammar was simpler than an older grammar. To this effect he proposed the constraint given in (4):

(4) **Paradigmatic Uniformity** (Kiparsky 1971: 598)

Allomorphy tends to be minimized in a paradigm.

This constraint evaluates the output of a grammar as a whole, and is in effect a restatement of Hermann Paul's "paradigmatic pressure" mentioned above.

#### 4. Optimality Theory

There are a number of reasons to believe why the recently introduced framework of **Optimality Theory** (OT) (Prince & Smolensky 1993, and much subsequent work) may shed further light on analogical change. First, the traditional assumption is that analogical change is caused by relations between different output forms. Optimality Theory is a theory of grammar in which outputs are selected on the basis of how well they observe general constraints, which are formalised over the outputs. No single output form will observe all constraints, but constraint interaction will ensure that an output form is mapped onto an input. Given this concern for the output, Optimality Theory may be a good candidate for dealing with an evidently output-related phenomenon such as analogical change. Second, analogical change represents a case of conflicting interests: phonological rules will bring about an alternation, and morphological wellformedness conditions (such as (4)) will try to level these out. In principle, such a state of affairs should be readily amenable to an Optimality-theoretic treatment, especially since in recent years proposals have been made to extend this framework with modules describing acquisition and variation (see e.g. Hayes 2000; Boersma 2000; Tesar 2000).

In the case of **analogical levelling**, exemplified in (1), a number of stages must be distinguished. The first stage is the introduction of an alternation as a result of the loss of the rounded semivowel before round vowels. This can be regarded as the effect of a constraint against the adjacency of two elements which are too much alike, the **Obligatory Contour Principle** (OCP, see e.g. McCarthy 1986). This constraint conflicts with a Faithfulness constraint which demands that underlying forms are changed as little as possible. Possibly after a stage of variation, in which the two constraints are somehow "balanced", the Obligatory Contour Principle constraint takes precedence and the semivowel will not be pronounced. Later, a labial consonant is restored. This could be formalised as a reversal of the ranking of the two constraints, on the assumption that underlying forms have remained the same, quite parallel to rule reordering in generative grammar. This begs the question of why the two constraints should be reordered, however. Recent analyses (e.g. Kenstowicz 1996) have

sought the explanation in the interplay of the two constraints mentioned above with a third one, a correlate of Paradigmatic Uniformity (4). The operation of such a constraint would favour outputs which have a uniform base form. An issue here is which direction of levelling is taken: there are always multiple ways in which a paradigm could be regularised, and the question is if we can predict which way will be taken in individual cases.

### 5. Conclusion

The term analogy is used for a wide variety of purposes. Here we focused on one of the most common types of analogical change, viz. analogical levelling. Since the times of the Neogrammarians, it has been suggested that pressure to keep paradigms as uniform as possible plays a role in eradicating phonologically-induced alternations. It turned out to be difficult to formalise this idea in traditional generative grammar. The framework of Optimality Theory may be better suited for the purpose in the light of its insistence on surface constraints and its explicit way of dealing with conflicting tendencies.

### 6. References

- Boersma, Paul (2000), “Learning a Grammar in Functional Phonology”. In: Dekkers et al. (eds.), 465–523
- Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York: Harper and Row
- Dekkers, Joost & van der Leeuw, Frank & van de Weijer, Jeroen (2000, eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*. Oxford: Clarendon Press
- Hayes, Bruce P. (2000), “Gradient Well-Formedness in Optimality Theory”. In: Dekkers et al. (eds.), 88–120
- Jankowsky, Kurt R. (1972), *The Neogrammarians: A Re-Evaluation of their Place in the Development of Linguistic Science*. The Hague: Mouton
- Joseph, Brian D. (1998), “Diachronic Morphology”. In: Spencer, Andrew & Zwicky, Arnold (eds.), *The Handbook of Morphology*. Oxford: Blackwell, 351–373
- Kenstowicz, Michael (1996), “Base-Identity and Uniform Exponence: Alternatives to Cyclicity”. In: Durand, Jacques & Laks, Bernard (eds.), *Current Trends in Phonology: Models and Methods, Vol. II*. Manchester: University of Salford Press, 365–395
- King, Robert D. (1969), *Historical Linguistics and Generative Grammar*. Englewood Cliffs/NJ: Prentice Hall
- Kiparsky, Paul (1971), “Historical Linguistics”. In: Dingwall, William Orr (ed.), *A Survey of Linguistic Science*. College Park: University of Maryland Linguistics Program, 576–649 [reprinted in: Kiparsky (1982)]
- Kiparsky, Paul (1972), “Explanation in Phonology”. In: Peters, Stanley (ed.), *Goals of Linguistic Theory*. Englewood Cliffs/NJ: Prentice Hall, 189–227 [reprinted in: Kiparsky (1982)]
- Kiparsky, Paul (1982), *Explanation in Phonology*. Dordrecht: Foris Publications
- Kuryłowicz, Jerzy (1949), “La nature des procès dits ‘analogiques’”. *Acta Linguistica* 5, 15–37
- Lehmann, Christian (1995), “Synsemantika”. In: Jacobs, Joachim et al. (eds.), *Syntax: Ein internationales Handbuch zeitgenössischer Forschung, Vol. II*. Berlin: de Gruyter (Handbücher der Sprach- und Kommunikationswissenschaft, 9.2), 1251–1266
- Manczak, Witold (1958), “Tendances générales des changements analogiques”. *Lingua* 7, 298–325, 387–420
- McCarthy, John J. (1986), “OCP Effects: Gemination and Antigemination”. *Linguistic Inquiry* 17, 207–263
- Meillet, Antoine & Vendryès, Joseph (1924), *Traité de grammaire comparée des langues classiques*. Paris: Champion
- Paul, Hermann (1920), *Prinzipien der Sprachgeschichte*. Halle: Niemeyer [<sup>1</sup>1880; <sup>2</sup>1886]
- Prince, Alan S. & Smolensky, Paul (1993), *Optimality Theory: Constraint Interaction in Generative Grammar*. Piscataway/NJ: Rutgers University (Technical Report of the Rutgers Center for Cognitive Science 2)
- Saussure, Ferdinand de (1983), *Course in General Linguistics*, translated and annotated by Roy Harris. London: Duckworth [<sup>1</sup>1916]
- Schindler, Jochem (1974), “Fragen zum paradigmatischen Ausgleich”. *Die Sprache* 20, 1–9
- Tesar, Bruce (2000), “Optimality and Strict Domination in Language Learning”. In: Dekkers et al. (eds.), 592–620

*Jeroen van de Weijer, Leiden  
(The Netherlands)*

## 149. Remotivation and reinterpretation

1. Characterization of the phenomena and definition of terms
2. Subclassifications and typical examples
3. Causes and hindering factors of secondary motivation
4. Effects of secondary motivation
5. Different approaches to secondary motivation
6. References

### 1. Characterization of the phenomena and definition of terms

Words which for some reason (non-native origin, nonce-morpheme, phonological structure) cause problems for memory processing and storage sometimes acquire secondary morphological and/or semantic structure; they thus change from **opaqueness** to **transparency**. This stands in contrast to the reverse process of **demotivation** and **obscuration**, by which originally motivated words lose their transparency (cf. Art. 150; see 3.1 for the reasons for the existence of these opposite processes). The process of **secondary motivation** or **remotivation** may take a number of different forms, for which different terms are used (there is, however, no universal agreement on the use of terminology).

**Reinterpretation** is a type of remotivation where no formal change has occurred; reinterpreted words show secondary semantic association with etymologically unconnected elements, which may effect semantic change (Fr. *jour ouvrable*, from Lat. *operari* '[to] work', now reinterpreted as connected with *ouvrir* '[to] open', 2.2.3). The term is also used for certain processes involving the analysis of inflectional affixes as word formation affixes (or vice versa), and for the reanalysis of syntagmatic structures (2.2.6, 2.2.7).

**Metanalysis** is another process which does not involve formal change. Here, the internal morpheme boundary of a word is shifted, so that semantic change as well as secondary analogies may result (standard example: Eng. *hamburg-er* → *ham-burger*, with resulting suffix {-burger}, which becomes productive in analogous formations like *cheese-burger*, *onion-burger*, etc.).

**Re-motivation** (as a specific process of remotivation) occurs when the opaqueness of the word, too, is secondary, i.e. when a word loses its original transparency and after a

period of opaqueness re-acquires it (e.g. through spelling pronunciation, as Eng. *fore-head* > /fɔrid/) > /fɔ:hed/) or (rarely) when a word acquires a new transparency different from the secondary one (= tertiary motivation, e.g. Germ. *Maul-wurf*, see 2.2.2).

**Folk etymology** or **popular etymology** are terms frequently used for secondary motivation, particularly of non-native words (e.g. Latinisms) or obsolete words or words with one obsolete element. The terms are, however, also used synonymously with *secondary motivation*. The term folk-etymology was coined by Ernst Förstemann in his article "Über deutsche Volksetymologie" (Förstemann 1852), but has often been criticized as inappropriate (Ullmann 1966: 34 f.; 1972: 101; Mayer 1962: 5–20; Miettinen 1965: 30; Koziol 1974: 118) because it disregards "erroneous" interpretation or intentional misetymologizing of words by medieval scribes, Renaissance humanists, etc. Such suggestions as *a posteriori motivation* (Ullmann 1959: 91), *associative etymology* (Ullmann 1966: 35), or *synchrone Etymologie* (Bergenholtz 1975) have however not been successful (cf. also Mayer 1962: 11, Fn. 2).

Phenomena related to secondary motivation are **malapropism** (the confusion of Latinisms), **agglutination/deglutination** (see 2.2.7), **back-formation** and **analogization** (see 2.2.8; cf. also Mayer 1962: 25–27), and certain types of **initialization** (Ungerer 1991; see also 2.2.8).

All processes of secondary motivation were for a long time seen as erroneous etymologies or (later) as rarely occurring reactions against the arbitrariness of the linguistic sign and therefore as exceptional if not "pathological" (this word is used in Saussure 1967: 210, orig. 1916: 241). Following Gilliéron (1922), they were later and are now usually viewed as normal and frequent processes which have their cause in the storage problems of human memory (Plank 1981: 196 f.). These processes are now also seen in connection with certain stages of child (first) language acquisition, at which children tend to "wrongly" analyze or overinterpret newly acquired words (Leopold 1949: 144 f. with additional literature; Berko 1958: 168–170; Karpf 1990: 129 f.; Art. 165).

Secondary motivation seems to be a fairly universal phenomenon in language change

and language contact. Examples quoted in the literature range from Ancient Greek, Latin, and Old English to most modern European and some African languages. The literature listed in a comprehensive bibliography (Schreiner 1987) concerns 48 different languages (see also the commented bibliography in Olschansky 1996).

## 2. Subclassifications and typical examples

### 2.1. Criteria of subclassification

A more systematic subclassification than the one implied in the different terms defined in 1 may use the following criteria (see also Olschansky 1996: 178–220 and Blank 1997: 304 f.):

- (a) etymological origin of the remotivated word. This word can either be of foreign origin (2.2.1) or a native word of which (usually) one element has become obsolete and is replaced by a still existing morpheme (2.2.2);
- (b) total or partial remotivation. If the remotivated word was originally opaque, total remotivation as compound or derivative is the normal case. This is the usual process with words of foreign origin, which however sometimes remain partially opaque (2.2.1). With words of native origin, which have usually remained semi-transparent, partial remotivation occurs, since only the obsolete or obscured element has to be remotivated (2.2.2);
- (c) change of phonetic and/or graphemic form, which may (2.2.4) or may not occur (2.2.3; cf. 1, “reinterpretation”; see also Mugdan 1984: 113 on Baudouin de Courtenay’s distinction between “inner” and “outer” folk etymology);
- (d) change of morpheme-structure, which may or may not occur. The remotivation of foreign words usually involves the acquisition of a new morphological structure (usually as compound, 2.2.1), while partial remotivation of indigenous words normally retains the old morpheme-structure (2.2.2). However, the following changes may occur: shift of morpheme-boundary (metanalysis), reanalysis of affix as root (2.2.4), reanalysis of root syllable or word formation affix as inflectional affix or vice versa (2.2.6);

- (e) re-motivation as simplex or as complex word (the different possibilities are listed in Koziol<sup>2</sup> 1972: 317–320);
- (f) other criteria (see the examples in 2.2.7).

### 2.2. Typical examples

In the following lists of examples only the information necessary for understanding the process of remotivation is given. Naturally, each word underwent a number of stages in pronunciation, spelling, and semantic interpretation, and some words show different stages of remotivation in the different varieties (e.g. dialects) of the language. To single out one example, Germ. *Friedhof* ‘peace-yard (cemetery)’ appears in Old High Germ. as *frit-hof*, in Middle High Germ. as *vrīt-hof* ‘fenced-in-yard (forecourt [of house or church])’. “Undisturbed” development would have resulted in Modern Germ. *Freit-hof*, which indeed is the form in certain dialects. Association of the meaning ‘church-yard’ with ‘Friede’ (‘peace’), however, was responsible for the remotivated form *Friedhof*, which is used in most varieties including Standard German. For detailed information concerning the development of the individual examples the reader is referred to the etymological dictionaries (Gamillscheg<sup>2</sup> 1969; Kluge<sup>22</sup> 1989; Onions 1966).

#### 2.2.1. Total remotivation of foreign words

This sub-section and the next contain the most frequently quoted examples of folk-etymology. Typically, in German and English, a polysyllabic word of Romance origin undergoes secondary motivation (other origins, however, are also possible); alternately, an obsolete element of an indigenous word is replaced by a still existing morpheme (see 2.2.2).

- Eng. *beak-iron* < Old Fr. *bicorne* ‘two-horn’;
- (American) Eng. *carry-all* < Fr. *carriole* ‘small cart’;
- Eng. *cause-way* < Fr. *chaussée* < Lat. *calciata (via)* ‘lime-road’;
- Eng. *Charter-house* < Fr. *chartreuse* < Anglo-Norm. *chartrouse* ‘Carthusian monastery’;
- Eng. *cut-lass* < Fr. *couteau* < Lat. *cultellus* ‘knife’;
- Eng. *goose-berry* (with obscured pronunciation) < Fr. *groseille*;
- Eng. *lance-knight* < Germ. *Lands-knecht* (with remotivation of *Lands-* as *lance*, and

- etymologically, but not semantically correct interpretation of *-knecht* as *-knight*) ;
- Eng. *leg-horn (hen)* < It. *Legorno* (now *Livorno*);
  - Eng. *man-drake* < Medieval Lat. *mandragola*;
  - Eng. *pent-house* < Old Fr. *appentis* < Medieval Lat. *appenditium* ‘that which is attached’;
  - Eng. *prim-rose* < Fr. *primerole* < Lat. *prim(er)ula* ‘firstling’;
  - Eng. *rose-mary* < Lat. *ros marinus* ‘sea-dew’;
  - Eng. *sparrow-grass* (dialectal) < Lat. *aspasagus* (see also 2.2.8);
  - Fr. *chou-croute* < Alsatian *sur-krut* (= Germ. *Sauerkraut*);
  - Fr. (*jeu de*) *l'âne salé* < Eng. (*game of*) *Aunt Sally*;
  - Germ. *Arm-brust* < Lat. *arcuballista* < *arcus* ‘bow’ + *ballista* ‘catapult’ (< Gk. *ballen* ‘[to] throw’);
  - Germ. *Fell-eisen* < Fr. *valise*;
  - Germ. *Hänge-matte* < Du. *hang-mat* (here the remotivation occurred) < Sp. *hamaca* < Taino (Haiti) *hamaka*;
  - Germ. *Murmel-tier* < *murmuntin* < Lat. *murem montis* ‘mountain-mouse’;
  - Germ. *Viel-fraß* < Norw. *fjell-fross* ‘mountain-cat’;
  - Germ. *Wetter-leuchten* < *Wetter-leichen* ‘weather-dancing’.

The following examples show only partial remotivation of foreign words, which has resulted in formations containing a nonce-morpheme:

- Eng. *cray-fish* < Fr. *écrevisse* < Middle High Germ. *krebis*;
- Eng. *cater-pillar* (with possible analogy to *caterwaul*) < Fr. *chatepeloze* ‘hairy cat’;
- Eng. *fe-male* < Fr. *femelle*;
- Eng. *mush-room* < Fr. *mousseron*.

## 2.2.2. Partial re-motivation of native words

Partial re-motivation is found in the following examples:

- Eng. *black-mail* < *black-māl* ‘black agreement’;
- Eng. *bride-groom* < *brīde-guma* ‘bride-man’;
- Eng. *mistle-toe* < *mistle-tān* ‘mistle-twigs’;
- Eng. *sand-blind* < *sam-blind* ‘half-blind’
- Eng. *shame-faced* < *seam-faest* ‘shame-bound’;
- Eng. *tit-mouse* < *tit-māse* ‘small bird’;

- Eng. *wed-lock* < *wed-lāc* ‘wedding gift, pledge’;
- Germ. *Fast-nacht* < *fase-nacht* ‘roving-night’;
- Germ. *Fried-hof* < *vrīt-hof* ‘fenced-in yard’;
- Germ. *Hage-stolz* < *hagu-stalt* ‘owner of (fenced-off) yard’;
- Germ. *Maul-wurf* < *molt-werf* ‘earth-thower’ < *mūwerf* ‘heap-thrower’ with re-motivation (tertiary motivation);
- Germ. *Lein-wand* < 17th c. *lein-wāt* ‘linnen-garment’;
- Germ. *Sünd-flut* (dialectal) < *sint-vluot* ‘great flood’.

The following example shows remotivation of a foreign element in a hybrid word:

- Germ. *Elfen-bein* < *helfent-pein* ‘elephant-bone’.

## 2.2.3. Remotivation without change of form (“reinterpretation”)

The examples in 2.2.1 and 2.2.2 all involve phonetic and/or graphemic change of the original word. However, examples of merely semantic reinterpretation without formal change also occur (see also the examples in Mayer 1962: 140–233):

- Eng. *auburn* < Lat. *alburnus* ‘whitish’, now associated with *burn*;
- Fr. *ouvrable* < Lat. *operari* ‘[to] work’, now associated with *ouvrir*;
- Germ. *irritieren* ‘[to] confuse, mislead in an annoying way’ < Lat. *ir-ritare* ‘[to] excite’, now associated with *irre-(machen* etc.).

This type may also involve homonymic elements, as in the association of Eng. *breakfast* with *fast* (adj.).

The following words are examples of remotivation through graphemic change (with no change in pronunciation):

- Eng. *island* (with learned introduction of <s> from Lat. *insula* (Old Eng. *īglānd*, Middle Eng. *īland*);
- Eng. *sovereign* (with introduction of <g> and association with *reign*);
- Eng. *posthumous* (with introduction of <h> and association with Lat. *humus*);
- Eng. *hiccup* (with association with *cough*).

Change from “obscured” to “transparent” pronunciation (with no change in spelling) is now common with a number of English

words whose spelling is still transparent. This phenomenon of “spelling pronunciation” is a type of re-motivation. Examples are *fore-head*, *waist-coat*, and *boat-swain*.

#### 2.2.4. Remotivation with change in morphemic structure

Examples of the type Eng. *beak-iron* and Germ. *Fell-eisen* (2.2.1) show acquisition or change of morpho-semantic structure (at least as far as the remotivating language is concerned), while those of the type Eng. *black-mail* and Germ. *Fast-nacht* (2.2.2) retain an already existing structure. Examples of change of morphemic structure in indigenous words are the following:

- (occasional) Eng. *cow's lip* < *cow-slip* < *cū-slyppe* ‘cow- dung’;
- Eng. *worm-wood* < *wermod* (monomorphemic Old English word);
- Germ. *Ein-öde* < *ein-oete* (suffix reinterpreted as root).

Remotivation of compound as derivative or vice versa occurs in:

- Germ. *weis-sagen* < *wīzagōn* ‘act as prophet’ (derivative of *wīzag* ‘knowing’, now interpreted as compound);
- Eng. *brid-al* < *bride-ale* < Old English *brīd-ealo*;
- Germ. *Junk-er* (with doubtful transparency) < *jungk-herr*;
- Eng. *fore-most*, *inner-most*, *utter-most* with reinterpretation of superlative suffix *-mest* as *most*.

#### 2.2.5. Remotivation as simplex

Most types of secondary motivation result in complex words (more frequently compounds than derivatives). The following are examples of monomorphemic remotivation:

- Eng. *jerusalem* (type of artichoke) < It. *girasole*;
- Germ. *Kater* (‘hangover’) < *Katarrh*.

Monomorphemic remotivation must be kept apart from word confusion and resulting semantic convergence (*ravel/rove*, *sheen/shine*, etc.), where two words of different origins, but similar form and already related meanings approach each other semantically even further.

#### 2.2.6. Remotivation involving inflexion

The reinterpretation of inflectional endings occurred with a number of Latin neuter plurals, which appear in French as feminine singulars, e.g.:

- Fr. *corne* < Lat. *cornu-a*;
- Fr. *levre* < Lat. *labr-a*;
- Fr. *voile* < Lat. *vel-a*.

Conversely, the Old Northern French singular form *cherise* was reinterpreted in English as the plural form *cherries*.

#### 2.2.7. Remotivation involving syntax

In some examples remotivation transcends the word boundary, as when the combination article + noun is reanalysed “wrongly” through association with another word (“agglutination” and “deglutination”), e.g.:

- Eng. *an aitch-bone* < *a nache-bone* < Fr. *nache* ‘buttock’;
- Eng. *a nick-name* < *an ēke-name* ‘an also-name’.

Secondary motivation may also affect syntagmas, which in the process undergo changes in syntactic structure or are reinterpreted as complex or simple words:

- Eng. *atone* < (*maken*) *at ūn* ‘make whole, reconcile’;
- Eng. *help-mate* < *help meet (for him)*, (from Genesis 2: 18,20, with misunderstanding of *meet* = ‘suitable’);
- Eng. *upside down* < *up so down*;
- Germ. *während des Sommers* < *währendes Sommers*.

Occasionally, an opaque word is reanalysed as a phrase, as the name of a coarse material, *mungo*, which in Yorkshire is associated with “it mun go” (it must go), allegedly said by the inventor of the machine for it (Weekley<sup>5</sup> 1961: 161).

#### 2.2.8. Other examples

The following examples show remotivation of a simplex (of either native or foreign origin) with expansion by addition of explanatory determinatum:

- Eng. *rein-deer*, Germ. *Renn-tier* < Icel. *hreinn*;
- Germ. *Lind-wurm* < *lint* ‘winding’;
- Germ. *Maul-beere* < Lat. *mūrus*;
- Germ. *Maul-tier* < Lat. *mūlus*;
- Germ. *Wind-hund* < *wint* ‘Vendic dog, greyhound’.

In some cases, remotivation operates simultaneously with other processes, particularly blending (cf. Art. 91). Eng. *run-a-gate* is both a remotivated word (from Fr. *renégat*) and a blend of this with *run-away*.

Formal analogizing, rather than remotivation, occurred in Eng. *lar-board* < *lade-board* ‘loading side’, which was brought in line with its opposite, *starboard* ‘steering side’, whereby partial loss of transparency was compensated with formal similarity to a frequent collocate. A quite different type of analogizing was the source of the semitransparent word-creation *mono-kini*, which owes its origin to the remotivated word *bi-kini* (originally from *Bikini Atoll*).

Initialization, i.e. the formation of **acronyms**, may lead to secondary motivation if the sequence of letters/phonemes results in a pronounceable word, particularly if this has some semantic relation with the initialized phrase. Thus such acronyms as *AIDS* (= acquired immune deficiency syndrome), *CALL* (= computer assisted language learning), *CARE* (= Cooperative for American Relief to Europe), and *WASP* (= White Anglo-Saxon Protestant) are primarily and secondarily motivated for some speakers (cf. Art. 92; for further discussion see Ungerer 1991).

Secondary motivation may operate in loan-translation, as is the case with Germ. *Tausend-giilden-kraut*. For the formation of this compound, the Latin word *centaureum* (< Gk. *kentauros*) was reanalyzed into the elements *centum* ‘a hundred’ and *aurum* ‘gold’ (see also 5.2).

Secondary motivation is sometimes restricted to a short period in the history of the word. Thus the remotivation of Eng. *coward* as *cow-heart* or *cow-herd* and *abundance* as *hab-undance* was only transitory, and the process was reversed. Remotivated forms may also persist, or exist only, in non-standard varieties: “the corruption of *asparagus* to *sparrow-grass*, which is now regarded as vulgar [...] was in good use in the eighteenth and early nineteenth centuries” (Greenough & Kittredge 1900: 334).

### 3. Causes and hindering factors of secondary motivation

#### 3.1. Causes

Secondary motivation is commonly attributed to the need of the language user for transparency and association with familiar elements (Koziol ²1974: 117), “the desire to motivate what is, or has become, opaque in language” (Ullmann 1972: 102) or an “Urbedürfnis des Menschen nach Deutung und Stützung der Wörter, die sich durch feh-

lende Motivation seinem Verständnis entziehen,” (Scheler 1977: 116, cf. also Leisi ¹1985: 76–79, where remotivation is discussed as a reaction against the “dissociation” of “hard words”). This explanation, however, is not entirely satisfactory in view of the existence of the reverse process of obscuration, which is almost as frequent as that of remotivation. The assumption of “two opposite tendencies [...] at work all the time in the development of language,” (Ullmann 1972: 93), or of “two rival factors, conventionality and motivation” (Ullmann ²1959: 92) is not convincing if it is not further explained. Somewhat more plausible is the assumption of two attitudes towards language, a naive, “non-reflective” one associated with the speaker’s interest in an economical use of language, and a “reflective” one resulting from the speaker’s need to make him/herself understood and from the hearer’s search for meaningful elements (Fill 1980: 25–27); these attitudes are thought of as present in both hearer and speaker, with one of them predominating depending on the situation – but each of them can also be thought of as having been prevalent in different periods, as for instance the reflective one, which favours secondary motivation, in the Renaissance period.

The opposite process of obscuration may even have favoured remotivation in some cases. Obscured and transparent forms of compounds frequently existed side by side, so that with words showing some formal relation to the opaque form hypercorrect formations began to appear. Thus the existence of the forms *gozzard* and *goose-herd*, *shepherd* and *sheep-herd* may have been an important factor in the remotivation of *coward* as *cow-herd* (Götz 1971: 22 f.).

Other explanations seek the need for transparency in the facilitation of language acquisition and word retention. These explanations are supported by evidence from the language of aphasics and from the linguistic associations of children. Examples: “*breakfast* is called *breakfast* because you have to eat it fast when you rush to school”, “*Friday* is a day when you have fried fish” (Berko 1958: 170), *chaotisch* – wenn jemand am Boden liegt und zusammengeschlagen ist < k. o. (Karpf 1990: 129), *chicken pox* – *chicken pots*, *ice cream cone* – *ice cream comb* (Bolinger 1968: 105), *forec'sle* reinterpreted as *fox-hole* by two children in Richard Hughes’ novel *A High Wind in Jamaica* (cf. Fill

1980: 26). Ad-hoc folk-etymologies of this kind are also provided by parents teaching new words to their children (Germ. *Bagger* explained as *Pack-er*, cf. Fill 1980: 26).

Frequently certain specific remotivations first appear in certain varieties of a language, e.g. in a specific dialect, or even in the idiolect of individual speakers. The association of Eng. *ballistic* with *ball*, of *mysticism* with *mist*, of *standard* with *stand* (cf. Waldron 1967: 124), of *hurricane* with *hurry* (cf. Mayer 1962: 197) and of Germ. *Gastronomie* (< Gk. *gaster* ‘stomach’) with *Gast*, *Dienstag* with *Dienst* and *Freitag* with *frei* (Weisgerber 1950: 60f.) is certainly still specific, not universal. Idiolectal remotivations may, however, spread and become “accepted”.

Another cause of remotivation concerns place names which conquering nations or new settlers fail to understand and try to adapt to their own language. “The first American hearing a French trapper in Colorado refer to the *Purgatoire River* judged that what he heard was *Picket Wire* with a foreign accent” (Bolinger 1968: 104f.). Other American examples of this type are *Smackover* < Fr. *Chemin Couvert* and *Low Freight* < *L'Eau Frais*. The Old Eng. place name *Eofor-wīc* (now York), ‘boar-town’ is remotivated from Celtic *Eburacum*, *Bear Park* in Durhamshire from Fr. *Beau Repas* (Görlich 1974: 84). Examples of remotivated place-names in German (*Mai-land* < *Mediolanum*, *Merse-burg* < Slavic *-bor* ‘wood’, etc.) are already given by Förstemann (1852: 15f.). For a discussion of folk-etymology in place-names see Vennemann (1999).

Until the beginning of the 19th century, the attempts of etymologists to explain the origin of words were frequently unscholarly guesses wide off the mark. Medieval writers, in particular, supplied examples of “learned folk-etymology” (Weekley<sup>5</sup> 1961: 156) or “guesstymology” (Bolinger 1968: 105) aimed at edification rather than presentation of the facts. Philippe de Thaun (12th century), for instance, explained Fr. *lundi* as ‘day of light’ (*lumière*), *mercredy* as ‘day of market’ (*marché*), *jeudi* as ‘day of joy’ (*joie*), *vendredi* as ‘day of truth’ (*vérité*), etc. (Weekley<sup>5</sup> 1961: 156). Wild etymologizing of the type “*lucus a non lucendo, canis a non canendo*” (= ‘grove from not being light, dog from not singing’) was also frequent in Classical Antiquity and in the Renaissance period. However, very few of these learned mis-etymolo-

gies left any traces in actual language use, so that as a cause of folk-etymology they can be disregarded.

### 3.2. Hindering factors

There are only few factors which hinder or prevent secondary motivation. Remarkably, it is not prevented or hindered by initial lack of semantic relation between the parts and the whole of the word. Analysis into familiar elements seems to be more important at first than the semantic connexion of these elements with the whole. With compounds, frequently one root element shows a semantic relation with the whole, the other one being completely arbitrary (cf. Eng. *mush-room* < *mousseron*, Eng. *cut-lass* < *coutelas*). Here, morphological transparency seems to be more important than constructional transparency, i.e. rule-dependent semantic relation between the elements and the whole (Plank 1981: 195f.). With non-compounds, a semantic relation is sometimes re-established through semantic change (at least on the connotative level) of the remotivated word (see 2.2.3).

The laws of sound-change likewise do not prevent or hinder secondary motivation. Baudouin de Courtenay already characterizes “outer” folk etymology as one of the changes which do not follow these laws (cf. Mugdan 1984: 113). Indeed, remotivation involving phonetic change almost by definition defies sound-laws: the need for analysis into existing elements seems to be stronger than the rules for semantic derivation or phonetic change. Phonotactics, however, may sometimes play a role as a cause of secondary motivation (Mayer 1962: 54–61).

A factor which has sometimes hindered or even reversed folk-etymological processes can be seen in the influence of (scholarly) etymologists, lexicographers, and orthoepists, particularly of the 19th century. Especially in German, a number of words which had undergone folk-etymological changes in previous centuries (e.g. *Sprichwort*, which in the 16th century had adopted the form *Sprüchwort* from *Spruch*) returned to their original form under this influence (Andresen<sup>4</sup> 1886: 17f.).

### 4. Effects of secondary motivation

Those effects of secondary motivation which are connected with its causes are self-evident: morphological and/or semantic transparency

are maintained or re-established, and problems of pronunciation are removed. However, the new morphological structure (or connexion) naturally influences the semantic structure of the word and may result in semantic changes on the denotative or the connotative level. Frequently, certain semantic elements are given more prominence than before the remotivation. This is the case, for instance, with 'Maul' in Germ. *Maulwurf* and *Maulesel*, with 'Sünde' in Germ. *Sündflut* or with 'bell' in Eng. *belfry* (< Old Fr. *berfray* < Middle High Germ. *ber(c)-vrit*). Other words develop an occasional meaning strongly influenced by a secondary element (Eng. *pundit*, also 'sophist' with influence from *pun*). The usual meaning of a word, too, can be affected by secondary motivation (e.g. Germ. *irritieren*).

Folk-etymology has frequently influenced popular beliefs concerning saints, sects, plants, and animals, and has given rise to anecdotes about the origin of customs and to legends about the origin of place-names (for examples see Weisgerber 1950: 56–66 and Kainz 1972: 366–374, who also discusses prejudices and superstitions originating in secondary motivation). A well-known English example of an anecdote inspired by remotivation is the story (told by Jonathan Swift of King James I) of the knighted loin of beef (*sir-loin*, really from Fr. *surlonge* 'above the loin', cf. Weekly 51961: 162; Greenough & Kittredge 1900: 331; Mayer 1962: 95).

**Child-etymology** has been shown to be an important phenomenon in child language acquisition, since it facilitates the acquisition of a more comprehensive vocabulary by establishing semantic links between words already stored in the brain and words to be newly acquired (Leopold 1949: 114). To what extent these early attempts at motivation persist and have an influence on the thinking of the grown-up still remains to be investigated.

Other effects concern intentional remotivation by writers and journalists. These effects can be either humorous or poetic. A humorous example is the formation *Vi-queens* to Vikings (Weekly 51961: 142; quoted from *Punch*). Comic effects are also intended by Shakespeare in *The Merchant of Venice* ("more than sand-blind, high-gravel blind" Act II, scene ii), and Friedrich von Schiller in *Wallensteins Lager* (the "Kapuzinerpredigt", scene 8). For didactic or poetic purposes remotivation has been used by many writers

from Bede (*Angli – Angeli*) to James Joyce (particularly in *Finnegans Wake*; see also the examples in Mayer 1962: 288–310). Effects of both comic relief and characterization are achieved by "malapropisms" in *The Rivals* by R. B. Sheridan and *Der Zauberberg* by Thomas Mann.

## 5. Different approaches to secondary motivation

Secondary motivation has been investigated with different approaches depending on the different levels of language involved, and from various theoretical positions adopted by the different schools of linguistics. The different linguistic levels from which it has been approached are the phonological (Ullmann 1972: 101; Mayer 1962: 54–79, 234–287), the morphological (Erben 1964, Koziol 21972: 317–320, Ullmann 1972: 101–105), the semantic (Waldron 1967: 124 f., 153 f.; Koziol 21974: 117–121; Mayer 1962: 127–233), and the psycho-pragmatic (Kainz 1972: 366–374). There are, however, no comprehensive treatments of remotivation which concentrate specifically on any one of these levels. The different approaches depending on linguistic school and theoretical background will be sketched out in the following sub-sections.

### 5.1. The historical approach

Secondary motivation, under the term *folk-etymology*, was already investigated in some detail in the second half of the 19th century. The interest of scholars was, however, chiefly historical. Three types of etymology were distinguished, the popular, the learned and the scholarly, of which the popular was held to be the oldest (Fürstemann 1852: 2). In accordance with the aims of Comparative Philology, examples of folk-etymology were investigated in many ancient languages (Ancient Greek, Latin, Gothic, Anglo-Saxon, Old Norse, etc.) and a few modern ones, particularly German, English, and the Scandinavian languages (Andresen 1876; Fürstemann 1877).

Folk-etymology was judged as "distortion of words" ("Verunstaltungen der Wörter", Fürstemann 1852: 7; "volksetymologische Entstellung", Andresen 41886: 16), as "false derivation" and "verbal corruption" (Palmer 1882: subtitle) or as "the tendency of the un-educated [...] to distort an unfamiliar or un-

intelligible word into some form which suggests a meaning" (Weekley<sup>5</sup> 1961: 105 f.). Evaluative judgments of this kind have occasionally persisted into the 20th century (in Saussure 1916 folk-etymology is discussed in the diachronic section as a rare and pathological phenomenon).

Since folk-etymology defies any claims concerning the regularity of sound-change (claims which were made in the 19th century by the Neogrammarians), it was frequently regarded from the onomasiological point-of-view, and remotivated words were categorized according to object groups or conceptual categories (proper names, place-names, words denoting plants, animals, etc., Andersen 1876). There was also a tendency to see folk-etymology largely as a thing of the past, at least as far as the standard language was concerned (Förstemann 1852: 14). In a number of publications, folk-etymology is discussed as one of the strange and curious adventures which words may undergo in the course of their romantic history (Greenough & Kittredge 1900; Weekley<sup>5</sup> 1961).

### 5.2. The synchronic approach

Since the advent of Structuralism and the possibility of a synchronic approach, the majority of scholars have taken a more tolerant and even positive attitude towards the phenomena under discussion. This change in attitude has also affected judgments concerning the frequency and statistical importance of the phenomena. While scholars adopting the historical approach tended to view folk-etymology as an exceptional and rare phenomenon restricted to a small number of examples of the type listed in 2.2.1 and 2.2.2, secondary motivation was later and is now regarded as frequent, natural, and indispensable for a living language: "Sekundäre Motivation offenbart sich als ein durchgängiges Prinzip der Sprache" (Mayer 1962: 348). Consequently, scholars interested in secondary motivation no longer restrict their research to previous stages of linguistic development (e.g. the period between the 14th and the 17th century), but increasingly take into account recent and contemporary instances. There is a growing awareness that remotivation and the phenomena related to it (e.g. rhyme association) are extremely widespread, particularly in colloquial speech and substandard varieties, and that the instances brought to the notice of researchers represent only a fraction of the actually existing cases (Miettinen

1965: 29). Interest also no longer focuses on the word alone, but scholars now consider syntagmatic and syntactic reinterpretation to be at least closely related to folk-etymology (Mayer 1962: 310–321).

With the synchronic approach and the introduction of the term "motivation", the phenomenon under discussion began to be seen as a "force" (Mayer 1962: 349) and an expression of the creativity of a language ("schöpferisches Mißverständnis", Knobloch 1968). This view is most strongly manifested in the so-called "inhaltsbezogene" and "leistungsbezogene Sprachbetrachtung" (Weisgerber 1950: 56–66), where folk-etymology is discussed as one of the powers which create and determine folk-customs and types of behaviour, and from which popular beliefs about animals (*Vielfraß*), plants (*Tausendgäuldenkraut*), patron saints (*St. Vincent – vin*), etc. are derived.

Folk-etymology is now seen as "acquisition of morphological and semantic motivation" (Ullmann 1972: 101–105). In an extension of this approach, the following types of secondary motivation have been distinguished (Mayer 1962; see also the criteria of subclassification in 2.1):

- phonological with mere adjustment on the level of sounds (e.g. Eng. *shamrock* < Ir. *seamróg* 'small clover');
- phono-semantic, roughly corresponding to Ullmann's "morphological motivation" (e.g. Eng. *beak-iron* < Fr. *bi-corne*);
- semantic motivation, i.e. semantic association without phonological change (e.g. Engl. *hurricane*, with association to *hurry*);
- onomatopoeic motivation, i.e. acquisition of sound-symbolism (e.g. Eng. *rush* < Fr. *reuser*);
- intentional secondary motivation (e.g. *Vi-queen*, her-story; see also the examples discussed in 4);
- parallel syntactic developments (Eng. *help-mate* < *help meet* (noun + adjective)).

Particularly with the inclusion of "semantic motivation" (in Mayer 1962 about 140 examples from English are discussed), secondary motivation becomes a frequent every-day phenomenon with idiolectal differences, or rather a dynamic process based on the "synchronic etymological competence" of the speaker (Augst 1975: 156–230; see also 5.3).

The generative-transformational approach to language has contributed to viewing secondary motivation as a process, which goes on in the individual, rather than as a historical fact. However, this approach has not played an important role in the further investigation of the phenomena. Folk-etymology is too “irregular” to admit the setting-up of rules. The products of secondary motivation do not normally show constructional transparency, since frequently only one element contributes to semantic motivation (Plank 1981: 195 f.). Besides, remotivated words do not follow one particular pattern of word-formation; rather, they show a variety of structures, from endocentric noun + noun structures (e.g. *Fried-hof*) to exocentric noun + noun (e.g. *Arm-brust*) or verb + noun (e.g. *cut-lass*) structures. In addition, most of them must be regarded as highly idiomatized complex words. As far as their generation is concerned, they are fairly isolated and lexicalized formations, for whose production no generalization is possible (Dressler 1976: 157).

**5.3. Psychological and sociological aspects**  
 Psychological aspects concerning the role of memory were first drawn attention to around the turn of the century (Paul 1909: 221 f.), when folk-etymology was also explained as a phenomenon related to a specific type of slip of the tongue, viz. substitution (Meringer & Mayer 1895: 76), slips being regarded as non-pathological phenomena stemming from constantly operating psychic forces which are subject to rules (Meringer & Mayer 1895: 9). In contrast to this, more recent psychological explanations of folk-etymology seek its causes in the “interpretative correction” of unfamiliar words and phrases (Bolinger 1968: 104).

The psychological aspects of secondary motivation have received special attention with the investigation of child language acquisition. Many children have private meanings for compounds, which they analyse in their own way, without being aware of or considering the history of the word (Berko 1958: 169). It is now generally accepted that child etymology is an important factor in language acquisition and concept-formation (Leopold 1949: 114 f.; Porzig 1967: 203; Fill 1976: 14 f.; Karpf 1990: 129). It has even been suggested that secondary motivation, which is thought to be based on synchronic etymological competence and to precede generative competence, can be practised to facilitate fur-

ther language acquisition (Augst 1975: 178). In child etymology, as in folk-etymology, the formal similarity need not be very close, as in *pretty coat* for *pettycoat* (Leopold 1949: 115). Likewise, constructional transparency (i.e. rule-based relation between form and meaning) is probably aimed at, but rarely achieved, as in the example “it is called *mushroom*, because it has rooms in it” (Fill 1976: 14). Social aspects of remotivation, which concern for instance its particular frequency in non-standard social and regional varieties of a language, have received attention from the very beginning (Förstemann 1852: 20–22; Andresen 1876). More recently, it has been shown that remotivation and the processes related to it are subject to variation and fluctuation in the different dialects and sociolects, and that no dividing line can be drawn between accidental formations due to carelessness or ignorance and intentional word associations created for comic and other effects (Miettinen 1965: 29 f.). Although dialects abound in remotivated formations, there is as yet no comprehensive study of the phenomena under discussion in specific dialects, let alone in larger groups of non-standard varieties.

## 6. References

- Andresen, Karl Gustaf (1876), *Über deutsche Volksetymologie*. Heilbronn: Henninger
- Andresen, Karl Gustaf (1886), *Sprachgebrauch und Sprachrichtigkeit im Deutschen*. Heilbronn: Henninger [1880]
- Augst, Gerhard (1975), *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*. Tübingen: Narr (Forschungsberichte des Instituts für deutsche Sprache 25)
- Bergenholtz, Henning (1975), “Volksetymologie oder synchrone Etymologie”. *Muttersprache* 85, 89–94
- Berko, Jean (1958), “The Child’s Learning of English Morphology”. *Word* 14, 150–177
- Blank, Andreas (1997), *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen: Niemeyer
- Bolinger, Dwight (1968), *Aspects of Language*. New York etc.: Harcourt, Brace & World
- Dressler, Wolfgang (1976), “Das Zusammenspiel verschiedener Ebenen und Prozesse in der diachronen Wortbildung: Thesen”. In: Panagl, Oswald (ed.), *Wortbildung diachron – synchron*. Innsbruck:

- Inst. für Sprachwissenschaft (Innsbrucker Beiträge zur Sprachwissenschaft 18), 155–157
- Erben, Johannes (1964), “Deutsche Wortbildung in synchronischer und diachronischer Sicht”. *Wirken des Wort* 14, 83–93
- Fill, Alwin (1976), “Synchrone oder diachrone etymologische Kompetenz”. *Klagenfurter Beiträge zur Sprachwissenschaft* 2.2–3 (*Etymologie* I, ed. Oswald Panagl & Hans-Dieter Pohl), 3–16
- Fill, Alwin (1980), *Wortdurchsichtigkeit im Englischen*. Innsbruck: Institut für Sprachwissenschaft (Innsbrucker Beiträge zur Sprachwissenschaft 24)
- Förstemann, Ernst (1852), “Über deutsche Volks-etymologie”. *Zeitschrift für Vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* (= *Kuhns Zeitschrift*) 1, 1–25
- Förstemann, Ernst (1877), “Über deutsche Volks-etymologie”. *Zeitschrift für Vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* (= *Kuhns Zeitschrift*) 23, 375–385
- Gamillscheg, Ernst (2<sup>nd</sup>1969), *Etymologisches Wörterbuch der französischen Sprache*. Heidelberg: Carl Winter Universitätsverlag [1928]
- Gilliéron, Jules (1922), *Les étymologies des étymologistes et celles du peuple*. Paris: Champion
- Görlach, Manfred (1974), *Einführung in die englische Sprachgeschichte*. Heidelberg: Quelle & Meyer
- Götz, Dieter (1971), *Studien zu den verdunkelten Komposita im Englischen*. Nürnberg: Carl
- Greenough, James Bradstreet & Kittredge, George Lyman (1900), *Words and Their Ways in English Speech*. New York: Macmillan [paperback reprint 1961]
- Kainz, Friedrich (1972), *Über die Sprachverführung des Denkens*. Berlin: Duncker & Humblot
- Karpf, Annemarie (1990), *Selbstorganisationsprozesse in der sprachlichen Ontogenese: Erst- und Fremdsprache(n)*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 352)
- Kluge, Friedrich (2<sup>nd</sup>1989), *Etymologisches Wörterbuch der deutschen Sprache*, ed. by Elmar Seibold. Berlin, New York: Walter de Gruyter [1883]
- Knobloch, Johann (1968), “Das schöpferische Mißverständnis”. *Lingua* 21, 237–249
- Koziol, Herbert (2<sup>nd</sup>1972), *Handbuch der englischen Wortbildungslehre*. Heidelberg: Winter [1937]
- Koziol, Herbert (2<sup>nd</sup>1974), *Grundzüge der englischen Semantik*. Wien, Stuttgart: Braumüller [1967]
- Leisi, Ernst (7<sup>th</sup>1985), *Das heutige Englisch*. Heidelberg: Winter [1955]
- Leopold, Werner F. (1949), *Speech Development of a Bilingual Child: A Linguist's Record, Vol. III: Grammar and General Problems in the First Two Years*. New York: AMS Press
- Mayer, Erwin (1962), *Sekundäre Motivation: Untersuchungen zur Volksetymologie und verwandten Erscheinungen im Englischen*. Köln: Universität Köln
- Meringer, Rudolf & Mayer, Karl (1895), *Versprechen und Verlesen: Eine psychologisch-linguis-tische Studie*. Stuttgart: Göschen
- Miettinen, Erkki (1965), “Beiträge zur deutschen Volksetymologie: Assoziative Umbildungen und Umdeutungen romanischer und lateinischer Entlehnungen”. *Neuphilologische Mitteilungen* 66, 28–91
- Mugdan, Joachim (1984), *Jan Baudouin de Courtenay (1845–1929): Leben und Werk*. München: Fink
- Olschansky, Heike (1996), *Volksetymologie*. Tübingen: Niemeyer
- Onions, C[harles] T[albut] (1966), *The Oxford Dictionary of English Etymology*. Oxford: Clarendon Press
- Palmer, Abraham Smythe (1882), *Folk-etymology: A Dictionary of Verbal Corruption or Words Perverted in Form or Meaning, by False Derivation or Mistaken Analogy*. London: Bell
- Paul, Hermann (4<sup>th</sup>1909), *Prinzipien der Sprachgeschichte*. Halle: Niemeyer [1880]
- Plank, Frans (1981), *Morphologische (Ir-)Regularitäten: Aspekte der Wortstrukturtheorie*. Tübingen: Narr (Studien zur deutschen Grammatik 13)
- Porzig, Walter (4<sup>th</sup>1967), *Das Wunder der Sprache*. Bern, München: Francke [1950]
- Saussure, Ferdinand de (1916), *Cours de linguistique générale*. Lausanne: Payot
- Saussure, Ferdinand de (1967), *Grundfragen der allgemeinen Sprachwissenschaft*, transl. by Herman Lommel. Berlin: de Gruyter
- Scheler, Manfred (1977), *Der englische Wortschatz*. Berlin: Schmidt (Grundlagen der Anglistik und Amerikanistik 9)
- Schreiner, Markus (1987), *Bibliographie zur Volksetymologie*. Münster: Inst. für Allgemeine Sprachwissenschaft (Studium Sprachwissenschaft 11)
- Ullmann, Stephen (2<sup>nd</sup>1959), *The Principles of Semantics*. Oxford: Blackwell [1951]
- Ullmann, Stephen (2<sup>nd</sup>1966), *Language and Style*. Oxford: Blackwell [1964]

- Ullmann, Stephen (1972), *Semantics: An Introduction to the Science of Meaning*. Oxford: Blackwell [¹1962]
- Ungerer, Friedrich (1991), “Acronyms, Trade Names and Motivation”. *Arbeiten aus Anglistik und Amerikanistik* 16, 131–158
- Vennemann, Theo (1999), “Volksetymologie und Ortsnamenforschung”. *Beiträge zur Namensforschung* N.F. 34, 269–322

- Waldron, R[oland] A. (1967), *Sense and Sense Development*. London: Deutsch
- Weekley, Ernest (¹1961), *The Romance of Words*. London: Murray [¹1912]
- Weisgerber, Leo (1950), *Die Muttersprache im Aufbau unserer Kultur*. Düsseldorf: Schwann

*Alwin Fill, Graz (Austria)*

## 150. Lexicalization and demotivation

1. Terminology
2. Lexicalization
3. Demotivation
4. References

### 1. Terminology

#### 1.1. Lexicalization

The term **lexicalization** (Germ. *Lexikalisiierung*) is of relatively recent origin; according to the OED (²1989) (see verb *lexicalize*, from which *lexicalization* has probably been derived) it was first attested in 1949. It has been used more widely from the 1960s onwards, but it is apparently still confined to the language of linguists. On its history see, e.g., Grimm (1991: 57–70) and Lipka (1992: 2–6). I shall briefly present some of the more important definitions that have emerged in the discussion.

##### 1.1.1. Accepting into the lexicon

The OED (²1989) defines *lexicalize* (1) as ‘to accept into the lexicon, or vocabulary, of a language’, and *lexicalization* as ‘the action or process of lexicalizing’. In this sense simple and complex words, native as well as loanwords can be lexicalized. Thus, Lyons (1968: 352) says “that the relationship of the transitive to the intransitive is ‘lexicalized’” in pairs like *kill* and *die* (*Bill died – John killed Bill*), i.e. the transitive (and causative) concept of ‘to cause someone to die’ is expressed by a separate word, *to kill* (*someone*). Quirk et al. (1985: 1525 f.) restrict lexicalization to words formed by word-formation processes, explaining it as the process of creating a new word (a complex lexical item) for a (new) thing or notion instead of describing this thing or notion in a sentence or with

a paraphrase. The use of words is more economical because they are shorter than the corresponding (underlying) sentences or paraphrases, and because they can be more easily used as elements of sentences. Thus one does not say ‘someone who writes a book [...] for someone else, who then often pretends it is their own work’, one says *ghost-writer* instead (see the definition in DCE ²1987). Quirk et al. (1985) also point out that the precise form of the lexicalization often cannot be predicted. Sometimes there are competing patterns: ‘to make national’ has been lexicalized as *to nationalize*, but ‘to make beautiful’ as *to beautify*, ‘to make broad(er)’ as *to broaden*, ‘to make narrow(er)’ as *to narrow*; sometimes different formations (synonyms) are used for the same concept, often with regional or stylistic differentiation: ‘a track round which horses race’ is a *race-track* in American English, but a *race-course* in British English (according to the DCE ²1987); ‘Saturday’ is *Samstag* in Southern German, but *Sonnabend* in Northern German.

##### 1.1.2. Replacement of a semantic configuration by a lexeme

Within the framework of generative semantics, lexicalization (or lexical insertion) refers to “a process in which a configuration of semantic elements in an abstract representation is replaced by a lexeme” (Lipka 2002: 111), e.g. [CAUSE BECOME NOT ALIVE] is lexicalized as *kill*, and in German [BEWIRKEN WERDEN SAUBER(ER)] is lexicalized as *säubern*, but I shall not pursue this aspect here.

##### 1.1.3. Formal and/or semantic changes

The third sense of lexicalization is now probably the most usual one; it is also the one which will be discussed in greater detail in

this article. It presupposes the first sense (inclusion of a word into the lexicon), however. Broadly speaking, lexicalization refers to the phenomenon that complex words and expressions are often not identical with the sum of their parts; instead they show idiosyncratic properties, i.e. formal (phonological, morphological) and/or semantic peculiarities which cannot be explained from their constituents or from the word-formation pattern. As Quirk et al. (1985: 1526) put it, the meaning of lexicalized words “is not recoverable from the form”, i.e. it cannot be deduced from the constituents, or at least not entirely. Thus *holiday* originally referred to a ‘holy day’, i.e. a church festival, but in Modern English it refers to ‘a day free from work (or school)’; Germ. *Hochzeit* originally referred to ‘any kind of festival’ but in Modern German it refers specifically to a ‘wedding’. Lexicalization in this sense is a complex phenomenon, and accordingly much more detailed definitions are possible (see Lipka 2002: 113; 1992: 7 f.).

Lexicalization overlays regular word-formation processes; Bauer says that a lexicalized lexeme has taken on a form “which it could not have if it had arisen by the application of productive rules” (Bauer 1983: 48). Lexicalization thus is a diachronic process.

Even though they did not use the term lexicalization, the phenomenon referred to here was known to and at least partly described by earlier scholars, in particular by Paul (1880) and Brugmann (1900).

Lexicalization stands in relation to a number of terms from which it is not always easy to distinguish. Therefore **idiomatization** and **institutionalization** must also be briefly discussed.

## 1.2. Idiomatization

Idiomatization (or its result, idiomaticity) also refers to semantic peculiarities of a complex word or a phrase which cannot or only partly be explained from its constituents, but whereas lexicalization can include phonological and morphological differences, too, idiomatization refers to semantic differences only. Moreover, lexicalization implies a diachronic process, e.g. change of meaning (cf. the examples of *holiday* and *Hochzeit* given in 1.1.3) or the addition of meanings, leading to polysemy, e.g. *reader* (a) ‘a person who reads (something)’, (b) ‘a person who reads books to put mistakes right before printing [i.e. a proofreader], or to decide whether to print them’; (c) ‘a senior British university

teacher just below the rank of professor’ (cf. DCE<sup>2</sup>1987), whereas idiomatization can refer to a synchronic process, whereby complex words or expressions are coined so as to present a specific meaning. Middle English *loveday*, for example, was probably coined to mean ‘day for the settlement of quarrels out of court’ (presumably as a loan translation of Lat. *dies amoris*) and did not mean something else first; the same applies to *ghostwriter* (see 1.1.1), also to phrasal idioms such as *golden wedding*/Germ. *goldene Hochzeit* ‘50th wedding anniversary’, *ins Gras beißen* ‘to die’. It is, of course, not always easy to ascertain whether a word or phrase was coined with a specific meaning or whether it developed this meaning only later; accordingly the distinction between lexicalization as a diachronic change of meaning and idiomatization as a synchronic specialization of meaning is sometimes difficult. Scholars are therefore not always consistent in their use of these terms. Consistency is especially difficult because idiomatized formations can then also be lexicalized in other ways: *chairman* /tʃeəmən/, for example, was probably coined to mean ‘a person who is in charge of a meeting’ or ‘a person who is the head of a large organization’ and did not simply have ‘man who sits in a chair’ as its original meaning, but rather ‘man who sits in a chair of authority’ (see OED<sup>2</sup>1989), but its second element *man* was, probably subsequently, weakened from /mæn/ to /mən/. But it is perhaps going too far to doubt the value of the distinction altogether, as Kastovsky (1982: 164 f.) seems to do.

Formations which are similar on the surface can nevertheless be idiomatized quite differently: *callgirl* ‘prostitute’ has the meaning ‘girl who is called (by men on the phone asking for paid sex)’, whereas the original meaning of *callboy* was ‘boy who calls (actors onto the stage)’ see, e.g., Kastovsky (1982: 165).

## 1.3. Institutionalization

This term was apparently made popular in morphology by Bauer (1983: 48). It also relates to the history of vocabulary items, but refers to an earlier stage than lexicalization. Bauer distinguishes three such stages, namely **formation**, **institutionalization**, and **lexicalization**.

### 1.3.1. Formation

Words are usually first coined in a particular context and in order to name a specific (new) thing or concept. Many are never taken over

by the speech community, they remain nonce-formations (**ad-hoc formations**, **hapax legomena**, Germ. **Augenblicksbildungen**, **Einzelprägungen**).

### 1.3.2. Institutionalization

Words are institutionalized when they enter the common vocabulary or at least the vocabulary of a certain group of speakers. The term institutionalization is thus roughly equivalent to (synonymous with) the definition of lexicalization as ‘accepting into the lexicon’ (see 1.1.1). Institutionalization can also be connected with the change from **type-familiarity** (recognizing a particular pattern of word-formation) to **item-familiarity** (knowing a particular word). In institutionalized words, possible ambiguity (polysemy) is often ignored and their use is restricted to one of their possible meanings. *Almsman*, for example, usually meant ‘man who receives alms’ and not ‘man who gives alms’, Germ. *Schüler* ‘pupil’ refers to ‘someone who is taught at school’ and not to ‘someone who teaches at school’ (the latter meaning is institutionalized as *Lehrer* in German).

On the other hand, formations with similar or partly identical elements can be institutionalized (and idiomatized) in different ways, as shown in 1.2 for *callboy* and *callgirl*, cf. also pairs such as *pork butcher* and *family butcher*, Germ. *Brillenträger – Hosenträger*, etc. But even in institutionalized formations, polysemy is sometimes retained (or re-introduced), e.g.: *fireman* can refer to ‘a person whose job is putting out fires’ or to ‘a person who looks after the fire in a steam engine or furnace’ (DCE 2<sup>1987</sup>); Germ. *Holzschuppen* can refer to a ‘shed made of wood’ or a ‘shed where wood is stored’ (or both at the same time).

### 1.3.3. Lexicalization

The third stage in the development of a lexical item is then lexicalization (in the sense of developing idiosyncratic formal and/or semantic features, i.e. change of meaning and/or form), as in *bonfire* ‘fire made for pleasure etc.’ (originally *bone-fire*), or Germ. *Drittel* (< *Dritteil* (der dritte Teil)). Not all institutionalized complex words become lexicalized, however: words such as *goldsmith*, *stone wall*, etc. have existed since Old English without any apparent change of meaning, and have undergone regular sound changes only, which affected the compounds as well as their ele-

ments. To cover both institutionalized and lexicalized formations Bauer (1983) also uses the term **established formations**.

### 1.4. Obscuration

This term refers primarily to the formal side of lexicalization. In **obscured** compounds, the elements show a different phonological (and consequently often also morphological) development in the compounds and in independent use. *Holiday* /hɒlədɪ/, /hɒlɪdeɪ/ is, for example, slightly obscured from its constituents *holy* /həʊlɪ/ and *day* /deɪ/, whereas *gospel* (from Old English *gōd* > Modern English *good* /gʊd/ and Old English *spell* > Modern English *spell* /spel/) is totally obscured (/gospɪl/) and has become monomorphemic. Obscured compounds often show semantic changes, too. On obscured compounds in English, see, e.g., Götz (1971), Faiß (1978), and Sauer (1992: 345–358).

### 1.5. Other terms

Some authors use terminologies quite different from the one favoured here. Lyons (1977: 535 ff.) calls the stage of the formation of words lexicalization (see 1.1.1). **Syntactic compounds** is his term for completely regular, non-lexicalized compounds which can be explained according to productive patterns of compounding. **Compound lexemes** is his term for semantically lexicalized compounds (i.e. with idiosyncratic semantic features) which he also terms petrified (1977: 547 f.) – for a similar use of petrification see Leech (1981: 225–227). From these Lyons distinguishes **fossilized compounds**; they are derived according to rules that are no longer productive in the language system (such as the so-called imperative compounds like *pick-pocket*, *turnkey*) – these are called morphologically lexicalized in our terminology (see 2.3.3). Lyons’s terminology does not seem very clear, however; for criticism see also Bauer (1983: 49 f.).

## 2. Lexicalization

### 2.1. Causes of lexicalization

Lipka (1981: 131) mentions several possible causes of lexicalization, which often work hand in hand. Lexicalization stresses the unified character of complex words and at the same time diminishes their character as syntagms. It should be borne in mind, however, that phenomena such as word-character,

naming function, and frequency of use can lead to lexicalization, but do not necessarily do so: Old English formations such as *grass-green* (Old English *gærsgrene*), *Sunday* (Old English *sunnandæg*) are still used in their original meaning and show only regular sound-changes which also affected their constituents; therefore they can be regarded as institutionalized, but not as lexicalized.

### 2.1.1. Word-character

In compounds unification of the constituents starts with the imposition of compound stress: syntactic group *bláck bóard* vs. compound *bláckbóard*, cf. Germ. *bláue Béere* vs. *Bláubéere*; this is, however, a fairly regular process and therefore not yet a case of lexicalization in the stricter sense. But unification can end in the coalescing of originally bimorphemic compounds into monomorphemic words, i.e. in total obscuration, e.g. Old English *scīrgerēfa* > Modern English *sheriff* (but Old English *scīr* > *shire*; *gerēfa* > *reeve*).

### 2.1.2. Naming function

A consequence of the naming function is **hypostasation** (Germ. *Hypostasierung*), i.e. the phenomenon that the existence of a (simple or complex) word makes people assume that there is a referent (thing, person, etc.) in reality (in the world) which the word refers to. Often this assumption is, of course, correct; there exist things or persons to which words like *bookcase*, *bookseller*, etc. refer. But this is not always the case – authors of science fiction stories, for example, invent new worlds with things and actions that do not exist in our world, such as to *uncreate*, to *unmurder* (cf. Kastovsky 1982: 152).

### 2.1.3. Frequent use

It can be assumed that frequent use was the reason for the lexicalization and eventual obscuration of a number of original compounds such as *daisy*, *gospel*, *lord*, *lady*, *sheriff*, *woman*, etc. It seems, however, impossible to decide precisely how often a formation must be used in order to become lexicalized, and, as stated in 1.3.3, there are a number of Old English compounds which are still commonly used in Modern English without having been lexicalized.

## 2.2. The scope of lexicalization

The majority of the examples given so far were compounds, but lexicalization can also occur with other types of word-formation,

for example: prefix-formations such as Germ. *abkratzen* in the sense ‘to die’; suffix-formations such as *reader* (see 1.2) or *sleepier* in the sense ‘heavy piece of wood, metal, etc., supporting a railway track’, Germ. *Zeitung* ‘newspaper’ (originally ‘message, news’); zero-derivations (conversions), e.g., *to corner someone* ‘to put someone into a corner so that he cannot escape’ (Leech 1981: 226), onomatopoeic formations such as Lat. *cuculus* ‘cuckoo’ > Fr. *cocu* ‘cuckold’, clippings such as *vamp* ‘seductive woman’ (from *vampire*). Syntactic groups (phrases) can be lexicalized, too; in this case they are usually called idioms. Examples are (cf. also 1.2): adjective + noun groups (*public school*, in British English ‘private school’; Germ. *geheimer Rat* [as a title]); groups of noun + preposition + noun (*man in the street*/Germ. *der Mann auf der Straße* ‘average person’); verb + object groups (*kick the bucket*/Germ. *ins Gras beißen* ‘to die’); verb + particle constructions (phrasal verbs: *to bottle up* ‘keep (one’s feelings) under control’; *to do in* ‘to kill’). The borderline between syntactic groups (phrases) and complex words (generated through word-formation processes) is not always easy to draw, anyway: syntactic groups can, for example, coalesce into complex words (e.g. Germ. *geheimer Rat* to *Geheimrat*; Old English *dōmes dæg* to Modern English *doomsday*); combinations with the first element in the genitive are regarded as compounds by some scholars but as syntactic groups by others (*craftsman*, *driver’s seat*, *bull’s-eye*, etc.); phrasal verbs are certainly lexical units, but since they are separable (*they done the old woman in* [Pygmalion]), they cannot be regarded as complex words.

## 2.3. Types of lexicalization

Lexicalization can affect all linguistic levels, see, e.g., Bauer (1983: 50–61), Grimm (1991: 71–121). Often, several of these types co-occur in a formation.

### 2.3.1. Graphemic lexicalization

Graphemic lexicalization usually reflects phonological changes, e.g. *tuppence* /tʌp(ə)ns/ for *twopence*.

### 2.3.2. Phonological lexicalization

This can be divided into segmental and suprasegmental.

(a) Segmental lexicalization: Lexicalization of segmental features occurs when the elements of a complex word develop phoneti-

cally different in independent use. This is especially frequent in compounds; here the causes may be loss of stress, shortening of an element, assimilation, etc.; often two or more of these factors are combined and lead to obscuration (opacity). *Woman* (from Old English *wif* > Modern English *wife* and *man* /man/ > *man* /mæn/), for example, shows shortening of the first element (before two consonants), loss of stress on the second element, and assimilation of the end of the first element to the beginning of the second element (regressive assimilation): Old English *wifman* > *wifman* > *wimman* > Modern English *woman* /womən/; Germ. *hohe Zeit* vs. *Hochzeit*. As stated in 1.4, there is a scale from slight obscuration (e.g. *cupboard* /kʌbəd/ vs. *cup* /kʌp/ and *board* /bɔ:d/) to total obscuration; totally obscured compounds such as *barn* (< Old English *bere-ern* ‘barley-house’), *lord*, *lady* (< Old English *hlāfweard*, *hlæfdige*), *lammas*, *gospel*, *gossip*, etc. have become monomorphemic, whereas partly or slightly obscured compounds can still be regarded as bimorphemic, e.g. *holiday*. Sometimes the pronunciation is obscured, but not the spelling, e.g. in *breakfast* (/brekfəst/ vs. /breɪk/ and /fə:st/). Phonological lexicalization can also occur with suffixes; in most of the formations with the (no longer productive) suffix *-th*, the underlying simplex has been changed in the derivation: *broad – breadth*, *deep – depth*, etc. (which can be explained historically as *i*-mutation caused by the suffix *-iθə* > *-th*). For words to take on a different form as the basis of derivations was relatively frequent in Old English, while it is rare in Modern English, apart from derivative alternants due to loan-influence, e.g. *consume* /kən'sju:m/ – *consumption* /kən'sam(p)fən/, *decide* /dɪ'said/ – *decision* /də'sizən/.

(b) Suprasegmental (prosodic) lexicalization: Compounds often have a different stress pattern from corresponding syntactic groups; in English and German, for example, syntactic groups typically have double stress (i.e. two main stresses), whereas compounds typically have a main stress on their first element and a secondary stress on their second element, e.g. *blåck båord* vs. *blåckbåord*, Germ. *schönes Wetter* vs. *Schönwetter*. Some loan suffixes retract the accent, e.g. *-átion* (*declaráre* /dɛkl're/ vs. *declarátion* /dɛklə'reiʃn/), *-ée* (*train* vs. *trainée*), Germ. *-ieren* (*Lack* – *lackieren*, *Marsch* – *marschieren*). Nouns de-

rived from phrasal verbs by zero-derivation (conversion), on the other hand, change the main accent from the second to the first element: *to pick up* > *a pickup*; *to lèt down* > *a létdown*. But all these phenomena are apparently synchronically productive processes and thus do not yet lead to lexicalization in the stricter sense. The stage of phonological lexicalization is reached, however, if, for example, the second element of a compound loses its secondary stress and is also reduced phonetically as a consequence, e.g. in some institutionalized compounds with *man*, such as *chairman* /tʃeəmən/, where the second constituent is pronounced as /mən/ instead of /mæn/ and thus approaches the character of a suffix. But *-man* is not a genuine suffix yet, because the weakening of /mæn/ to /mən/ probably took place after the formation of these compounds. In some cases, however, suffixes originated from the weakened second elements of compounds, e.g. the suffix *-ful* /f(ə)l/ from the word *full* /fʊl/ (*careful*, *grateful*, *thankful*); frequently lexicalized elements thus became productive in a different function. In other cases weakening of the second element due to loss of the (secondary) stress apparently was the first step to obscuration, especially in a number of older compounds.

### 2.3.3. Morphological lexicalization

This type of lexicalization can also be classified into a number of subtypes.

(a) Linking elements, e.g. Germ. *Arbeit* but *Arbeitsvertrag*. Bauer (1983: 53 f.), who regards the insertion of linking elements as one kind of morphological lexicalization, points out that often their addition follows synchronically productive rules and thus cannot be regarded as lexicalization in the strict sense, but there can also be diachronic changes resulting in genuine morphological lexicalization: Germ. *Schwanenhals*, for example, retains the old form of the genitive, whereas in independent use *Schwan* now has the *s*-genitive (*des Schwans*); something similar applies to Lat. *pater familias*, which also retains the old genitive, whereas in classical Latin, the genitive of *familia* is *familiae*. While linking elements are fairly frequent in German, they are rare in English (*-in-* in *nightingale* < Old English *nihtegale*, cf. Germ. *Nachtigall*), apart from the *-s*, which can, however, usually be regarded as the genitive *-s*, sometimes combined or merged with the plural *-s*, cf. *ladies' room*, *salesman*, *townsman*, etc.

(b) Unproductive word-formation patterns or elements: Words formed according to unproductive patterns or containing unproductive affixes are also morphologically lexicalized. In English this applies to the so-called imperative compounds (*pickpocket*, *dolittle*, *speakeasy*), cf. (Lyons 1977: 547f.), or to derivations with the suffix *-th* (*warmth*, *growth*) – formations like *length*, *strength* are thus lexicalized phonologically as well as morphologically.

English moreover has a large number of loan-words where the suffix then also became productive in English. Words with these suffixes fall into two groups: (i) loans where the base does not occur independently in English, e.g. *atheism*, *atheist* (but no \**athe-*), *baptism*, *baptist* (but no \**bapt-*), – these could be regarded as lexicalized, but usually they are just regarded as loan-words; they show, however, that the distinction between loan-words and English word-formation is also sometimes blurred. (ii) Formations where the base also occurs independently in English, e.g. *escape* > *escapism*, *escapist*, but these are not lexicalized (at least not morphologically).

(c) Blocked morphemes: Blocked (unique) morphemes are elements of compounds which occur just in one compound and which do not or no longer occur as independent words, e.g. many first elements of the days of the week such as *Tues-* in *Tuesday*/Germ. *Dienstag*, or *rasp-* in *raspberry*, Germ. *Brom-* in *Brombeere*, -*galel-gall* in *nightingale*/Germ. *Nachtigall* (cf. also Sauer 1992: 340–345). Blocked morphemes are frequent in English place-names, e.g. the first elements of *Manchester*, *Gloucester*, *Worcester*, *Salisbury*, etc.

(d) Obscured compounds: Morphological lexicalization has, of course, also taken place in obscured compounds which have become monomorphemic, e.g. *nostril* (< Old English *nos-hyrel*, lit. ‘nose-hole’), see 2.3.2(a).

(e) With partly obscured compounds the question arises of whether the obscured elements should still be regarded as allomorphs of the corresponding independent words, e.g. whether *Mon-* /mʌn/ in *Monday* should be seen as an allomorph of *moon* /mu:n/. The answer probably depends on whether the speakers still feel a semantic connection between the forms, which is probably not the case in *moon* and *Mon(day)*, while it is probably true (due to the spelling) in *cup(board)* /kʌ-/ and *cup* /kʌp/.

### 2.3.4. Semantic lexicalization

Semantic lexicalization entails change of meaning. This is often subdivided into addition of semantic features, which results in restriction (narrowing) of meaning, loss of semantic features, which conversely results in extension (widening) of meaning, and a mixture of both. These changes also happen in simplex words, e.g. *fowl* (Old English *fugol* ‘bird’ > Modern English *fowl* ‘farmyard bird’) and conversely *bird* (Old English *bridd* ‘young bird’ > Modern English *bird* ‘bird’), but in connection with lexicalization we are only concerned with semantic changes in complex words. Examples are:

(a) Addition of semantic features, e.g. Germ. *Hochzeit* (see 1.1.3). Here belong probably also formations containing implied features in addition to the stated features (in the terminology of Warren 1978): Middle English *hand-whil* means ‘time needed to turn one’s hand’, i.e. ‘very short time, moment’; Germ. *Augenblick* has the same implied feature ‘very short time’. Strictly speaking, *hand-whill* Germ. *Augenblick* are probably rather cases of idiomatization than of lexicalization. Addition of features also occurs if one element of a phrase is lost through ellipsis and the remaining element in addition takes on the meaning of the lost element, e.g. *private (soldier)*, Germ. *Süddeutsche (Zeitung)*. For the distinction between general (systematic) features such as in *sleepwalker*/Germ. *Schlafwandler* [+HABITUAL] and idiosyncratic features such as in *pushchair* [+FOR CHILDREN], see also 2.4.

(b) Loss of semantic features, e.g. *arrive* (< Fr. *arriver* < Lat. \**adripare*, \**arripare*) ‘come to a place’, originally ‘land at the shore’. Cases where a constituent loses its meaning entirely are treated under demotivation.

(c) Mixture (combination) of addition and loss of features, e.g. *holiday* (has lost the feature ‘religious’ and added the feature ‘free from work or school’), *blackboard* (has lost the feature ‘black’, because *blackboards* are nowadays usually green or white, and added the feature ‘made to write upon’). This phenomenon seems to be more frequent than simple loss or addition of features. Here belong probably also instances of metaphoric transfer, e.g. the use of *bottleneck* ‘neck of a bottle’ for ‘narrow part of a road which slows down traffic’ (DCE 21987). Demotivation can also result in change of semantic features:

a *Junggeselle* (Germ.) usually is no longer young, but he is unmarried.

(d) As pointed out in connection with *reader*, the same formation can be semantically lexicalized in several ways. Examples are numerous among the phrasal verbs, e.g. *to hold out* 'to offer', 'to continue to exist', 'to continue in spite of difficulties, endure'.

The examples discussed so far seem fairly clear, but semantic description in connection with the question of lexicalization and idiomatization also has to face a number of problems (cf. Bauer 1983: 56–59); I shall mention two of them.

(e) In connection with added semantic features it is sometimes unclear where the borderline between a semantic (linguistic) description of a word and an encyclopedic description of the thing or being concerned lies. A *typewriter* can be linguistically explained as an 'instrument that writes with types' and as such is not lexicalized (but rather institutionalized), but of course a typewriter has many more elements to it than just types. The same problem often arises with plant and animal names: There are many animals that hop around in the grass (e.g. hares), but only a certain kind of insect is called *grasshopper*/Germ. *Grashüpfer*.

(f) In other cases the distinction between semantic processes (lexicalization or idiomatization) and word-formation processes is not easy to make. The so-called **exocentric compounds**, with the subgroups of **bahuvihi (possessive) compounds** like *redcoat* 'person who has/a red coat', *paleface*/Germ. *Bleichgesicht*, etc. and of **imperative compounds** like *dolittle*, *pickpocket*, Germ. *Vergißmeinnicht*, *Fürchtegott* 'person who performs the action indicated by the verb' all contain the element 'person' (or 'animal' as in *stickleback*, or 'thing' as in *greenback*, *scarecrow*), which is, however, not expressed on the surface. There are two possible explanations for this, one in semantic terms and one in terms of word-formation: These formations could be seen as a kind of **metonymy** with a part standing for the whole, or the action standing for the agent, and the feature 'person' (or 'animal' or 'thing') could then be regarded as an additional semantic feature, which would make them lexicalized (or idiomatized) compounds. But since this added feature can also be regarded as typical of the entire patterns of word-formation concerned, it could be ar-

gued with Kastovsky (1982: 167 f.), who follows Marchand (2<sup>nd</sup> 1969), that all formations based on these types (*paleface*, *pickpocket*) are regular formations with a complex determinant and a zero-morpheme as a determinatum, e.g. *redskin*/ $\emptyset$  'someone who has a red skin' *pickpocket*/ $\emptyset$  'someone who picks (other people's) pockets'. Marchand (2<sup>nd</sup> 1969) and Kastovsky (1982) classify the zero-morpheme as a zero-suffix with the systematic meaning 'person' (or 'animal' or 'thing'), apparently as a parallel to explicit suffixes such as *-er* (with the systematic meaning 'person', as in *baker*, or 'instrument', as in *bottle-opener*). This would, moreover, classify the exocentric formations as (zero) derivations rather than as compounds.

### 2.3.5. Syntactic lexicalization

This is the most problematic of the types. At least in English and German, there are fewer clear cases of syntactic lexicalization than of semantic, phonological and morphological lexicalization.

(a) Bauer (1983: 59) says that syntactic lexicalization occurs when a syntactic pattern according to which complex lexemes have been formed falls into disuse. An example of this would be Germ. *Vergißmeinnicht* 'forget-me-not', derived from the sentence *vergiß mein(er) nicht*, whereas in Modern German the construction is *vergiß mich nicht*. In this connection Bauer also mentions the so-called imperative compounds (*pickpocket*, *scarecrow*, to which *Vergißmeinnicht* belongs, too). As stated in 1.5, this pattern seems to be unproductive in present day English and is thus morphologically lexicalized, but it is not syntactically lexicalized (*Vergißmeinnicht* being an exception) because the underlying sentence pattern verb + object, e.g. '(to) scare crows (away)' is still usual in Modern English.

(b) Many compounds show a change of word order as compared with the underlying (or paraphrase) sentence or phrase, e.g. *as green as grass* → *grassgreen*, *the door at the front* → *the front door*. But since this change of word order is a fairly regular phenomenon in compounding, it should not be regarded as indicative of lexicalization.

### 2.3.6. Mixed lexicalization

Often two or more types of lexicalization are combined in one word. This has become clear from some of the examples given above: *holi-*

*day* and Germ. *Hochzeit* (see 1.1.3, 1.2, 2.3.4) show phonological as well as semantic lexicalization (different pronunciation of at least one constituent and change of meaning). Bauer points out “that once a form is lexicalized in one way it is easier for it to become lexicalized in others” (Bauer 1983: 61). This is certainly true for many of the totally obscured English compounds (*barn*, *lord*, *lady*, *lammas*, *gospel*, *sheriff*), which show phonological changes with resulting morphological changes (originally bimorphemic, now usually monomorphemic), and semantic changes (idiomatization or semantic lexicalization).

#### 2.4. Degrees of lexicalization

There is a gliding scale from slight lexicalization to total obscuration (or demotivation). Examples of phonological and morphological obscuration have been given in 1.4 and 2.3.2 (relatively slight: *cupboard*, *holiday*; medium: *bonfire*; total: *lord*, *lady*). Here I want to concentrate on degrees of semantic lexicalization (idiomatization). Usually, a distinction is made in the literature between the addition of general and systematic semantic features and the addition of specific and idiosyncratic features (see, e.g. Lipka 1981: 125–129; Kastovsky 1982: 195 ff.; Sauer 1992: 46–50, 384–388).

(a) General (systematic) additional features: Here belong, e.g., [+HABITUAL], e.g. *crybaby*, [+PROFESSIONAL], e.g. *baker*/Germ. *Bäcker*. Instruments as a rule have the feature [+PURPOSE] and they can be paraphrased as ‘B is (made) for A’, e.g. *bottle-opener*/Germ. *Flaschenöffner*, *candlestick*/Germ. *Kerzenständer*.

This kind of analysis is not without its problems, however. Features like [+HABITUAL] and [+PROFESSIONAL] are certainly inherent in some formations, e.g. [+PROFESSIONAL] in *baker*, *miller*/Germ. *Bäcker*, *Müller*, whereas in others they seem to depend on the context. *Writer*/Germ. *Schreiber* in *the writer of this letter*/*der Schreiber dieses Briefes* is a pure agent nominalization (‘the person who writes or has written this letter’) without any additional semantic features; in *she is a regular letter-writer (or writer of letters)* the feature [+HABITUAL] is indicated by *regular*. Authors like G. B. Shaw and Hemingway were professional writers, but Caesar and Chaucer were writers in their pastime.

(b) Specific (idiosyncratic) additional features: The elements of the historical term

Germ. *Morgengabe*/Old English *morgengifu* just state that it is ‘a gift given in the morning’, but the precise meaning is ‘a gift (present) given by the husband to his wife in the morning after the wedding night’. Cf. also examples discussed in 1.2, such as *ghostwriter*, *reader*, etc. Leech (1981: 225–227) calls the addition of such specific and unpredictable semantic features petrification. As far as I can see they are impossible to classify.

(c) Additional semantic features can, of course, be combined: Compounds like *crybaby* and *sobsister* have [+HABITUAL, +DE-ROGATORY]. There are also combinations of added general features and added specific features: In formations like *playboy*, *callgirl* the general features [+HABITUAL] or [+PROFESSIONAL] are combined with added (or deleted) specific features: a *playboy* is usually not a boy, but a man; he plays in a very special way, he is rich and does not work. A *crybaby*, moreover, usually is no longer a baby, a *sobsister* is not necessarily a sister, etc.

(d) As has been shown before (cf. *reader*), the same formation can also be non-lexicalized (have no additional semantic features) in one sense and be heavily lexicalized (have specific features) in another sense.

(e) It is not always clear where the general additional features end and the specific (idiosyncratic) features begin. *Crybaby*, *sobsister*/Germ. *Heulsuse* also have the feature [+DE-ROGATORY], which, according to Marchand (1969: 73 f.), occurs relatively frequently in formations of the type verb + noun. [+DE-ROGATORY] could therefore be regarded as a fairly general semantic feature, too, but it is not usually mentioned as such in the literature, although it is usually given as a stylistic label in dictionaries.

(f) It is not even always clear where exactly lexicalization (or idiomatization) begins. If one follows the distinction made by Bauer (1983), the use of a formation in just one of its possible meanings to the exclusion of others is a case of institutionalization but not yet of lexicalization. Institutionalization entails that in the paraphrase only one verb (or a group of synonymous verbs) can be used, e.g. for *stone wall* ‘the wall consists/is made of stone’, and not, e.g. \*‘the wall produces stone’ or \*‘the wall resembles stone’. Lipka (2002: 114) mentions the historical *highwayman* among the unpredictable idioms with rather specific semantic features. It can, how-

ever, be paraphrased as ‘man who robs (others) on the highway’ (cf. DCE 2<sup>1987</sup>), and according to this analysis it would be a simple case of institutionalization, excluding meanings such as \*‘man who meets others on the highway’ or \*‘man who helps others on the highway’. On the other hand, while one author listed *country house* among the compounds without additional features, Lyons (1977: 535–544) spends several pages to point out that, at least in British English, it means not just ‘house in the country’, but rather ‘large house in the country owned by a member of the aristocracy, as opposed to his town house’ (cf. Sauer 1992: 119).

(g) Features which are additional in some formations can already be part of an element in others. [+DEROGATORY] or [+NEGATIVELY EVALUATED] is additional in *crybaby*, *playboy*, etc., but in Middle English *flescheslust*/Germ. *Fleischeslust* it is probably inherent because the flesh was (is) regarded as inherently weak and sinful by the church. Similarly, in formations such as Germ. *Langweiler*, *Trödler*, the feature [+DEROGATORY] is not added but inherent in one of the constituents (*Langeweile*, *trödln*).

### 3. Demotivation

Demotivation has been described as the loss of the sign-character of one or both constituents of a complex lexeme. Seen from the semantic aspect it indicates that one or both elements have lost their original meaning (and in some cases have acquired a new meaning). Thus demotivation partly overlaps with the phenomenon of semantic lexicalization. But before discussing demotivation in some more detail, the underlying and antonymous term **motivation** should be briefly explained.

#### 3.1. Motivation

Since F. de Saussure it has been fairly generally accepted that the linguistic sign is on principle arbitrary, i.e. unmotivated, because as a rule there is no natural connection between the *signifiant* (expression side, phonetic form) and the *signifié* (content side, meaning). This is clear from several facts: Both sides can change, the expression side through sound-changes (e.g. Old English *stān* > Middle English *stōn* > Modern English *stone* / *stəʊn/), and the content side through semantic change, i.e. change of meaning (e.g.*

Germ. *Gift*, originally ‘gift, present’ [as still in *Mitgift* ‘dowry’] → ‘poison’). Moreover, the same meanings or denotata can be expressed by different *signifiés* in different languages, e.g. Lat. *arbor*, Eng. *tree*, Germ. *Baum*, as well as in the same language; in the latter case they are called synonyms and are often dialectally or stylistically differentiated, e.g. *begin*, *commence*, *start* or British English *railway* vs. American English *railroad*. But there are a number of exceptions to this general principle; words can be phonetically, morphologically or semantically motivated (see Ullmann 1962: 82–93; Fill 1980: 14–16).

(a) Phonetic motivation: Onomatopoeic words, such as *miaow*/Germ. *miauen*, are imitations of animal cries and thus phonologically motivated; phonological motivation also plays a role in words formed with the help of sound-symbolic elements, such as Eng. -ash (indicating quick movement (ending suddenly)) in *clash*, *crash*, *dash*, *flash*, *lash*, etc.; for details see Marchand (2<sup>1969</sup>: 397–428).

(b) Morphological motivation: Complex words, especially compounds, prefix- and suffix-formations are morphologically motivated, i.e. if one knows the meaning of the constituents (which are unmotivated in themselves), one can predict or at least make an informed guess at the meaning of the entire combination. If the speaker/hearer knows for example, that *re-* means ‘again, once more’ and *tell* ‘narrate’ he will also know that *retell* means ‘to tell again’; this aspect has been called type-familiarity.

(c) Semantic motivation: This applies in cases of semantic transfer by metaphor or metonymy; if one speaks, for example, of *the foot of a mountain*/Germ. *Fuß des Berges*, or if one calls a narrow part of a road a *bottle-neck*. Compounds used metaphorically are thus morphologically and semantically motivated; the same applies to metaphorically used syntactic groups like *wet blanket* (cf. Lyons 1977: 548). As stated earlier, exocentric nouns (bahuvrihi nouns such as *paleface* and imperative nouns such as *dolittle*) can generally be seen as metonymic formations (‘person with a pale face’, ‘person who does little’, etc.). Semantic transfer, of course, also occurs in simplexes, e.g. in *fox*/Germ. *Fuchs* ‘clever person full of tricks’ (in Germ. also ‘horse with a reddish colour’, in Eng. also ‘sexually attractive woman’).

### 3.2. Loss of motivation

Motivation on all the levels mentioned can also be lost; I shall concentrate on loss of morphological motivation. There are several ways of classifying this phenomenon, e.g. according to extralinguistic/intralinguistic changes or according to the number of constituents affected (cf. Lipka 1981: 127f.; Sauer 1992: 49f.).

(a) Extralinguistic changes are changes in the cultural background or in our knowledge of the world. *Shoemakers/Germ. Schuhmacher, Schuster* and *watchmakers/Germ. Uhrmacher* usually no longer make shoes or watches, but repair them; *blackboards* were originally black, but are nowadays usually green or white; a *Schreibfeder* (Germ.) is now usually made of steel and no longer of feathers (quills). Strictly speaking, in all these cases it is not the meaning of the compounds or derivations that has changed, but the activity or structure of the referent; *shoemaker* can still be paraphrased as ‘someone who makes shoes’, as opposed to Germ. *Junggeselle*, which cannot be paraphrased as \*‘der Geselle ist jung’ or \*‘junger Geselle’.

(b) Intralinguistic changes are due to a change of meaning in one or both constituents. In *mincemeat* ‘mixture of fruit and spices ...’ *meat* does not have its modern meaning ‘edible flesh’, but retains the older meaning ‘food in general’. Germ. *Mitgift* ‘dowry’ retains the old meaning of *Gift*, namely ‘gift, present’, whereas the modern meaning of *Gift* is ‘poison’.

(c) Loss of meaning in one constituent: The first constituent has lost its original motivation and therefore its original meaning in, e.g. *blackboard*, *Whitsunday* (originally, white garments were worn on this Sunday), Germ. *Bockbier, Bockwurst*, the second constituent in *bluebottle* (a kind of blue fly), *mincemeat, sweetmeat*, Germ. *Handschuh, Bahnhof, Lehrkörper*. The original motivation has often been lost in plant and place names: speakers usually do not know what, for example, horses have to do with *horse-chestnuts* or *horsetistles*; many towns are still called *Newtown, Newton/Germ. Neustadt*, although many of them were founded in the Middle Ages. A routine formula such as *how do you do* has also partly lost its motivation.

Here belong also blocked morphemes, which do not (or no longer) carry any meaning in themselves, but serve to differentiate

one compound from another; *Tues-/Germ. Diens-* differentiate *Tuesday/Dienstag* from other days of the week, *rasp-/Germ. Him-* differentiate *raspberry/Germ. Himbeere* from other berries (*blueberry, blackberry, Germ. Blaubeere, Stachelbeere*).

(d) Loss of meaning in both constituents: Here belong, e.g., animal names such as *butterfly* (which is neither a fly nor is it known what its connection with butter should be), *understand* (connected neither with *under* nor with *stand*), Germ. *Junggeselle* (who is neither young nor a *Geselle* ‘companion’, but unmarried), *Schornstein* (with a blocked morpheme as its first element) – cf. phrases like *alter Junggeselle*. A French example is *beaucoup* ‘a lot’ (no longer ‘a nice blow’). It has been said that formations such as *butterfly, Junggeselle, etc.* have a misleading motivation (cf. Fill 1980: 70). A number of place names and surnames have also lost their original motivation: *Oxford/Germ. Ochsenfurt* were originally settlements at a ford where oxen could cross; the towns are still there, but the fords for oxen have gone. People called *Fowler, Forester, Germ. Bartenschlager, Strohschneider* have quite different occupations or professions than their ancestors had. Obscured compounds which have become monomorphemic (*barn, gospel, lord, sheriff*) have, of course, also lost their motivation completely. Loss of motivation in both elements is furthermore frequent in phrasal idioms, e.g. *white elephant, red tape*, including phrasal verbs such as *to do in* ‘to kill’, etc.

### 3.3. Degrees in the loss of motivation

As with lexicalization in general, there are also degrees in the loss of motivation. Formations such as *shoemaker/Germ. Schuhmacher*, are still analyzable and probably represent only slight cases of demotivation, since we still know about the cultural background that existed during their formation. Formations with blocked morphemes or where one constituent has lost its meaning (*raspberry, Germ. Brombeere, Handschuh*) show a middle degree of demotivation, and formations where both elements have lost their meaning as well as entirely obscured compounds show total loss of motivation (*butterfly, Germ. Junggeselle; lord*), cf. Fill (1980: 69–73). The loss of motivation and accordingly the degree of idiomatization is stronger in *Großmutter* than in *Großstadt*: a *Großstadt* is a large city (although the minimum size of a *Großstadt*

has been officially defined), whereas the *groß* in *Großmutter* has lost the meaning ‘large’ (see Fleischer<sup>5</sup> 1982: 13). There will probably always be difficult or disputed cases. Some dictionaries say that the reason for (i.e. the motivation of) the plant name *strawberry* is not known, which would make it a partly demotivated compound (with loss of meaning in *straw*). I believe, however, that strawberry fruits were (as they still are) protected from decay by surrounding them with straw or fine shavings, which would make *strawberry* a fully motivated compound (‘berry which is protected with straw’).

### 3.4. Remotivation

Native as well as loan words that have partly or entirely lost their motivation are sometimes remotivated.

(a) If the spelling is still intact this can happen through spelling pronunciation, e.g. in *waistcoat* /'weskət/ > /'weirstkəut/; this seems to be frequent in place-names, e.g. *Cirencester* /'sɪsɪtə/ > /'saɪrənsestə/.

(b) In other cases it can happen through popular etymology (folk etymology, secondary motivation; cf. Olschansky 1996). Here an unmotivated word or the unmotivated part of it (e.g. a blocked morpheme) is changed so as to coincide with existing morphemes. This can be done with loan-words which were never motivated in the receiving language: Lat. *levisticum* became *Liebstöckel* in German and *lufestice* ‘love-stitch’ in Old English, cf. Modern English *lovage*; *asparagus* (monomorphemic) was turned into *sparrowgrass* (bimorphemic). It can also happen with native words which lost their motivation partly or entirely: The second element of Old English *brydguma* lit. ‘bride-man’ was exchanged so as to yield *bridegroom*; the first element of Germ. *Sintflut* (‘big flood’ with blocked *Sint*-) was exchanged so as to yield *Sündflut* (‘flood [as a punishment] for the sins’). But the new morphological motivation is often just superficial, once more a kind of misleading motivation: *aspargus/sparrowgrass* is neither a kind of grass nor have sparrows anything to do with it.

(c) Sometimes the same combination survives in different stages of transparency or lexicalization: the Old English syntactic

group *hālig dæg* ‘holy day’ survives as syntactic group *holy day*, as lexicalized compound *holiday* and as still more lexicalized (obscured) compound *Halliday* (only used as a surname).

## 4. References

- Bauer, Laurie (1983), *English Word-Formation*. Cambridge: Cambridge University Press
- Brugmann, Karl (1900), “Über das Wesen der sogenannten Wortzusammensetzung” [reprinted in: Lipka & Günther (1981, eds.), 135–178]
- DCE (<sup>2</sup>1987) = *Longman Dictionary of Contemporary English*, Della Summers (editorial director). Harlow: Longman [<sup>1</sup>1978] [now <sup>4</sup>2003]
- Faiß, Klaus (1978), *Verdunkelte Compounds im Englischen*. Tübingen: Narr
- Fill, Alwin (1980), *Wortdurchsichtigkeit im Englischen*. Innsbruck: Institut für vergleichende Sprachwissenschaft
- Fleischer, Wolfgang (<sup>5</sup>1982), *Wortbildung der deutschen Gegenwartssprache*. Leipzig, Tübingen: Niemeyer [<sup>1</sup>1969]
- Götz, Dieter (1971), *Studien zu den verdunkelten Komposita im Englischen*. Nürnberg: Hans Carl
- Grimm, Ursula (1991), *Lexikalisierung im heutigen Englisch am Beispiel der -er Ableitungen*. Tübingen: Narr
- Kastovsky, Dieter (1982), *Wortbildung und Semantik*. Düsseldorf: Schwann-Bagel
- Leech, Geoffrey (<sup>2</sup>1981), *Semantics: The Study of Meaning*. Harmondsworth: Penguin
- Lipka, Leonhard (1981), “Zur Lexikalisierung im Deutschen und Englischen”. In: Lipka & Günther (eds.), 119–132
- Lipka, Leonhard (1992), “Lexicalization and Institutionalization in English and German”, *Linguistica Pragensia* 1, 1–13
- Lipka, Leonhard (2002), *English Lexicology*. Tübingen: Narr [<sup>1</sup>1990 under the title *An Outline of English Lexicology*]
- Lipka, Leonhard & Günther, Hartmut (1981, eds.), *Wortbildung*. Darmstadt: Wissenschaftliche Buchgesellschaft
- Lyons, John (1968), *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press
- Lyons, John (1977), *Semantics*, Vol. 1–2. Cambridge: Cambridge University Press

- Marchand, Hans (1969), *The Categories and Types of Present-Day English Word-Formation*. Munich: Beck [1960]
- OED (1989) = *The Oxford English Dictionary*, ed. by Sir James Augustus Henry Murray et al.; 2nd ed. by John Andrew Simpson & Edmund S. C. Weiner. Oxford: Clarendon Press [1933]
- Olschansky, Heike (1996), *Volksetymologie*. Tübingen: Niemeyer (Germanistische Linguistik 175)
- Paul, Hermann (1880), *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer [1920]
- Quirk, Randolph & Greenbaum, Sidney & Leech, Geoffrey & Svartvik, Jan (1985), *A Comprehensive*
- Grammar of the English Language*. London: Longman
- Sauer, Hans (1992), *Nominalkomposita im Frühmittelenglischen*. Tübingen: Niemeyer
- Ullmann, Stephen (1962), *Semantics: An Introduction to the Science of Meaning*. Oxford: Blackwell
- Warren, Beatrice (1978), *Semantic Patterns of Noun-Noun Compounds*. Göteborg: Acta Universitatis Gothoburgensis (Gothenburg Studies in English 41)

Hans Sauer, Munich (Germany)

## 151. Change in productivity

1. Introduction
2. Types of changes in productivity
3. Synchronic vs. diachronic productivity
4. Linguistic factors affecting the productivity of bases
5. References

### 1. Introduction

Most languages reveal fluctuations in the **productivity** of word formation patterns over time. Despite this, the diachronic aspect of productivity has been little investigated in a systematic fashion. A discussion of changes in productivity depends, first of all, however, on a coherent definition of the notion of productivity (cf. Art. 33). Different scholars have used the term in different ways (cf. Botha 1968: 149), although common to most definitions is an implicit notion of **frequency** of occurrence, as in Bauer (1983):

“Basically, any process [...] is said to be productive if it can be used synchronically in the production of new forms, and non-productive if it cannot be used synchronically in this way.” (Bauer 1983: 18)

Thus, in order for some process to be productive, new words must be generated. It follows that a greater degree of productivity should be reflected in a greater number of new forms. Thus, in German, for instance, the suffix *-bar* ‘able’ is both productive and frequent, e.g. *trinkbar*. Forms that are non-productive or lexically idiosyncratic must be listed in the lexicon, and appear before the level of lexical insertion.

### 2. Types of changes in productivity

Within the group of productive processes, some are more productive than others, while non-productive processes can be accounted for in terms of a list of the bases to which they have applied. Changes in productivity may affect roots as well as affixes. It is well known that in English, for instance, a number of new affixes not existing in Old English were borrowed, e.g. *-ment* as in *involvement*, *-ity* as in *purity*, *mal-* as in *malformed*, etc., while other native ones disappeared or declined in use. Although words containing the suffix *-ment* are still analyzable, the suffix is no longer productive as can be seen from the fact that new words are not being coined. Other nominalizing suffixes are preferred instead and existing forms with *-ment* have become lexicalized. Similarly, all nominalizations from adjectives which end in *-th* are now lexicalized, e.g. *warmth*, *depth*, *length*, etc., and most also have unproductive root forms. No new forms can be created with *-th* so that *\*realth*, for instance, is not synchronically possible. Moreover, new suffixes are created all the time by analysis of existing lexemes, e.g. *forecast* led to *funcast* (i.e. the weather outlook for sporting activities), *quickcast* (i.e. a brief summary of the weather), *travelcast* (i.e. forecast for travellers), etc.

Productive processes may become less productive over time, while non- or semi-productive processes can become more productive. The upper limits on productivity are

vague. Gruber (1976: 322) has suggested that to be fully productive, an affix must be usable with all [bases] definable by some semantic, syntactic or phonological property. Few, if any, affixes would meet this criterion.

Many scholars have, however, suggested that the suffix *-ness* is fully productive (cf. Jespersen 1924). Its productivity extends to both major and minor categories:

- major categories: *goodness*, *jackassness*, *wellness*;
- minor categories: *muchness*, *threeness*, *betweeness*, *whyness*;
- phrasal/semi-phrasal: *broken-heartedness*, *fed-up-ness*, *up-to-dateness*, *know-nothingness*.

Some of these formations provide counterexamples to generalizations about word formation, namely that only major categories can be the product of word formation rules and that only form classes can be used as bases in the formation of derivatives (cf. Aronoff's principles of word formation; cf. Aronoff 1976: 21). In addition, however, *-ness* managed to attach itself to native as well as non-native roots and therefore was able to compete with French suffixes such as *-ity*, cf. e.g. *stableness* vs. *stability*; *oddity* vs. *oddness*, etc.

Looking at the cognate suffix *-nis* in the history of German, we see that the opposite state of affairs prevails in the modern language, where many of the Old High German formations have disappeared, and most of the modern formations were established by 1800. As in English, there is a competing French suffix *-ität*, which applied mainly to foreign bases. Unlike English, however, *-nis* never really extended itself to foreign bases and most of the formations are deverbal. There are a total of 428 derivations in Matter's (1971) reverse dictionary, more than 3/4 of which are compounds or derivatives of only 29 bases. Word forms ending in *-ität* account for nearly as many of the total abstract noun formations (48 percent in *-ität* as compared to 52 percent in *-nis*), as shown in Table 151.1.

	<i>-ness/-nis?</i>	<i>-ityl/-ätäl/-i)té</i>
English	33 percent	67 percent
German	52 percent	48 percent

Tab. 151.1: Formations in *-ness/-nis* and *-ityl/-ätäl/-i)té* in English and German

From a comparative historical perspective it is clear that not only has *-ness* met a more productive locus in English, but the same is true of English *-ity* compared with the productivity of its cognate in French. In French there is a total of 797 words ending in the suffix *-ité/té*. In English there is a total of 1926. Nouns ending in *-ity* in English account for 33 percent of the total number of *-ness* and *-ity* formations, while those ending in *-ness* account for 67 percent (cf. Lehnert 1971; Juilland 1965). The difference between English and French on the one hand and English and German on the other is quite striking.

Scholars such as Thompson (1975: 347) have drawn a distinction between **productivity** and **analogy**. An analogical formation is a new formation clearly modelled on one already existing lexeme and does not give rise to a productive series. The problem in drawing an absolute distinction between productivity and analogy is that an analogical formation may provide the impetus for a series of formations. Thus, at one extreme would be cases which are more clearly analogical formations resulting in the creation of only one new form which cannot be accounted for by rule, e.g. *wargasm*, from *war* + *orgasm* (see Bauer 1983: 96, 100). This would be the limiting case of productivity. At the other end, however, are cases such as *sit-in* providing the model for *teach-in*, *swim-in*, and *hamburger* providing the model for *cheeseburger*, *fishburger*, etc. (see Art. 148).

A somewhat different example of a change in word formation patterns which involves an alteration in productivity can be seen in German. In Middle High German compounds with first elements ending in *-keit* were formed without any linking element by adding one stem to another, e.g. *Gerechtigkeitsliebe* 'love of justice' consists of *Gerechtigkeit* 'justice' + *Liebe* 'love'. In modern German, however, all words ending in *-keit* add a linking element *-s* (i.e. Kompositionsfugenelement) when they enter into compounds. Lexemes which were institutionalized before the change in linking element either have now added *-s* in modern German, e.g. *Gerechtigkeitsliebe*, or have become lexicalized in their old form since they can no longer be generated by synchronic productive rules. It is also possible for the same two lexemes to form compounds at two different time periods with different linking elements. In such cases the compounds may constitute minimal pairs

which differ in meaning, e.g. *Wassernot* ‘drought’ vs. *Wassersnot* ‘flood’ (cf. Bauer 1983: 53 f.).

### 3. Synchronic vs. diachronic productivity

The extent to which it is possible to distinguish between synchronic and diachronic productivity depends at least partly on maintenance of a distinction between synchrony and diachrony which has become increasingly less tenable. Most scholars recognize that a rigid boundary between the two is difficult to defend and that synchronic speech forms reflect diachronic processes. Matthews (1972: 191, n.2) notes that the “introduction of diachronic evidence ... in no way implies that it is a diachronic and not a synchronic fact that it is evidence for.” If this view is accepted, then there can be no clear-cut boundary between diachronic and synchronic productivity. Words occur all the time with varying degrees of probability in both speech and writing. A large proportion of complex word forms heard everyday are nonce formations, and do not become established in the language. It is only with hindsight that we can determine if a word has become established and thus contributed to the productivity of a morphological process over the long term rather than just momentarily. The rules allowing speakers to create new words by applying patterns of word formation can be thought of as a matter of synchrony while the subsequent fate of established lexemes is a matter for diachrony. Although some have said that processes such as **back formation** are strictly diachronic (cf. Art. 60), there must be some synchronic rule which allows analogies to be made between apparently complex forms with other cases where the suffixed and non-suffixed forms are both lexemes, e.g. *laze* is derived from the earlier form *lazy*, possibly by analogy with forms such as *craze* and *crazy* (cf. Bauer 1983: 64 f.).

It is possible to gain some retrospective idea of productivity, and hence some insight into its diachronic dimension by looking at a dictionary, as in the comparisons made above among German, French and English with respect to *-ität*, *-ité/té* and *-ity* (cf. 2). In providing a list of the established lexemes of a given language, a dictionary includes the existing products of various word formation rules which have applied in the past to produce conventionalized and accepted words.

One way to study productivity diachronically is to look at the output of various word formation rules as recorded in dictionaries at different points in time. It is particularly instructive to compare word formation processes which compete for the same bases. In such cases the factors constraining productivity become clearer, and it is evident that synchronic restrictions on productivity are essentially the result of diachronic changes.

Romaine (1983: 186 f.) examined a list of 100 adjectives in the contemporary *Oxford English Dictionary* (OED) and compared these with lists obtained from two earlier dictionaries, Samuel Johnson’s *Dictionary of the English Language* (1755) and Thomas Sheridan’s *A General Dictionary of the English Language* (1780) to see how many *-ity* and *-ness* formations were listed, e.g. *pompous*: *pomposity/pompousness*. Over time there has been an increase in the number of words taking both suffixes. For example, the OED records 60 words out of the list of 100 which may take either *-ity* or *-ness*, compared to only 22 for Johnson’s dictionary and 18 for Sheridan’s. Similarly, more words take neither suffix in Sheridan’s and Johnson’s dictionaries (15 and 21, respectively) than in the OED (1). There are only minor differences among the dictionaries with regard to those words which are listed as taking *-ness* only. The OED records 32, Samuel Johnson’s dictionary, 35, and Thomas Sheridan’s, 42. The differences are greater for *-ity*. The OED lists only seven words which can take *-ity* only, while Sheridan has 25 and Johnson, 20. This method could be used with other affixes, and to look at individual authors writing in earlier time periods.

This procedure, however, gives only a partial glimpse of productivity because dictionaries list only actual, attested and not possible words (cf. Art. 171). It must not be forgotten that dictionaries are in some sense arbitrary collections of the words in a language at a given time and will not necessarily reflect the actual written or spoken usage of a speech community (cf. Mugdan 1984: 240–259). Thiel (1973: 379) found that 62 percent of the compound words in an issue of the German popular news magazine, *Die Zeit*, were not listed in dictionaries. Morphological processes may be exceedingly productive and never result in established lexemes which are recognized by a dictionary as being legitimate words of the language. One well known example is found in the English use of infixes

Base Type	Formations in <i>-ness</i>		Formations in <i>-ity</i>	
	Number	Example	Number	Example
<i>-ous</i>	<i>generousness</i>	514	<i>generosity</i>	94
<i>-ive</i>	<i>transitiveness</i>	391	<i>transitivity</i>	96
<i>-able</i>	<i>reasonableness</i>	346	<i>reasonability</i>	200
<i>-al</i>	<i>musicalness</i>	169	<i>musicality</i>	315
<i>-ible</i>	<i>fallibleness</i>	99	<i>fallibility</i>	156
<i>-ile</i>	<i>fragileness</i>	17	<i>fragility</i>	75
<i>-ic</i>	<i>domesticness</i>	15	<i>domesticity</i>	63

Tab. 151.2: Formations in *-ness* and *-ity* in relation to base type

forms such as *absobloodylutely*. Here the **in-fixes** are potential word forms rather than bound morphs like most other English affixes. Although infixated forms such as these have been in colloquial use for some time, it is unlikely any of them will ever appear in a dictionary due to their taboo status. Just because such words do not appear in the dictionary, this does not mean that infixation is non-productive, or that the forms resulting from them are not established lexemes in the mental lexicon.

#### 4. Linguistic factors affecting the productivity of bases

Just as it is impossible to talk about productivity in a meaningful sense synchronically without taking into account other factors, e.g. phonological, syntactic, it is impossible to do so diachronically. Although some have assumed that the morphological form of the base is the most important factor affecting productivity (cf. 4.1), for others, semantic considerations are primary (cf. 4.4). The successful spread and productivity of certain suffixes such as *-ness* can be accounted for in terms of the interaction of these factors (see also Baayen 1989).

##### 4.1. Morphological form of the base as a factor restricting productivity

It is obvious that the morphological form of the base plays an important role in determining productivity (cf. Aronoff 1976: 53 f.). Compare the difference in the number of *-ness* and *-ity* formations which obtain when the suffixes are attached to two distinct classes of adjectives, those ending in *-ive*, and

those ending in *-ile*. There are approximately five times as many words ending in *-iveness* as there were ending in *-ivity*. The second base, however, yields different results. The number of words ending in *-ility* overwhelmingly exceeded those ending in *-ileness*. Table 151.2 shows some results for different base types from Romaine (1983: 182) using Lehner's (1971) dictionary. The boxes indicate where the locus of productivity lies for each suffix. Nouns ending in *-ness* are most productively formed with bases ending in *-ous*, *-ive*, and *-able*, while *-ity* occurs most frequently with *-al*, *-ible*, *-ile*, and *-ic*.

Although a consideration of the morphological type of base imposes a considerable refinement on the analysis of productivity, it still does not deal with the issue of recurrent vs. possible lexemes, and whether the productivity of existent types bears a regular relationship to that of possible types. Aronoff (1976: 37) claims that speakers of a language have intuitions about productivity which entail the notion 'likelihood of being a word of the speaker's active vocabulary'. He furthermore speculated that if pairs such as *perceptiveness* and *perceptivity* are presented to native speakers of English, they will say that although both are possible, *perceptiveness* is preferable. Romaine (1983), however, did not find this to be the case in an experiment conducted to test speakers' intuitions about the productivity of *-ness* and *-ity* formations.

Comparing the results with those from Table 151.1 based on the dictionary, Romaine found that in all cases where we would expect *-ness* to be more productive than *-ity*, it was not; nor was *-ity* more productive than *-ness* in all those cases where we would expect it

to be. This indicates a significant discrepancy between productivity measured in terms of existent lexemes as recorded by dictionaries and possible lexemes as indicated by speaker judgements. The suffix *-ity* turned out to be more productive than *-ness* over all the base types. These results suggest that the factor of morphological type of base is itself a dependent and not an independent variable in productivity. There is also wide variation among the words belonging to a particular type, which indicates a greater degree of lexical idiosyncrasy than Aronoff allows.

#### 4.2. Lexical restrictions on productivity

Productivity may also be affected by idiosyncratic lexical restrictions. For example, the suffix *-ric* occurs only with *bishop*, e.g. *bishopric* (cf. Bauer 1983: 93). In some *-ity* nominalizations from adjectival bases ending in *-ous*, the *-ous* becomes *-os-*, e.g. *curiosity*, while in others it is deleted, e.g. *voracity* from *voracious* (cf. Aronoff 1976: 40). Because these processes are not predictable, these exceptions must be dealt with on an ad hoc basis in the lexicon.

Another type of conditioning can be seen in the preference of suffixes such as *-ee* to apply more productively to transitive verb bases, e.g. *employee* (cf. Bauer 1983: 98).

#### 4.3. Phonological restrictions on productivity

Phonological factors affecting the productivity of bases include stress shift and vowel and consonant alternations. Speakers tend not to use words whose pronunciation they are unsure of. In order to add the suffix *-ity* to the base *electric* to produce the derived form *electricity*, a speaker must apply rules of stress shift and velar softening. In experiments testing the alleged psychological reality of these kinds of processes as formulated by Chomsky & Halle (1968) in their generative model of English phonology, it was found that when speakers were asked to apply such rules to bases which normally did not undergo them, e.g. *methanity* (*methane* + *-ity*), they did not use them (cf. Steinberg & Krohn 1975; Baldi et al. 1985).

In other cases the phonological shape of the base determines whether it may be used as the input to a word formation rule. In French the diminutive suffix *-ette* is not normally added to bases ending in /t/ or /d/, while in English the adverbial suffix *-ly* is not productively added to adjective bases already

ending in *-ly*, e.g. *\*sillily*. The suffix *-eer* shows a preference for bases ending in /t/, e.g. *profiteer*, *racketeer*, etc.

In still other cases productivity is sensitive to syllable structure. Thus, the English prefix *mal-* may apply only to polysyllabic bases such as *adapted*, e.g. *maladapted*. Another case where syllable structure impedes productivity can be found in the infixes mentioned earlier. In most cases the infixes must be inserted immediately before the syllable of the base which bears the lexical stress so that forms such as *\*abbloodysolutely* are ruled out. Stress is also a relevant conditioning factor since it is difficult to insert such infixes in words with stress falling on the first syllable with no subsequent subsidiary stress, e.g. *solid*. Most commonly infixes occur in words of three or more syllables (cf. Bauer 1983: 89–91).

#### 4.4. Semantic restrictions on productivity

There is one general semantic factor which limits the productivity of word formation rules, which can be called **hypostatization**. This means that words will not be formed to denote non-existent entities (cf. Bauer 1983: 85; Art. 80). The use of a word presupposes the existence of a relevant category.

Some have suggested there is a simple relationship between productivity and meaning: a morpheme whose meaning is more semantically coherent will be more productive (cf. Aronoff 1976: 39). Comparing *-ness* and *-ity*, it can be easily shown that the semantics of *-ness* nominalizations has been highly coherent throughout history. The meaning of X-ness can be stated in terms of a choice among three operations on the base X:

- (a) ‘the fact that someone/something is X’, e.g. *His kindness* (‘the fact that he is kind’) *amazes me*.
- (b) ‘the extent to which someone/something is X’, e.g. *His kindness* (‘the extent to which he is kind’) *amazes me*.
- (c) ‘the quality or state of being X’, e.g. *Kindness is a virtue*.

By contrast, the semantics of *-ity* derivatives is not nearly so coherent. Many of the established lexemes ending in *-ity* have strayed from their original abstract sense to denote concrete objects and now have a collective or technical sense (cf. Romaine 1985). The root sense of many *-ity* formations is abstract, i.e. a noun denoting a state, quality or condition,

e.g. *nobility* in the sense of ‘nobleness’, the quality or state of being noble. Some of these formations then became used as nouns denoting occupations, offices, positions or titles signifying the abstract qualities of the occupations, e.g. *nobility* in the sense ‘class of nobles’. Many of the formations ending in *-ty* (from French *-té/té*) denote offices or titles of occupations, e.g. *royalty*, *admiralty*, *mayorality*, etc. A number of these denote occupations or offices associated with feudalism and are no longer in use, e.g. *squiralty*, *vassality*, etc. Some *-ity* formations developed a sense of collectivity or totality of people belonging to an occupation or who have the quality of being associated with it, e.g. *humanity*, *Christianity*, *polity*, *society*, *nationality*, *fraternity*, etc. Some also came to be used as scientific or technical terms, e.g. *gravity*, *velocity*, and others to denote things having the abstract quality of the adjective base associated with them, e.g. *oddity*, *antiquity*, *curiosity*, etc. Thus, nouns ending in *-ity* have drifted away from the semantics of their derivational base.

This semantic itinerary is established diachronically for *-ness* too, although the number of *-ness* formations which have strayed from their original sense is far smaller. In Old English there were some cases where a noun ending in *-ness* denoted something concrete, e.g. *smirines* ‘oil’, or the still surviving *wilderness* or *witness* used to refer to both the person offering evidence and the evidence or testimony itself. There is an inverse relationship between **productivity** and **lexicalization**: the most productive patterns are not lexicalized and fully lexicalized processes are not productive (cf. Bauer 1983: 88). Formations such as *witness*, *business*, *wilderness* and *likeness* are examples of this. There were also cases where *-ness* formations became used as titles connected with occupations, e.g. *highness*, *holiness*. The meaning of *-ness* in this sense is apparently still available to speakers synchronically since it is used humorously in Agatha Christie’s novel *Death on the Nile*, when an Indian character addresses the famous detective as “Your Sleuthness”. Finally, the word *business* refers to a collectivity.

Where competing *-ness* and *-ity* nominalizations exist, they occupy different semantic spaces. Thus, *nobleness* refers only to the capacity or state of being noble, while *nobility* also refers to the capacity or state of being noble. *Humanness* refers to the quality of being human, while *humanity* refers to the col-

lectivity of humans; *antiqueness* refers to the state of being antique, while *antiquity* refers to an old object or time period, etc. Lexicalized *business* referring to a collectivity has the counterpart *busyness* denoting the state of being busy. The semantic lexicalization of *business* is accompanied by phonological lexicalization in that the final vowel of *busy* is lost. Many words spread themselves over more than one slot in semantic space. For example, *curiosity* can denote both the condition or state of being curious and also an object having the quality or state associated with the adjective but *curiousness* can be used only to denote the quality or state of being curious. *Oddity* and *oddness* contrast in the same way. The fact that the semantics of *-ity* formations is less predictable and less regular is implicated in productivity since speakers will tend not to use words whose meaning is unclear to them.

Van Marle (1988: 141 f.) has discussed a similar case in Dutch, where the greater semantic heterogeneity of deverbal adjectives ending in the suffix *-lijk* has led to its decline in productivity, while *-baar* has increased its domain of application. He also suggests that the loss of productivity of one of two competing word formation rules must be explained by internal factors, i.e. in terms of factors pertaining to the properties of the category of formations produced by that process. In Dutch it seems that there are non-standard dialects where *-lijk* has declined in productivity, but has not been replaced by *-baar*. Thus, *-baar* has not been responsible for the decline in productivity of *-lijk*. An important factor in this case is the large-scale occurrence of non-literal, metaphorical readings of words ending in *-lijk*, e.g. *onberispelijk* ‘perfect’ from the verb *berispen* ‘to rebuke, reprehend’. The meaning of the underlying verb has completely disappeared and the derivative occurs only in combination with the negative prefix *on-*. This represents a drift away from the core or inherent properties of the verbs to which the suffix is attached (cf. van Marle 1988: 144 f.). Once this happens, the semantics of the derivatives becomes less transparent.

Semantic considerations are also probably responsible for the recent increased productivity of *-free* over *-less* (cf. Slotkin 1990). Originally, *-free* and *-less* were semantically equivalent, and in some formations, still are, e.g. *sugarfree* and *sugarless*. However, in other formations they are not, e.g. *childless*

vs. *childfree*. The adjective ending in *-less* has negative overtones, while the one in *-free* is positive. Thus, to describe someone as *childless* implies that having children is desirable, but to be *childfree* imputes a positive value of being free from the burden of having children. The negative connotations of *-less* may have been acquired through the conventionalization of compounds such as *feckless*, *hopeless* and *ruthless*, which denote undesirable qualities. The original meaning of *careless* to indicate someone free from care is now expressed by *carefree* and *careless* has undergone pejoration. Thus, *careless* means ‘not taking care’ and the result of carelessness is generally negative. Now *-less* seems to be attached primarily to nouns to indicate undesirable connotations and has declined in productivity. The suffix *-free*, however, has increased its productivity and is used increasingly in a variety of nonce denominal adjective formations.

**Blocking** conditions on productivity have been formulated in terms of meaning so they too may be seen as a type of semantic constraint. Aronoff (1976: 56) predicts that it will be impossible for there to be two words with the same meaning and the same root in one person’s lexicon at the same time. Others have also suggested that some possible formations will be blocked if there is already an existing lexeme with that meaning (e.g. *Regel der besetzten Stelle*, or preemption by synonymy; cf. Art. 85). Thus, we would predict that *stealer* is not likely to become conventionalized due to the existence of *thief*. There is no nominalization ending in *-eur* from *voler* ‘to fly’ in French due to the prior existence of a homophonous form *voleur* meaning ‘thief’. Similarly, the presence of nominals such as *glory* and *fury* blocks the application of a word formation rule suffixing *-ity* to the adjectives *glorious* and *furious*. Despite this, however, nominalizations ending in *-ness* are possible, i.e. *furiousness*, *gloriousness*. Blocking does not in any case prevent coining. It acts only as a brake on conventionalization. Formations in *-ness* often appear in the semantic slots of nouns formed with other abstract deadjectival noun forming suffixes. The individual limits and differences in semantic distinguishability between pairs such as *sincereness/sincerity*, *chasteness/chastity*, *stability/stableness*, etc. remain to be investigated.

#### 4.5. Pragmatic factors restricting productivity

Closely related to such meaning considerations affecting changes in productivity are pragmatic factors having to do with non-referential associations such as the stylistic and social connotations associated with a particular pattern of word formation. Since they are optional, word formation rules may be sensitive to the same kinds of constraints which affect other sociolinguistic variables, e.g. age, sex, social class, region, etc. From a diachronic perspective it is possible to establish that certain patterns of word formation vary with time-depth and style.

Nonce formations are often glossed as such in written texts. In modern writings new terms are often put in inverted commas. When new formations are introduced, they may be glossed, e.g. *cherite þet is lufe* (Ancrene Riwle) or used with a synonym, e.g. *sikernesse and surete*. The repetition of two synonymous forms, one native and one foreign, is still used in legal registers, where there is a high proportion of Latin and French vocabulary, e.g. *life and limb*, *break and enter*, *legal and lawful*, etc.

Whether a new term gains in currency will depend on many factors, one of which is the status of the person who uses it. Where new concepts are concerned, e.g. upon the invention of the television or internet, new words are required, but prestige is often a factor in motivating new formations. Nowadays popular media, in particular, the world wide web, are an important influence. Historically speaking, almost as soon as French and Latin words were introduced into English, native prefixes and suffixes were added to them, e.g. *stableness*. As soon as sufficient lexical material was borrowed for foreign word formation patterns to be transparent and isolable, they could be used productively with both native and newly borrowed words, e.g. *oddity*. For the educated in society, knowing the derivational systems of both French and English made mixing foreign and native systems of derivational morphology possible. Authors such as Chaucer and Wyclif experimented with these types of formation very early. Wyclif, for example, uses doublets such as *feersness* (cf. Modern English: *fierceness*) and *feerste* (cf. Modern English: *fiercity* and *ferocity*) and *bareyness* (cf. Modern English: *barrenness*) and *bareynte* (cf. Modern English: *\*barrenty*). Once massive borrowing in the Middle English period had disrupted the

prevailing native patterns of word formation, there were numerous derivatives which are closely related semantically and possibly also etymologically, but which have different roots, e.g. *regal* and *royal*, *eat* and *edible*, etc. In many cases only one member of these sets of roots remains productive in modern English. It does not seem possible to create new lexemes in English by adding suffixes to the root *ed-*, but *eat* has been productive in new forms such as *eater*.

English is remarkable in its ability to accommodate and integrate the derivational morphological systems of Romance with its own native Germanic one. Neither Czech nor German normally mixes native bases with foreign affixes or foreign bases with native affixes. Thus, German *\*Sterbation* is ungrammatical, while English *starvation* is not. In so far as there are some restrictions on the compatibility of foreign and native elements, these can be considered to be of the same type as others which are due to the morphological form of the base (cf. 4.1). Aronoff (1976: 51 f.) proposes that bases have to be marked as [+latinate] or [-latinate]. Some foreign suffixes have never combined with native roots, but over time such restrictions have been relaxed. It is not clear what factors are responsible for this diachronic change in status.

Speakers' reactions to novel expressions may also affect productivity. When speakers notice a new form, they often oppose it and write letters to the newspaper complaining about the use of the new form on etymological, grammatical, semantic or aesthetic grounds. One example is *Time* magazine's article (cf. *Time* 1962) entitled "The Nesselrode to ruin" complaining about the productivity of *-ness*. The article regarded the prolifness of *-ness* formations such as *inwardness*, *thereness*, *matter-of-factness*, etc. as a dangerous tendency which was bringing about a decline in standards of English usage. English speakers have also complained about the extension of *-wise* meaning 'with respect to' or 'in regard to' in formations such as *moneywise*, *I'm o.k.* which means 'I have enough money' (cf. *Houghton* 1968). This increase occurred sometime around the middle part of this century. Earlier in the century, it looked as if this suffix was unproductive since it existed primarily in a few lexicalized forms such as *likewise*, *otherwise* and *clockwise*.

#### 4.6. Other factors restricting productivity

Among the other possible factors limiting the productivity of word formation processes is simply time itself. The longer a process of word formation has been used productively, the more likely it is that a sizable number of the bases to which the rules can be applied will become institutionalized lexemes. This means that over time there will be fewer bases available for the process to apply to productively (see Bauer 1983: 98). The relationship between productivity and lexicalization is still not well understood. Nor is the interaction of all the factors already discussed. Bauer (1983: 99) speculates that potential formation of a particular word is blocked by a cumulation of factors whose precise weightings we are not in a position to assess at the moment. Van Marle (1988: 147) however, argues that semantic factors are primary and are in effect the root cause of productivity. Semantic coherence is a requisite for productivity. This implies that all word formation processes which are productive are semantically coherent, while those which display semantic coherence are not necessarily productive. This is supported by the existence of non-productive processes such as the denominal female personal suffix *-in* in Dutch. All existing formations in Dutch are regular, e.g. *leeuw* 'lion' and *leeuwin* 'lioness', etc., but newly coined formations are decidedly jocular in character.

### 5. References

- Aronoff, Mark (1976), *Word Formation in a Generative Grammar*. Cambridge/MA: MIT Press (Linguistic Inquiry Monograph 1)
- Baayen, Harald (1989), *A Corpus-based Approach to Morphological Productivity*. Ph.D. dissertation, Vrije Universiteit Amsterdam
- Baldi, Philip & Broderick, V. & Palermo, David S. (1985), "Prefixal Negation of English Adjectives: Psycholinguistic Dimensions of Productivity". In: Fisiak, Jacek (ed.), *Historical Semantics; Historical Word-Formation*. Berlin: Mouton, 33–57
- Bauer, Laurie (1983), *English Word Formation*. Cambridge: Cambridge University Press
- Botha, Rudolf P. (1968), *The Function of the Lexicon in Transformational Generative Grammar*. The Hague: Mouton
- Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York: Harper & Row

- Gruber, Jeffrey S. (1976), *Lexical Structures in Syntax and Semantics*. Amsterdam: North Holland
- Houghton, Donald E. (1968), "The Suffix *-wise*". *American Speech* 43, 209–215
- Jespersen, Otto (1924), *The Philosophy of Grammar*. London: Allen and Unwin
- Juillard, Alphonse (1965), *Dictionnaire inverse de la langue française*. The Hague: Mouton
- Lehnert, Martin (1971), *Rückläufiges Wörterbuch der englischen Gegenwartssprache*. Leipzig: VEB
- Marle, Jaap van (1988), "On the Role of Semantics in Productivity Change". *Yearbook of Morphology* 1, 139–154
- Mater, Erich (1971), *Rückläufiges Wörterbuch der deutschen Gegenwartssprache*. Leipzig: VEB
- Matthews, Peter H. (1972), *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press
- Mugdan, Joachim (1984), "Grammatik im Wörterbuch: Wortbildung". In: Wiegand, Herbert Ernst (ed.), *Studien zur neuhochdeutschen Lexikographie IV*. Hildesheim etc.: Olms (Germanistische Linguistik 1–3/83), 237–308
- Romaine, Suzanne (1983), "On the Productivity of Word Formation Rules and the Limits of Variability in the Lexicon". *Australian Journal of Linguistics* 3, 177–200
- Romaine, Suzanne (1985), "Variability in Word Formation Patterns and Productivity in the History of English". In: Fisiak, Jacek (ed.), *Papers from the 6th International Conference on Historical Linguistics* [Poznan 1983]. Amsterdam: Benjamins; Poznan: Adam Mickiewicz Univ. Press (Current Issues in Linguistic Theory 34), 451–465
- Slotkin, Alan R. (1990), "Adjectival *-less* and *-free*: A Case of Shifting Institutional Currency". *American Speech* 65, 33–49
- Steinberg, Danny & Krohn, Robert (1975), "The Psychological Reality of Chomsky and Halle's Vowel Shift Rule". In: Koerner, Konrad (ed.), *The Transformational Generative Paradigm and Modern Linguistic Theory*. Amsterdam: Benjamins, 233–259
- Thiel, Gisela (1973), "Die semantischen Beziehungen in den Substantivkomposita der deutschen Gegenwartssprache". *Muttersprache* 83, 377–404
- Thompson, Sandra A. (1975), "On the Issue of Productivity in the Lexicon". *Kritikon Litterarum* 4, 332–349
- Time (1962) = "The Nesselrode to Ruin". Volume 79, May 11, 70

*Suzanne Romaine, Oxford  
(United Kingdom)*

## 152. Morphologische Entlehnung und Lehnübersetzung

1. Zum Gegenstandsbereich
2. Die außersprachlichen Bedingungen für Interferenz
3. Grundtypen der Interferenz
4. Gliederung der Phänomene
5. Zitierte Literatur

### 1. Zum Gegenstandsbereich

Die Veränderung natürlicher Sprachen ist nicht nur eine Angelegenheit der betreffenden Sprachen selbst, sie geht zu erheblichen Teilen auf den Kontakt mit anderen Sprachen und die dadurch ausgelösten Interferenzen zurück (einen allgemeinen Überblick geben Boretzky & Igla 1994). Im folgenden sollen einige Begriffe eingeführt werden. Leider muß man sagen, daß es keine allgemein akzeptierte Terminologie gibt und teilweise auch verschiedene Ansichten über den Cha-

rakter der im Kontakt auftretenden Prozesse bestehen, bedingt auch dadurch, daß manches nicht mehr empirisch zugänglich ist. Wir wollen uns deshalb auf das Notwendigste beschränken.

Viele Linguisten meinen mit **Interferenz (interference)** die in der *parole* zweisprachiger Individuen zustandegekommenen Abweichungen von der intendierten Sprache. Wir wollen im folgenden darunter die Gesamtheit der durch Sprachkontakt bewirkten Phänomene verstehen. Die aktuelle Übertragung von Elementen aus einer Sprache in die andere, also den Prozeß selbst, nennen wir **Transfer (transfer)**, den Übergang von transferierten Elementen und Relationen aus der *parole* in die *langue* aber **Entlehnung (borrowing)**. Im folgenden wird in der Sache fast nur von Entlehnung die Rede sein. Wie die empirische Forschung gezeigt hat, bleibt

grundsätzlich keine Ebene oder Unterebene der Sprache von fremden Einflüssen ganz verschont. Es gibt also kaum absolute universale Restriktionen für Interferenz; Versuche, solche Beschränkungen zu formulieren (vgl. Moravcsik 1978: 110–113), haben die außersprachlichen Umstände der jeweiligen Situation zu wenig berücksichtigt.

Interferenzphänomene fallen uns vor allem im Bereich der Lexik und in etwas geringerem Umfang im Bereich der Wortbildung (s. 4.1) auf, aber auch die Übertragung von Lauten (mit oder ohne Konstituierung neuer Phoneme) und von syntaktischen Regeln ist nicht ungewöhnlich. Viel seltener begegnet uns die Übertragung von Flexionsmorphologie (s. 4.2.3), weshalb auch schon früh die Behauptung ausgesprochen wurde, diese Art von Interferenz existiere so gut wie nicht (Meillet 1921: 82; auch noch Oksaar 1972: 492). Die Erfahrung zeigt jedoch, daß auch Flexionsmorpheme übertragbar sind, wie bereits Schuchardt (s. Spitzer 1928: 195) betont hat. Häufiger finden wir jedoch im engeren grammatischen Bereich, also dem Bereich der grammatischen Kategorien und ihrer Ausdrücke, andere Interferenzphänomene: Entlehnung grammatischer Partikeln oder (flektierter) Auxiliare (s. 4.2.1); **Lehnübersetzungen (loan translations, calques)** von Partikeln, Auxiliaren (s. 4.2.2) bzw. umfangreicheren peripherastischen grammatischen Konstruktionen, d. h. sinngemäße Wiedergabe, sofern die genannten Elemente eine eigenständige Bedeutung erkennen lassen. Auch Entlehnungen aus dem Bereich der kleinen Wortklassen beinhalten mehr als syntaktische Phänomene, ohne doch direkt die Morphologie zu berühren. Sie können hier jedoch nicht behandelt werden. Dies alles sind interferenzbedingte **Aufbauprozesse**. Morphologie kann aber auch unter fremdem Einfluß abgebaut werden (s. 4.2.2).

Es sei noch angemerkt, daß mit Entlehnung gebundener Morphologie am ehesten in solchen Sprachen zu rechnen ist, die selbst schon vergleichbare Morphologie haben.

Von Transfer und Entlehnung gebundener Morphologie ist zu unterscheiden die **morphologische Adaptation** lexikalischer Elemente. Im ersten Fall werden gebundene Flexions- und Wortbildungsmorpheme entlehnt und mit Erbwörtern kombiniert, im zweiten Fall werden Wörter entlehnt und in der Regel mit ererbter Flexionsmorphologie versehen bzw. mit ererbten Wortbildungsmorphemen kombiniert. Hierbei ist z. B. zu untersuchen:

bis zu welchem Grad morphologische Adaptation erfolgt; nach welchen Kriterien Nomina in Genussprachen ein Genus zugeordnet wird; in welche Flexionsklassen die Nomina eingeordnet werden; welche Form der Verben (Wurzel bzw. Stamm, Präsensstamm, Präteritalstamm usw.) als Grundlage der morphologischen Adaptation dient; ob für Lehnwörter besondere Flexionsklassen entwickelt werden (wie z. B. im europäischen Romani; vgl. Boretzky 1989), u. a. m. Auch diese Thematik kann aus Raumgründen hier leider nicht weiter verfolgt werden.

## 2. Die außersprachlichen Bedingungen für Interferenz

Interferenz zwischen zwei oder mehreren Sprachen verläuft normalerweise nicht ungesteuert, sondern in einer bestimmten Richtung, die durch das Verhältnis der am Kontakt beteiligten Gruppen diktiert wird. Die Qualität sprachlicher Strukturen ist dabei von geringer Bedeutung, entscheidend sind vielmehr außersprachliche Faktoren, die man grob mit dem Stichwort ‚Prestige‘ umreißen kann – jedenfalls, soweit Entlehnung im engeren Sinne gemeint ist. Entlehnungen können nützlich sein, indem sie im Sinne einer Sprachökonomie (s. Martinet 1955) dazu beitragen, Lücken im grammatischen System zu füllen, Strukturen zu parallelisieren u. a. m., als notwendig wird man sie jedoch kaum je bewerten können. Zwar trifft es sehr oft zu, daß Wörter aus einer Kultur entlehnt werden, die neue Begriffe zu bieten hat, die in gewissem Sinne also notwendig sind; für die anderen Ebenen, insbesondere für die Morphologie, kann man dies jedoch schwerlich behaupten.

Als Voraussetzung für die Entlehnung von Wortbildungsmorphemen und noch mehr von Flexionsmorphemen gilt, daß vorher bereits Interferenz auf anderen Ebenen stattgefunden hat, zunächst auf der lexikalischen, dann auf der syntaktischen und phonologischen (vgl. Moravcsik (1978: 110); auch von Thomason & Kaufman (1988: 20 f.) für normale Entlehnung, nicht für Substratwirkung akzeptiert). Wie mir scheint, darf man dies jedoch weniger als eine linguistisch-strukturelle Voraussetzung verstehen. Eher sind hier Faktoren wie Kulturgefälle, Intensität und Dauer des Kontakts und wohl auch die Frequenz der verschiedenen Morphemtypen entscheidend, die den Boden für weitergehende

Entlehnungen psychologisch vorbereiten. Der durch solche Faktoren begünstigte hohe Grad an Zweisprachigkeit nimmt den fremden Strukturen viel von ihrer Fremdartigkeit. Es sei eingeräumt, daß es auch rein linguistische Vorbedingungen für Entlehnung von Morphologie gibt: Wenn Wortbildungsmorpheme übernommen werden, dann ist diesem Prozeß gewöhnlich die Entlehnung von Wortmaterial, in dem solche Morpheme vorkommen, vorausgegangen. Für reine Flexionsmorphologie gilt dieser Gesichtspunkt in viel geringerem Maße, denn der Transfer von Wörtern mit ihrer Flexion ist nur selten zu beobachten, am ehesten im Stadium des **Sprachverfalls** (*language decay*; s. Sasse (1990: 45); Art. 153; vgl. auch Problem "Mischsprachen" 3.3).

### 3. Grundtypen der Interferenz

Verschiedene Arten von Interferenz haben unterschiedliche Konsequenzen für die beeinflußten Sprachen (s. Thomason & Kaufman 1988). Entlehnung als Übergang einzelner Elemente und Relationen aus einer Sprache in eine andere ist verständlicherweise überall beteiligt, wir können aber je nach Art und Ausmaß vier Typen von Interferenz unterscheiden:

#### 3.1. Entlehnung

Der bekannteste Typ ist die **Entlehnung** oder, wie wir in Ermangelung eines besseren Terminus sagen wollen, die normale Entlehnung, bei der keine der beteiligten Sprachen ihre Identität verliert. Die beeinflußte Sprache, die **Kopiesprache** (*replica language*) übernimmt Elemente verschiedener Ebenen aus der **Modellsprache** (*source language, model language*). Wenn die Interferenz größeren Umfang annimmt, kann sich der Charakter der Sprache wesentlich ändern; es können Strukturen akzeptiert werden, die den ererbten Strukturen zunächst durchaus fremd waren. Die Sprache bleibt jedoch "dieselbe", sie wird in ununterbrochener Folge im natürlichen Spracherwerbsprozeß tradiert.

#### 3.2. Sprachwechsel

Interferenz kann auch durch **Sprachwechsel** (*language shift*) zustandekommen. Eine Sprechergruppe wechselt ihre Sprache, meist innerhalb eines relativ kurzen Zeitraumes, und nimmt aus ihrer Muttersprache oder ihrer ersten Sprache verschiedene Phänomene mit in die neu angenommene Sprache (**Zielsprache**,

**target language**) hinüber. Die aufgegebene Sprache wird dabei gewöhnlich als **Substrat** (**substratum**) bezeichnet. Gibt eine Gruppe ihre Sprache auf, ohne die Chance gehabt zu haben, die Zielsprache ausreichend zu erlernen, entstehen **Kreolsprachen** (v. a. in der Kolonialzeit entstandene Sprachen mit gebrochener Entwicklungslinie und Zügen aus verschiedenen Sprachen, deren genaue Entstehung umstritten ist; s. auch Art. 153). Die Unterscheidung zwischen normaler Entlehnung und Interferenz durch Sprachwechsel ist deshalb wichtig, weil sich die Ergebnisse deutlich unterscheiden: Nur im ersten Fall sind viele Lehnwörter aus der Modellsprache zu erwarten, im zweiten Fall findet man eher Lehnübersetzungen nach dem Muster des Substrats, gerade auch für den grammatischen Bereich. Dies hat vielleicht damit zu tun, daß Lehnwörter für die sprachwechselnden Sprecher leichter als fremde Elemente zu erkennen sind als die viel weniger als solche ins Bewußtsein tretenden Lehnübersetzungen und deshalb besser eliminiert werden können.

Bezüglich der Morphologie und der grammatischen Kategorien kann man konstatieren, daß bei normaler Interferenz sowohl Morpheme als auch abstrakte Strukturen von grammatischen Kategorien übertragen werden, während beim Sprachwechsel zwar mit der Nachahmung von grammatischen Kategorien, aber kaum mit dem Transfer von morphologischem Material zu rechnen ist. Die Gründe scheinen dieselben wie bei der Vermeidung von Lehnwörtern zu sein.

#### 3.3. Mischsprachen

Ein weitreichender Fall von Interferenz, der früher in Abrede gestellt wurde, hat zu **Mischsprachen** (*mixed languages*; s. auch Bakker & Mous 1994, eds.) geführt. Im Unterschied zu normaler Entlehnung handelt es sich hier um Sprachen, die genetisch verschiedenen Wurzeln entspringen, also keine eindeutige Abstammungslinie haben. Sie gleichen darin den Kreolsprachen, zeichnen sich zusätzlich aber dadurch aus, daß einzelne Sprachebenen, ja selbst Unterebenen der Morphologie, verschiedener Herkunft sein können. Im Gegensatz dazu beziehen Kreolsprachen ihre Lexik im wesentlichen aus einer Sprache, ihre Kategorien bauen sie jedoch selbst (zum Teil unter Substrateinfluß) neu auf.

Beispiele für Mischsprachen:

- (a) In Europa gehören dazu verschiedene, von Zigeunern gesprochene Idiome wie Angloromani und Caló (in Spanien). Hier

ist die Lexik zum überwiegenden Teil erbt, fast die gesamte Grammatik, Morphologie wie Syntax, entstammt aber dem Englischen bzw. dem Spanischen. Hier ein Beispielsatz aus dem Angloromanen, der keine Abweichungen von der englischen Morphosyntax zeigt (englische Elemente fett; zu diesem Typ s. Hancock 1984; Boretzky 1985):

- (1) *once apre a chairus a*  
 einmal auf INDEF Zeit INDEF  
*Rommany chal chor-ed a*  
 Roma Kerl stehl-PRÄT INDEF  
*rani chillico.*  
 Frau Vogel  
 'Es war einmal ein Rom, der einen Truthahn stahl (vgl. engl. *once upon a time a Romany guy stole a turkey*)'

- (b) In Kanada wird das Mitchif oder Métif von Indianern und Mischlingen gesprochen. Das Nominalsystem stammt oberflächenstrukturell aus dem Französischen, das Verbsystem aber oberflächlich wie tiefenstrukturell aus dem Cree (s. Bakker 1989).
- (c) Die Media Lengua aus Ecuador kombiniert spanische Lexik mit gebundener Quechua-Morphologie und -Syntax (s. Muysken 1981).
- (d) Im Ma'a oder Mbugu aus Kenia und Tansania ist ein südkuschitischer Grundwortschatz mit einer durchweg bantuierten Grammatik vereint (s. Thomason 1983).

### 3.4. Diffusion

Die Verbreitung bestimmter grammatischer Phänomene über große geographische Regionen, in denen miteinander nicht verwandte Sprachen gesprochen werden und die auch kaum durch lexikalische Entlehnung verbunden sind, läßt die Existenz einer weiteren Art von Interferenz vermuten. Wir stellen z. B. fest, daß die meisten afrikanischen Sprachen den Komparativ semantisch einheitlich mithilfe eines Verbs 'passieren; übertreffen' ausdrücken. Hier muß es eine Art **Diffusion** (Sapir 1921: 203, 206) abstrakter grammatischer Strukturen geben, die möglicherweise ohne lexikalische Entlehnung abläuft. Vielleicht ist eine Bedingung für diese Art von Interferenz, daß es zwischen den Sprechergruppen kein nennenswertes Prestige- und Kulturgefälle gibt. Dieser Typ kann als eine Unterart von Entlehnung (s. 3.1) angesehen werden.

## 4. Gliederung der Phänomene

Zunächst werden die Phänomene behandelt, die häufiger vorkommen und weniger gravierende Einbrüche fremder Strukturen darstellen, es folgen dann die selteneren und ungewöhnlicheren. Die Entlehnung von Wortbildungssuffixen ist insgesamt häufiger anzutreffen als die von Flexionsmorphemen, offenbar weil erstere eher konkrete Bedeutungen haben und kaum in festgefügte Paradigmata eingebunden sind. Wenn ganz neue Affigierungstypen entlehnt werden, stellt dies bereits einen stärkeren Eingriff in die vorhandene Sprachstruktur dar. Beispiele für die Entlehnung von Flexionsmorphologie finden sich eher in kleinen Sprachen oder in Randdialekten von geringer sozialer Bedeutung, die uns besser bekannten westeuropäischen Sprachen sind fast frei davon.

### 4.1. Wortbildung

Die Übertragung von Wortbildungssuffixen auf muttersprachliches Material wird meist als mittelbare oder vermittelte Entlehnung rekonstruiert: Die Affixe gelangen mit Lehnwörtern in die Replika, werden dort abstrahiert und dann mit Erbwörtern kombiniert (Weinreich 1970: 31). Dies mag in der Mehrzahl der Fälle so sein; man sollte die Vorstellung jedoch nicht auf alle Fälle verallgemeinern, denn es gibt Kontaktsituationen, für die dieses Szenario eher unwahrscheinlich ist. Das Albanische hat z. B. eine große Zahl von Suffixen aus dem Slawischen entlehnt, z. B. *-ac*, *-aq*, *-ak*, *-icë*, *-ik*, *-inë*, *-ishtë* u. a. (s. Xhuvani & Çabej 1962), ohne daß sich slawische Lehnwörter in nennenswertem Umfang finden würden. Meines Erachtens reicht weitgehende **Zweisprachigkeit** oder **Bilingualismus** (**bilingualism**) aus, um solche Phänomene zu erklären. Die Entlehnung von Wörtern mit solchen Suffixen und die Übertragung von diesen Suffixen auf ererbtes Wortmaterial können parallel ablaufen, die Sprecher brauchen nicht unbedingt auf Lehnwörter zurückzugreifen, wenn ihnen die Kontaktsprache voll zugänglich ist. Wir sollten uns nicht nur den strukturellen, sondern auch den psycholinguistischen Aspekt solcher Situationen klar machen. Wenn es bereits eine große Menge von Lehnwörtern mit den entsprechenden Morphemen gibt, die wichtige Begriffe besetzen und frequent sind, dann ist die Annahme gerechtfertigt, die Morpheme seien von diesen Lehnwörtern auf ererbtes Wortmaterial übertragen worden. Wo es jedoch nur eine

Handvoll von Lehnwörtern des entsprechenden Typs gegeben hat, die u. U. eher *ad hoc* gebraucht werden, ist es unwahrscheinlich, daß sich die Sprecher an ihnen ausgerichtet haben. Bei Lehnübersetzungen oder semantischen Nachbildungen von Wortbildungs- und Flexionsmorphemen (s. 4.2.2) wäre die hier in Zweifel gezogene Position ohnehin schlecht haltbar; wir müßten ja dann mit innersprachlichen Übersetzungen rechnen.

Beispiele für Präfixe: englisch *re-* < französisch/lateinisch *re-*: *re-write*, *re-draw*, *re-ify*, evtl. über Fälle wie *re-turn* < französisch *retourner*; schwedisch *be-* < deutsch *be-*: *be-vittna* ‘bezeugen’, *be-gagna* ‘gebrauchen’, *be-teende* ‘Verhalten’, evtl. über Lehnwörter wie *be-fria*, *be-döva* ‘betäuben’ vermittelt.

Rumänische und slawische Verbpräfixe sind in verschiedene Romanidialekte eingedrungen; vgl. *do-lav* ‘nehmen’ < slawisch *do-* ‘zu-’, *za-bistrava* ‘vergessen’ < slawisch *za-*. Hierbei ist das Romaniverb nach serbisch *zaboraviti* geformt worden. Es handelt sich um eine **Verstärkung (reinforcement)**, da *bistrava* bereits ‘vergessen’ bedeutet. Weiter *des-mekljol* ‘auftauen’ < rumänisch *des-* ‘ent-’ u. ä. (eigene Erhebungen).

In diesen Fällen haben wir es eher mit Innovationen, Bereicherungen zu tun, da vor der Entlehnung keine direkten Entsprechungen vorhanden waren. Bei den neuen Präfixen im Romani ist sogar eine neue Wortbildungsklasse mit einem neuen Morphemtyp entstanden, denn Romani kannte keine Präfigierung.

Anders verhält es sich in Fällen, in denen ein neues Suffix ein altes substituiert, z. B. bei deutsch *-er* < lateinisch *-arius*, in *Bäck-er*, *Töpf-er* usw., wofür es ältere Bildungen wie mittelhochdeutsch *becke*, neuhochdeutsch *Beck* gab; oder bei dem maskulinen Deminutivsuffix *-ici* im Romani < griechisch *-ítsi*, z. B. in *bobici* (eine Bohnensorte) zu *bobo* ‘Bohne’. Hier gibt es ein ererbtes Äquivalent *-oro*, das allerdings meist nur an Erbwörter tritt (eigene Erhebungen).

Eine Besonderheit bietet der folgende Fall: das Istrorumänische, eine dem Rumänischen nahestehende Sprachform auf Istrien, hat aus dem Kroatischen eine Reihe von Verbaffixen übernommen. Dieser Vorgang geht jedoch über bloße Wortbildung hinaus und betrifft grammatische Kategorien, denn die Affixe dienen dem Aufbau eines Aspektsystems, also der Schaffung von imperfektiv-perfektiven Verbpaaren nach slawischem Muster, (zu den Details und weiterer Literatur s. Sala 1988).

Zwischen Wortbildung und Flexionsmorphologie stehen die folgenden, in der Geschichte des Romani zu beobachtenden Phänomene. Lehnwörter werden im Romani anders als die voreuropäischen Erbwörter, und zwar meist nach griechischen Mustern, flektiert; so gehen ererbte Maskulina mehrheitlich in der Grundform Nominativ Singular auf *-ó* aus, z. B. *čav-ó* ‘Junge’, entlehnte Maskulina aber mehrheitlich auf unbetontes *-os*, das aber als *-ós-* auch in den obliquen Kasus erhalten bleibt, weshalb *-os* nicht einfach als Morphem ‘NOM.SG.M’ bezeichnet werden kann, sondern auch Wortbildungsfunktion hat. Als Beispiel sei *dúb-os* < slawisch *dub* ‘Eiche’ genannt. Dieses *-os* geht auf griechisch *-os* und *-ós* zurück, die im Griechischen gleich häufig sein mögen. Aus diesem Fall ersieht man, daß die Replika hier eine aktive Auswahl getroffen hat, indem sie ein Allomorph gewählt hat, das von dem ererbten *-ó* möglichst verschieden ist. Interessanterweise finden sich unter den Gräzismen des Romani außerhalb Griechenlands so gut wie keine Wörter, die im Griechischen selbst auf *-os* oder *-ós* ausgehen. Das deutet darauf hin, daß sich die Roma bei solchen Entlehnungen am ehesten nach Frequenzverhältnissen des Griechischen und anderen Faktoren gerichtet haben.

Auch bei den Affixen gibt es neben der direkten Entlehnung von Morphemen als lautive Gebilde den Fall der Übersetzung bzw. semantischen Nachformung. So ist schwedisch *genom-* ‘durch-’ nach deutsch *durch-*, z. B. in *genom-föra* ‘durch-führen’, geprägt worden; *genom* war im Schwedischen zunächst nur Präposition. Das Sorbische hat Verbalpartikeln (Präverbien) nach deutschem Muster entwickelt (s. Schuster-Šewc 1977: 466), obwohl es als slawische Sprache über eine große Zahl von Verbpräfixen verfügt. So finden wir das Element *horje* als Entsprechung von deutsch ‘hinauf’ in *horje-nosić* ‘hinauftragen’. Wie im Deutschen kommt es aber auch in der Position am Ende des Prädikats vor:

- (2) *ja nošu te měchi horje*  
ich trage die Säcke hinauf

#### 4.2. Flexionsmorphologie und Syntax

##### 4.2.1. Entlehnung grammatischer Partikeln und Auxiliare

Solche Prozesse betreffen zunächst die Morphosyntax, teilweise aber auch die Morphologie. Auch wenn diese Elemente unmittelbar

bei dem Wort stehen, das sie determinieren, sind sie leichter isolierbar als echte Flexionsmorpheme, denn sie haben oft eine gut greifbare Bedeutung und werden wohl deshalb leichter entlehnt. Zunächst ein Fall aus dem Anglo-Irischen (Hibernian English), einer vom irischen Substrat beeinflußten Variante des Englischen. Hier ist ein Element *be* [bi] < irisch *bí* übernommen worden, das zum Ausdruck eines habituellen Präsens dient (s. (3)).

- (3) *we be often wondering where he gets the money*  
 ‘wir pflegen uns (oft) zu wundern, woher er das Geld nimmt’ (Henry 1957: 170)

Wie man sieht, steht *be* nicht unmittelbar beim Verb, hat also nicht den Status eines Flexionsmorphems angenommen.

In Romanidialekten des Kosovo ist *tuj/tyj* u. ä. < albanisch *tuel/tye* u. ä. für den Ausdruck des Progressivs übernommen worden. Anders als im Albanischen verbindet sich *tuj* mit dem finiten Verb, nicht mit einem Partizip oder sonstigem Infinitum.

- (4) *avel tuj rov-ol*  
 er.komm PART wein-3.SG  
 ‘er kommt weinend’ (Boretzky 1989: 367)

Hier ist der Marker bereits zu einem Verbpräfix und zu einem Flexionsmorphem geworden, obwohl er dies im Albanischen nicht war. Da ältere indogermanische Verbflexion normalerweise suffigierend ist, werden solche Fälle oft falsch kategorisiert.

Auch bereits bestehende Kategorien können auf diese Weise eine Ausdruckerneuerung erfahren. Nicht selten tritt dann die Partikel an die Stelle älterer suffigierter Flexion. So hat das Albanische lateinisch *magis* als *mâ/më* zur Bildung des Komparativs übernommen (Meyer 1881); vgl. *mâ (i) mirë* ‘besser’ (bzw. ‘der bessere’). Zwischen Partikel und Adjektiv können Artikelformen treten, so daß hier kein Flexionsmorphem entstanden zu sein scheint. Es ist aber auch möglich, Komparativmarker und Artikel als präfigierte Flexionskette zu interpretieren. Vor dem lateinischen Einfluß muß es synthetische Bildungen für die Komparation gegeben haben. Hier dürfte nicht nur das Komparationselement, sondern das ganze Verfahren aus einer spätlateinischen Vorlage stammen (vgl. auch rumänisch *mai* für den Komparativ).

#### 4.2.2. Entlehnung grammatischer Strukturen durch Lehnübersetzung

Möglicherweise noch häufiger als die Entlehnung von Partikeln, Auxiliaren u. a. in ihrer lautlichen Form ist die **Lehnübersetzung (loan translation, calque)**, die Nachahmung fremder Kategorien mit eigenen Mitteln. Naturgemäß läßt sich hier der Beweis für Interferenz weniger leicht führen als bei direkter Entlehnung von Morphemen, weshalb solche Fälle oft von Verfechtern eines (nur) innersprachlich bedingten Wandels in Zweifel gezogen werden. Wo aber eine innersprachliche Erklärung mit einer kaum wahrscheinlichen Häufung von Zufällen rechnen müßte, sollten wir doch Entlehnung annehmen. Kategorien, die auf diese Weise eingeführt werden, haben eher peripherastische als morphologische Struktur. Im Gefolge einer innersprachlichen Entwicklung kann der Ausdruck aber morphologisiert werden. Als Beispiel für letzteres kann z. B. der bulgarische postponierte Artikel *-āt, -ta, -to, -te* dienen, der von einem Demonstrativpronomen stammt. Das Verfahren dürfte aus balkanromanischer (und evtl. albanischer) Quelle stammen; vgl. bulgarisch *žena-ta*, rumänisch *femei-a*, albanisch *grua-ja* ‘die Frau’ (s. Sandfeld 1930: 165–173). Wir wissen nicht genau, wie und wo die Artikelformation im Bulgarischen eingesetzt hat, fest steht jedoch, daß das Altbulgarische keinen Artikel hatte. Der Artikel ist heute im Bulgarischen eine echte Flexionskategorie.

Im Neopersischen finden sich Possessiv-Suffixe (neben selbständigen Elementen) statt -pronomina wie sonst im Indogermanischen; vgl. *pidar-am, -at, -aš* usw. ‘mein, dein, sein Vater’ usw. wie türkisch *baba-m, -n, -sı* usw. (s. Doerfer 1967: 59). Die persischen Suffixe stammen von enklitischen Pronomina. Auch hier ist echte Flexion entstanden. Entlehnung ist sehr wahrscheinlich, weil das Persische noch viele andere türkische Einflüsse aufweist (s. Doerfer 1967).

Diese Art von Interferenz reicht verschieden weit, je nachdem ob sie nur das morphologische Muster betrifft oder die Übernahme einer Kategorie zusammen mit dem morphologischen Muster. Im Prinzip ist es nicht einmal ausgeschlossen, daß fremder Druck zur Entwicklung einer Kategorie führt, ohne daß im Ergebnis eine strukturelle Ähnlichkeit zwischen Modell und Replika zu erkennen ist. Der Beweis für Interferenz ist hier besonders schwer zu erbringen. Fremder Einfluß wird dann wahrscheinlicher, wenn es sich um eine in den Sprachen der Welt selten anzu-

treffende Kategorie handelt, die in zwei nicht näher verwandten Nachbarsprachen vorkommt. So muß beim bulgarischen Narrativsystem (System der indirekten, wiedergegebenen Rede) mit türkischer Herkunft gerechnet werden, obwohl die Ausdrücke nur schwache strukturelle Ähnlichkeit miteinander aufweisen (zu den Details s. Roth 1978: 55 ff.)

Bisher sind nur Fälle angeführt worden, in denen durch Interferenz neue Strukturen geschaffen wurden. Interferenz kann sich aber auch negativ auswirken und zur Beseitigung ererbter Strukturen führen. Naturgemäß ist auch hier der Interferenzbeweis nicht leicht zu führen, weil Kategorien auch im Zuge von internem Sprachwandel schwinden können. Bei einem Fall aus dem baltisch-finnischen Sprachkontakt dürfte aber Interferenz die Ursache sein: in den livonischen (tamischen) Dialekten des Lettischen, die auf livischem, also finnischem Substrat erwachsen sind, gibt es starke Tendenzen, das Femininum und damit das ganze 2-Genus-System zu beseitigen. Dies wirkt sich nicht nur dahingehend aus, daß die Kongruenz maskulin wird, sondern daß die ursprünglichen Feminina direkt maskuline Kasusformen annehmen; vgl. *saimniec laimig-s* (statt fem. *laimig-a*) ‘die Wirtin ist glücklich’; *vist-s* ‘Huhn’ (statt *vist-a*); im Dativ Plural *mät-am* statt *mät-em* ‘Mutter’ (s. Endzelin 1922: 342). Da dieser Prozeß nur in Mundarten auf finnischem Substrat zu beobachten ist, kann bloßer Zufall als Erklärung kaum herhalten.

#### 4.2.3. Entlehnung von Flexionsmorphemen

Flexionsmorpheme werden zugegebenermaßen recht selten übernommen. Sprachen unterziehen sich diesem Einfluß wohl am ehesten, wenn sie unter besonders starkem fremdsprachigen Einfluß stehen. Solche Fälle werden auch deshalb so selten entdeckt, weil man eher in den Standards und zu selten in der *parole* nach ihnen sucht (Weinreich 1970: 33). Es hat den Anschein, als ob bei dieser Art von Interferenz meist fremde an die Stelle eigener Morphologie tritt, also die Zahl der Allomorphe eines Flexionsmorphems vermehrt wird, ohne daß irgendwelche zusätzlichen Funktionen ausgedrückt oder gar neue grammatische Kategorien eingeführt würden. Strukturelle Faktoren können allenfalls da eine Rolle spielen, wo in einer Kategorie ein schwächer markiertes Allomorph durch ein stärker markiertes ersetzt wird (vgl. Weinreich 1970: 33) und dadurch im Sinne der Natürlichkeitstheorie (Prinzip des Konstruk-

tionellen Ikonismus; vgl. Mayerthaler 1981: 23) innerhalb einer Kategorienopposition die stärker markierte Kategorie auch den merkmahaltigeren Ausdruck erhält.

Recht häufig finden wir die Entlehnung fremder Pluralmorphe. Der folgende Fall ist wahrscheinlich anders, nämlich als ein Wandel in der Flexionsklasse, zu verstehen. Im Jiddischen wird der maskuline Plural *-im*, der hebräischer Herkunft ist, auf einige Nomina deutscher Herkunft übertragen; z. B. *pojer-im* ‘Bauern’, *doktojr-im* ‘Doktoren’ (s. Weinreich 1970: 31). Da es im Jiddischen aber hebräische Wörter in großer Zahl gibt und diese dahin tendieren, die hebräische Flexion zu behalten, muß man *-im* als ein zum System des Jiddischen gehöriges Allomorph betrachten und den Fall als innersprachliche Expansion eines Allomorphs interpretieren.

Auch der nächste Fall könnte eine ähnliche Geschichte haben. In einem türkisch beeinflußten Romanidialekt, der heute in Athen gesprochen wird, sind die 1. und 2. Person Plural durch türkisch *-muz* bzw. *-nuz* verstärkt worden; z. B. *gel-am* ‘wir gingen’ zu *gela-mus*, *gel-en* ‘ihr gingt’ zu *gela-nus*. Dieser Dialekt weist aber türkische Lehnverben in großer Zahl auf, die durchweg türkische Präsens- und Präteritalflexion beibehalten. Es ist daher wahrscheinlich, daß es sich hauptsächlich, wenn auch nicht ausschließlich, um einen innersprachlichen Übertragungsprozeß handelt. Die Übertragung ist wohl auch durch die Ähnlichkeit der Morpheme begünstigt worden (ausführlicher s. Igla 1996).

Die weiteren Fälle lassen sich eher nicht in der Weise deuten. Das ambigene Plural-Allo morph des Rumänischen *-uri* wird in Romanidialekten an Lehnwörter verschiedener Herkunft angefügt, und zwar in der Form *-ur(j)* oder *-ur(j)-a*, d. h. um das Plural-Allo morph *-á* des Romani erweitert; vgl. *taljanurja* ‘Italiener’. Wie das Beispiel zeigt, tritt das Allomorph anders als im Rumänischen auch an Belebte, was wieder eine innersprachliche Expansion ist. Es ist aber weniger sicher, daß *-urj(a)* an sich aus rumänischen Lehnwörtern isoliert worden ist, denn die Roma folgen im allgemeinen konsequent der Regel, Lehnwörter grammatisch voll zu adaptieren und fremde Flexionsmorpheme nur dort beizubehalten, wo sie ganze Syntagmen aus der Kontaktsprache ad hoc übernehmen. Solche Prozesse fallen unter Interferenz, sind aber nicht als Entlehnung zu werten. Es wäre wohl eine verkürzte strukturalistische Sicht, wenn man bei allen diesen Fäl-

len den Faktor Wortentlehnung allein in Betracht zöge, ohne den Faktor der geläufigen Zweisprachigkeit zu berücksichtigen. Das deutsche Plural-Allomorph *-en* ist in rätoromanische Dialekte eingeführt worden; vgl. *vacch-en* ‘Kühe’ (Sala 1988: 8). Im Albanischen tritt türkisch *-lar* in begrenztem Umfang an albanische Wörter, besonders an Familiennamen; vgl. *Belko-llarë* ‘die Belkos’ (s. Boretzky 1975: 238). Die Regeln der Modellsprache wurden also nur selektiv kopiert. Das Baskische bildet regelmäßig eine infinite Verbform auf *-tu*, für die allgemein lateinisch-romanische Herkunft (< *-tus*, *-ta*, *-tum*) angenommen wird (s. Hurch 1989: 12, mit weiterer Literatur). Der genaue Hergang des Prozesses ist leider nicht bekannt.

Im Fall des Meglenorumänischen aus Nordgriechenland, das vor allem slawische Sprachkontakte hatte, sind an die 1. und 2. Person Singular Präsens *-u*, *-i* einer Reihe von Verben die slawischen Morpheme *-m*, *-š* angefügt worden, z. B. *aflu-m* ‘ich finde’, *aflis-š* ‘du findest’ (s. Capidan 1925: 94; Weinreich 1970: 33). Da die ursprünglichen Morpheme nicht wegfallen, ist dies ein Fall von Verstärkung. Die eher seltenen slawischen Lehnverben werden in diesem Dialekt nicht in ihrer slawischen Form flektiert, weshalb die Verstärker-Elemente direkt aus dem Slawischen übernommen sein dürften.

Das Albanische hat vom Türkischen das Adverbmorphem *-çe* übernommen; vgl. *vend-çe* ‘nach heimischer Art’, *malcor-çe* ‘wie die Bergbewohner’ (s. Boretzky 1975: 268). Im Albanischen finden wir zwar *-aslazi* als Adverb-Marker, aber sie werden vornehmlich nicht mit Nomina kombiniert. Daher ist auch eine grammatische Bereicherung mit der Entlehnung verbunden.

Nordtadžikische (d. h. persische) Dialekte haben u. a. aus dem Özbekischen, einer Türkssprache, verschiedene Kasussuffixe übernommen, z. B. den Ablativ *-dan*, vgl. *yakom klasaš-dan* ‘aus der ersten Klasse’. Daneben besteht aber weiter die Möglichkeit, Präpositionalphrasen nach iranischem Muster zu bilden (s. Doerfer 1967: 54). Diese Dialekte nähern sich dem Typ der Mischsprachen (s. 3.3).

Im folgenden Fall ist die fremde Kategorie nur unvollkommen adaptiert worden; gleichwohl handelt es sich um eine funktionelle Neuerung: In 4.1 war die Entlehnung slawischer Verbalpräfixe in Romanidialekte erwähnt worden. Im Slawischen sind diese Präfixe nicht nur wortbildend, sondern liefern u. a. den Ausdruck für den perfektiven

Aspekt des Verbs. Nach diesem Muster ist in russischen Romanidialekten nun in Ansätzen ein Verbalaspekt, also eine neue Kategorie, entstanden. Es fehlt allerdings ein Imperfektivierungsverfahren (zu den Details s. Boretzky 1989: 367 f.).

Noch stärker als die Möglichkeit der Entlehnung einzelner grammatischer Morpheme ist die ganzer Paradigmata in Abrede gestellt worden. Indessen gibt es selbst für solch weitreichende Interferenz Belege.

Mednyj Aleut, eine ursprünglich aleutische, mit dem Eskimoischen verwandte Sprache, hat seine komplizierte Verbalflexion durch die des Russischen ersetzt; vgl. für das Präsens *ujuč-im*, *-iš*, *-jat* ‘sitz-PRÄS1.SG, -PRÄS2.SG, -PRÄS3.SG’, und für das Präteritum *ujuči-l* ‘sitz-PRÄT.M’ und *ujuči-la* ‘sitz-PRÄT.F’, d. h. es sind die Formantien der sog. *i*-Konjugation angetreten. In derselben Weise ist auch die Bildung des Infinitivs und des Futurs (mit *budu* + Infinitiv) auf das Aleut übertragen worden. Morphologisch haben hier russische Suffixe aleutische substituiert; man kann jedoch nicht sagen, daß die Strukturen des Aleut dabei wirklich erhalten geblieben sind (zu den Details s. Menovščikov 1969: 132 und Thomason 1986: 277 f.). Im Ergebnis hat sich hier eine Übergangsstufe zu den Mischsprachen eingestellt (s. 3.3).

## 5. Zitierte Literatur

- Bakker, Peter (1989), “Relexification: The Case of Métif (French Cree)”. In: Boretzky et al., 119–137  
 Bakker, Peter & Mous, Maarten (1994, eds.), *Mixed Languages: 15 Case Studies in Language Intertwining*. Amsterdam: IFOTT  
 Boretzky, Norbert (1975), *Der türkische Einfluß auf das Albanische, Teil 1: Phonologie und Morphologie der albanischen Turzismen*. Wiesbaden: Harrassowitz  
 Boretzky, Norbert (1985), “Sind Zigeunersprachen Kreols?” In: Boretzky, Norbert & Enninger, Werner & Stolz, Thomas (Hrsg.), *Akten des 1. Essener Kolloquiums über ‘Kreolsprachen und Sprachkontakte’ vom 26. 1. 1985 an der Universität Essen*. Bochum: Brockmeyer, 43–70  
 Boretzky, Norbert (1989), “Zum Interferenzverhalten des Romani”. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42, 357–374  
 Boretzky, Norbert & Enninger, Werner & Stolz, Thomas (1989, Hrsg.), *Vielfalt der Kontakte. Beiträge zum 5. Essener Kolloquium über ‘Grammati-*

- kalisierung: Natürlichkeit und System-Ökonomie” vom 6. 10.–8. 10. 1988 an der Universität Essen, Bd. I.* Bochum: Brockmeyer
- Boretzky, Norbert & Igla, Birgit (1994), *Kontakt-initierter Wandel, 1. Interferenz* (= B. Jeßing (Hrsg.), Sprachdynamik. Auf dem Wege zu einer Typologie sprachlichen Wandels, Bd. III. Bochum: Brockmeyer)
- Capidan, Theodor (1925), *Meglenoromăni I*. Bucureşti: Cultura Naţională
- Doerfer, Gerhard (1967), *Türkische Lehnwörter im Tadschikischen*. Wiesbaden: Steiner (Abhandlungen für die Kunde des Morgenlandes 37.3)
- Endzelin, Jan (1922), *Lettische Grammatik*. Riga: Gulbis
- Hancock, Ian (1984), “Romani and Angloromani”. In: Trudgill, Peter (Hrsg.), *Language in the British Isles*. Cambridge: Cambridge University Press, 367–383
- Henry, Patrick Leo (1957), *An Anglo-Irish Dialect of North-Roscommon*. Zürich: Aschmann & Scheller
- Hurch, Bernhard (1989), “Hispanisierung im Bas-kischen”. In: Boretzky et al. (Hrsg.), 11–35
- Igla, Birgit (1996), *Das Romani von Ajia Varvara: Deskriptive und historisch-vergleichende Darstellung eines Romanidialekts*. Wiesbaden: Harrassowitz
- Martinet, André (1955), *Economie des changements phonétiques: Traité de phonologie diachronique*. Bern: Francke
- Mayerthaler, Willi (1981), *Morphologische Natürlichkeit*. Wiesbaden: Athenaion
- Meillet, Antoine (1921), *Linguistique historique et linguistique générale*. Paris: Champion
- Menovščikov, Georgij Alekseevič (1969), “O nekotorych social'nych aspektach evoljucii jazyka”. *Voprosy social'noj lingvistiki*. Leningrad: Nauka, 110–134
- Meyer, Gustav (1881), *Etymologisches Wörterbuch der albanesischen Sprache*. Strassburg: Trübner
- Moravcsik, Edith (1978), “Language Contact”. In: Greenberg, Josef H. (Hrsg.), *Universals of Human Language, Vol. I*. Stanford: Stanford University Press, 93–122
- Muysken, Pieter (1981), “Halfway between Quechua and Spanish: The Case for Relexification”. In: Highfield, Arnold & Valdman, Albert (Hrsg.), *Historicity and Variation in Creole Studies*. Ann Arbor: Karoma, 52–78
- Oksaar, Els (1972), “Bilingualism”. In: Sebeok, Thomas A. (Hrsg.), *Current Trends in Linguistics, Vol. IX: Linguistics in Western Europe*. The Hague: Mouton, 476–511
- Roth, Juliana (1978), *Die indirekten Erlebnisformen im Bulgarischen*. München: Sagner
- Sala, Marius (1988), *El problema de las lenguas en contacto*. México: Universidad de México
- Sandfeld, Kristian (1930), *Linguistique balkanique*. Paris: Champion
- Sapir, Edward (1921), *Language*. New York: Harcourt, Brace & World
- Sasse, Hans-Jürgen (1990), *Theory of Language Decay und Contact-induced Change: Similarities and Differences*. Institut für Sprachwissenschaft – Universität zu Köln. Arbeitspapiere Nr. 12 (Neue Folge)
- Schuster-Šewc, Heinz (1977), “Die deutschen Lehnübersetzungen im Obersorbischen und ihre Stellung im System der obersorbischen Lexik und Wortbildung”. *Zeitschrift für Slavistik* 7, 455–468
- Spitzer, Leo (1928), *Hugo Schuchardt-Brevier: Ein Vademeum der allgemeinen Sprachwissenschaft*. Halle/Saale: Niemeyer
- Thomason, Sarah Grey (1983), “Genetic Relationship and the Case of Ma'a (Mbugu)”. *Studies in African Linguistics* 14, 195–231
- Thomason, Sarah Grey (1986), “Contact-induced Language Change: Possibilities and Probabilities”. In: Boretzky, Norbert & Enninger, Werner & Stolz, Thomas (Hrsg.), *Akten des 2. Essener Kolloquiums über “Kreolsprachen und Sprachkontakte” vom 29. und 30. 11. 1985 an der Universität Essen*. Bochum: Brockmeyer, 261–284
- Thomason, Sarah Grey & Kaufman, Terence (1988), *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press
- Weinreich, Uriel (1970), *Languages in Contact*. The Hague: Mouton
- Xhuvani, Aleksandër & Çabej, Ekrem (1962), *Prapashtesat e gjuhës shqipe*. Tiranë: Universiteti Shtetëror

Norbert Boretzky, Bochum (Deutschland)

## 153. Pidginization, creolization, and language death

1. Introduction
2. Reduction processes: pidginization and language death
3. Expansion processes: creolization
4. References

### 1. Introduction

In the field of pidgin and creole studies the main question is how exactly a new language can come into existence, and how the particular grammatical properties of the newly formed languages, **pidgins** and **creoles**, are related to the way in which they have emerged. A pidgin language is generally defined as a strongly reduced linguistic system that is used for incidental contacts between speakers of different languages, and that is the native language of nobody (DeCamp 1971). A creole language is a language that has emerged when a pidgin has acquired native speakers. The social circumstances that lead to the emergence of a **jargon** (a very primitive contact system) and consequently of a more stable pidgin generally involve the migration of a socially dominated group. This can be in the context of slavery or of contract labour in a colonial setting. Often trade carried out on an unequal footing is involved. A group of people is forced by the circumstances to develop a new communication system, to be used with foreigners who do not speak their own language.

Even though various languages are involved when a jargon or pidgin emerges, the vocabulary of a pidgin generally derives from one language: the language that is socially or politically dominant in the original contact situation. Because most pidgins have resulted from the European (and later more generally Western) colonial expansion, starting in the 15th century, the vocabulary of most pidgins and creoles is derived from a European language (Portuguese, English, Spanish, French, Dutch).

Most pidgins still in existence are spoken in Africa and in the Pacific Ocean. Creole languages are found predominantly in the Caribbean, in West Africa, in the Indian Ocean, and in the Far East. The most well-known of the creoles are spoken, of course, by the descendants of the black slaves who were brought to the Caribbean to work in the sugar plantations.

**Language death** is obviously the falling out of use of a language, either because its original speech community is dissolved or because the community shifts to another language. Language death may (but need not, this is a point of debate) involve gradual reduction of the resources of the language in question, lexical, morpho-syntactic, stylistic. Language death occurs in every part of the world; at present it is particularly frequent in the Americas and in Australia, where many native languages are in the process of disappearing.

When looking at pidginization, creolization, and language death from the perspective of morphology, it is necessary to consider both reduction and expansion.

### 2. Reduction processes: pidginization and language death

From the point of view of morphology, the most important question concerning pidginization and language death is: can morphological elements survive the processes of morphological reduction inherent in these developments, and if so, which ones? This question will first be discussed in relation to pidginization, and then to language death, before a more general, comparative perspective is introduced.

#### 2.1. Pidginization

Most Caribbean creoles appear to have lost the largest part of the suffixes of their lexical source languages. This is particularly true of inflection. A case in point is the Haitian Creole verb, which is invariant in form, whereas French, the lexical source language, has an extensive paradigm with tense, mood, and person inflections: invariant *reme* 'love' versus *ɛ:mə*, *ɛ:mõ*, *ɛ:me*, etc.

The Mediterranean contact language *Lingua Franca* (Schuchardt 1909) exemplifies a pidgin Romance with invariant verb forms:

- (1) *mi star contento mirar per ti*  
 I be glad see for you  
 "I am glad to see you"

Here the source language infinitive endings are generally used on the invariant verb forms. There is one exception to the invariance in verbal morphology:

- (2) *mi mira-to per ti*  
 I see-PF for you  
 "I have seen you"

Here the past participle is used to mark perfective aspect.

The loss of morphology in pidgins could result from a principle such as "Express each separately intuited element of meaning by a separate form" (Naro 1978) if speakers of the dominant languages are responsible for the reduction (in a kind of foreigner talk strategy), or from properties of early second language learning processes (in which words are acquired first in an invariant form), if the speakers of the subordinate languages are (Ferguson 1975; Andersen 1983, ed.).

Nonetheless, before attempting to explain this supposed absence of inherited morphology, one should ask whether it is real. It turns out that in the Caribbean creoles, many derivational affixes from the European lexical source languages did survive (Holm 1989; see also DeGraff 2001). An example is Papiamentu, spoken on some of the Netherlands Antilles and on Aruba, where the action nominalizer *-shon* (< Ibero-Romance *-ción*) and the agentive marker *-dō* (< Ibero-Romance *-dor*) are used productively.

Non-European-based pidgins and creoles (Heine 1978) provide further evidence against the supposed loss of morphology and show that it is risky to make generalizations on the basis of pidgin and creole derived from European colonial languages. Thus Nubi Arabic has lost the Arabic binyan system, but preserves a number of Arabic compounds, prefixes, and suffixes (Heine 1982). Similarly, in Kituba there has been loss of subject and object clitic pronouns and partial loss of tense/aspect morphology, but preservation of derivational morphology.

What explains this partial loss and partial retention? Consider first some data from Amazon Quechua (Ecuador). Later a tentative general answer in terms of morphological typology will be given. A simple example can be used to illustrate some of the properties of Amazon Quechua, a resultant of ethnolinguistic restructuring concomitant with colonization and drastic population decline in the Amazon basin. Many different small groups of Indians are forced to disband and they form new, more stable, communities in which the peasant language of the Andean highlands is adopted.

- (3) *riku-u-ni yura-ma pishku-una*  
 see-PRES-1 tree-to bird-PL  
*tiya-nau-nga-chu ni-sha*  
 be-PL-3.FUT-INT say-SR  
 'I am looking at the trees to find out if there are any birds (lit. saying: will there be any birds).'

Amazon Quechua presents many features that may be attributed to pidginization and creolization:

- (a) the loss of morphophonemic adjustment rules such as the vowel lowering rule (Adelaar 1977);
- (b) loss of nominal person marking;
- (c) reduction from 17 to 7 modal suffixes on the verb;
- (d) formation of a periphrastic future/modal;
- (e) reinterpretation of the subordinator *-pti-* as *-k-pi* 'agentive-locative', and extension to other nominalizers, yielding *-shka-pi* and *-na-pi*;
- (f) development of a periphrastic construction involving the complementizer *ni* + *sha* 'saying' (this may be due to Jivaroan substrate influence).

Nonetheless, morpheme counts (for 100 words of text) suggest that Amazon Quechua is morphologically less complex than e.g. Cuzco Quechua (Peru), but certainly has not lost all of its morphology:

	traditional narrative	syntactic field notes
Cuzco Quechua:	2.01	2.54
Amazon Quechua:	1.74	2.20

Tab. 153.1: Morpheme/word proportions in two Quechua varieties

To comprehend that the pattern of reduction is not random, consider the set of Quechua person markers in Cuzco Quechua and Amazon Quechua (see Table 153.2).

A perusal of the forms that have survived shows that: (a) the first person plural inclusive/exclusive distinction has been lost for subjects; (b) the only object forms that have survived are those that can be fit into the scheme {1.SG OBJECT} – {SUBJECT}. The syntactically irregular forms have disappeared.

The fact that so many affixes have survived at all in pidgin/creole varieties of languages such as Quechua as well as in Swa-

Cuzco	Amazon	
-ni	-ni	1.SG
-nki	-ngi	2.SG
-n	-n	3.SG
-y-ku		1.PE
-n-chis		1.PI
	-nchi	1.PL
-nki-chis	-ngichi	2.PL
-n-ku	(-nau..)-n	3.PL
-yki		SBJ.1.SG&OBJ.2.SG
-wa-nki	-wa-ngi	SBJ.2.SG&OBJ.1.SG
-wa-n	-wa-n	SBJ.3.SG&OBJ.1.SG
-su-nki		SBJ.3.SG&OBJ.2.SG
-wa-nchis		SBJ.3.SG&OBJ.1.PI
-yki-ku		SBJ.1.PE&OBJ.2.SG
-yki-chis		SBJ.1.SG&OBJ.2.PL
-wa-nki-ku		SBJ.2.SG&OBJ.1.PE
-wa-nki-chis	-wa-ngichi	SBJ.2.PL&OBJ.1.SG
-wa-nchis-ku		SBJ.3.PL&OBJ.1.PI
-su-nki-chis		SBJ.3.SG&OBJ.2.PL
-su-nki-ku		SBJ.3.PL&OBJ.2.SG
-wa-n-ku		SBJ.3.SG&OBJ.1.PE

Tab. 153.2: Quechua person markers

hili, for instance, no doubt has to do with their agglutinative character. This leads to the issue of morphological typology (cf. Anderson 1985).

At least five dimensions need to be distinguished in a morphological typology, with respect to the survival of specific morphological patterns:

- (a) The type of concept that can be expressed by morphemes (Sapir 1921). In most of the pidgin and creole languages only derivational concepts are expressed morphologically, but it is not clear whether these are inherited morphemes.
- (b) The extent to which words are complex. Complexity by itself does not determine whether a form is maintained or lost in pidginization processes.
- (c) The extent to which there are morphophonological relations between the components of a word. Consider in this respect **allomorphy**, the type of fusion found in Semitic systems, etc. It will be clear that this dimension plays an important role. Only morphological patterns that rate low on this dimension can survive a process of reduction.
- (d) The extent to which there are non-syntactic selectional restrictions between the

components of a word. Relevant in this respect are elements sensitive to the declension class of their base, or subject to lexical specification. Again, elements low on this dimension have a better chance to survive a process of reduction, and **agglutination** can be defined as the absence of morphological selectivity.

- (e) The extent to which the meanings of complex forms are paradigmatically rather than syntagmatically defined. The case of Amazon Quechua person marking suggests that paradigmaticity could play a role: only those elements that are syntagmatically interpretable survive. The loss of Indo-European inflectional morphology in the Caribbean creoles can be interpreted in the same light: involved are highly paradigmatized morphological subsystems.

It is beyond the scope of this article to further develop the notions sketched here. What is important is the need for a much more differentiated view of the process of morphological reduction (see Kusters 2003).

## 2.2. Language death

Although highly frequent, the processes of language **attrition**, language **obsolescence**, and **language death** have been little studied. In part this is because it is a subject fraught with methodological problems. To mention but two: (a) what is the reference point for the process of loss or decay investigated; and (b) how to account for the discrepancy that many speakers show between active and passive competence in the language undergoing loss?

What follows is based on two of the best-known studies in this field: one on Scottish Gaelic (Dorian 1981) and one on the Australian aboriginal language Dyirbal (Schmidt 1985). The studies compare the speech of the older, more traditional speakers with that of the younger speakers, primarily using oral sentence translation tasks.

Although the differences between East Sutherland fishing communities and the Jambun communities in northeast Queensland, Australia, could not be larger, there are some striking similarities in the results of the two studies. Language death is not directly apparent in a decreasing fluency, but rather in morphological and syntactic changes, which could also have been the result of ordinary linguistic change, but are more numerous

and proceed at a faster rate. Both studies document a gradual decrease in use of some traditional morphological elements of the language, though not of others (Dorian 1981: 146–151; Schmidt 1985: 76–78). For both languages, case endings are highly unstable; it may well be that the alternative signalling function of word order and borrowed prepositions (both languages are giving way to varieties of English) is a factor here.

For the rest, typological differences make a detailed comparison impossible. In Gaelic verbal number is also unstable as a category, but nominal number and gender are stable. Tense and passive morphology are maintained by younger fluent speakers but decline in frequency in the speech of semi-speakers of Gaelic. In young people's Dyirbal, almost all derivational affixes survive, as well as some mood and aspect markers. Consider an example such as (Schmidt 1985: 217):

- (4) *oh she baji-baji-yarra-nyu down*  
*oh she fall-RDP-ASPECT-NFUT down*  
 “Oh! she started to fall down!”

In addition to case markers, some tense markers and other verbal inflectional affixes tend to disappear.

Both studies stress the role of **borrowing** in language attrition and death, but borrowing is not a reliable diagnostic for it. What makes a fragment of spoken Quechua, for instance, show signs of attrition is not the amount of Spanish borrowing: much more fluent speakers also use many Spanish words. Rather it is the lack of variation in the syntax and morphology: only a few suffixes are used, even if they are used frequently.

What is urgently needed is a comparative and controlled study of morphological loss and reduction in the context of language death, where the relative weight of formal linguistic factors (e.g. the dimensions referred to in 2.1), frequency, phonological saliency, functional load, etc. is studied.

### 2.3. A comparison between reduction due to pidginization and to language death

Both studies of language death briefly discussed stress the differences between language death and pidginization (Dorian 1981: 152–156; Schmidt 1985: 214–218). In addition to the numerous functional differences, the main difference mentioned is the amount of morphology retained from the source language.

As seen in 2.1, morphological simplicity *per se* is not a necessary concomitant of pidginization, when the source language is agglutinative, for instance. Indeed, one could imagine a morphologically complex pidgin-Dyirbal, given the transparent suffixal nature of its morphology. For Gaelic, the situation is quite different, however. Here the morphological distinctions are highly fusional and allomorphic, involving initial lenition and nasalization, and final consonant mutation as well. It is hard to conceive of a pidgin that would have maintained even a small portion of Gaelic morphology.

We may conclude that the evidence points to a qualitative difference between the two processes: the type of disruption in the transmission of linguistic forms is much more radical in the case of pidginization.

### 3. Expansion processes: creolization

The creole languages that emerged out of pidgins with various degrees of elaboration and complexity reconstituted the lexical richness characteristic of the native language of a speech community in various ways. Here several of the **lexical expansion** strategies used will be surveyed: compounding, phrasal compounding, affixation, reduplication, tone- or stress-shift, conversion. The survey can only be preliminary because creole morphology is a sorely neglected area of study (due to the wide-spread belief that creoles have no morphology to speak off).

Three preliminary general remarks are in order. First, the reconstituted system is neither in its organization nor in its formal aspects necessarily a replica of what is found in the European colonial lexical source language. A case in point is Papiamentu, where no less than seven different forms replaced the Ibero-Romance “reflexive” clitic *se* (Muysken 1993; Muysken & Smith 1994: 285 f.):

- (5) *paña* (< Port. *pano*, Span. *pañ* ‘cloth’)  
*mi ta bisti paña* (< Span. *vestirse*)  
 ‘I dress (myself)’
- (6) *kurpa* (< Port. *corpo*, Span. *cuerpo*  
 ‘body’)  
*lanta kurpa* (< Span. *levantarse*)  
 ‘get up’
- (7) null reflexive  
*diskulpá* (< Span. *disculparse*) ‘excuse oneself’

- (8) possessive + *kurpa*  
*skonde su kurpa* (< Span. *esconderse*)  
'hide'
- (9) object pronoun  
*sinti e tristi* (< Span. *sentirse*) 'feel sad'  
*haña e* (< Span. *hallarse*) 'find oneself'
- (10) object pronoun + *mes* (< Port. *mesmo*  
'self')  
*e ta komport'e (mes)* (< Span. *comportarse*) 'he behaves himself'
- (11) possessive pronoun + *mes*  
*hasi su mes malu* (< Span. *hacerse mal*)  
'hurt oneself'

Examples like these illustrate the morpho-lexical autonomy of the creoles with respect to their lexical source languages, even in a case such as Papiamentu, which would seem to be rather close to its lexical source languages Spanish and Portuguese.

Second, particularly in the domain of tense/mood/aspect, European verbal affixes have been replaced by separate, grammaticalized, pre-verbal particles (Holm 1989). Thus the progressive marker *ta*, as in *mi tā hóndi pingo* 'I hunt boar' in the Surinam maroon language Saramaccan derives from Eng. *stand* (Smith 1987). These particles have a fixed position, but may not be considered verbal prefixes. When subject to sandhi rules, they tend to form a unit with the preceding subject rather than with the following verb.

Third, lexical expansion also occurs in pidgins, to some extent. Consider the question words of Chinese Pidgin English:

	forms	analysis
WHO	<i>who (-man)</i>	'who (-man)'
WHAT	<i>wat ting</i>	'INT-thing'
WHEN	<i>wat-time</i>	'INT-time'
WHERE	<i>wat-side</i>	'INT-side'
WHY	<i>wat-for</i>	'INT-for'
HOW	<i>how (-fashion)</i> <i>wat-fashion</i>	'how (-fashion)' 'INT-fashion'

Tab. 153.3: The question words of Chinese Pidgin English (Bisang 1985, as analyzed in Muysken & Smith 1990)

This system is highly transparent in that there is a stable question particle derived from English *what*.

### 3.1. Compounding

The primary means of lexical expansion in almost all, if not all, creoles is compounding, which tends to be highly productive. Com-

pounding is used for all word classes. The creole language of coastal Surinam Sranan has *mofoneti* 'mouth night (midnight)' and *bobimofo* 'breast mouth (nipple)', Haitian has *bouch kabrit* 'mouth goat (Cassandra)' (Hall 1953: 41). Where Papiamentu has *bula bay* 'fly go (fly away)', a compound derived from a serial verb construction, Haitian has *magne manje* 'touch eat (just eat a little)' (Hall 1953: 42). Serial verbs are verbs that can be concatenated within a single clause.

Compounding is also used to form new function words. Some examples from Papiamentu and Chinese Pidgin English were already given in 3. Compounding is also illustrated in an example from the reflexives in Saramaccan and Sranan (cf. Muysken & Smith 1994: 276):

	Saramaccan	Sranan
myself	<i>mi-séei</i>	<i>mi-srefi</i>
yourself	<i>juli-séei</i>	<i>ju-srefi</i>
himself	<i>en-séei</i>	<i>en-srefi</i>
ourselves	<i>wilu-séei</i>	<i>wi-srefi</i>
yourselves	<i>unu-séei</i>	<i>unu-srefi</i>
themselves	<i>den-séei</i>	<i>den-srefi</i>

Tab. 153.4: Reflexives in Saramaccan and Sranan

At first sight, it would seem that these forms are direct reflexes of the English forms. Notice, however, that they are directly based on Saramaccan and Sranan pronouns, respectively. Only for the third person forms a direct inheritance could be postulated.

We find that quantifiers, different kinds of pronouns, question words, complex prepositions, etc. are all formed through compounding in many creole languages.

### 3.2. Phrasal compounding

Very common in some creole languages is **phrasal compounding**. It involves (a) cases where one of the members of the compound (in fact, the non-head) is a phrase rather than a word, and (b) cases where the structure of the compound itself reflects a syntactic rather than a morphological pattern.

An example of the first type is Saramaccan agentive compound formation, involving the noun *ma* 'man'. It is attached to verbs or verbal complexes:

- (12) *pai-ma*  
bear-man 'pregnant woman'

- hóndi-ma*  
hunt-man ‘huntsman’  
*tjá-búika-ma*  
carry-mouth-man ‘messenger’  
*pai-ku-mujée-ma*  
bear-with-woman-man ‘midwife’

Notice first of all that the meaning of *ma* is no longer exclusively masculine but rather ‘person that ...’, i.e. the form is somewhat grammaticalized (cf. 3.3). Second, the left-hand member can contain nouns (*búika* ‘mouth’), prepositional phrases (*ku mujée* ‘with woman’), and even clauses and serial constructions.

Similar perhaps are cases in Haitian where verb + object noun combinations function as nouns (Hall 1953: 41) (cf. 3.6):

- (13) *pase raj* pass rage ‘exotic dance’  
*lävi mouri* want die ‘imprudent person’  
*pote mak* bear mark ‘he who is scarred’

An example of the second type is Papiamentu, where most compounds include a linking morpheme *dili* ‘of’ (Dijkhoff 1987):

- (14) *palu di garganta* stick of neck ‘neck bone’  
*kabes di boto* head of boat ‘lift’  
*barba di yònkuman* beard of young.man ‘herb’

It can be shown that these forms, in spite of their syntactic appearance, behave as lexical islands for pluralization, adjectival modification, extraction, etc. (Dijkhoff 1987).

### 3.3. Affixation

Affixation is more limited in the morphological systems of creoles. It is not easy for a language to develop new affixes. There are two ways in which this can be done:

- (a) by borrowing; in 2.1 the case of Papiamentu nominalizers was mentioned as an example of retention from the source language;
- (b) reanalyzing part of a compound; it is possible that the righthand member of a compound is semantically bleached and phonologically weakened so that it can develop into a suffix.

A possible example may be Saramaccan *ma* mentioned in 3.2. Similar Saramaccan examples are *fási* compounds:

- (15) *tjalí fási* sorrow fashion ‘sorrowful’  
*tooká fási* change fashion ‘change in appearance’  
*bínu fási* good fashion ‘generous’

Other Saramaccan nouns which may be turning into affixes in a similar way are *se* ‘side’, *kamia* ‘place’, *libi* ‘life’, *sembe* ‘somebody’, *soni* ‘something’, and *te* ‘time’.

Another way to introduce affixation is through borrowing. Two nominal affixes in Berbice Dutch creole are borrowed from the African language Eastern Ijo: *-je* ‘NR’ and *-apu* ‘PL’ (Kouwenberg 1991; Smith et al. 1987). The form *-je* occurs in forms such as:

- (16) *di kalijé* ‘the small one’  
*en gu seteje* ‘a big fat one’  
*di eshtije* ‘the first one’

Notice that the category plural, when attached to terms for individuals (‘Hilda’ below) can have the meaning of an associative plural (see Art. 100), something common in creoles:

- (17) *matj-ap* ‘friends’  
*gugu-j-apu* ‘the big ones’  
*hild-apo* ‘Hilda and her family’

The nominalizing element *-je* in Berbice Dutch may be an enclitic noun, or a phrasal affix with nominal status. It is possible to add this element more than once, as can be seen from:

- (18) *kali-je-apu-je-apu* [kalijapjapu]  
 small-NR-PL-NR-PL  
 ‘the ones that belong to the small ones’

A more complex form of borrowing involves cases where the lexical source language and a substrate donor language have a morpheme similar in form in both languages. The adoption in the creole has been analyzed as the result of conflation (Kihm 1988) or convergence (Kouwenberg 1992). In Guinea-Bissau Portuguese Creole pairs occur such as:

- (19) (a) *firma* ‘stand’  
*firma-n-ta* ‘raise, put on its base’  
 (b) *sibi* ‘go up’  
*sibi-n-ti* ‘make go up’  
 (c) *yentra* ‘go in’  
*yentra-n-da* ‘make go in’

The causative ending could have two sources (Kihm 1988): (i) the rather unproductive Portuguese derivational suffix *-nt-* ‘causative, inchoative’ that occurs in:

- (20) (a) *ferver* ‘boil/be boiling’  
*aserve-nt-ar* ‘boil/bring to boil’
- (b) *quebrar* ‘break’ *quebra-nt-ar* ‘break something’

(ii) a causative suffix in the neighbouring West-African languages:

- (21) (a) Manjaku:  
    *-lenp* ‘work’  
    *-lenpandan* ‘make work’
- (b) Mandinka  
    *sonka* ‘quarrel’  
    *sonkandi* ‘cause to quarrel’

The process of conflation would involve a mutual reinforcement of these two.

The Berbice Dutch Creole perfective marker *-te* (Kouwenberg 1992) can be linked both to a similar Ijo form and to the Dutch past tense suffix *-ta*. Again, presumably its presence in both contributing languages favored its adoption in the Creole.

### 3.4. Reduplication

Many Europeans will immediately associate **reduplication** with pidgins and creoles, because it is part of stereotypical foreigner talk, “primitive” language use. However, the idea of reduplication as an iconic means of intensification or pluralization, still adopted by Schuchardt (1914), for instance, does not necessarily correspond to the way this device is used in most creoles. Reduplication occurs frequently with adjectives and verbs in Berbice Dutch (Kouwenberg 1991), but it has different meanings with each category. With adjectives it can be:

- (22) (a) intensifying  
*kal-kali* ‘tiny’
- (b) emphatic  
*eshti-eshti* ‘the very first’
- (c) distributive  
*do pote-pote kenap* ‘the old people (one by one)’

With verbs it is iterative, which has a different nuance of interpretation with different examples of verbal usage:

- (23) (a) iterative-habitual  
*futel-futel* ‘converse regularly’
- (b) iterative-aimlessness  
*kap-kap* ‘cut here and there’

- (c) iterative-repetitive  
*mu-mu* ‘to keep going (step by step)’
- (d) iterative-distributive  
*findi-fint* ‘to open one by one’

Reduplication is recursive, as can be seen from examples such as:

- (24) (a) *eni bu-bu-bu-bu-te wer*  
they drink-RDP-RDP-RDP-PF again  
‘they kept drinking on and on again’
- (b) *ek hafu*  
I have.to  
*bjonto-bjonto-bjonto-bjonto*  
remember-RDP-RDP-RDP  
‘I have to keep on remembering’

In this case reduplication does have its iconic intensifying meaning. It may be the case that reduplication has two sources: as an iconic device, not fully grammaticalized, and as a special type of affixation, in which an unspecified affix is added which is then spelled out as a copy of the base form.

In Saramaccan there are many kinds of reduplication (Bakker 1987). Productive, grammaticalized reduplication can have a meaning completely different from the iconic one. With verbs it is used to form resultative adjectives:

- (25) *di lai-lai goni* ‘the loaded gun’  
(*lai* ‘to load’)
- di dee-dee koosu* ‘the dried cloth’  
(*de* ‘to dry’)

Adjectives and nouns can also be reduplicated, yielding the meaning ‘X-ish’ or ‘X-like’:

- (26) *geligeli* ‘yellowish’ (*geli* ‘yellow’)  
*baafubaafu* ‘soup-like’ (*baafu* ‘soup’)

### 3.5. Stress and tone shift

Though not as common as other morphological devices, there are instances of **stress-** and **tone-shift** in creoles. One illustration is the verb *duna* ‘give’ in Papiamentu. Its citation form has penultimate stress and a low-high tone pattern. In the imperative, however, there is high-low stress with bisyllabic verbs of the *duna* class (Römer 1983). The past participle form has shift of stress to the last syllable, but maintains low-high tone.

The same shift pattern occurs with some deverbal nouns:

- (27) (a) *piska* ‘to fish’  
*piská* ‘a fish’
- (b) *pika* ‘to stab, to burn’  
*piká* ‘a stab, a burn’

### 3.6. Multifunctionality

In all creoles lexical items are multifunctional in that they can belong to several word classes at a time. Assuming there to be a base category for each item, a morphological process of **conversion** or **zero-derivation** can be postulated (cf. Art. 90).

Voorhoeve (1981) has studied this problem in some detail for Sranan. He posits rules of conversion, on the basis of regular patterns of correspondence. Thus the word *siki* can be the adjective 'ill', the noun 'illness', the transitive verb 'make (someone) ill' and the intransitive verb 'be ill'. The same pattern holds for *hebi* 'heavy', *blaka* 'black', etc. while *dede* 'dead', *bigi* 'big', and *bun* 'good' cannot have the interpretation of being a transitive verb, while sharing the other three options. *Wasi* 'wash' can only be a transitive and an intransitive verb, but not a noun.

The major problems posed by forms such as these are (a) how to establish a base form; (b) how to formulate rules that account for the regularities in what is possible and what is not?

A complex case of **multifunctionality**, where the category of the element involved is contextually determined, involves adjectives /stative verbs in Sranan. These elements function as stative verbs when used predicatively and cannot have a preceding copula (*de* in Sranan):

- (28) (a) *a liba bradi*  
‘the river is wide’
- (b) \* *a liba de bradi*

*Bradi* is a verb, and hence a copula is not allowed, presumably. Consider now a case where *bradi* is preceded by the adjectival modifier *so*:

- (29) (a) \* *a liba so bradi*
- (b) *a liba de so bradi*  
‘the river is so wide’

Here the copula suddenly is obligatory. The same holds when *bradi* is question with the particle *o* 'how':

- (30) (a) \* *o bradi a liba*
- (b) *o bradi a liba de*  
‘How wide is the river?’

When *bradi* has a pre-head specifier, it is an adjective. Similar, it must be viewed as an adjective in attributive position:

- (31) *a bradi liba*  
‘the wide river’

### 3.7. Concluding remarks

Almost the full spectrum of morphological techniques or strategies available has been called upon in the process of morphological expansion, even if compounding, reduplication, and multifunctionality play a more central role than affixation or stress-shift, for instance.

## 4. References

- Adelaar, Willem (1977), *Tarma Quechua*. Leiden: Brill
- Andersen, Roger (1983, ed.), *Pidginization and Creolization as Language Acquisition*. Rowley/MA: Newbury House
- Anderson, Stephen A. (1985), “Typological Distinctions in Word Formation”. In: Shopen, Timothy (ed.), *Language Typology and Language Description, Vol. III: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press, 3–50
- Bakker, Peter (1987), “Reduplications in Saramaccan”. In: Alleyne, Mervyn C. (ed.), *Studies in Saramaccan Clause Structure*. Amsterdam: University of Amsterdam; Mona/Jamaica: University of the West Indies (Caribbean Culture Studies 2), 17–40
- Bisang, Walter (1985), *Das Chinesische Pidgin Englisch*. Amsterdam: University of Amsterdam
- DeCamp, David (1971), “Introduction: The Study of Pidgin and Creole Languages”. In: Hymes, Dell (ed.), *The Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press, 13–39
- DeGraff, Michel (2001), “Morphology in Creole Genesis: Linguistics and Ideology”. In: Kenstowicz, Michael (ed.), *Ken Hale: A Life in Language*. Cambridge, Mass.: MIT Press, 53–121
- Dorian, Nancy (1981), *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press
- Dijkhoff, Marta (1987), “Complex Nominals and Composite Nouns in Papiamentu”. In: Maurer, Philip & Stolz, Thomas (eds.), *Varia Creolica*. Bochum: Brockmeyer, 1–10
- Ferguson, Charles J. (1975), “Towards a Characterization of English Foreigner Talk”. *Anthropological Linguistics* 17, 1–14
- Hall, Robert A. Jr. (1953), *Haitian Creole: Grammar, Texts, Vocabulary*. American Anthropological Association Memoir 74)
- Heine, Bernd (1978), “Some Generalizations on African-Based Pidgin-Languages”. *Afrika und Übersee* 61, 219–229

- Heine, Bernd (1982), *The Nubi Language of Kibera: an Arabic Creole*. Berlin: Dietrich Reimer
- Holm, John (1989), *Pidgins and Creoles, Vol. I: Theory and Structure*. Cambridge: Cambridge University Press
- Kihm, Alain (1988), "Conflation as a Directive Process in Creolization". In: Boretzky, Norbert & Enninger, Werner & Stolz, Thomas (eds.), *Beiträge zum 4. Essener Kolloquium über "Sprachkontakt, Sprachwandel, Sprachwechsel, Sprachtod"*. Bochum: Brockmeyer, 111–38
- Kouwenberg, Silvia (1991), *Berbice Dutch Creole*. Ph.D. dissertation, University of Amsterdam
- Kouwenberg, Silvia (1992), "The Convergence Hypothesis in the Genesis of Berbice Dutch Creole". *Lingua* 88, 263–300
- Kusters, Wouter (2003), *Linguistic Complexity. The Influence of Social Change on Verbal Inflection*. Ph. D. dissertation, Leiden University
- Muysken, Pieter (1993), "Reflexes of the Ibero-Romance Reflexive Clitic in Papiamentu". In: Byrne, Francis & Winford Donald (eds.), *Focus and Grammatical Relations*. Amsterdam: Benjamins, 285–302
- Muysken, Pieter & Smith, Norval (1990), "Question Words in Pidgin and Creole Languages". *Linguistics* 28, 883–903
- Muysken, Pieter & Smith, Norval (1994), "Reflexives in Pidgin and Creole Languages". In: Arends, Jacques & Muysken, Pieter & Smith, Norval (eds.)
- Pidgins and Creoles: An Introduction*. Amsterdam: Benjamins, 271–288
- Naro, Anthony J. (1978), "A Study on the Origins of Pidginization". *Language* 54, 314–347
- Römer, Raul G. (1983), "Papiamentu Tones". *Papiamentu: Problems & Possibilities*. Zutphen: Walburg, 85–96
- Sapir, Edward (1921), *Language*. New York: Houghton
- Schmidt, Annette (1985), *Young People's Dyirbal: An Example of Language Death from Australia*. Cambridge: Cambridge University Press
- Schuchardt, Hugo (1909), "Die Lingua Franca". *Zeitschrift für romanische Philologie* 33, 441–461
- Schuchardt, Hugo (1914), *Die Sprache der Sarakkañeger in Surinam*. Amsterdam: Johannes Muller (Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Nieuwe Reeks, Deel 14, No. 6)
- Smith, Norval (1987), *The Genesis of the Creole Languages of Surinam*. Ph.D. dissertation, University of Amsterdam
- Smith, Norval & Robertson, Ian & Williamson, Kay (1987), "The Ijo Element in Berbice Dutch". *Language in Society* 16, 49–90
- Voorhoeve, Jan (1981), "Multifunctionality as a Derivational Problem". In: Muysken, Pieter (ed.), *Generative Studies on Creole Languages*. Dordrecht: Foris, 25–34

Pieter Muysken, Nijmegen  
(The Netherlands)

## 154. Morphological reconstruction

1. Introduction: basic questions and methods
  2. Reconstruction: The first steps
  3. Further guiding principles
  4. Going beyond simple reconstruction of forms
  5. Pushing the limits: reconstructed states as real languages
  6. Conclusion
  7. References
1. Introduction: basic questions and methods

Morphology is above all concerned with the forms (morphemes and words) of a language. These forms include roots and affixes, as the

basic building blocks of words, as well as the patterns of combination – both derivational and inflectional – by which words are built up with these elements, and the phonological adjustments that apply within words as part of the overall word-formation process. Moreover, of interest too are the categories that these forms express and mark, e.g. gender, number, person, animacy, tense, mood, etc. All of these aspects of morphology can be reconstructed for earlier stages of a language and/or proto-language, once one has appropriate data and methods to work with and some guiding principles.

The basic methods are those that serve reconstruction in other domains (especially

phonological). That is, they include primarily the **Comparative Method** and **Internal Reconstruction**, as well as whatever other means one can employ to reach a reasonable set of assumptions about the paths of development from reconstructed elements to attested outcomes (e.g., observation of attested changes and models derived from them). These methods are illustrated, and expanded upon with other useful principles, in the sections that follow.

## 2. Reconstruction: The first steps

For the most part, the **reconstruction** of the formal side of morphological units can be fairly straightforward, drawing on the successes linguists have had with the reconstruction of sounds and sound systems. That is, once a set of valid sound correspondences holding among (related) languages have been worked out, and phonological units for a proto-language for those languages have been reconstructed based on those correspondences, one can exploit those reconstructions in the reconstruction of the form of morphemes. Essentially, all one has to do is observe the sound correspondences in a comparison of actual morphemes – as opposed to simply matching up sounds that are, so to speak, disembodied – and string together the reconstructible segments to give whole reconstructed forms (a practice Anttila 1972: 351 has called doing morphological reconstruction by “applied phonology”). This is a reasonable step to take, since the correspondences themselves emerge out of a comparison of morphemes in the first place.

For example, the comparison sets *p/p/p/f*, *īlālālā*, *t/t/t/d*, etc. across Sanskrit, Greek, Latin, and Gothic, respectively, in *pitar-/pater-/pater/fadar*, and the recognition that each set is a regular, i.e. valid and well-established, sound correspondence that licenses the reconstruction of a Proto-Indo-European phonological unit, together allow one to put the reconstructed segments together sequentially to yield a Proto-Indo-European form *\*pəter-*. Similarly, putting *a/eleli* and *s/s/s/s* together across these same languages, along with *t/t/t/t* (in the environment *s\_*) and *i/i/Ø/Ø* (in the environment *\_#*), in *astil/éstilest/ist* warrants a reconstruction for Proto-Indo-European of word *\*esti* (actually, more properly *\*H<sub>1</sub>esti* in laryngealistic terms). With a less obvious but no less compelling set of cor-

respondences, the equation of Skt. *ch*, Hittite *sk*, Lat. *sc*, and Alb. *h* in verbal forms permits one to reconstruct a Proto-Indo-European morpheme *\*-sk'*, given that the sound correspondences in evidence here are impeccable (cf. Skt. *chāya-* and Alb. *hije* ‘shadow’, where Gk. *skia* gives a clear indication of *\*sk'* as the starting point for this set).

Within the individual languages, several of the forms that these reconstructions are based on fall into paradigmatic sets with other forms that are equatable across the languages, e.g. along with *asti* in Skt. there is *asmi*, along with Gk. *esti*, there is *emmi* (in the Aeolic dialect), and so on, and these together license a reconstruction *\*esmi* (actually *\*H<sub>1</sub>esmi*). To the extent that these reconstructed forms *\*esmil/\*estii* constitute members of the same paradigm – as their outcomes in the various languages do – segmentation of the forms into constituent morphemes *\*es-* and *\*-til/\*mi* is possible, and thus inferences about the order of morphemes in the reconstructed words become possible.

What has not been addressed here is the meaning to be assigned to these reconstructed forms, yet that is a crucial step in morphological reconstruction, to the extent that morphemes are seen to be pairings of both form and meaning. In the simplest case, involving lexical items with a clear meaning, as with *pitar-*, etc. above, all of which mean ‘father’, or *as-*, etc., all of which mean ‘be’, the decision is straightforward; thus a form-meaning nexus *\*pəter-* ‘father’ or *\*es-* ‘be’ can safely be assumed for Proto-Indo-European. Similar considerations also hold for non-root elements, so that the observation that the reflexes of *\*-ti* all mark third person singular and those of *\*-mi* all mark first person singular forms in their respective languages licenses one to assign those values to the proto-morphemes in question. It is also therefore safe to assume that the person markings were suffixes, since they are suffixal in all of the languages represented.

In somewhat trickier cases, the meanings of the elements in the comparison set do not match perfectly, and thus more difficult decisions need to be made. To some extent, reconstruction in such cases becomes a matter of historical semantics and semantic reconstruction, not morphological reconstruction per se, but the ramifications for reconstructed proto-morphemes should be clear. Thus, Skt. *mātar-/Gk. māter-/Lat. māter*, etc. mean ‘mother’ but reconstructing

that meaning for the Proto-Indo-European word must take the Alb. cognate *motēr*, meaning ‘sister’, into account; conceivably the original of the word was broader (e.g. ‘female in nuclear family’) or else Albanian alone innovated (perhaps through the involvement of an original *dvandva* (coordinative) compound ‘sister-and-mother’ as a merism (defined by Watkins 1995: 9 as “a two-part figure which makes reference to the totality of a single higher concept”) for female kin). At the grammatical level, semantic/functional mismatches for cognate markers can pose problems for reconstruction, especially in the absence of clear bases for understanding how category markings can themselves mutate and be altered. Thus, in the *\*-sk'*-case mentioned above, the Hittite morpheme marks iterative action, the Latin morpheme marks inchoative (beginning) action, the Sanskrit morpheme is simply one of several markers for present tense system stems, and the Albanian morpheme marks non-active voice (with passive, reflexive, reciprocal, and stative functions) but only in the present tense system. Thus there is an element of presentability in most of the reflexes of *\*-sk'*-, but one needs to ask whether the original function of this morpheme was just to mark present, or instead was more specific, with one of the attested values as the starting point in Proto-Indo-European. Any answer here runs the risk of seeming to be arbitrary, but without an answer, the reconstruction for this morpheme is incomplete; reconstructing the form alone without the meaning or function is satisfying only half of the burden.

The above cases draw on the Comparative Method, the mainstay of historical reconstruction for nearly two centuries, but other methods can be employed as well in reconstructing morphology. In particular, Internal Reconstruction can be used, where, in essence, alternations within a single synchronic stage of a language are “undone” as it were, and an earlier state without the alternation is reconstructed. Morphological reconstruction in such a case consists in the positing of a unique form for a given alternation found in a later stage, generally taking the later allomorphy to be the result of conditioned sound changes. Thus the allomorphy seen in the Modern English plural marker *-s/-z/-əz* allows for a hypothesis of an earlier marker with a single undifferentiated form, e.g. *-əz*, with the variants having arisen, as is so often the case, via sound changes, e.g. syncope and

voicing assimilation. But even this method can break down; there is no way to reconcile the plural marker *-en* (restricted in Modern English just to *oxen* and *brethren*, though *children* contains it too) with the *-sl/-z/-əz* forms, and that is as it should be, given that its distribution is lexically idiosyncratic and not rule-governed and not tied to phonological conditioning in any way. But even cases involving clear phonological conditioning of the allomorphy can be problematic for reconstruction; for instance, the Korean nominative markers are in complementary distribution: *-i* occurring after a consonant-final noun, *-ka* after a vowel-final stem, but the phonetic distance between the two forms makes a reconstruction of a single nominative marker that gave rise to both of these alternants most unlikely. The third method, or rather a principle (cf. 1), namely ensuring that there is a reasonable path of development from the earlier reconstructed stage to the attested ones, must be invoked here, so that one does not take *-i* and *-ka* back to the same proto-form, internally reconstructing by brute force, as it were. Presumably they each have an independent origin and have come, by various developments, to stand as phonologically distributed functionally equivalent alternants.

### 3. Further guiding principles

More can be said about these methods and principles. Critical to positing a reconstruction that requires only a reasonable set of changes for the development of any given element is an understanding of what the expected processes of language change are. What is thus especially important here, besides recognizing the regular nature of sound change (which gives the regular sound correspondences drawn on in doing morphological reconstruction as “applied phonology”, as above, cf. 2), is allowing for the workings of other processes of change, such as analogy. For instance, the consonantal matchings between the root nouns (where the root with nothing added constitutes the stem) seen in Gk. *pod-* and Lat. *ped-*, both meaning ‘foot’, are perfectly regular, but the vowels do not match up as expected, as Gk. *o* usually corresponds to Lat. *o* (reflecting PIE \**o*) and Lat. *e* usually corresponds to Gk. *e* (reflecting PIE \**e*). Armed with the knowledge that there are Indo-European languages in which

reflexes of \**e* and reflexes of \**o* alternate in grammatically determined environments (e.g. English present tense *sing* from \**sengʷʰh-*, versus past tense *sang* from \**songʷʰh-*), rather than reconstructing a different vowel for the *o* ~ *e* correspondence, a step which would ultimately lead one to reconstruct a different vowel for every such correspondence (and there are others, especially when languages other than Greek and Latin are included), one can instead reconstruct a process by which \**o* and \**e* alternate within the grammatical forms of ‘foot’ in the proto-language. One can then treat the Greek and Latin forms with fixed vowels as the result of analogical change within the paradigm as the individual languages took shape. The exact conditions for the \**e* /\**o* alternation may be unclear, but the method of allowing for the “undoing” of the potential effects of analogy leads one to the reconstruction of a morphological process (of vowel gradation) for the proto-language, that is, an aspect of a Proto-Indo-European word-formation process. It can be noted as well that the inferences about morpheme order derivable from the comparison of \**esti* and \**esmi* similarly reflect reconstructive assumptions made about word-formation processes of the proto-language, i.e. suffixing (at least for person/number marking).

Moreover, another useful principle can be invoked to illuminate the nature of the reconstructed \**e*/\**o* alternation. In particular, it is known that the relics of processes or earlier states that were once productive can be found embedded in compounds; for instance, English *with-* retains its original adversative meaning (‘against’) only in composite forms such as *withstand* (‘stand against’, not ‘stand alongside of’) with the synchronically regular comitative meaning of *with*). Therefore, one can look to compounds for some insights into aspects of earlier formations. The occurrence of \**e* in the genitive case of an old root noun found in the Greek composite form *des-pótēs* ‘master, owner’, from \**dems-potēs*, literally ‘dwelling:GEN-master’ as opposed to the consistent \**o* found in derivatives related to the root nouns, such as Lat. *domus* ‘house’, invites the suggestion that \**e* was proper, originally, to the oblique cases (e.g. genitive) whereas \**o* was found in the direct cases (e.g. nominative) of root nouns in the proto-language. Therefore, one can infer an original paradigm with nominative stem \**pod-* and genitive (oblique) stem \**ped-*.

A corollary of the use of relic forms for guidance in reconstructing earlier states is what can be called the **Meillet Principle** (after Antoine Meillet, based on work discussed in Arlotto 1981: 144 f.), namely to reconstruct from synchronic irregularities and isolated forms, that is any sort of unproductive material present in a language at a given time. The rationale is that synchronically unproductive material is exactly what demands a historical explanation, whereas synchronically productive forms could in principle have been created at any time by means of the productive and regular processes, and thus are not an indicator of the presence of some element in the proto-language. For example, when confronted with Skt. *bhr-ta-* and Old Irish *breth* as past passive participles of \**bher-* ‘carry’, one can mechanically reconstruct \**bhr-to-* as a Proto-Indo-European pre-form underlying them historically, even though the \*-*to-* participle formation is clearly the productive one for Proto-Indo-European and on into the offspring languages, to judge from its widespread occurrence in all of the branches of the family. Adding in Lat. *lā-tus* and Gk. *ois-tós* as the participles to the outcomes of \**bher-* in these languages (present tense *ferō/phérō*, respectively) complicates the picture somewhat, though. The Latin form seems clearly to be an import from the regular participle of a semantically related verb *tollō* ‘pick up; take on’ (*lātus* from \**tlātus*, formed regularly with \*-*to-* from a root \**telH₂-*) and thus presumably simply shows the substitution of a participle form from a different paradigm; such a substitution is understandable, as *tollō* supplied the form *tult* that functions as the perfect tense associated with *ferō*. However, which form did *lātus* replace, an outcome of \**bhr-to-* or something else again? Gk. *oistós* provides the answer, since it is synchronically isolated within Greek (the root *ois-* occurs elsewhere only in a suppletive future of *phérō*) and thus can only be explained historically, as a relic of an earlier irregular (i.e., suppletive) state. We thus reconstruct the Proto-Indo-European participle as \**oistos* and explain the Sanskrit and Irish forms as simply creations within those branches using the productive pattern.

Meillet’s Principle thus licenses the reconstruction of suppletion, that is, a proto-language morphological irregularity, and in that sense can lead to a different solution than one might reach with internal reconstruction

where irregularities are taken back to earlier regular states. Similarly, allowing for the reconstruction of suppletion means that not every paradigm that one might want to reconstruct has to be fully articulated, with all members intact (i.e., all cells filled), in keeping with the observations from attested language states that there can be defective paradigms and that paradigms can be built up piece-meal (see Watkins 1962 for discussion of how paradigms can be reconstituted).

The overriding factor here, as always, is to give the best account of the facts. In a sense, then, Meillet's Principle is akin to the **Accountability Principle**, known from sociolinguistics (cf. Labov 1982: 30; Winford 1990: 227), in that it requires that all the forms in a set of comparanda be accounted for – in this case it is not enough to account for just the Sanskrit and Old Irish forms with a Proto-Indo-European form that does not allow for a straightforward account of the Greek suppletion. That is, an explanation is possible for why *bhrta-* is to be found in Sanskrit or Celtic: at any time due to productivity of \*-to- participial formation, such a form could have come into being; the Gk. *ois-*, on the other hand, demands a different kind of account, e.g. continuing an inherited suppletive paradigm, as suggested above.

#### 4. Going beyond simple reconstruction of forms

What else can be reconstructed of the morphology in a proto-language? Clearly, once one makes assumptions about the functions of particular reconstructed items, one has also made inferences about the relevant grammatical categories for the proto-language, e.g. number, person, tense, etc., based on \*-ti and \*-mi in Proto-Indo-European. However, it is also possible to go beyond these categories, and engage in some internal reconstruction on the value of categories reconstructed for a proto-language.

For example, the suppletion that was reconstructed in the paradigm of \**bher-*, as discussed above (cf. 3), is found with other verbs in Proto-Indo-European, most often focusing on present versus past tense forms (as with Lat. *ferō/tulī*, noted above (cf. 3), or Alb. *ha* ‘eats’, from a root \*(*H<sub>1</sub>*)*ed-* vs. (*hēn*)*gra* ‘ate’, from the root \**gʷʰrō-* (with prefixes \**Ho-en-*). One way to make sense of the relatively widespread occurrence of apparent present/past

suppletion is to assume that it has to do with some inherent properties of the roots in question, derivable from their respective basic meanings. However, present versus past seems to be an unlikely distinction to be associated with the lexical meanings of particular roots (after all, how could an action be inherently associated with past time). Thus, it is more likely that the original distinction was aspectual in nature, e.g. imperfective (durative) versus perfective (completive), since that is a lexically encodable property (what is sometimes referred to as “Aktionsart”). Presumably, these aspectual distinctions were rearranged and altered somewhat on the way to the attested languages. Internal reconstruction, therefore, attempts to rationalize the irregularity of reconstructed suppletion by projecting it back to a previous state where it is not unmotivated but rather follows from some other property, in this case, aspect as determined by the semantics of a given root. This exercise thus leads to some reasonable inferences about the pre-Proto-Indo-European state of affairs regarding verbal categories.

Internal reconstruction can also be carried out on the reconstructed morphemes themselves. For instance, the relation between the reconstructed third person singular present ending \*-ti for Proto-Indo-European (as above, 2) and the comparable past ending \*-t (cf. Skt. -t# = Lat. -d#), as well as \*-mi vs. \*-m in the first person, invites the analysis that the \*-i itself marks present time, and that moreover the present ending was diachronically derived from the endings \*-t and \*-m through the addition of the \*-i. It would follow, therefore, that the original value for \*-t and \*-m was just to mark person (and number), not to mark past time reference directly; only once the opposition of \*-ti with \*-t emerged was a present/nonpresent distinction relevant.

Hypotheses such as these made by internally reconstructing from reconstructed proto-language elements are not subject to confirmation the way that internal reconstruction on an attested state is (e.g. reconstructing /-əz/ for the English plural allomorphy is verified by examining earlier English s-plurals). However, they are compelling scenarios to the extent that they are based on what is known about language in general.

In a similar vein, and as noted above, to the extent that one can be sure of various trends in language change in general, they

can be employed in reconstruction. Thus, although it is not impossible for affixes to turn into full-fledged words (see Joseph & Janda 1988, Campbell 1991, and Janda 2001 for discussion), it is nonetheless true that the opposite development, in which words develop into affixes, is by far the more common direction for changes involving words and bound morphemes. That is, examples such as Old English *hād* ‘condition/state’ giving the Modern English derivational suffix *-hood* (cf. *child/child-hood*) from original compounds (‘the state of being X’) or the English negative adverb *not* giving, at least in part via phonological reduction in an unstressed position, the inflectional affix *-n’t* (see Zwicky & Pullum 1983), are relatively common, while examples such as the Old English bound genitive case-suffix *-(e)s* giving, via a reanalysis and aligning with the pronoun *his*, the less-bound and somewhat word-like possessive marker *’s* in Modern English (less bound in that it attaches at the end of a phrase, as in *the King of England’s hat*) are significantly rarer, though definitely attested.

This observation means that in some cases, we can reasonably locate the history of a given affix in a word, and this holds whether the basic methodology being employed is the Comparative Method or Internal Reconstruction. For instance, an internal comparison of the English adjectival and adverbial suffix *-ly* with the similarly used and phonologically similar free word *like*, as in *quickly/quick-like, friendly/friend-like*, might suggest an historical derivation of the suffix from a reduction of the free word, perhaps under conditions of low accentual prominence, and a cross-linguistic comparison of the French adverbial suffix *-ment* with the more word-like element in Spanish, *-mente* (word-like in that a single instance is distributable across two conjoined adverbs, e.g. *clara y rápidamente* ‘clearly and rapidly’), permits the reasonable inference that the suffix derives from a once less-bound element. In the case of *-ly/like*, earlier English evidence confirms the reconstruction (cf. the use of *-lic* in Old English), as does Latin evidence in the case of *-ment/mente* (regarding the use of the ablative *mente* of the noun ‘mind’ in adverbial phrases, e.g. *obstinatā mente* ‘with an obstinate mind’, i.e. ‘obstinately’). Still, such hypotheses are not iron-clad, and synchronic resemblances can be misleading – the *-less* of *friendless*, for instance, has nothing to do

with the independent word *less* (from Old English *lēs(sa)*) deriving instead from the Old English preposition *lēas* ‘without’.

### 5. Pushing the limits: reconstructed states as real languages

Some of the successes of morphological reconstruction discussed in the previous sections involve simply applying analytic techniques that are well-known in linguistics, and in some instances, applying them to a reconstructed proto-language, essentially treating it as just another language, a synchronic state that can be analyzed and, among other things, subjected to internal reconstruction. As a final example of morphological reconstruction, an example is presented of how one can reconstruct morphophonemic rules for a proto-language, based on the Proto-Indo-European paradigm for the present of the verb ‘be’. This example serves as a suitable conclusion, as it draws on comparative methodology, internal reconstruction, and Meillet’s Principle, as well as general principles of morphological analysis.

Besides the first and third person forms *\*esmi* and *\*esti*, it is possible to reconstruct a second person form as well in the singular. The comparison can be made of Skt. *asi* ‘you are’ with Gk. *ei* ‘you are’, both of which are irregular within their respective languages (e.g., the non-occurring *\*\*assi* might be expected in Sanskrit, all things being equal). The “applied phonology” methodology (cf. 2) leads to a reconstruction for Proto-Indo-European of *\*esi*, and the shared irregular status of the Sanskrit and Greek forms ensure that this form is to be posited for the proto-language; that is, the paradigm was not a defective one in Proto-Indo-European lacking a second person singular form. Moreover, this form itself is an irregularity within the Proto-Indo-European verbal system, since *\*\*essi* would be expected, based on the secure reconstruction of a root *\*es-* (cf. 2)) plus the second person singular ending *\*-si* (seen, for instance, in Skt. *bhara-si* ‘you carry’, Gothic *bairi-s*, etc.). Given the difference between the form expected on morphological grounds, *\*essi*, and the reconstructed form based on the comparative evidence, *\*esi*, it is reasonable to reconcile the two forms by treating *\*essi* as the (morphologically motivated) underlying form for Proto-Indo-European and *\*esi* as the surface form,

also for Proto-Indo-European; this step means that a morphophonemic rule converting an underlying /-ss-/ into a surface [-s-] must be posited for the proto-language. Moreover, it licenses the inference that the underlying form was the actual pre-Proto-Indo-European form, and that a (sibilant) degemination sound change operated between pre-Proto-Indo-European and Proto-Indo-European to give the surface form \**esi*, and create the reconstructed irregular singular paradigm with \**esmilesilesti*.

## 6. Conclusion

In a sense, then, morphological reconstruction is not significantly different in its goals, methods, and guiding principles from phonological reconstruction. It is thus not surprising that the topic (often under the rubric of “grammatical reconstruction”) is not discussed at great length in any standard textbooks, except insofar as internal reconstruction leads to results that have consequences of a morphological (most usually morphophonological) nature; still, interested readers should consult Anttila (1972; 1989), Fox (1995), Hock (2<sup>1991</sup>), and Trask (1996) for additional general discussion and examples.

## 7. References

- Anttila, Raimo (1972), *An Introduction to Historical and Comparative Linguistics*. New York: MacMillan
- Anttila, Raimo (1989), *Historical and Comparative Linguistics*. Amsterdam: John Benjamins [revised edition of Anttila 1972]
- Arlotto, Anthony (2<sup>1981</sup>), *Introduction to Historical Linguistics*. Washington, D.C.: University Press of America [<sup>1</sup>1972, Boston: Houghton Mifflin Co.]
- Campbell, Lyle (1991), “Some Grammaticalization Changes in Estonian and Their Implications”. In: Traugott, Elizabeth & Heine, Bernd (eds.), *Approaches to Grammaticalization, Vol. I*. Amsterdam: John Benjamins, 285–299
- Fox, Anthony (1995), *Linguistic Reconstruction: An Introduction to Theory and Method*. Oxford: Oxford University Press
- Hock, Hans Henrich (2<sup>1991</sup>), *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter [<sup>1</sup>1986]
- Janda, Richard D. (2001), “Beyond ‘Pathways’ and ‘Unidirectionality’: On the Discontinuity of Language Transmission and the Counterability of Grammaticalization”. *Language Sciences* 23 (Special Issue: Campbell, Lyle (ed.), *Grammaticalization: A Critical Assessment*), 265–340
- Joseph, Brian D. & Janda, Richard D. (1988), “The How and Why of Diachronic Morphologization and Demorphologization”. In: Hammond, Michael & Noonan, Michael (eds.), *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press, 193–210
- Labov, William (1982), “Building on Empirical Foundations”. In: Lehmann, Winifred & Malkiel, Yakov (eds.), *Perspectives on Historical Linguistics*. Amsterdam: John Benjamins, 17–92
- Trask, R. L[awrence] (1996), *Historical Linguistics*. London: Arnold
- Watkins, Calvert (1962), *Indo-European Origins of the Celtic Verb*. Dublin: Institute for Advanced Studies
- Watkins, Calvert (1995), *How to Kill a Dragon: Aspects of Indo-European Poetics*. Oxford: Oxford University Press
- Winford, Donald (1990), “Copula Variability, Accountability, and the Concept of ‘Polylectal’ Grammars”. *Journal of Pidgin and Creole Languages* 5, 223–252
- Zwickly, Arnold M. & Pullum, Geoffrey K. (1983), “Cliticization vs. Inflection: English *n’t*”. *Language* 59, 502–513

*Brian D. Joseph, Ohio (U.S.A.)*

## XVIII. Morphologischer Wandel II: Fallstudien

### Morphological change II: Case studies

#### 155. From Old English to Modern English

1. Introductory remarks
2. Inflection
3. Word-formation
4. References

##### 1. Introductory remarks

###### 1.1. Language periods

The linguistic history of English may be divided into the following periods: Old English from the 7th century to about 1100, Middle English from ca. 1100 to ca. 1500, Early Modern English from ca. 1500 to ca. 1700, Late Modern English covering the 18th and 19th centuries, Modern English leading to Present-Day English.

###### 1.2. Dialects

Four Old English dialects can be distinguished: West Saxon, Kentish, Mercian, Northumbrian. The last two are called Anglian. In the Middle English period, they roughly correspond to Southern (or Southwestern), Kentish (or Southeastern), Midland and Northern respectively. Northern is also referred to as Northumbrian, the others as Southumbrian. The Midland area may be further subdivided into West and East, the East itself into North and South. Additional subdivisions, like Northwest Midland, are applied when necessary.

The variety of English spoken in Lothian and Lowland Scotland developed from Northumbrian, being originally called *Inglis*, renamed *Scottis* in the 15th century.

The dialect which came to be accepted as standard in Early Modern English was that of the London area, where political, military and economic factors contributed to its acquisition of this status. But most important was William Caxton's publication in 1477 of the first book to be printed in England. He

based his spelling on that of manuscripts which had previously appeared there, particularly official documents, thus helping to spread the use of this dialect into other areas of England.

##### 2. Inflection

###### 2.1. From inflection to non-inflection

Old English is highly inflected, whereas Middle English shows the transition to a nearly non-inflected language, which was gradually reached in Early Modern English, Middle English being the time-span when the language changed from a (relatively) synthetic to a (relatively) analytic one. This will be illustrated by outlining the history of the inflectional systems of the noun and the verb. For adjectives and pronouns cf. Faß (1989: 128–214; 1992: 21–38).

###### 2.2. Nouns

###### 2.2.1. Gender, Number, Case, Declension

Like the German **noun**, the Old English one has three grammatical genders, i.e. masculine, feminine, neuter. The grammatical gender disappeared in the 14th century. Again, like in German, Old English nouns have the number-distinction singular – plural and four cases: nominative, genitive, dative, accusative. The instrumental, which differed occasionally from the dative in the earliest texts only, may be disregarded.

Both languages, German and Old English, share two main **declensions**, strong (or vocalic) and weak (or consonantal), according to whether the stem originally ended in a vowel or *-n* respectively. Stems originally ending in other consonants than *-n* are usually subsumed under minor declensions (cf. Wright & Wright 1908: 160).



Map 155.1: Old English Dialects

### 2.2.2. Strong declensions

An example of the strong masculine declension is OE. *eorl* 'nobleman', one of the strong feminine declension OE. *ār* 'honour' (see Table 155.1).

The strong neuter declension is different from the strong masculine one in that it has *-ul-o* in the nominative/accusative plural, if the noun has a short stem-syllable like *hof* 'dwelling', *scip* 'ship', but *-Ø*, if it is long as in *bān* 'bone', *hors* 'horse', *nett* 'net' (cf. Faiß 1992: 16f.).

### 2.2.3. Weak declensions

An example of the weak masculine declension is OE. *lida* 'sailor' (see Table 155.2).

	masculine		feminine	
	sing.	plural	sing.	plural
nom.	<i>eorl</i>	<i>eorlas</i>	<i>ār</i>	<i>āre, -a</i>
gen.	<i>eorles</i>	<i>eorla</i>	<i>āre</i>	<i>āra, -(e)na</i>
dat.	<i>eorle</i>	<i>eorlum</i>	<i>āre</i>	<i>ārum</i>
acc.	<i>eorl</i>	<i>eorlas</i>	<i>āre</i>	<i>āre, -a</i>

Tab. 155.1: Strong masculine and feminine declensions

Weak feminine and neuter nouns like *tunge* 'tongue' or *ēage* 'eye' are inflected the same way.



Map 155.2: Middle English Dialects

	singular	plural
nom.	<i>lida</i>	<i>lidan</i>
gen.	<i>lidan</i>	<i>lidena</i>
dat.	<i>lidan</i>	<i>lidum</i>
acc.	<i>lidan</i>	<i>lidan</i>

Tab. 155.2: Weak declensions

	singular	plural
nom.	<i>fōt</i>	<i>fēt</i>
gen.	<i>fōtes</i>	<i>fōta</i>
dat.	<i>fēt</i>	<i>fōtum</i>
acc.	<i>fōt</i>	<i>fēt</i>

Tab. 155.3: Minor declensions

#### 2.2.4. Minor declensions

An example of one of the minor declensions is OE. masculine *fōt* ‘foot’ (see Table 155.3).

The morphonemic alternation /fō:t/~/fe:t/ is due to *i*- (or palatal) **mutation**, cf. the original Germanic forms of the dative singular,

\**fōti*, and of the nominative plural, \**fōtiz*, whence the *i*-mutated vowel spread into the accusative plural, whose original form was \**fōtunz* (cf. Brunner 1965: 226; Wright & Wright 1908: 192).

### 2.2.5. Plural and genitive

The Old English inflectional system of the noun was significantly reduced on its way to Present-Day English (cf. Faiß 1989: 109–127). The only markers to have been retained for inflection are in fact *-(e)s* for plural and genitive as well as *-(e)n* and vowel change (or /replacive/) for plural as, e.g., in *oxen* and *feet* (cf. Faiß 1992: 15f.). The other markers had been dropped by the 14th century. Since then the noun has had a two-case system with a common (or unmarked) case and a genitive (or possessive or marked) case together with the important number-distinction singular – plural.

Among the inflectional suffixes retained, *-(e)s* has become productive. As a plural indicator it goes back to *-as* of the Old English strong masculine declension. *-as* yielded ME. *-es*, which was common by the end of the 14th century. It was gradually replaced by *-s* from about 1500 onwards except after sibilants. The modern distribution of the phonemically conditioned allomorphs of {PLURAL}, /ɪz, z, s/, established itself during the 17th century. Genitival *-(e)s* stems from the genitive singular of the Old English strong masculine and neuter declensions and became generalized in the 14th century. The ousting of the Middle English successors to the Old English genitive plural markers came to an end in the 14th century too, so that, apart from a few exceptions, *-(e)s* expressed both singular and plural genitive. The tendency towards generalisation proved stronger than the resistance of individual nouns. As to the spelling, <’s> was not fixed until the 1690s. Obviously, <’s> is not attested before the 2nd half of the 18th century. The pronunciation and the allomorphic distribution of the *s*-genitive are identical to the *s*-plural. (For the history of mixed plurals like *loaf-loaves*, *house-houses*, *mouth-mouths*, which show voicing of the final consonant of the singular, and for plurals of nouns from Greek and Latin such as *alumnus*, *alga*, *velum*, *crisis*, *oxymoron* and many others cf. Faiß 1992.)

### 2.2.6. *of*-Genitive, *to*-Dative

Clear signs of the language’s adoption of an **analytic** character are the *of*-genitive and the *to*-dative. They originated in Old English from constructions consisting of the prepositions *of* and *tō* plus noun phrase, which were preferably used to denote origin and direction/goal respectively. Yet it was not before

the Middle English period that they took over the syntactic-semantic functions of the inflected (or synthetic) genitive and dative. The question remains whether French constructions with the prepositions *de* and *à*, which had replaced the Latin synthetic genitive and dative, encouraged the expansion of the corresponding English ones.

### 2.3. Verbs

Old English **verbal inflection** may be compared to that of German, Greek, Latin, Italian, Spanish, to a lesser extent to that of French, with respect to the number of inflectional signs denoting person, number, tense and mood. The verb in Present-Day Standard English, on the contrary, has but one such sign at its disposal, i.e. *-(e)s* for 3rd person singular present indicative as in *he pull-s /he push-es*.

Old English strong verbs comprise eight classes. The two last ones are generally put together in one with a subdivision into a) and b) or 1. and 2. Old English weak verbs are generally grouped in three classes.

#### 2.3.1. Old English strong verbs

A characteristic feature of the Old English strong verbs is **ablaut**, representing some kind of internal inflection which has been retained in the corresponding modern stem-forms. They were reduced from four to three in the course of time, with but one exception: the verb *be* has still four, namely *be-was-were-been*. The Old English forms which are usually adduced are: infinitive – 1st/3rd person singular preterite indicative – plural preterite indicative – past participle (cf. Tab. 155.4)

#### 2.3.2. Preterite and past participle

The nowadays common distribution of the phonemically conditioned allomorphs of {PRETERITE} and {PAST PARTICIPLE}, /ɪd, d, t/, gradually established itself in Early Modern English. The usual Middle English ending was *-ed*. It is the merger of OE. *-ed(-)*, which indicates **preterite** and **past participle** of certain members of the 1st weak class, and of West Saxon *-od(-)*, non-West Saxon *-ad(-)*, as typical indicators of the 2nd weak class.

The Old English strong verbs were marked by *-en* for past participle and also by *-anl-onl-un* in the non-West Saxon dialects. These suffixes merged into ME. *-en*. *-n* denoted the past participles of *bēon* ‘be’, *dōn* ‘do’, *gān* ‘go’. Middle English is characterized by a

	PrE.	infinitive	preterite		past participle
			singular	plural	
1st class		<i>rīdan</i>	<i>rād</i>	<i>ridon</i>	<i>ridden</i>
2nd class	PrE.	<i>ride</i>	<i>rode</i>	<i>sc(e)uton</i>	<i>ridden</i>
		<i>scēotan</i>	<i>scēat</i>		<i>sc(e)oten</i>
3rd class a)	PrE.	<i>shoot</i>		<i>shot</i>	<i>shot</i>
		<i>bindan</i>	<i>band/bond</i>	<i>bundon</i>	<i>bunden</i>
3rd class b)	PrE.	<i>bind</i>		<i>bound</i>	<i>bound</i>
		<i>helpan</i>	<i>healp</i>	<i>hulpon</i>	<i>holpen</i>
3rd class c)	PrE. (weak)	<i>help</i>		<i>helped</i>	<i>helped</i>
3rd class c)	PrE. (weak)	<i>ceorfan</i>	<i>cearf</i>	<i>curfon</i>	<i>corfen</i>
3rd class c)	PrE.	<i>carve</i>		<i>carved</i>	<i>carved</i>
4th class	PrE.	<i>feohtan</i>	<i>feaht</i>	<i>fuhton</i>	<i>fohten</i>
		<i>fight</i>		<i>fought</i>	<i>fought</i>
5th class	PrE.	<i>brekan</i>	<i>bræc</i>	<i>brācon</i>	<i>brocen</i>
		<i>break</i>		<i>broke</i>	<i>broken</i>
6th class	PrE.	<i>giefan</i>	<i>geaf</i>	<i>gēafon</i>	<i>gesen</i>
		<i>give</i>		<i>gave</i>	<i>given</i>
7th class a)	PrE.	<i>sc(e)acan</i>	<i>sc(e)ōc</i>	<i>sc(e)ōcon</i>	<i>sc(e)acen</i>
		<i>shake</i>		<i>shook</i>	<i>shaken</i>
7th class b)	PrE. (weak)	<i>lētan</i>	<i>lēt</i>	<i>lēton</i>	<i>lēten</i>
		<i>let</i>		<i>let</i>	<i>let</i>
		<i>cnāwan</i>	<i>cnēow</i>	<i>cnēowon</i>	<i>cnāwen</i>
	PrE.	<i>know</i>		<i>knew</i>	<i>known</i>

Tab. 155.4: Strong verbs

high degree of variation between past participles with and without *-n*, cf. *iwurden*, *ihalde*, *ofslage*, *idon*. Variation between *-(e)n* and *-Ø* still existed in Early Modern English. From the 17th century onwards, the present distribution has been established, but not obligatorily, as is proved by modern variants like *beaten/beat*, *bidden/bid* and others.

The prefix *i-* is descended from OE. *ge-*, cf. Germ. *ge-*, which was optionally attached to past participles. Also written <y> in Middle English, it disappeared towards the end of that period, lingering on, however, as an archaism in the following periods, above all in poetry.

### 2.3.3. Present participle

The common marker for **present participle** in Old English was *-ende*, varying with *-ande* in Northumbrian, which mainly occurred with verbs of the 2nd weak class. *-ende* continued to exist in Middle English, especially in the Southeast Midland. *-ande* appeared as *-and* in Northumbrian and Scottish after the loss of unstressed final *-e* in Northern Early Middle English. A third form arose in the 12th century, *-inde*, showing [ɪ] through the raising

of unstressed [e] before [nd]. This suffix was widely used, yet most frequently in the West Midland and the South. PrE. *-ing*, though, does not go back to any one of them. In Old English the predecessor of this suffix, *-ing*, served as a means of deriving mostly feminine abstract nouns from verbs belonging to the 1st weak class such as *nering* ‘protection, defence’ from *nerian* ‘save’. The first instances of present participles with *-ing* can be detected about 1200 in some southern regions. It was rapidly spreading becoming generalized in the whole Southumbrian area at the end of the 14th century, with some sporadic *-inde* relics surviving in 15th century texts. It was in this century that *-ing* entered into Northumbrian and Scottish, there competing with *-and*, which was eventually ousted except for some dialectal traces in Southern Scots.

### 2.3.4. Infinitive

Old English {INFINITIVE} has the allomorphs /an/, ian, n/. /an/ occurs with weak verbs of the 1st and 3rd classes and with strong verbs, cf. *dēman* ‘deem’, *libban* ‘live’, *rīdan* ‘ride’, /ian/ with weak verbs of the 1st class having stem-final *-r* and with those of the 2nd class,

cf. *nerian* ‘save’, *endian* ‘end’, /n/ with contracted verbs such as *tēon* ‘draw’ (strong), *bēon* ‘press’ (weak) and with *bēon* ‘be’, *dōn* ‘do’, *gān* ‘go’. /an/ and /ian/ generally merged into /ən/ <en> in the 12th century, while /n/ remained as such.

There was a great deal of variation between *n-* and *n-less* infinitives in Early Middle English. Owing to the loss of final *-n* already in Late Old Northumbrian, followed by that of final *-e* in the 13th century in its dialectal successors, the inflectionally unmarked infinitive was the rule in 14th century Northumbrian and Scottish, cf. *accord*, *disclar*, *suceid*. In Southumbrian the infinitive grew unmarked in the 15th century, cf. *do*, *reduce*, *translate*.

### 2.3.5. Person, number, tense, mood

Old English verbs are inflected for **person**, **number**, **tense**, and **mood**, often showing a combination of internal and external inflection, as do German verbs. In keeping with Verner’s Law stem-final consonants can alternate, and so can the stem-vowels in the 2nd and 3rd persons singular present indicative on account of *-i-* in the original inflection of these forms. An example of this kind of verbs is *cēasan* ‘choose’ (cf. Tab. 155.5; imperative singular *cēos*, plural *cēosab*).

Some peculiarities need explaining: *-t* in *cīest* and *-þ* in *cīesp* came into existence through vowel deletion in *-est* and *-eþ* respectively, cf. *dēm(e)st*, *dēm(e)þ* (1st weak class). *-est*, Northumbrian *-es*, is also present in the 2nd person singular preterite indicative of weak verbs as in *pū dēmdest*. The verbs of the 2nd weak class have *-ie* in the 1st, *-ast* in the 2nd, *-aþ* in the 3rd person singular and *-iaþ* in the plural present indicative, cf. *lōcie*, *lōcast*, *lōcap*, *lōciap* of *lōcian* ‘look’.

During the Middle English period, the rather sophisticated inflectional system, which, by the way, was still more complex on account of marked dialectal variation, became radically simplified (cf. Faib 1989: 215–297). The only overt person-denoting inflectional signs typical of Early Modern English were in fact *-(e)st* and *-(e)þl-(e)s* for the 2nd person singular present and preterite and 3rd person singular present indicative respectively. In the standard language *-(e)st* dropped out in the 18th century, when *thou* became replaced with *you*. *-(e)th* yielded to *-(e)s* around the middle of the 17th century, being distributionally equal to plural and genitive *-s*. With the exception of the *were*-subjunctive in the 1st and 3rd persons singular, *-Ø* in the 3rd person singular and *be* for all persons in the present subjunctive, this mood is no longer recognizable formally, cf. if *I/he were, he goes* vs. *he go, I ... be*.

The history of *-s* for 3<sup>rd</sup> person singular present indicative has been long debated. Yet the following may be rather safely concluded. The starting-point is the Late Northumbrian dialect (2nd half 10th century). Here the 2nd person present indicative is mainly marked by *-s*, which is a general characteristic of early texts but became superseded by *-st* in other dialects. *-s* first spread into the plural present indicative gradually replacing *-þ*, before it appeared in the 3rd person singular present indicative (cf. Berndt 1956: 184–196, 202–213). The linguistic fact itself is best explained by a vivid tendency towards the generalization of one single inflectional sign, which is not only obvious in this case, but is also illustrated by, e.g., the *s*-plural and the *s*-genitive. In Middle English, *-s* dominated in Northumbrian and Scottish, whence it made its way slowly via the Midlands to the

	present		preterite	
	singular	plural	singular	plural
<b>indicative</b>				
1st person	čc cēose	wě cēosab	čc cēas	wě curon
2nd person	pū cīest	gě cēosab	pū cure	gě curon
3rd person	hē cīesp	hīe cēosab	hē cēas	hīe curon
<b>subjunctive</b>				
1st/2nd/3rd persons	čc/pū/ hē cēose	wě/gě/ hīe cēosen	čc/pū/ hē cure	wě/gě/ hīe curen

Tab. 155.5: Inflection

South. There, however, it had to fight it out against *-th*, even in London, where the latter persisted to the middle of the 17th century. After being ousted from the standard language by *-s*, *-th* has survived in literary style and dialectally, as has *-st* for the 2nd person singular. In contrast to *-st* or *-s* as the successor to Northern *-s* of the 2nd person singular present indicative as well as to Early Modern English reduced *-st*, *-th* is practically restricted to Eastern Cornwall and Southern Devonshire.

#### 2.4. Conclusion

Unlike that of German and to a lesser extent those of the Romance languages the inflectional history of English is characterized by a clear-cut tendency towards generalization. This necessarily means a radical minimizing of inflectional signs of the parts of speech discussed. Thus *-s* for plural, genitive and 3<sup>rd</sup> person singular present indicative, *-ing* for present participle/gerund, *-ed* for preterite and past participle have become generalized. As a consequence new devices had to be invented to express syntactic relations. This was, on the one hand, the fixing of word-order and on the other the use of prepositions, as is proved by the *of*-genitive and the *to*-dative.

The developments described set in mostly in Middle English. Some of them can first be detected in the northern dialects making their way more or less rapidly to the South and into the standard language.

### 3. Word-formation

The instances below have been mainly selected from BT (1976), BTS (1973) (Old English), MED (1954 ff.) (Middle English), Marchand (1969), OED (1989), more recent ones from Ayto (1989), Barnhart et al. (1990) as well as from newspapers and the like (cf. also Faiß 1992: 54–111).

#### 3.1. Old English

Old English basically shares the word-formation types of the Germanic languages: affixation (prefixation and suffixation), compounding and zero-derivation (cf. Krahe & Meid 1967; Art. 37, 87, 89, 90).

##### 3.1.1. Affixation

Some Old English **prefixes** can still be observed in Present-Day English.

*for-* ‘forward, away’ (often intensifying) as in *forbærnan* ‘burn (away)’, *forheard*

‘very hard’, though no longer productive, has been retained, as in *forbid*, *forget*.

*un-*<sub>1</sub> with negative meaning as in *unfæger* ‘not fair, ugly’, *unfrif* ‘unpeace, war’ is highly frequent in Old English (cf. Marchand 1969: 201), yet only few coinages have survived, like *unclean*, *unripe*.

*un-*<sub>2</sub> with reversative and privative meaning as in *unfealdan* ‘unfold’ and *ungeocian* ‘unyoke’ respectively formed a lot of verbs in Old English, of which but a few still exist such as *undo*, *untie*.

Old English **suffixes** have been preserved to a greater extent.

*-ed* ‘provided with, having’ as in *hilted* ‘hilted’, *hringed* ‘ringed’ has gained nearly unlimited productivity with bases consisting of, e.g., adjective + noun: *beautiful-gowned*, *high-heeled*, and noun + noun: *diesel-engined*, *eagle-eared* (cf. Faiß 1987: 124–129). Many were coined in Early Modern English. Shakespeare has, among others, *light-winged*, *nimble-footed* and *honey-tongued*, *iron-witted* (cf. Scheler 1982: 118).

*-er* ‘one having to do with’ used to form agent nouns (cf. Kastovsky 1971) such as *bæcere* ‘baker’ from *bacan* ‘bake’ (de-verbal) and *bōcere* ‘scribe’ from *bōc* ‘book’ (denominal). Predominantly from the 16th century onwards, *-er* has come to derive instrument nouns like *blotter*, *nutcracker*, *sailer*. The suffix is extraordinarily productive in Present-Day English owing to its capability of being appended to practically any base (cf. Faiß 1987: 118–121).

*-isc*, PrE. *-ish*, ‘belonging to’ originally derived ethnic adjectives with *i*-mutated stem-vowels, such as *frenscisc* ‘French’, *englisc* ‘English’, *scytisc* ‘Scottish’. The meaning of the suffix changed to ‘nature of, condition’ as in OE. *ceorlisc* ‘churlish’, *cildisc* ‘childish’. In Middle English, it developed a disparaging tinge as in *foolish* (before 1300), *swinish* (ca. 1200). Nowadays *-ish* is widely used in colloquial and journalistic language (cf. Faiß 1987: 124).

*-nes*, PrE. *-ness*, ‘state, condition’ mainly derived adjectives and participial adjec-

tives as *æbelnes* ‘nobility’ from *æbele* ‘noble’, *cýþnes* ‘testament’ from *cýþed*, past participle of *cýþan* ‘make known’. In Present-Day English, the suffix may be appended to almost any base except verbs, thus to adjectival composites as in *selfconceitedness*, *square-tonguedness*, adjectival compounds as in *airsickness*, *homesickness*, and adjectival phrases as in *get-at-ability*, *little-boyishness* (cf. Faiß 1987: 123).

### 3.1.2. Compounding

**Compounding** is a productive process not only in Old English but in all Germanic languages. Old English compounds are predominantly nouns. The majority of them follow the patterns noun + noun: *níht-waco* ‘night-watch’, *scip-ráp* ‘ship-rope’; adjective + noun: *eald-fáder* ‘grandfather, ancestor’, *héah-clif* ‘high cliff’; verb(-stem) + noun: *ete-lond* ‘pasture land’, *hwet-stán* ‘whet-stone’; verbal noun in -ungl-ing + noun: *langung-hwíl* ‘time of longing, weariness’, *stemping-ísern* ‘stamping-iron’.

Old English has its share of compound adjectives as well. Productive patterns are: noun + adjective: *ælmes-georn* ‘alms-eager, generous’, *mere-wérig* ‘sea-weary’; adjective/adverb + adjective: *wíd-cúb* ‘widely known’, *wís-hygende* ‘wise-thinking’.

Additive copula compounds are rare in Germanic in general, being still rarer in Old English. The numerals *thirteen* to *nineteen*, e.g., belong here: *þréotýne*, *féowertýne*, *fiftýne*, *siextýne*, *siofontýne*, *eahtatýne*, *nigon-týne*. Productivity has considerably increased in recent years: *poet-critic-humanist*, *saxo-phonist-composer-bandleader*, *soldier-statesman-author-orator*.

The development of *þréotýne* to *thirteen* and *fiftýne* to *fifteen* leads to the problem of obscured compounds (cf. Faiß 1978), meaning that original compounds can no longer be analyzed as such from the point of view of modern English.

### 3.1.3. Zero-derivation

Derivation by means of a **zero-morpheme** is a productive process in Old English. It occurs most often with verbs yielding nouns (cf. Kastovsky 1968) and with nouns yielding weak verbs of classes 1 and 2. Examples of deverbal nouns are: from strong verbs: *drinc* ‘drink, draught’ from *drincan* ‘drink’, *stand* ‘stand, pause, delay’ from *standan* ‘stand’; from weak verbs: *hete* ‘hatred’ from *hatian*

‘hate’, *ræd* ‘council, advice’ from *rædan* ‘advise’. Denominal verbs are: class 1 (with *i*-mutation): *blēdan* ‘bleed’ from *blōd* ‘blood’, *dēman* ‘deem’ from *dōm* ‘doom’, *fēdan* ‘feed’ from *fōda* ‘food’; class 2: *beddian* ‘prepare/make a bed’ from *bed(d)* ‘bed’, *cealfian* ‘calve’ from *cealf* ‘calf’, *horsian* ‘provide with (a) horse(s)’ from *hors* ‘horse’.

Verbs from adjectives are: class 1 (with *i*-mutation): *fýlan* ‘befoul’ from *fūl* ‘foul’, *syl-lan* ‘fill’ from *full* ‘full’; class 2: *deorciān* ‘grow dark’ from *deore* ‘dark’, *ealdian* ‘grow old’ from *eald* ‘old’. Class 2 verbs from particles are: *forbiān* ‘send forth’ from *forb* ‘forth’, *in-nian* ‘get within’ from *in* ‘in, within’; from verbs: *fandian* ‘examine’ from singular preterite *fand* of *findan* ‘find’, *lifian* ‘live’ from preterite stem *lif-* of *libban* ‘live’.

The productivity of zero-derivation has been on the constant increase throughout the history of English. This was favoured by the fact that, with no grammatical suffixes to denote infinitive since around 1500 at the latest, words easily shifted from one lexical category to the other.

Mention has to be made of exocentric or pseudo-compounds, i.e. zero-derived composites (cf. Marchand 1969: 386–389). Most Old English formations of this type are adjectival **bahuvihi** compounds with the patterns adjective + noun, numeral + noun and noun + noun like *bær-fót* ‘barefoot(ed)’, *ān-ēage* ‘one-eyed’; *wulf-heart* ‘wolf-hearted’. Already in Old English -ed coinages were strong competitors. They succeeded in replacing those patterns in Early Middle English, generally speaking, with the exception of *barefoot*. Concerning nominal bahuvihis, Old English has some loan-translations of Latin animal and plant names: *ān-horn* for *unicornis*, *fíflēaf* for *quinquefolium*, genitival *hundes tunga* for *cynoglossum*, to which may be added *dæges-ēge* ‘daisy’, *oxan-slyppe* ‘oxlip’ (plants) and coinages such as *lady’s bedstraw*, *lady’s finger* (plants), *bull’s-eye*, *cat’s-eye* (objects), which are paralleled by genuine compounds like *doomsday* and *driver’s seat* (cf. Marchand 1969: 65–69). They are equally attested in Old English: *Tíwes-dæg* ‘Tuesday’ and *Wōdnes-dæg* ‘Wednesday’ are but two instances.

## 3.2. Middle English

### 3.2.1. Affixation

Many **prefixes** which are still productive like *de-*, *dis-*, *in-*, *non-* came over from French during the Middle English period. Among

them *non-* ‘not’, which represents Lat. *nōn*, Anglo-Norm. *noun-*, Central Fr. *non-*, has developed a combinatory flexibility which is unequalled by other prefixes. Although having achieved its full productivity after the middle of the 17th century, it can be regarded as fairly productive in the 15th and 16th centuries. The first instances are direct borrowings from French: *noumper* ‘umpire’ (ca. 1350), *nonpower* ‘non-power’ (before 1387) (nouns), *noncerteyn* ‘non-certain’ (probably before 1400), *nounparalle* Fr. ‘non-pareil’ (ca. 1450) (adjectives), and from Latin: *noun laudabile* (before 1398).

As to Eng. *non-*, the following may be observed. In the 15th century, it predominantly combines with deverbal nouns denoting action, activity, condition, quality. The pattern has been documented by numerous examples to the present day, with the exception of nouns in *-ing*, whose connection with *non-* obviously dropped out of use in the 19th century. Some instances are: *non-Punyshyne* (1422), *Non-finding* (1525), *non-syllogizing* (1651), *non-repealing* (1791), *non-understanding* (1833), *non-labelling* (1901); *noun obeisance* (1447), *non-distribucioun* (1540), *non-committance* (1650), *non-submission* (1763), *non-pronunciation* (1842), *nontransitivity* (1974). Combinations with agent nouns have become frequent since the 19th century, some instances being: *non housholder* (1439), *non-subscribers* (1599), *Non-contributors* (1643), *Nonjuresses* (1723), *non-smoker* (1846), *non-achiever* (1972). Coinages with other nouns are not as frequent: *noncredibility* (before 1450), *non-religion* (1600), *non-freemen* (1760), *non-form* (1886), *nonfriends* (1971). If used attributively, though, the pattern appears to be rather free from restrictions, which is corroborated by lots of examples from the 19th and 20th centuries like *non-factory* (1835), *non-market* (1884), *Nonkernel* (1963), *nontitle* (1973). These centuries also abound with adjectival coinages: *non-adherent* (1870), *non-analytic* (1890), *nonearthly* (1970), *nonpermissive* (1972). *non-* has easily combined with adjectival present and past participles since the 17th century: *Non-preaching* (1622), *non-reading* (1797), *non-enquiring* (1864), *non-fattening* (1971); *Non-incarnated* (1671), *non-perforated* (1776), *non-galvanized* (1849), *non-palatalized* (1956). Since the 19th century, it has been attached to adjectival compounds whose determinata are present participles, such as *non-officeholding* (1888), *non-English-speaking* (1949).

It has been used with adverbs since the 17th century: *non-attendingly* (1678), *non-voritically* (1882), *nonbiologically* (1974). Though reluctantly, it has been prefixed to verbs since the 17th century as well: *Non-act* (1645), *non-concur* (1703), *non-apply* (1846), *non-intervene* (1969); with infinitives it has been employed since the 1920s to form an attributive phrase as in *non-crush* (1924), *non-shrink* (1962). Since the late 1960s it has acquired the additional meaning ‘not being true, genuine; pseudo-’ as in the nouns *non-actor* (1970), *non-newspaper* (1973).

The impact of French on English **suffixes** is about the same as that on prefixes (cf. Kastovsky 1985: 227). Among those arising in Middle English can be found *-able* and *-ess*.

*-able* ‘fit for doing, fit for being done’, with the variant *-ible*, forms adjectives from verbs and nouns like *determinable*, *understandable*; *honourable*, *pesible* ‘peaceable’. Deverbal adjectives have always been more frequent than denominal ones, in Middle English as well as in later periods. Instances are: *bombable* (1930), *demountable* (1909), *flyable* (1936); *carriageable* (1702), *fissionable* (1945).

*-esse*, PrE. *-ess*, forms [+female] nouns parallelly to [+male] personal nouns such as *charmeresse*, *daunceresse*, *frendesse*. The suffix is most productive in the 16th and 17th centuries slowly declining afterwards. The 19th century has *squires*, *visitress*, *waitress*; *rebeless* dates from 1912.

### 3.2.2. Compounding

Early Middle English continues the **compounding** traditions of Old English, with new types and patterns emerging (cf. Sauer 1992).

In inversion compounds the usual sequence determinant – determinatum is permuted under French influence: *knight erraunt* (ca. 1400) from *chevalier errant* ‘knight errant’, *herbe Ive*, *herbe Robert* (plant names; before 1300). Further instances are: *court-martial*, *cousin german*, *heir apparent*.

Gender-denoting attributive copula compounds with *he-* and *she-* as determinants came into being about 1300, originally designating but animals: *hee-catte* ‘he-cat’, *he lambe* ‘he-lamb’; *she-as* ‘she-ass’, *shee ape* ‘she-ape’. Words for persons have been recorded since the 16th century, those with *she-* being more frequent: *he-friend*, *he-lover*; *she-*

*baker*, *she-favorite*, *she-friend*, 20th-century being: *he-male* (1909); *She Guardian* (1937), *she-robot* [+human] (1991).

Compounds with particles like *over*, OE. *ofer*, as in *overlord* or *overbysy* ‘overbusy’, and *out*, OE. *ut*, as in *outerie* ‘outray’, *outlep* ‘outleap’ remain common in Middle English.

### 3.2.3. Zero-derivation

In Middle English, the number of nominal bahuvrihi-compounds increases considerably. The paraphrase of such a combination is usually ‘X has/is characterised by what is expressed by the composite’. Instances include *Curtmauntel* ‘C(o)urtmantle’ (surname of Henry II), *long-here* ‘person with long hair’, *redrest* ‘redbreast’ (bird), *stykylbak* ‘stickleback’ (fish). Person names and nicknames are most frequent: *Henheued* ‘hen-head’, *Hertblod* ‘heart-blood’, *Irnefot* ‘iron-foot’ (cf. Jönsjö 1979: 111, 115; Erlebach 1979: 71–87, 93–99 for names of French origin such as *Bonquer* ‘bon cœur’ and *Cordelion* ‘œur de lion’, i.e. ‘Lionheart’). In the following centuries, many [+human] as well as [−human] nouns have been coined. The meaning of the former is mainly pejorative, that of the latter only occasionally. As in Middle English, the patterns adjective + noun and noun + noun occur most frequently, cf. *wetback* ‘illegally immigrated Mexican’, *whitethorn* (plant), *whitethroat* (bird); *bricktop* ‘red-haired person’, *cottontail* ‘rabbit’, *paperbelly* ‘person unable to drink much’. This highly restricted selection may be completed by the weak patterns numeral + noun as in *four-eyes* ‘person wearing glasses’, *four-wheel* (vehicle) and reduced adjectival past participle + noun as in *crackbrain*, *scatterbrain* ‘stupid person’. The latter obviously did not come into existence before the 16th century.

A different pattern originating in Middle English consists of verb and complement. It arose under the influence of French imperative phrases and is hence sometimes termed **imperative compound** (but cf. Giurescu 1975: 69–71). Early Middle English coinages follow French conventions in that they mostly denote personal agents: *pyke-purs* ‘pickpurse’, *trayle-bastoun* ‘s.o. who carries a club or a cudgel’ (a criminal), *wesche-disch* ‘dishwasher’. Numerous person names and nicknames, a few attested in the 2nd half of the 11<sup>th</sup> century already (cf. Reaney 1976: xi), are formed this way, like *Schaketaille* ‘shake-tail’, *Schaketrot* ‘shake-trout’, *Schakspær* ‘shake-spear’ (cf. Jönsjö 1979: 156). All

of them have a pejorative tinge, which is not inherent in the rare [−animate] nouns dating from the 14th and 15th centuries such as *breakfast* and *turnpike*.

### 3.3. Early Modern English and Later

The transition from Middle to Early Modern English and then to Late Modern English was not as drastic as that from Old to Middle English. No significant breaks occurred in word-formation processes, which nonetheless were enriched by new concepts.

#### 3.3.1. Affixation

Some modern affixes of note became productive in Early Modern English. Among the **prefixes** are *counter-*, *de-*, *inter-*, *pre-*, *re-*; **suffixes** include *-ation*, *-ee*, *-ive*, *-ment*, as in *counterevidence* (1665), *depopulate* (1545), *intermarry* (1574), *pre-elect* (1570), *reconsider* (1571); *derivation* (1530), *debtee* (1531), *trustee* (1647) (patient nouns, agent nouns such as *escapee* (1875–76) are later and particularly of American origin), *persuasive* (1589), *retirement* (1596).

Besides, Early Modern English experienced the rise of combining forms, which have become highly productive. They combine not only with free morphemes, but also with other combining forms. Early instances are: *-cacy* ‘rule, power’ from *aristocracy* (1561), *-logy* ‘doctrine, study of’ from *theology* (as early as 1362), *poly-* ‘many’ from *poly-tragick* (1605), some current formations being: *Eurocracy*, *mediocracy*, *meritocracy*; *escapology*, *mafiology*, *urbanology*, with reference made to the study of political situations or events as in *Kremlinology*, *Pekin(g)-ology*, *Sovietology*; *polydrug*, *polymorphous*, *polynuclear*.

Further combining forms being productive nowadays are: *bio-* ‘life’ as in *bio-ethics*, *biometric*; *eco-* ‘ecology, ecological’ as in *ecotourist*, *ecotoxicology*; *Euro-* ‘Europe, European’ as in *Europerson*, *Euroslit*; *mega-* ‘extremely; very large; one million’ as in *megaboring*, *mega-busy*; *megastore*; *megatanker*.

#### 3.3.2. Compounding

*Fowel of flight* (ca. 1300) is the first instance of a new **compound** pattern the determinant of the inversion compound being a prepositional phrase with *of*. Productivity is achieved in Early Modern English with instances like *Bill of Exchange* (1579), *House of Commons* (1621), *House of Lords* (1672).

### 3.3.3. Zero-derivation

The pattern verb + complement has become frequent since Early Modern English (cf. Uhrström 1918), although, in our century, it has lost much of its productive force. Person nouns are: *lack-love*, *mumble-news*, *please-man* from Shakespeare, further instances being *killjoy* (1776), *turnkey* (1654), *turn-penny* (1824). Animal nouns are: *shearwater* (ca. 1671), *shufflewing* (1829), *wagtail* (1510), plant names *catch-fly* (1597), *cover-shame* (1693), *heal-all* (1853), others (tools, devices, garments, etc.) *breakwater* (1721), *cover-shame* (1629), *stopgap* (1533).

A different kind of zero-derivation is represented by the patterns *showoff* and *blackout*. Both gave rise to numerous coinages especially in the 19th century and afterwards, the frequency of non-animate *blackout* nouns being higher than that of mainly [+human] *showoff* nouns, which did not come into existence till the 2nd half of the 16th century, while members of the other group, though rarely, appeared a century earlier. Some instances are: *go-between*, *look-alike*, *runabout*, *runaway*; *build-up*, *cutback*, *cut-down* (cf. Marchand <sup>2</sup>1969: 380–386).

### 3.3.4. Shortening

The shortening processes clipping and blending (cf. Art. 91; 92), developed to a certain degree in Middle English, intensified during the centuries to follow. Together with acrony whole (cf. Art. 92) they have become outstanding features of the modern language.

**Clipping** got firmly established in Early Modern English. Examples comprise: *mere* ‘siren, mermaid’ (1225) from *mere-min*, *gent* (1564) from *gentleman*, *mas* (1575) from *master* (back-clippings); *lepi* ‘single’ (ca. 1303) from *onlepi*, *wig* (1675) from *periwig*, *winkle* (1585) from *periwinkle* (fore-clippings). Back-clipping has turned out to be the most productive among clipping processes. Some instances are: *celeb* from *celebrity*, *cig* from *cigarette*, *fax* from *facsimile machine* (back-clippings); *burger* from *hamburger*, *liner* from *airliner* (fore-clippings); *fridge* from *refrigerator*, *Soo* from *Lake Superior* (fore- and back-clippings, rare); *Ten-Tom* from *Tennessee-Tombigby-Waterway*, *Tex-Mex* from *Texas-Mexican* (multiple back-clippings). Clipping also affects compounds. Consisting of one clipped and one non-clipped free constituent (either the first or the last), they are called “clipping compounds”. An early instance is

*blatterature* (ca. 1512) from *blatter* and *literature*. From various spheres are: with the first free constituent clipped: *compucenter*, *docudrama*, *helihop*; with the second one clipped: *beefalo*, *happenseance*, *steelionaire*.

**Blending** does not seem to be frequent until the 19th century: *lemeren* ‘glisten, glimmer’ (probably before 1400) from *lemen* ‘glow, shine’ and *glimeren* ‘gleam’, *blurt* (1573) from *blow* and *spurt*, *Barsolistor* (1888) from *barrister* and *solicitor*, *clantastical* (1803) from *clandestine* and *fantastical* are some instances (cf. Bergström 1906: 51–70). In Present-Day English, though, blending is a productive means of expressing the mixture of two concepts, such as languages in contact: *Engdish* from *English* and *Yiddish*, *Frenglish* from *French* and *English*, *Japlish* from *Japanese* and *English*.

**Acrony whole** is the latest addition to shortening in the English language. It obtained some productivity in the 19th-century special language of chemistry (cf. Marchand <sup>2</sup>1969: 453), but the first real surge of acronyms did not come before World War I: *ANZAC* ‘Australian New-Zealand Army Corps’, *DORA* ‘Defence of the Realm Act’ are but two examples.

### 3.3.5. Back-derivation

Often considered a sub-class of shortening, **back-derivation** (or back-formation) is the coining of new words by the deletion of actual or supposed suffixes in longer ones (cf. Marchand <sup>2</sup>1969: 391–395). Most back-derivations are verbs, as *laze* (before 1592) from *lazy*, *dizz* (1632) from *dizzy*; *infract* (1798) from *infraction*, *book-keep* (1886) from *book-keeping*, *caretake* (1893) from *caretaker*; *team-teach* (1976) from *team-teaching*, *back-calculate* (1987) from *back-calculation*.

### 3.3.6. Reduplication

Being sparsely attested in Middle English, **reduplication** (cf. Art. 57) has become more common since Early Modern English and is highly productive in the modern language.

Rhyming combinations are: *handy-dandy* ‘covert bribe or present’ (1362); *hocus-pocus* ‘jugglery, trickery’ (1624), *hurly-burly* ‘tumult, confusion’ (1539); *hubble-bubble* ‘confused talk’ (1720), *mumble-jumble* ‘speak indistinctly’ (1833), *quavery-wavery* ‘undecided’ (1749). The 20th century has *clankie-tankies* ‘members of the Royal Tank Regiment in the Gulf War’, *heeble-jeebies* ‘state of being scared’, *hotsy-totsy* ‘fine’.

Ablaut combinations have quickly gained ground: *pitter-patter* ‘rapid repetition of words’ (ca. 1425); *bibble-babble* ‘idle talk’ (1532), *clitter-clatter* ‘chatter’ (1535), *dilly-dally* ‘waste time’ (before 1610); *twiddle-twaddle* ‘foolish chatter’ (1798). From the 19th century onwards have been coined *drizzle-dazzle* ‘continuous rain’, *fizz-fuzz* ‘Coca-Cola’, *rickety-rackety* ‘unsteady’.

Reduplicative compounds (or repetitions), infrequent in Middle English, remained rare to the 18th century. Since the 19th, however, their frequency has been increasing considerably (cf. Thun 1963: 49–200 *passim*). Some sound-imitating coinages are: *cuccu* ‘cuckoo’ (ca. 1240); *diddle-diddle* (fiddle) (1523), *jingle-jingle* (jingling sound) (1664); *tick-tick* (1774) (clock), *too-too* (flute) (1812), *woo-woo* (wind) (1841), *zip-zip* (light sharp sound) (1875). *choo-choo* (train), *chuff-chuff* (engine), *clop-clop* (hooves/wooden sandals) are 20th century.

### 3.4. Conclusion

The basic Old English word-formation types affixation, compounding and zero-derivation have generally persisted to the present day. As in the case of inflection, Middle English takes a crucial position. Not only did it experience the influx of French prefixes and suffixes and the coming out of use of Old English ones, but it also laid the foundation for new patterns such as the zero-derivations *cutpurse* and *blackout*, new compound types and the shortening processes clipping and blending. The most important contribution of Early Modern English to word-formation is doubtless the combining form, which has gained high productivity during the following centuries showing an ever-increasing frequency in Present-Day English. Of Early Modern English origin is furthermore the zero-derivational pattern *showoff*, whereas acronyms was not en vogue until World War I. Since then, however, its productivity has been growing steadily, as is illustrated by numberless coinages.

## 4. References

- Ayto, John (1988), *The Longman Register of New Words*. Harlow: Longman  
 Barnhart, Robert K. & Steinmetz, Sol & Barnhart, Clarence L. (1990), *Third Barnhart Dictionary of New English*. n.p.: Wilson

Bergström, Gustav A. (1906), *On Blending of Synonymous or Cognate Expressions in English: A Contribution to the Study of Contamination*. Ph.D. dissertation. Lund: Ohlsson

Berndt, Rolf (1956), *Form und Funktion des Verbums im nördlichen Spätaltenglischen*. Halle/S.: Niemeyer

Brunner, Karl (³1965), *Altenglische Grammatik*. Tübingen: Niemeyer [¹1942]

BT (1976) = Bosworth, Joseph, *An Anglo-Saxon Dictionary*, ed. and enl. by T. Northcote Toller. London etc.: Oxford Univ. Press [repr., ¹1898]

BTS (1973) = Bosworth, Joseph, *An Anglo-Saxon Dictionary. Supplement* by T. Northcote Toller with rev. and enl. addenda by Alistair Campbell. London etc.: Oxford Univ. Press [repr., ¹1921]

Erlebach, Peter (1979), *Die zusammengesetzten englischen Zunamen französischer Herkunft*. Heidelberg: Winter (Anglistische Forschungen 137)

Faiß, Klaus (1978), *Verdunkelte Compounds im Englischen*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 104)

Faiß, Klaus (1987), “Words, Words, Words: Compound or Suffixation in German and English”. In: Lörtscher, Wolfgang & Schulze, Rainer (eds.), *Perspectives on Language in Performance: Studies in Linguistics, Literary Criticism, and Language Teaching and Learning. To Honour Werner Hüllen on the Occasion of His Sixtieth Birthday*, Vol. I. Tübingen: Narr, 110–132

Faiß, Klaus (1989), *Englische Sprachgeschichte*. Tübingen: Francke

Faiß, Klaus (1992), *English Historical Morphology and Word-Formation. Loss versus Enrichment*. Trier: wwt (Fokus Bd. 8)

Giurescu, Anca (1975), *Les Mots Composés dans les Langues Romanes*. The Hague, Paris: Mouton (Janua Linguarum Series Practica 228)

Jönsjö, Jan (1979), *Studies in Middle English Nicknames. I: Compounds*. Lund: Gleerup (Lund Studies in Eng. 55)

Kastovsky, Dieter (1968), *Old English Deverbal Substantives Derived by Means of a Zero Morpheme*. Ph.D. dissertation, Univ. Tübingen

Kastovsky, Dieter (1971), “The Old English Suffix -ER(E)”. *Anglia* 89, 285–325

Kastovsky, Dieter (1985), “Deverbal Nouns in Old and Modern English: From Stem-Formation to Word-Formation”. In: Fisiak, Jacek (ed.), *Historical Semantics. Historical Word-Formation*. Berlin etc.: Mouton, 221–261

Krahe, Hans & Meid, Wolfgang (1967), *Germanische Sprachwissenschaft*, Vol. III: *Wortbildung*

dungslehre. Berlin: de Gruyter (Sammlung Göschens 1218/1218a/1218b)

Marchand, Hans (²1969), *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*. München: Beck [¹1960 Wiesbaden: Harrassowitz]

MED (1954 ff.) = *Middle English Dictionary*, Vol. 1ff. Ann Arbor/MI: Univ. of Michigan Press [predominantly] [various editors]

OED (²1989) = *The Oxford English Dictionary*, Vol. 1–20, prepared by J. A. Simpson & E. S. C. Weiner. Oxford: Clarendon Press [¹1933]

Reaney, P. H. (²1976), *A Dictionary of British Surnames*. London, Boston: Routledge & Kegan Paul [¹1958]

Sauer, Hans (1992), *Nominalkomposita im Frühmittelenglischen: Mit Ausblicken auf die Geschichte der*

## XVIII. Morphologischer Wandel II: Fallstudien

englischen Nominalkomposition. Tübingen: Niemeyer (Buchreihe der Anglia 30)

Scheler, Manfred (1982), *Shakespeares Englisch: Eine sprachwissenschaftliche Einführung*. Berlin: Schmidt (Grundlagen der Anglistik und Amerikanistik 12)

Thun, Nils (1963), *Reduplicative Words in English: A Study of Formations of the Types Tick-tick, Hurly-burly and Shilly-shally*. Uppsala: Bloms

Uhrström, Wilhelm (1918), *Pickpocket, Turnkey, Wrap-Rascal, and Similar Formations in English: A Semasiological Study*. Stockholm: Bergvall

Wright, Joseph & Wright, Elizabeth Mary (1908), *Old English Grammar*. London etc.: Frowde, Oxford Univ. Press

Klaus Faiß, Mainz (Germany)

## 156. Vom Althochdeutschen zum Neuhochdeutschen

1. Periodisierung der deutschen Sprachgeschichte
2. Synchrone Heterogenität und historische Homogenität
3. Forschungsgeschichte
4. Entwicklungsgeschichtliche Grundzüge
5. Flexion
6. Wortbildung
7. Zitierte Literatur

### 1. Periodisierung der deutschen Sprachgeschichte

Der theoretischen Vielfalt der Kriterien für eine sprachgeschichtliche Periodisierung entspricht die tatsächliche Vielfalt der Periodisierungsversuche der deutschen Sprachgeschichte (vgl. die Zusammenstellung bei Roelcke 1995; 2001). Oberhalb aller Feindifferenzierungen gilt seit Wilhelm Scherer vor allem eine Unterteilung in die vier Sprachstufen Althochdeutsch (Ahd.), Mittelhochdeutsch (Mhd.), Frühneuhochdeutsch (Frnhd.) und Neuhochdeutsch (Nhd.) (vgl. Scherer ²1878: 13–15). Soweit die Unterteilung erststellig durch innersprachliche Kriterien begründet wird, handelt es sich dabei zumeist um Kriterien phonetisch-phonologischer Art; in nur einem Fall wird eine (flexions)morpho-

logische Entwicklung als für die Periodisierung wesentlich benannt: die den Übergang des Frühneuhochdeutschen in das Neuhochdeutsche markierende Ausgleichung des präteritalen **Numerusablauts** (**Ablautausgleich**) bei den starken Verben (ahd./mhd./frnhd. *fand – funden* > frnhd./nhd. *fand – fanden*; vgl. Behaghel ⁵1928: 149; Moser 1929: 1). Diese primäre Orientierung an phonologischen Prozessen zeigt die weiterwirkende **junggrammatische** Bindung der Sprachgeschichtsschreibung: Aufgrund vorgängig physiologisch-lautgesetzlicher Veränderungen werden auch grammatische Strukturen zerstört; erst in der Folge werden die defekten Strukturen aufgrund von Assoziationen der Sprecher (als **Analogie**) wieder in eine sinnvolle, neue Ordnung gebracht (vgl. Kohrt 1984: 392; Werner ²1998: 575). Damit ist eine erste Ursache auch des morphologischen Wandels primär physiologisch-lautgesetzlich erklärt, die im Bereich der Morphologie wirksame Analogie bekommt “eine nachgeordnete komplementäre Erklärungsfunktion” (Putschke ²1998: 476). Es folgt somit der junggrammatischen Logik, die wesentlichen Periodengrenzen über die als Ursache der grammatischen Veränderungen gedachten Lautveränderungen zu markieren.

Im Althochdeutschen (frühes 7. Jahrhundert bis Mitte 11. Jahrhundert) beginnt die geschichtliche und d.h. schriftsprachliche Zeit des Deutschen (vgl. Wolf 1981: 66). Da als sprachgeschichtlich wesentliche Entwicklung die „Transponierung und Bewältigung der Latinität in die Volkssprache“ gilt, meint „Althochdeutsch“ in erster Linie „die Sprachverwendung und erst in zweiter Linie das Sprachsystem“ (Wolf 1981: 72). Als innersprachliches Abgrenzungskriterium gegenüber dem chronologisch vorausliegenden Südgermanischen gilt die **2. oder althochdeutsche Lautverschiebung**, bei der u.a. aus germanischen Tenues neue Spiranten/Affrikaten entstanden. Da die Lautverschiebung nicht alle (süd)germanischen Stammesdialekte des späteren deutschen Sprachraumes ergriffen hat (Ausnahme besonders das Sächsische), liefert sie zugleich auch das bis in die Gegenwart hinein gültige Kriterium zur dialektgeographischen Abgrenzung eines hochdeutschen (hd.) und eines niederdeutschen (nd.) Sprachraums.

Als wesentliches Abgrenzungskriterium des Mittelhochdeutschen (Mitte 11. bis Mitte 14. Jahrhundert) gegenüber dem vorausliegenden Althochdeutschen gilt die in einheitlicher Graphie -e- aufscheinende Abschwächung der vollen Endsilbenvokale (ahd. *ich bitt-u* > mhd. *ich bitt-e*, ahd. *er bitt-it* > mhd. *er bitt-et*), daneben auch die Durchführung des Sekundärumlauts (ahd. *wahsit* > mhd. *wähset*; vgl. Wolf 1981: 64); die jüngere Grenze zum Frühneuhochdeutschen hin markieren weitere Lautwandel, so besonders die sogenannte (früh)neuhochdeutsche Diphthongierung/Monophthongierung (z.B. mhd. *wīn* > nhd. *Wein*; mhd. *bruoder* > nhd. *Bruder*). Zum ersten Mal in der deutschen Sprachgeschichte wird im schriftsprachlichen Bereich eine funktional (literarisch) wie diastratisch (höfisch-ritterlich) bestimmte Vorbild-Varietät herausgebildet. Diese als höfisches – klassisches – und auch als Normalmittelhochdeutsch bezeichnete Sprachform (vgl. Sondergger 1979: 167) stellt jedoch nur einen regional wie funktional begrenzten Ausschnitt der Sprachwirklichkeit des Zeitraumes dar, so daß in neuerer Forschung der Sprachbegriff Mittelhochdeutsch teilweise bewußt vermieden wird (vgl. Wells 1990: 106).

Im Frühneuhochdeutschen (Mitte 14. bis Übergang 18. Jahrhundert; vgl. besonders Reichmann 1988) tritt die Varietätenvielfalt des Deutschen erstmals deutlich in die Überlieferung ein. Die schriftdialektale Bindung und d.h. eine in unterschiedlicher Ausprägung bestehende landschaftliche Variation auch in der Schriftlichkeit wird zum charakteristischen Merkmal der gesamten Periode (vgl. Penzl 1984: 12 f.). Im Neben- und Miteinander der verschiedenen Varietäten findet im Verlauf des 14. bis 18. Jahrhunderts auf der Ebene der Schriftlichkeit ein in Stufen verlaufender Aus- und Angleichungsprozeß statt: Klar bestimmbare sonderlandschaftliche Schreibsysteme werden in eine überregionale und oberhalb der gesprochenen Mundarten liegende Schriftsprache überführt; dabei sind im Verlauf des 15./16. Jahrhunderts mindestens 12

verschiedene Schreib-/Kanzlei- und Druckersprachen unterscheidbar (vgl. Besch 1980: 590–592). Diese Schriftsprache wird zur Grundlage der in der Folge einsetzenden Kodifizierung und Normierung: Das Frühneuhochdeutsche wird zur entscheidenden Phase für die Herausbildung der neuhochdeutschen Schriftsprache. Da insbesondere im 16. Jahrhundert die eigenständige niederdeutsche Schriftsprache des norddeutschen Raumes durch die hochdeutsche Schriftlichkeit verdrängt wurde (vgl. Sodmann 2000), ist damit erstmals in der deutschen Sprachgeschichte eine oberhalb aller dialektalen Differenzierungen liegende schriftsprachliche, überdachende Varietät des gesamten deutschsprachigen Raumes herausgebildet: Die die historische Einzelsprache *Deutsch* ausmachende Existenzform eines differenzierten Diasystems ist in seinem Grundgerüst herausgebildet.

## 2. Synchrone Heterogenität und historische Homogenität

Die sprachgeschichtliche Skizze zeigt, daß eine geschichtliche Morphologie des Deutschen bereits hinsichtlich des Sprachbegriffs ambivalent ist: So meint „Deutsch“ eine konkrete Sprachform, die innerhalb eines historisch-sozial bestimmten Kommunikationsraumes eine allgemeine Geltung besitzt, konkret die (neu)hochdeutsche Schrift- oder Hochsprache; diese Sprachform ist erst seit dem 18. Jahrhundert entwickelt. Zugleich eint „Deutsch“ eine abstrakte und inner- wie außerlinguistisch bestimmte Zusammengehörigkeit einer Reihe konkret vorhandener Sprachformen innerhalb des Gültigkeitsbereichs der Standardsprache (z.B. Dialekte), die eine jede nur über eine eingeschränkte Geltung verfügt. Innerhalb der unterscheidbaren Varietäten/Subsysteme ist die Standardform (das „Hochdeutsche“) selbst nur eine Varietät neben anderen. Ihre Bedeutung liegt jedoch u.a. darin, daß erst ihre (geschichtliche) Herausbildung und Stabilisierung die zu ihr komplementäre Bestimmung der Varietäten ermöglicht. Der Zustand dialektischer Gliedertheit sowie das konkrete Varietätengefüge sind Ergebnis einer insbesondere im Frühneuhochdeutschen vollzogenen Entwicklung, die in besonderer Weise mit dem neuzeitlichen Prozeß der Herausbildung nationalstaatlicher Verkehrs- und Kommunikationsräume verbunden ist. Dabei zeigt sich, daß die mundartlichen Grundlagen der überdachenden Standardvarietät allein innerhalb des südlichen, hochdeutschen Sprachraums liegen (Teilhabe an der 2. oder ahd. Lautverschiebung), daher die in den Epochen-

bezeichnungen (z. B. "Althochdeutsch") vorliegende Eingrenzung. Somit ist der Begriff "Hochdeutsch" innerhalb des Varietätsystems ebenfalls doppeldeutig: normativ im Sinne der Normvarietät, historisch und sprachgeographisch im Sinne der durch die Lautverschiebung definierten Sprachstufen und Mundarten des Deutschen (vgl. Sonderegger 1979: 166). Gemäß der **Varietätendifferenzierung** stellt auch das historisch entwickelte System der gegenwärtigen standardsprachlichen Flexions- und Wortbildungsmorphologie nur eines unter anderen deutschen Systemen dar. Dabei hat die Morphologie jeder der unterscheidbaren Varietäten ebenso ihre eigene Geschichte wie auch die besondere diasystemische Gegliedertheit der deutschen Morphologie insgesamt. Damit ist das grundlegende Problem der diachronen Betrachtung des Deutschen angesprochen: Es geht um ihre Identität in Geschichte und Raum (vgl. Mattheier 1998: 825). Tatsächlich kann von einem "Wandel" sprachlicher Einheiten/Strukturen nur innerhalb eines als identisch zu bestimmenden funktionalen, sozialen und regionalen Gültigkeitsbereichs gesprochen werden. Das System der Flexions- wie Wortbildungsmorphologie der im späten 18. Jahrhundert schon weitgehend erreichten "Normalvarietät" des Neuhochdeutschen wäre somit zu vergleichen mit dem System entsprechender Varietäten der vorausliegenden Sprachstufen: Eine "Normalvarietät" ist aber weder für das Althochdeutsche, noch für das Mittelhochdeutsche oder das Frühneuhochdeutsche anzunehmen; entsprechende Paradigmen der historischen Grammatiken, auf denen der sprachgeschichtliche Überblick gründen muß, haben ausschließlich heuristischen Wert. Die relativierende Formulierung erweist sich besonders bezüglich der geschichtlichen Genese der neuhochdeutschen Standardvarietät als bedeutsam. Denn im Unterschied zu anderen europäischen Kultursprachen weist das Deutsche nur begrenzt eine historische Homogenität auf: Im deutschen Reich fehlte ein klar bestimmbarer Zentrum, dessen Varietät sich im sprachgeschichtlichen Prozeß zur allgemeinen und d. h. Standardvarietät hätte qualifizieren können. Somit ist auch das im Neuhochdeutschen entwickelte morphologische System in wesentlichen Teilen nurmehr ein Produkt der Überlagerung und eklektizistischen Addition verschiedener und varietätendifferenziert herausgebildeter Teilsysteme/-strukturen/-prinzipien. Auch und gerade außersprachliche Ur-

sachen haben in der Sprachgeschichte des Deutschen eine wesentliche Rolle gespielt (zur strukturellen Disposition sowie den weiteren Prinzipien des im Frühneuhochdeutschen erfolgten Sprachausgleichs vgl. Besch 2003).

Die Berücksichtigung der Varietätendifferenzierung hat Folgen für die sprachgeschichtliche Argumentation: Die gemeinhin als "Wandel" oder "Fortentwicklung des Systems" klassifizierte Veränderung z. B. mhd. *gabe* > nhd. *gebe* (1. Sg. Ind. Präs.) ergibt sich nur aus dem unmittelbaren Vergleich der neuhochdeutschen Standardvarietät mit der in den entsprechenden Grammatiken ausgewiesenen mittelhochdeutschen "Normalform". Tatsächlich kann die Entwicklung adäquater beschrieben werden als Perpetuierung und anschließende Verallgemeinerung einer schon in mittelhochdeutscher Zeit neben *gabe* im mitteldeutschen Sprachraum parallel existierenden Form *gebe*. Damit ist *gabe* nicht durch einen "Wandel" verändert, sondern aufgrund einer sprachlandschaftlich veränderten Fokussierung nurmehr innerhalb der Schriftlichkeit verdrängt; die Form *gabe* ist in oberdeutschen Mundarten unverändert bis in die Gegenwart hinein gebraucht.

Die deutsche Sprachgeschichte und damit auch die Geschichte des morphologischen Systems erweist sich zu einem nicht unbedeutenden Teil als die chronologisch aufeinanderfolgende Fokussierung mal der einen, mal der anderen Varietät; insofern liegt nicht zwingend eine Entwicklung des Systems, sondern die unterschiedliche Wirksamkeit varietärer Ausprägungen vor.

### 3. Forschungsgeschichte

Der wesentliche Aspekt der Forschungsgeschichte liegt in der Ausdifferenzierung der **Wortlehre** (i.e. Morphologie) in eine jeweils eigenständige Flexions- und Wortbildungslære (Solms 1998). Für die Frühphase einer Grammatikographie des Deutschen, die der geschichtlichen Herausbildung einer Standardvarietät im Übergang des Frühneuhochdeutschen zum Neuhochdeutschen parallel geht (im 17. Jahrhundert insbesondere Justus Georg Schottel, im 18. Jahrhundert besonders Johann Christoph Gottsched und Johann Christoph Adelung), gilt eine noch weitgehend undifferenzierte Thematisierung der Wortlehre. Eine solche Differenzierung gelang erst durch eine adäquate Berücksichtigung der Funktion und d. h. der durch Flexion und Wortbildung jeweils realisierten spezifischen "Bedeutung": Während die Fle-

xionslehre “es der Syntax überlässt, die Bedeutung und Verwendung” der Flexive zu bestimmen, ist jene der Wortbildungsmittel allein durch eine eigene, die Wortbildungslehre zugleich unmittelbar konstituierende “Funktionslehre” zu beschreiben (Paul 1981: 17 f.).

Den Weg zu einer solchen Unterscheidung wies zuerst Jacob Grimm, indem er die Differenzierung von Wortbildung und Flexion als eine im Fortgang der sprachgeschichtlichen Entwicklung erst herausgebildete erkannte: Grimm erkannte innere Wortbildung (*trinken, der Trank*) und Wurzelflexion (*trinken – trank*) als ursprünglich ungeschieden. Im darin wirksamen Ablaut sah Grimm das wesentliche und d. h. historisch wie genetisch grundlegende Form- und Bauprinzip der deutschen Sprache: er erkannte es als ihr ältestes Prinzip, wobei die Wurzeln aller Wörter mit dem Grundsatz des Ablaut sowie “dieser mit der natur der wurzeln wesentlich in gemeinschaft” stehen (Grimm 1826: 4–7). So wenig somit innere Wortbildung und Wurzelflexion zu unterscheiden sind, so wenig kann man “die flexion, d. h. die dem verbum anwachsende pronominalform, die dem nomen anwachsenden geschlechtszeichen und partikeln strenge nicht vom begriffe der wortbildung ausschließen” (Grimm 1826: 86). Erst in der geschichtlichen Entwicklung werden die formalen Ausdrucksmittel und -muster funktional profiliert und als Kategorie faßbar, so z. B. die von Grimm so bezeichneten “eigentlichen” Komposita (z. B. *Berg-werk* ohne Fugen-element versus *Krieg-s-geschrei* mit Fugen-element):

“Zur charakteristischen Compositionsform wurden diese altertümlichen Gefüge [erst] durch die Ausbildung der Flexion. Denn eben dadurch, dass man sich gewöhnte, die Beziehungen zwischen Wörtern durch Flexionen zu bezeichnen, erschienen die Verbindungen mit der reinen Stammform als Composita.” (Wilmanns 1906–1909: 4)

Eine konsequente Berücksichtigung des damit formulierten relationalen Charakters der Unterscheidung von Wortbildung und Flexion und ihrer historischen Entwicklung ist bisher nicht geleistet. Für Grimm gehört die Morphologie neben der Lautlehre zum Kern von Sprachvergleichung wie Sprachgeschichte. Diese Gegenstandskonstitution bleibt auch im späteren junggrammatischen Sprachwandelkonzept bestehen (vgl. Eichinger 1984: 436). Indem nun jedoch die bei Grimm vorherrschende Überzeugung, Sprachgeschichte erweise in ihren historischen Stu-

fen/Epochen den fortschreitenden “Verfall” der ursprünglichen Klarheit und Systematizität, überwunden wird, werden nun durch den junggrammatischen Ansatz eines zu allen Zeiten “nach denselben strengen phonetischen und psychologischen Regeln” (Werner<sup>2</sup> 1998: 575) erfolgenden Sprachwandels alle Sprachstufen als prinzipiell gleichwertig erkannt; auf der Grundlage eines jeweils reichen empirischen Materials eröffnet dies grundsätzlich die Möglichkeit zum Systemansatz und d. h. für eine Beschreibung der jeweiligen Sprachstufen als eigenständiger und hinsichtlich übereinstimmender Merkmale homogener Synchronien (vgl. Eichinger 1984: 436). Der junggrammatische Ansatz führt zu einer Reihe historischer Grammatiken (nicht nur zum Deutschen), die bis heute ihren Wert als Standardwerke nicht eingebüßt haben (vgl. zusammenfassend Werner 1998: 575 f.); einige Grammatiken beschreiben den Gesamtzusammenhang der geschichtlichen Entwicklung des deutschen Flexionssystems vom vorgermanischen Ausgangspunkt bis in das Neuhochdeutsche hinein (Wilmanns 1906–1909; Paul 1916–1920), andere bleiben auf jeweils eine der definierten Sprachstufen begrenzt (Braune<sup>14</sup> 1987; Weinhold<sup>2</sup> 1883; Paul<sup>24</sup> 1998; Michels<sup>5</sup> 1979). Mit Ausnahme der periodenübergreifenden Darstellungen gerät die Wortbildung nicht in den Blick, unter der Formenlehre wird nun ausschließlich Flexionslehre betrieben. In eben dieser Begrenzung hatte auch Virgil Moser (1929: XII) seine “frühneuhochdeutsche Formenlehre” geplant; die Gleichsetzung von Morphologie und Flexion gilt auch für neuere Darstellungen als eine gegebene (vgl. Kern & Zutt 1977: 1; vgl. auch Kienle<sup>2</sup> 1969 und Russ 1978). Der Schwerpunkt der epochenbezogenen Grammatiken liegt in der Dokumentation der jeweiligen Zustände; auf systematische Umstrukturierungen, auf mögliche Ursachen oder funktionale Verschiebungen wird kaum eingegangen. Damit geht für das Alt- und Mittelhochdeutsche eine mehr oder weniger konsequente Konzentration auf ein hypothetisiertes “Normalsystem” und damit eine nur ansatzweise Berücksichtigung u. a. der diatopischen Varietäten einher. In der jüngeren und einem nicht mehr junggrammatischen, sondern strukturalistischem Ansatz verpflichteten Frühneuhochdeutschen Grammatik ist der Versuch einer Korpusgrammatik unternommen, bei der die Varietätendifferenzierung zum Prinzip der Darstellung geworden ist; ein Normalsystem wird nicht

mehr konstruiert. Eine nur annähernd entsprechende Darstellung der Wortbildung, die zudem die funktionalen Verschiebungen innerhalb des Inventars der Wortbildungsmittel erfaßt, steht jedoch auch für das Frühneuhochdeutsche noch aus. Hier muß man bis auf weiteres auf schon ältere und zumeist etymologisch orientierte Darstellungen von Willmanns (1906–1909), Paul (¹1920), Kluge (²1925) und Henzen (³1965) zurückgreifen. Aus der Vielzahl einzelner Untersuchungen und epochenbezogener Überblicksdarstellungen sei zusätzlich zu den historischen Grammatiken verwiesen auf die einschlägigen Beiträge in Besch et al. (²1998; ²2000, Hrsg.), Schwarz (1986), Bergmann et al. (1987, Hrsg.), Müller (1993), Habermann (1994), Winkler (1995), Joeres (1995), Prell & Scheben-Schmidt (1996); zur bibliographischen Erschließung vgl. Ronneberger-Sibold (1989).

#### 4. Entwicklungsgeschichtliche Grundzüge

Der geschichtliche Zeit- und Varietätenraum des Deutschen zeigt eine Vielzahl morphologischer Einzelveränderungen: Einige davon profilieren wortartübergreifende und die Wortbildungs- und/oder Flexionsmorphologie insgesamt prägende Prozesse; andere prägen nurmehr die Flexion/Wortbildung einer Wortart; wiederum andere zeigen sich als wesentlich in der **Profilierung** klar(er)er subsystemischer Flexionsstrukturen innerhalb einer Wortart, z. B. nur der sogenannten **starken Verben** oder nur der sogenannten **schwachen Substantive**; andere schließlich fügen sich nur schwer in einen überordnenden Zusammenhang und bleiben als Einzelveränderung singulär. Oberhalb der Vielzahl einzelner Veränderungen ordnet die sprachgeschichtliche Perspektive die Einzelercheinungen zu Entwicklungstendenzen, die sich als historisch konstante (in allen Sprachstufen und bis in die Gegenwart wirkende) oder als inkonstante (historisch nur teilweise aufscheinende) zeigen (vgl. Sonderegger 1979: 241–255, 340–345). Vom neuhighochdeutschen Ausgangspunkt geurteilt werden die Einzelergebnisse zu Indizien einer auf das Neuhighdeutsche ziellenden Teleologie, sie werden entsprechend als “entwicklungsgeschichtliche Grundzüge” bestimmt (Hotzenköcherle 1962). Solche erweisen sich im Ab- sowie dem paradigmatisch verbleibenden Aus- und Umbau des flexivischen Materials:

#### XVIII. Morphologischer Wandel II: Fallstudien

- (a) Als eine konstante und in allen Sprachstufen beobachtbare Entwicklungstendenz erweist die Flexionsgeschichte eine “kontinuierliche Reduktion und Umstrukturierung der Flexionsmorpheme” (Sonderegger 1979: 243); die Umstrukturierung erfolgt paradigmatisch sowie durch Ausweichen auf die syntagmatische Ebene. Indem eine Konzentration der flexivischen Ausdrucksmittel auf bestimmte morphosyntaktische Kategorien erfolgt, werden diese flexivisch profiliert. Wortartbezogen ergibt sich somit bei den Verben eine “Profilierung der Tempuskategorie” (Besch 1980: 594; vgl. Hotzenköcherle 1962: 331), eine “Numeruskristallisierung beim Substantiv” (Hotzenköcherle 1962: 331), eine “Genusprofilierung beim Adjektiv” (Solms & Wegera 1991: 317f.). Bei der Flexion der Adjektive zeigt sich zudem die Durchsetzung einer syntaktisch vom bezogenen Substantiv und vom sonstigen Substantivbegleiter aus organisierten “Formregel” (Hotzenköcherle 1968: 3).
- (b) Strukturelle Konsequenz einer **flexivischen Profilierung** ist eine entsprechende Nivellierung bestimmter Kategorien: insbesondere die des Kasus bei den Substantiven (vgl. Sonderegger 1979: 248f.), insbesondere die des Modus bei den Verben.
- (c) Eine gänzliche oder teilweise Ersetzung von Formativen durch markanter ausgezeichnete Formen tritt bei “funktionaler Untauglichkeit” ein (vgl. Wegera & Prell ²2000: 1594). Eine solche ist primär phonetisch-phonologisch verursacht: vom Althochdeutschen zum Mittelhochdeutschen die in einheitlich -e- aufscheinende Abschwächung der vollen Nebensilbenvokale, vom Mittelhochdeutschen zum Frühneuhochdeutschen eine Fortsetzung der Nebensilbenabschwächung in der Apokope. So geht z. B. in der Wortbildung die althochdeutsche Möglichkeit zur ausdrucksseitig eindeutigen Markierung der **Nomina-agentis-Bildung** durch das Derivationssuffix -o (z. B. kebo, ‘der, der gibt’) verloren, es wird durch das lat. Lehnsuffix -ari (nhd. -er) ersetzt; so wird in der Flexion die vokalisch volle Endungsflexion des Althochdeutschen zerstört, Oppositionen werden aufgehoben (z. B. ahd. Nom. Sg. hirti, Nom. Pl. hirta > mhd. Sg./Pl. hirte). Dadurch kommt es zum Mittelhochdeutschen hin zu eingreifenden morphemischen Strukturveränderungen (vgl. Wolf 1981: 202), die sich in einer paradigmatischen Umstrukturierung des Systems oder auch einer verstärkten Nutzung entsprechender Peripherasen zeigen. Die über das Mittelhochdeutsche hinaus weitergehende Apokope erzeugt bei den Substantiven schließlich den gänzlichen Zusammenbruch des alten Flexionssystems, das im Frühneuhochdeutschen neu aufgebaut wird.
- (d) Die Funktionsverlagerung in die **Periphrase** (Wandel vom synthetischen zum analytischen

- Prinzip) erscheint in Flexion wie Wortbildung. So werden die semantisch-syntaktischen Wortbildungsfunktionen der germanischen Ableitungssuffixe **schwacher Verben** sowohl durch die **Präfigierung** als auch durch entsprechende Peripherasen übernommen: Es erscheinen z. B. statt der im Germanischen möglichen Wortbildung mittels eines Suffixes \*-jan, das ein kausatives Verb deriviert (z. B. gotisch *diupjan* ‘vertiefen’ aus Adjektiv *diups* ‘tief’ + Wortbildungssuffix -jan), seit dem Althochdeutschen sowohl Präfixbildungen z. B. mit *ir-* (*irtrocknen* ‘trocknen machen’) als auch Peripherasen des Basiswortes mit *tuon* oder *machen*. Als dann weitergehende Tendenz zur ‘Periphrase in der Morphemstruktur’ zeigt sich die Funktionsablösung von Simplicia durch Präfixverben seit dem Mittelhochdeutschen (z. B. *jmdn. weinen* > *jmdn. beweinem*). Bei der Flexion weist die Tendenz zu periphrastischen Umschreibungen auf die kontinuierliche Veränderung innerhalb der im Deutschen genutzten drei **Flexionsprinzipien**: die aus dem Vordeutschen übernommene “**regressive**” (von der Endung her erfolgende Steuerung mittels **Relationsmorphen**, z. B. 3. Sg. Ind. Präs. *suo-ch-it* vs. 3. Sg. Konj. *suo-ch-e* im Althochdeutschen), die ebenfalls vordeutsche Stamm- oder Wurzelflexion (untrennbarer Einheiten lexikalischer und grammatischer Bedeutung, besonders der **Vokalwechsel** “**Umlaut**” und “**Ablaut**; Ablaut als morpholexische Erscheinung, vgl. Durrell 1980:23; ähnlich Simmler 1987:121 f.; z. B. 1. Pl. Präs. *geben* vs. 1. Pl. Prät. *gäben* im Mittelhochdeutschen), die “**progressive**” Steuerung durch vorangestellten Begleiter (z. B. 1. Pl. Ind. Prät. *wir gäbun* vs. 3. Pl. Ind. Prät. *sie gäbun* im Althochdeutschen). Ein jedes dieser Prinzipien erscheint nur selten ausschließlich, zumeist kommen sie in Kombination miteinander vor (z. B. ahd. Nom. Sg. *gast* vs. Pl. *ge-sti*; nhd. Nom./Akk. Sg. *das Rad* vs. Pl. *die Räder*). Die deutsche Sprachgeschichte zeigt eine jeweils unterschiedliche und die einzelnen Sprachstufen charakterisierende Nutzung dieser genannten drei Prinzipien: vorwiegend “**regressive**” oder Endungsflexion im Althochdeutschen, vermehrter Ausbau von Wurzelflexion und progressiver Steuerung im Mittelhochdeutschen, entscheidende Verlagerung im Neuhochdeutschen zur progressiven Steuerung (vgl. Sonderegger 1979: 243). Peripherasen werden vor allem bei den Verben für Perfekt, Futur, Passiv und auch Konjunktiv genutzt; sie finden sich besonders seit dem 15. Jahrhundert mit einem Schwerpunkt im oberdeutschen Sprachraum.
- (e) Der tendenzielle Ausbau der Wurzelflexion mittels Wechsel des Wurzelvokals zeigt die **Morphologisierung** und dann auch Systematisierung phonetisch-phonologisch bedingter und schon im Voralthochdeutschen eingetretener **Alternationen** (i. e. regelmäßiger und syn-

chronisch funktionierender **Lautwechsel** in etymologisch verwandten Wörtern, hier besonders Umlaut) bei den Verben, den Substantiven sowie auch den Adjektiven (Komparation). Der Vokalwechsel wird besonders zum Mittel der Flexionsmorphologie entwickelt, weniger zu einem der Wortbildung.

- (f) Die phonetisch-phonologische Verursachung überlagernd, verstärkend und ersetzend sind weitere Entwicklungsmotive abzuleiten: so ein Streben nach Ökonomie, Redundanzsteuerung (vgl. Koenraads 1953; Lüdtke 1988: 1633). Dazu gehört, daß isolierte/lexemspezifische Besonderheiten innerhalb einer Gruppe/Klasse von Lexemen oder einer Wortart (nicht zwingend) abgebaut werden: der **grammatische Wechsel** bei der Flexion der Verben (z. B. Sg./Pl. Prät. *was* – *wären* > *war* – *waren*; vgl. Sonderegger 1979: 253), **suppletive Formen** bei den Pronomen (*dirre* > *dieser*; vgl. Ebert et al. 1993: § M60), die **rückumlautende Verbflexion** (*stellen* – *stalite* > *stellen* – *stellte*), die starke Verbflexion sowie ein Teil der ehemaligen **Präteritopräsentien**. Tendenzielles Ergebnis eines Abbaus von “Besonderheiten” ist die weitere Profilierung und Bestätigung einer “Normalität”. Ein markantes Beispiel für den Abbau systemischer und lexemdistribuierter Varianz ist die Herausbildung der ehedem schwachen und somit spezifischen Verbalflexion zur allgemeinen und somit “Normalflexion” des Verbs im Deutschen.
- (g) In der Profilierung spezifischer Flexionsweisen erweist sich auch der besondere Einfluß, den die Erweiterung des Wortschatzes durch Entlehnung oder Wortbildung auf die Entwicklung des (geschlossenen) Flexionssystems hat. Denn jedem neu in den Wortschatz aufgenommenen Wort wird unmittelbar eine der möglichen Flexionsweisen zugeordnet. In diesem konkreten Sprachverhalten erweist sich die Produktivität (Analogie) bestimmter Muster (Art. 148). Insofern solche Muster somit durch zusätzliche Wortschatzelemente gestützt werden, kann der Sog dieser Muster so weit verstärkt werden, daß variante Muster ihre Wirkung verlieren. Die ehemals schwachen Verben sind auch hier ein Beispiel; mit ganz wenigen Ausnahmen (z. B. *schreiben* im Althochdeutschen oder *preisen* im Mittelhochdeutschen) sind alle entlehnten Verben automatisch als schwache Verben in den Wortschatz integriert worden; auch dadurch ist der Wortschatzanteil starker Verben relativ gesunken.
- (h) Eine “Tendenz zur formalen Trennung von Flexions- und Derivationsmorphemen” (Wegera & Prell 2000: 1594) und damit auch eine Tendenz zur Trennung von Flexion und Wortbildung zeigt sich sowohl hinsichtlich bestimmter Prinzipien morphologischer Informationsbindung als auch hinsichtlich der funktionalen Trennung der Relationsmorpheme. Der in der vordeutschen Wortbildung wesent-

liche Vokalwechsel (z. B. Ablautbildungen *Trunk* oder *Trank* zu *trinken*, *Fluss* oder *Flotte* zu *fließen*) wird seit dem Althochdeutschen primär flexivisch genutzt (Ausbau der Umlautnutzung bei den Substantiven). Bis zum Neu-hochdeutschen hin verläuft ein Prozeß der deutlichen Trennung einzelner Flexions- und Wortbildungsmorphe (vgl. Erben<sup>2</sup>1983: 138): z. B. ist das in der Wortbildung ehedem produktive und inzwischen nurmehr aktive Wortbildungssuffix *-e* (Femininabstrakta z. B. *Rinne*) nurmehr als Flexiv produktiv (dient insbesondere der Markierung des Plurals bei den Substantiven); z. B. hat sich das Präfix *ge-*, das bis in das Mittelhochdeutsche hinein aspektuelle Differenzierungen ermöglichte (z. B. 3. Sg. Ind. Präs. *er geas* mit perfektiver Bedeutung ‘er hatte gegessen’) und daneben als ein die Verbseman-tik modifizierendes Präfix (“Aktionsartenaus-fächerung”) fungionierte (z. B. Präfixverb *ge-heilen* ‘gesund werden’ zum Simplexverb *heilen* ‘gesund machen’), im Verlauf des Frühneu-hochdeutschen nurmehr zum Bestandteil des diskontinuierlichen Part. Prät.-Morphems her-ausgebildet (vgl. Solms 1991).

- (i) In der morphologischen Profilierung und funktionalen Bindung von Flexiven und Formativen zeigt sich eine “Tendenz zur deutlichen strukturellen Trennung der Wortarten” (Wegera & Prell<sup>2</sup>2000: 1594), “unverkennbar” zwischen Substantiv und Adjektiv (Erben<sup>2</sup>1983: 138).

## 5. Flexion

### 5.1. Flexion der Verben

Die diachron motivierte Klassifizierung der deut-schen Verben unterscheidet neben den schon im Germanischen ausdifferenzierten Hauptklassen der “starken” und “schwachen” Verben eine weitere Gruppe flexivisch heterogener **besonderer Verben**. Kriterium der Unterscheidung der zwei Hauptklassen ist die ohne oder mit einem Dentalsuffix *-t-* er-folgende Flexion des Präteritums (schwach: ahd. *suothen* – *suocta*; stark: ahd. *grifan* – *graiflgrifen*). Daneben werden als ebenfalls eigenständige Gruppe die sogenannten “Präteritopräsentien” (Prät. Präs.; z. B. mhd. *wizzen* oder *gunnen*) unter-schieden, ursprünglich entstanden aus dem indo-germanischen Perfekt, der teils Vergangenheit und teils Gegenwart bezeichnete (aus letzteren Fällen entstehen Präteritopräsentien, aus ersten “nor-male” starke Verben; vgl. Hirt 1932: § 115). Ent-sprechend zeigen die Präteritopräsentien im Präsens (ablautende) Präteritalformen, zu denen dann schwache Präteritalformen neu gebildet sind. In ei-ner synchronen Klassifizierung gilt die Flexion der Präteritopräsentien schon seit dem Althochdeut-schen als Besonderheit nur weniger Lexeme. Dies rechtfertigt ihre Gruppierung zusammen mit einer nur kleinen und hinsichtlich ihres Inventars nicht konstanten Restgruppe auffälliger und individuell

## XVIII. Morphologischer Wandel II: Fallstudien

flektierter Verben: ahd. *uuellen* ‘wollen’, *tuon* ‘tun’, *sîn* ‘sein’, *gân/gênl/gangan* ‘gehen’ und *stân/stênl/stantan* ‘stehen’. “Restgruppe” wie “Präteritoprä-sentien” werden in den historischen Grammatiken zumeist in einer nicht für jede Sprachperiode iden-tischen Gruppe “besonderer Verben” gefaßt (vgl. Braune<sup>14</sup>1987: § 370–385; Paul<sup>24</sup>1998: § 267–288): ihre Flexion ist als jeweils einzellexematische Kom-bination verschiedener Stamm- (teilweise supple-tiv) wie auch Relationsmorphie zu beschreiben.

Die vom Alt- bis zum Ausgang des Frühneu-hochdeutschen strukturell wirksamen Veränderungen zeigen sich teilweise als Fortführun-gen schon im Germanischen begonnener Tendenzen, zum Teil als später einsetzende Entwicklungen; die Veränderungen bleiben partiell subsystemisch, z. B. nur auf die star-ken Verben begrenzt, partiell führen sie in ihrer Gesamtheit zur Veränderung des gesamten verbalen Flexionssystems: zur Herausbil-dung der schwachen von einer germanisch besonderen zu der neu-hochdeutsch regelmäßi-geen Normalflexion. In diesen Zusam-menhang gehören Angleichungen innerhalb der schwachen Flexionsklassen, Angleichungen der Endungsflexivik der starken an die der schwachen Flexion, der Wechsel zahlreicher starker zu den schwachen Verben, die Über-nahme der schwachen Flexionsweise auch bei einigen besonderen Verben sowie auch die automatische Übernahme der schwachen Flexionsweise bei neu in den Wortschatz auf-genommenen Verben. Die Entwicklung der schwachen zur neu-hochdeutsch “regelmäßi-geen” oder “Normalflexion” bedeutet für den im Frühneu-hochdeutschen vollzogenen Pro-zeß bei den starken Verben, daß es sie als ei-genständige Klasse spätestens seit der Auf-klärung nicht mehr gibt (vgl. Augst 1975: 263); sie haben als Ausnahmen zu gel-ten. Die noch alt-/mittelhochdeutsche For-menvielfalt der besonderen Verben wird im Frühneu-hochdeutschen reduziert, in einigen Fällen wird auch die schwache Flexion durchgesetzt (bei mhd. *tugen*, *gunnen*, *sollen*), teilweise werden die Besonderheiten in ande-rer Weise aufgegeben (*beginnen* ist nurmehr stark). Besonders die Veränderungen bei den besonderen Verben sowie auch die Reduktion des “Grammatischen Wechsels” und stärkere Systematisierung des Ablauts bei den starken Verben zeigt eine (partielle) Aufhebung von Suppletion zugunsten geregelter Alternatio-nen; hier liegt ein als Stamm- oder **morpholo-gisches Prinzip** bezeichneter Prozeß zur ein-heitlichen Wurzel-/Stammstruktur in allen Flexionsformen vor.

Die Herausbildung einer schwachen Verbflexion ist eine das Germanische vom Indo-germanischen typologisch unterscheidende Neuerung (vgl. Schmidt 2<sup>1998</sup>: 995 f.). Das schwache (endungsflexivische) Prinzip wird zum Merkmal der zumeist derivierten Verben; das starke (wurzelflexivische) Prinzip des Ablauts (seltener auch der Reduplikation) erweist sich als das zunehmend abgebaute Prinzip des zumeist schon dem Indo-germanischen angehörenden Grundbestandes an Verben. Die (Wort)Bildung der neuen und schwach flektierten Verben des Germanischen erfolgt suffixal: Die ausdrucksseitig unterschiedenen **Stämme** (durch die stammbildenden Suffixe \*-ja, \*-ō, \*-ē, \*-na) definieren im Germanischen vier Klassen schwacher Verben. Dabei ist jedoch schon für das Althochdeutsche fraglich, inwieweit dieses “Stammelement” noch “als Kennzeichen einer paradigmatischen Struktur” (Simmler 1987: 135) segmentierbar ist. Wenn für das Althochdeutsche trotzdem noch schwache **Stammklassen** (vgl. Sonderegger 1979; <sup>2</sup>2000; Braune <sup>14</sup>1987) unterschieden werden, dann zeigt sich das deliminative Kriterium (Übereinstimmung in der phonologischen Substanz des Auslauts, d.h. des ursprünglichen Stammelements in allen Wortformen) jedoch nur noch teilweise; die übliche Klasseneinteilung des Althochdeutschen ist somit vorwiegend historisch und nicht synchron motiviert. Es werden unterschieden: die -ien/-en-Verben (< germ. \*-jan) als schwache Verben der Klasse I (z.B. deverbal-kausativ: *se(z)zen* zu *si(z)zen*), die -ōn-Verben als Klasse II (z.B. denominal-instrumental/ornativ: *salbōn* zu *salba* ‘Salbe’), die -ēn-Verben (unter Einschluß der germanischen \*-nan-Verben) als Klasse III (z.B. denominal-ingressiv/inchoativ: *altēn* ‘alt werden’). Dabei erweisen einzelne Oppositionen das noch teilweise Funktionieren der alten Stammklassenelemente (vgl. Wolf 1981: 200): ahd. *rōten* ‘rot machen’ (< \*rōt-jan) neben *rōtēn* ‘rot werden’ (< \*rot-ē-n); darüber hinaus zeigen sich die einzelnen Klassen als mehr oder weniger produktiv: Alle lateinischen Lehnwörter schließen sich der Klasse II an (vgl. Russ 1987: 336). Möglicherweise begründet hier die hohe Type-Anzahl dieser Verben ihre historische Produktivität: Von den ca. 1798 althochdeutsch schwachen Verben (vgl. Russ 1987: 336) gehören zu Klasse I 860 Lexeme, zu Klasse II 660 Lexeme, zu Klasse III nur 278 Lexeme. Formalmorphologisch scheinen die ursprünglichen Ableitungs- oder Stammsuffixe

im Althochdeutschen eindeutig: -i-, -ō-, -ē gelten als jeweiliger Bindevokal (vgl. Braune <sup>14</sup>1987: § 306) zwischen der nur historisch/etymologisch zu fassenden Wurzel und dem jeweiligen Endungsflexiv. Diese Wertung enthält jedoch ein Problem, da -ō/-ē in allen Formen des jeweiligen Verbs erscheinen; sie erweisen sich als Bestandteile eines paradigmatisch nicht mehr in Wurzel oder Stamm (bzw. Stammelement oder Flexiv) segmentierbaren Morphs. Gegenüber dem Germanischen wird die Klassifizierung im Althochdeutschen verkompliziert, da aufgrund phonologisch zu fassender Veränderungen bei einem Teil der Verben aus Klasse I der “Bindevokal” (-i-) innerhalb des Präteritums und des flektierten Part. Prät. zum Althochdeutschen hin ausgesunken ist; als Folge daraus (nicht eingetretener Umlaut) hat sich der einige schwache Verben bis heute charakterisierende “Rückumlaut” (*brennen* vs. *brannt*) herausgebildet. Aus der (sekundären) Berücksichtigung dieser **apophonischen** Varianten (mit Vokalwechsel) sowie in erster Linie der historisch/etymologisch bestimmten Stammklassen und der ihnen entsprechenden Bindevokale (im Präteritum und flektierten Part. Prät.) wird für das Althochdeutsche ein Fünfklassensystem der schwachen Verben definiert (vgl. Sonderegger 1985: 1023), die sich jedoch endungsflexivisch, von wenigen Ausnahmen abgesehen, identisch verhalten:

- I.1: mit Vokalwechsel/ohne Bindevokal (*brennen*, *branta*);
- I.2: ohne Vokalwechsel/ohne Bindevokal (*hören*, *hōrtā*);
- I.3: ohne Vokalwechsel/mit Bindevokal -i- (*legen*, *legita*);
- II.: ohne Vokalwechsel/mit Bindevokal -ō- (*machōn*, *machōta*);
- III: ohne Vokalwechsel/mit Bindevokal -ē- (*tagēn*, *tagēta*).

Die zum Mittelhochdeutschen eintretende Abschwächung der vollen Nebensilbenvokale verändert die Klassifizierung; das -e- der Nebensilbe wird als Bestandteil des Flexivs gewertet. Das Vorhandensein/Fehlen dieses Bindevokals im Präteritum wird zum zentralen Klassifizierungskriterium (vgl. Paul <sup>24</sup>1998: § 265):

- I.1: mit Vokalwechsel/ohne Bindevokal (*brennen*, *branta*);
- I.2: ohne Vokalwechsel/ohne Bindevokal (*teilen*, *teilte*);
- II: ohne Vokalwechsel/mit Bindevokal (*legen*, *legete*).

Da ein ausdrucksseitiges Kriterium für den Gebrauch des *-e-* nicht vorliegt (mit Ausnahme der Kopplung an den Vokalwechsel), ist das mittelhochdeutsche Inventar beider Klassen nur noch partiell durch die althochdeutsche Zugehörigkeit bestimmt. Dies ergibt sich allein schon durch das immense Anwachsen des schwach flektierten Verbwortschatzes (durch Wortbildung, Entlehnung) auf mehr als 5000 Verben (vgl. Sonderegger 1979: 258). Doch auch das jeweils ererbte Flexionsverhalten der einzelnen Lexeme ist aufgrund mehrerer Ursachen nicht klassenstabil. Zwar wird als maßgeblich für das mittelhochdeutsche Vorhandensein des Bindevokals die Quantität der Silbe genannt (vgl. Hempen 1988: 254), doch wird unabhängig davon *-e-* „vielfach“ synkopiert (Paul<sup>24</sup>1998: § 265). Dies führt zur Vermischung der althochdeutsch getrennten Klassen I.2. und II. Auch das Inventar der apophonisch funktionierenden Lexeme stimmt aufgrund der Wirkungen des „Sekundärumlauts“ sowie „anorganisch“ eintretender Rückumlauten (z. B. *kēren – karte* oder *enden – ante*) nicht mehr mit dem der althochdeutschen Gruppe überein (vgl. Paul<sup>24</sup>1998: § 259–266). Obwohl die Gruppe der Verben mit einer apphonischen Wurzelflexion im Mittelhochdeutschen stark anwächst (über 200), nimmt ihr relativer Anteil am Gesamt aller schwachen Verben stark ab. Dieser Prozeß wird im Frühneuhochdeutschen noch wesentlich verstärkt. Es verschmelzen während des Frühneuhochdeutschen auch die im Mittelhochdeutschen noch hinsichtlich des Bindevokals unterschiedenen Klassen. Die Entwicklung verläuft wiederum über eine Vermischung der historischen Gruppen, so z. B. *teilete* (mittelhochdeutsch ohne Bindevokal), aber *legte* (mittelhochdeutsch mit Bindevokal). So erweist sich schon zu Beginn des Frühneuhochdeutschen der Gebrauch des Nebensilben-*e* nurmehr deutlich phonologisch distribuiert. Trotz eines deutlichen Anstiegs in der Verwendung des *-e-* vor allem im Oberdeutschen des 16. Jahrhunderts ist die neuhochdeutsch gültige Regelung eines beibehaltenen *-e-* nur nach Dental in der zweiten Hälfte des 17. Jahrhunderts usuell (vgl. Ebert et al. 1993: § M90.2.); es gibt nurmehr eine Klasse schwacher Verben (mit nach Dental eintretendem Bindevokal *-e-*). Die (neuhochdeutschen) Ausnahmen innerhalb der schwachen Flexion bestimmen sich durch den als Rückumlaut bezeichneten Vokalwechsel z. B. *brenn-* (Inf./Präs./Part. Präs.) zu *brann-*

(Prät./Part. Prät.). Im ausgehenden Frühneuhochdeutschen funktionieren gemäß diesem Muster zusätzlich zu den neuhochdeutsch erhaltenen (*denken* und synchron analog auch *bringen, brennen, nennen, rennen, senden, wenden*; daneben einige semantisch isolierte und erstarre Formen, z. B. *abgeschmackt* zu *schmecken, behaftet* zu *heften, getrost* zu *trösten*) auch noch z. B. *legen, trennen, schenken, schätzen, setzen, zerren* (vgl. Hoffmann & Solms 1987: 53 f.). Innerhalb des **Ausgleichsprozesses** zugunsten der stamminvarianten Flexion ist in nur wenigen Fällen ein Ausgleich auch zugunsten der (rückumlautenden) Prät./Part. Prät.-Form eingetreten (z. B. mhd. *bestellen* – *bestalte* > nhd. Inf. *bestallen* neben *bestellen*). Das Gesamtinventar der mittelhochdeutsch rückumlautenden Verben umfaßt noch zum mindesten 221 Lexeme (vgl. Weinhold<sup>25</sup>1883: § 383). Die Ausgleichung findet dann erst im Frühneuhochdeutschen statt, begünstigt durch die schon im älteren Mittelhochdeutschen und in der ersten Hälfte des Frühneuhochdeutschen gültige sprachlandeskundliche Differenzierung (vgl. Stärck 1912: 308): *brennen – brante – gebrant* im Mitteldeutschen versus *prennen – prant – geprennet* im Oberdeutschen. Insbesondere vom Oberdeutschen ausgehend („oberdeutscher Präteritumschwund“, d. h. Rückgang des Gebrauchs finiter Prät.-Formen und Zunahme der periphrastischen Tempusform) greift im Verlauf des Frühneuhochdeutschen der Ausgleich zugunsten der Form des Inf./Präs. (und oberdeutsch auch Part. Prät.). Die Entwicklung zeigt sich als Rückbau einer flexivischen „Übercharakterisierung“ (durch Dentalsuffix und zusätzlichen Vokalwechsel) des Präteritums (Schirmunski 1962: 501) zugunsten einer als morphologisches Prinzip bezeichneten Stabilität des Stammes in allen Positionen des Paradigmas.

Das im Althochdeutschen paradigmatisch wie regional noch differenzierte System der Flexionssuffixe erfährt bis zum ausgehenden Frühneuhochdeutschen eine Vereinfachung, die eine Tilgung teils paradigmatisch und teils auch regional isolierter Flexive darstellt. Die differenzierte Flexivik des Althochdeutschen bestimmt jedoch keine ausgeprägten klassenbezogenen und/oder morphosyntaktisch bezogenen Paradigmen: So liegen einzelne klassenbezogene Unterschiede z. B. bei der 1. Sg. Ind. Präs. vor (starke sowie schwache Verben der Klasse I mit *-u*, schwache Verben der Klassen II/III mit *-m*); klassenindifferente Unterschiede hinsichtlich der morphosyntak-

tischen Merkmale liegen z. B. bei der Modusdifferenzierung der 1. Sg. Präs. vor (Ind. *-u/-m* zu Konj. *-e*); klassendifferente Unterschiede hinsichtlich der morphosyntaktischen Merkmale liegen z. B. vor bei der 2. Pl. Präs. (Variation *-etl/-ent* bei starken sowie schwachen Verben der Klasse I, *-t* bei schwachen Verben der Klassen II/III). Dabei zeigt schon das Althochdeutsche klassen- wie merkmalsübergreifenden Ausgleich, so z. B. von der Konjunktivform schwacher Verben ausgehende Verallgemeinerung von *-st* in der 2. Sg. (beibehaltene Besonderheit: *-i* im Präteritum starker Verben). Die Entwicklung zum Mittelhochdeutschen ist einerseits geprägt durch den Zusammenfall der Nebensilbenvokale in *-e* (Aufhebung der ausdrucksseitigen Oppositionen z. B. 1. Sg. Ind. Präs. *-u* und 1. Sg. Konj. Präs. *-e*); andererseits kommt es zur An- und Ausgleichung innerhalb der wenigen eine Variation aufweisenden morphosyntaktischen Merkmale (z. B. 1. Sg. Ind. Präs.): So weit es sich um eine auch klassendifferente Variation handelt (1. Pl. *-mēs, -en*), ist die Variation schon zum Ende des Althochdeutschen aufgegeben; dort, wo die Variation zugleich eine Klassendifferenzierung berührt, verläuft die Entwicklung bis in das Frühneuhochdeutsche hinein (z. B. flexivische Besonderheit der 2. Sg. Ind. Prät. bei den starken Verben). Bis auf wenige Positionen (3. Pl. Ind. Präs. *-ent* zu 3. Pl. Konj. Präs. *-en* sowie bis in Neuhochdeutsche erhaltener Unterschied bei 1./3. Sg. Ind. Prät. starker/schwacher Verben: z. B. *gab-* versus *sagt-e*) erscheint die Endungsflexivik im Normalmittelhochdeutschen sowohl zwischen starken und schwachen Verben als auch hinsichtlich Präs./Prät. oder Ind./Konj. weitgehend ununterscheidbar (vgl. Paul<sup>24</sup>1998: § 239, 257). Eine notwendige Berücksichtigung der Varietäten in mittelhochdeutscher Zeit erweist die Verhältnisse jedoch als nicht ganz so eindeutig (z. B. erscheint Flexiv-*n* der 1. Sg. Ind. Präs. besonders westoberdeutsch und -mitteldeutsch klassenunabhängig und schwindet weitgehend erst im 16. Jahrhundert; vgl. Paul<sup>24</sup>1998: § 367, 395; Ebert et al. 1993: § M88). Im Verlauf des Frühneuhochdeutschen werden durch Ausgleicherscheinungen die schon im Mittelhochdeutschen (paradigmatisch oder regional) isolierten Endungsflexive aufgegeben. Gänzlich erfolgt dies hinsichtlich der Flexion der 2. Sg. Ind. Prät. starker Verben durch einerseits *-e* und andererseits die (umgelautete) Wortform des Prät. Pl. (*gaeb-e* > *gab-st*); nurmehr zeitlich und regional be-

grenzt hinsichtlich der 1./3. Sg. Ind. Prät. der starken Verben, die insbesondere in der zweiten Hälfte des Frühneuhochdeutschen oft mit Relationsmorph *-e* erscheint ("lutherisches *-e*"; vgl. Hoffmann & Solms 1987: 42–46). Die Flexion der 3. Pl. Ind. Prät. (mittelhochdeutsch *-ent*) ist überlagert von einer bis in das 17. Jahrhundert reichenden Entwicklung "mit sprachlandschaftsinterner und teilweise erheblicher zwischenlandschaftlicher Variation", durch die spezifisch landschaftliche Paradigmen herausgebildet werden (Ebert et al. 1993: § M74; Tendenz z. B. zum Einheitsplural *-et*).

Schon für das Althochdeutsche zeigt sich die definierende Eigenschaft des starken Verbs, seine "Tempusstämme nur durch den Vokalwechsel zu unterscheiden" (Braune<sup>14</sup>1987: § 324), als Ergebnis einer vorausliegenden Entwicklung. Während die Unterscheidung im Indogermanischen und partiell noch im Germanischen auch aufgrund konsonantischer Variationen des Stammes, suffizialer und auch präfixaler Elemente erfolgt, sind solche Möglichkeiten im Althochdeutschen weitgehend geschwunden. Entsprechend den vier überkommenen Ablautstufen werden für den Großteil starker Verben vier primäre Tempusstämme ("Stammformen", bei Berücksichtigung der indogermanischen/germanischen Sprachgeschichte ist statt vom Stamm besser von "Wurzel" zu sprechen) unterschieden: z. B. *rītan* (Präs./Inf.) – *reit* (1./3. Sg. Ind. Prät.) – *riten/riti* (1. Pl. und 2. Sg. Ind. Prät.) – *giritan* (Part. Prät.). Innerhalb des Präteritums der Klassen I–V dient der qualitative/quantitative Ablaut zusätzlich auch als Numerusablaut; Klasse VI und VII (ehemals **reduplizierende Verben**) haben im Singular und Plural einen jeweils identischen Stamm (z. B. VI: ahd. *wahsan* – *wuohs* – *wuohsun/wuohsi* – *giwahsan*). Die Divergenz zwischen 4 Ablaut-, aber nur 3 Zeitstufen stellt ein funktionales Mißverhältnis dar (vgl. Augst 1975: 251), das über das Althochdeutsche hinaus bis in das Frühneuhochdeutsche vorliegt. Eine Entwicklung setzt erst im Frühneuhochdeutschen ein, es werden die jeweiligen Numerusablauten weitgehend seit der zweiten Hälfte des 15. Jahrhunderts zumeist zugunsten einer der beiden Tempus-Alternanten ausgeglichen: z. B. ahd./mhd. (Sg./Pl.) *reit/riten* wird nhd. *ritt/ritten*, ahd./mhd. *fand/funden* wird nhd. *fand/fanden*. Der Ablaut hat nun erst zum Neuhochdeutschen hin "seine wahrhaft funktionelle Geltung erlangt" (Paul<sup>5</sup>1920: 211): Die morphologische Kategorie

des Tempus erscheint flexivisch profiliert. Eine über den Ablautausgleich bestimmte Periodengrenze des Frühneuhochdeutschen zum Neuhochdeutschen hin ist nicht eindeutig zu ziehen, da der konkrete Ausgleichsprozeß durch mehrere chronologisch versetzt wirksame Prinzipien gesteuert ist. Während der Ablautausgleich bei den starken Verben der Klassen I (z. B. *rīten*) und II (z. B. *bieten*) schon zum Ende des 17. Jahrhunderts lange abgeschlossen ist, bleibt die Entwicklung bei Klasse III und hier vor allem bei Verben mit Nasalverbindung nach dem Wurzelvokal (Subklasse IIIa, z. B. *finden*) zum Ende des 17. Jahrhunderts noch unabgeschlossen. Für den raschen Ausgleich in Klasse I und II wird eine strukturelle Disposition darin gesehen, daß der Stammvokal des Partizip Präteritums mit jenem des Singular oder Plural Präteritums übereinstimmte. Hier liegt ein **intraparadigmatischer Ausgleich** vor, der ohne Einfluß seitens einer anderen Ablautmöglichkeit eintrat (vgl. Behaghel 1928: §436). Eine identische Disposition weisen auch die Verben der Klasse IIIa auf (mhd. *finden* – *fand* – *fund* – *gefunden*). Ihr Ablautausgleich ist jedoch zum Ende des Frühneuhochdeutschen nicht nur nicht entschieden, sondern die Richtung des im Neuhochdeutschen letztlich erreichten Ausgleichs (*fand* – *fanden* statt *\*fund* – *\*funden*) widerspricht vielmehr auch der Annahme einer wirksamen und das Part. Prät. einschließenden strukturellen Disposition: Die zugunsten des Sg. Prät.-*a* ausgleichende Entwicklung ist als eine interparadigmatische und an das Ablautmuster der starken Verben der Klasse IV/V (z. B. *nēmen/gēben*) angelehnte Ausgleichung zu werten. Das Ergebnis ist jedoch weniger das Ergebnis eines unbewußt vollzogenen Wandels, als vielmehr des bewußten und die Entwicklung beeinflussenden Eingreifens der normierenden und einem Konzept der Analogie verpflichteten Grammatiker des 18. Jahrhunderts (Gottsched und Adelung; vgl. Solms 1993). Im Rahmen dieses schon intentionalen Prozesses werden weitere, im fröhneuhochdeutscher Zeit begonnene und das Flexions- und Klassifikationssystem der starken Verben verändernde Entwicklungen weiter forciert. So kommt es insbesondere zum ausgehenden Frühneuhochdeutschen zum internen **Klassenwechsel** einiger starker Verben; es erscheinen dominante Ablautmuster herausgebildet (**interparadigmatischer Ausgleich**), wodurch die noch althochdeutsche Systematik der starken Verben endgültig aufgehoben ist.

Über den Ablautausgleich hinaus findet in einigen Fällen ein Ausgleich zugunsten der -o- aufweisenden Stammform des Part. Prät. statt, z. B. bei *melken*, *fechten*, *glimmen*. Hier erscheint die Reduzierung und Konzentration auf nurmehr eine Stammform für die Tempusinformation „Vergangenheit“ (sei es finit oder periphrastisch mittels des Part. Prät.) konsequent fortgesetzt; vor dem Hintergrund des im Oberdeutschen seit dem 15. Jahrhundert beobachteten „Präteritumswundes“ zugunsten der Peripherase ist die intraparadigmatische Orientierung am Part. Prät. nur die konsequente Fortsetzung der flexivischen Profilierung des Tempus. Darüber hinaus verlieren einige weitere Lexeme ihren ursprünglichen und partiell bis über das 18. Jahrhundert erhaltenen Ablaut und übernehmen in einem interparadigmatischen Ausgleich ebenfalls den Stammvokal -o- (z. B. mhd. *heben* – *huob* – *geheben* > nhd. *heben* – *hob* – *gehoben*). Zum ausgehenden Frühneuhochdeutschen werden somit dominante Ablautungen herausgebildet, die ihrerseits zum Zielpunkt von Veränderungen werden: Die dominanten Ablautungen mit -i/e- (z. B. *schreiben* – *schrieb* – *geschrieben*) oder -o- umfassen zum Neuhochdeutschen hin ca. 51 Prozent aller starken Verben (vergleichsweise mittelhochdeutscher Anteil ca. 31 Prozent; vgl. Solms 1984: 71 f.).

Eine Vereinfachung sowohl hinsichtlich der Anzahl flexivischer Muster als auch hinsichtlich der Verallgemeinerung des einfachsten Musters zeigt die Stammalternation in der Präsensflexion der starken Verben. Die aufgrund vorgängig phonologischer Entwicklungen zum Althochdeutschen hin entstandenen drei Muster „sekundärer“ und d. h. nicht durch Ablaut erzeugter Stämme (ohne Alternation; Vokalwechsel im Ind. Sg., z. B. Inf. *wērdan* mit 1. Sg. Ind. *wirdū*, 2. Sg. Ind. *wirdis*, 3. Sg. Ind. *wirdit*; Vokalwechsel nur in 2./3. Ind. Sg., z. B. *varan* mit *varu*, *veris*, *verit*) wird im Verlauf des Frühneuhochdeutschen auf nurmehr zwei Muster reduziert (ohne Alternation; Vokalwechsel in 2./3. Ind. Sg, z. B. auch *werden* mit 1. Sg. Ind. *werde*, *wirdest*, *wirdet*).

Während die Anzahl starker Verben vom Althochdeutschen (349) bis zum Mittelhochdeutschen (339) weitgehend gleich bleibt, ist zum Neuhochdeutschen (169) hin ein deutlicher Schwund eingetreten (vgl. Solms 1984: 23, Anm. 33): starke Verben gehen zum Neuhochdeutschen hin unter (z. B. *biten*) oder nehmen schwache Flexion an (z. B. *gīgen*; vgl.

Solms 1984: 27 f.). Die Übernahme der schwachen Flexion zeigt sich strukturell disponiert bei Identität des Stammes von Inf./Präs. mit dem des Part. Prät., so z. B. bei Verben der Klasse VII wie *valten* – *vield* – *gevalten* (*falten* – *faltete* – *gefaltet*) (vgl. Schirmunski 1962: 504; Solms 1984: 323 f.).

## 5.2. Flexion der Substantive

Die diachron motivierte Klassifizierung der deutschen Substantiv unterscheidet seit dem Althochdeutschen „starke“ (differenzierte flexivische Realisierung der morphosyntaktischen Kategorien) und „schwache“ (flexivisch weniger differenziert, im Neuhochdeutschen nur -en in allen obliquen Kasus) Substantiv. Die Unterscheidung überlagert die für das Germanische über Stammsuffixe definierte Stammklassengliederung, deren Gültigkeit das Inventar der althochdeutsch vollvokalisch erhaltenen Endungsflexive sowohl am partiell markierten Nom. Sg. (z. B. kann Endung -o nur mask. n-Stamm sein, z. B. *hano*) als auch über zahlreiche paradigmatische Oppositionen (z. B. Instr. Sg. *tagu* zu Gen. Pl. *tago* beim mask. a-Stamm) noch erweist.

Bereits im Althochdeutschen sind genusbezogene Hauptklassen profiliert: die „schwache“ n-Deklination (alle Genera), die „starke“ a-(Maskulinum, Neutrum), ô- (Femininum) und i-Deklination (Maskulinum, Femininum); neben den Hauptreihen existieren einige kaum repräsentierte Restklassen, deren Inventar zumeist bereits von den Hauptreihen aufgesogen ist (besonders -iz/-az-Stämme in die a-Deklination; vgl. Wolf 1981: 197). Über die flexivische Angleichung verschiedener Klassen ist im Althochdeutschen gegenüber dem Germanischen zugleich die Tendenz deutlich, flexivische Kasusunterscheidungen einzubauen, z. B. in der weitgehenden Aufhebung der Distinktion Nom. versus Akk. Sg. bei den germanischen a-Stämmen (gotisch Nom. Sg. *dag-s* und Akk. Sg. *dag*, ahd. Nom./Akk. *tag*; vgl. Hirt 1932: § 38), z. B. auch in der weitgehenden Aufgabe des Instrumentals. In der Folge erweisen die weiteren flexivischen Entwicklungen eine vorwiegende **Numerusprofilierung** bei gleichzeitiger **Kasusnivellierung**; dabei wird die Wurzelflexion aus-, die Möglichkeit der regressiven Flexion stark abgebaut. Die Flexion bleibt im Althochdeutschen noch weitgehend regressiv gesteuert, der (Primär)umlaut erscheint lautgesetzlich verteilt und somit stammklassenbezogen sowohl im Singular als auch im Plural (z. B. Dat. Sg. und auch Nom./Akk. Pl. *ensti* zum fem. i-Stamm *anst* ‘Gunst’). Die Umlautformen werden teilweise schon im Verlauf

des Althochdeutschen aufgegeben, so daß die für das Deutsche formulierte Tendenz zu umlautlosem Singular und allenfalls umgelautesem Plural bereits im Althochdeutschen greift (vgl. Sonderegger 1979: 309). Die wesentliche Umgestaltung des althochdeutschen Flexionssystems wird dann durch den zum Mittelhochdeutschen eingetretenen Zusammenfall der vollen Nebensilbenvokale in -e bewirkt, wodurch ein strukturell grundsätzlich neues Flexionssystem entsteht (vgl. Stopp 1974): Es werden die meisten paradigmatischen Distinktionen eingeebnet (z. B. Zusammenfall der Plural-Flexivreihe der a- und i-Deklination, z. B. Nom/Akk. Pl. *tag-/gest-e* aus *-a/-i*), so daß das Inventar der segmentierbaren Flexive von 52 auf 16 im Mittelhochdeutschen reduziert ist (vgl. Sonderegger 1979: 246). Strukturell wirkt sich dieser Vorgang in mehrfacher Weise aus. Einerseits wird nun zusätzlich zur regressiven Endungsflexion auch die Wurzelflexion (Umlaut) genutzt. Der Umlaut erscheint analog übertragen auch bei Lexemen, bei denen er historisch nicht entwickelt ist, so z. B. bei mask. a-Stämmen (z. B. Nom. Pl. *nagele* neben *negele*) oder den angegliederten Restgruppen (z. B. r-Stamm *veter(e)* neben *vater(e)*; vgl. Paul<sup>24</sup> 1998: § 177; Wolf 1981: 198); damit erhält der Umlaut eine deutlichere flexivische Signifikanz als Kennzeichen des Plurals. Andererseits ist durch die eingetretene Uniformierung der Nebensilbe “automatisch die Grenze zwischen Stamm und Flexionsendung hinter die alte Flexionsendung” (Wegera 1987: § 11) verschoben: Es kommt zu der für das Mittelhochdeutsche wesentlichen Veränderung einer nurmehr zweigliedrigen Wortstruktur von Stamm und Flexionszeichen, durch die der Nom. Sg. zur unmarkierten Grundform wird (vgl. Paul<sup>25</sup> 1989: § 174; Klein 1987: 155). Relational dazu wird das ursprüngliche Stammbildungssuffix -er (germ. \*-iz/-az) zum Plural-Kennzeichen profiliert und nun auf eine Reihe weiterer, zum Teil im Mittelhochdeutschen erstmals belegter Lexeme übertragen (z. B. *cleid* – *cleider*). Ebenso wird nun bei allen Substantiven, deren Plural im Althochdeutschen durch vokalisches Flexiv gebildet wird (z. B. Nom. Pl. *taga*, *gebâ*, *listi*), das -e zu einem der vorwiegenden mittelhochdeutschen Plural-Flexive, das insbesondere im Mitteldeutschen auch auf bisher nicht markierte Fälle übertragen wird (besonders neutr. a-Stämme, z. B. Nom./Akk. Pl. *wort* > *worte*). Damit ist schon im Mittelhochdeutschen die Tendenz zur eindeutigen

Kennzeichnung der Numerusopposition deutlich (vgl. Wolf 1981: 199). Dieser Prozeß wird allerdings durch die seit dem 13. Jahrhundert zuerst im Oberdeutschen eintretende Apokope und die damit eintretende Zerstörung der über -e geleisteten Plural-Kennzeichnung überlagert. Dies prägt die schreibsprachliche Entwicklung bis in das 16. Jahrhundert hinein. Dabei lassen sich die fröhneuhochdeutschen Entwicklungen insgesamt als zwei Großprozesse beschreiben: Über eine Umorganisation vorhandener flexivischer Mittel (z. B. -er, -en, Umlaut; vgl. Wegera 1987: § 63) wird der bisher unbezeichnete (mittelhochdeutsche neutr. a-Stämme, z. B. *wort* – *wort*) oder durch Apokope unbezeichnet gewordene Plural „durch deutliche und sichere Flexive“ markiert (Wegera 1987: § 11); durch eine (flexivische) Nivellierung der Kasus (z. B. Abbau der -en-Flexive in obliquen Kasus bei Feminina oder gegenläufig auch Übertragung des -en in den Nom. Sg.; vgl. Wegera 1987: § 48), die über andere analytisch-syntaktische Möglichkeiten zum Ausdruck gelangen, erfolgt paradigmatisch eine Profilierung des Numerus (Plural). Da beide Prozesse (zumindest teilweise) genussgesteuert verlaufen, zeigt sich neben Kasusnivellierung und Numerusprofilierung im Verlauf des Fröhneuhochdeutschen implizit ebenfalls eine **Genusprofilierung** (vgl. Wegera 1987: § 11, 18). Schon im 16. Jahrhundert deutet sich aus dem von der Apokope weitgehend unberührt gebliebenen Obersächsischen kommend eine Restituirung des Plural-e an, das im 17. Jahrhundert zunehmend wieder in der Schriftsprache verwendet wird. Hier spielt die Bibelübersetzung Luthers und ihre Verbreitung eine herausragende Rolle („lutherisches -e“); es wirkt jedoch auch die „strukturelle Überlegenheit des omd. Systems“: Hinsichtlich einer „konsequenteren Numerusmarkierung durch das Plural-e“ entsprach dieses schreiblandschaftlich ausgebildete System „der Tendenz zur Numerusprofilierung besser als die nur ansatzweise durchgeführten Numerusunterscheidungen mit Hilfe anderer Flexive“ (Wegera 1987: § 95). So wird das Plural-e „vielfach wieder zur bevorzugten Pluralbildung bei Lexemen, die in davorliegenden Zeiträumen andere Pluralbildungen angenommen hatten“ (Wegera 1987: § 72). Zum Neuhochdeutschen hin ist -e zum wichtigsten Numerusflexiv innerhalb des Kernwortschatzes herausgebildet (vgl. Augst 1975: 37). Daneben hat die Kombinierbarkeit des Umlauts mit suffixalen Flexiven zum

Neuhochdeutschen deutlich zugenommen (vgl. Sondergger 1979: 246); mit dem Umlaut nicht kombinierbar sind neuhochdeutsch allein die nicht-fremdsprachlichen Pluralflexive -s und -en.

### 5.3. Flexion der Adjektive

Die Flexionsgeschichte der Adjektive ist insbesondere innerhalb der gesamten Substantivgruppe als Durchsetzung einer Artikel- sowie **Gruppenflexion** zu beschreiben, die sich sowohl an konkreten flexivischen Veränderungen als auch in der Aufgabe einer Flexion des prädikativen Adjektivs zeigt: **Artikelflexion** meint einerseits die grammatische Abhängigkeit vom und andererseits die Funktionalisierung auf das bezogene Substantiv hin (das Adjektiv realisiert die morphosyntaktischen Merkmale des Substantivs); dabei wird das Adjektiv in Abhängigkeit von der ausdrucksseitig vorhandenen Markiertheit der Substantivgruppe „determinierend“ (genusdifferenzierend im Nom. Sg.: *ein lieber Junge*, *ein liebes Kind*) oder „indeterminierend/attribuierend“ (genusindifferent im Nom. Sg. -e: *der liebe Junge*, *das liebe Kind*) flektiert. „Gruppenflexion“ meint die tendenzielle Flexivdistribution in Abhängigkeit besonders von der Markiertheit des Substantivs selbst: „starke“ Adjektivflexion bei fehlendem substantivischem Kasusflexiv (*groß-er Liebe fähig sein*), „schwache“ Adjektivflexion bei vorhandenem substantivischem Kasusflexiv (*groß-en Mutes fähig sein*).

Die das Adjektiv innerhalb der Hauptwortarten charakterisierende Artikelflexion ist bereits im Althochdeutschen usuell, die Regelung wird im Fröhneuhochdeutschen endgültig fest. Diese „Formregel“ („mechanische Regelung“, Behaghel 1923: § 10) löst eine frühere „Sinnregel“ ab, bei der über die Flexion des Adjektivs primär eine ko- oder kontextuelle (Un)Bestimmtheit signalisiert wurde. Dabei stimmen die vordeutschen Flexionsparadigmen der Adjektive weitgehend mit denen der Substantive überein (vgl. Kern & Zutt 1977: 109). Aufgrund jedoch einer den Pronomen identischen Verwendungsweise treten schon im Germanischen zunehmend auch pronominale Formen neben die ursprünglichen nominalen (vgl. Kern & Zutt 1977: 110 f.). Die Durchsetzung dieser pronominalen Flexive ist beim Adjektiv mit dem 16. Jahrhundert abgeschlossen; sie führt zur flexivischen Angleichung von Adjektiv und Pronomen, die sich ähnlich auch bei den Pronomen zeigt, z. B. Dat. Pl./Gen. Pl. des Per-

sonalpronomens *in* > *inen*, *ir* > *irer* oder Gen. Sg./Pl. *des* > *dessen*, *der* > *deren*, *wes* > *wessen* (vgl. Walch & Häckel 1988: § 12.2., Anm. 34). Eine Angleichung der adjektivischen und pronominalen Flexion erweist sich schließlich insbesondere “in der Durchsetzung der Regelung des grammatisch determinierenden/indeterminierenden Flexionsprinzips auch beim Poss. Pron., den Pnominaladj. (und sehr selten auch Pnominalsubst.)” (Ebert et al. 1993: § M60). In diesem Zusammenhang ist die Adjektivflexion als eine Gruppenflexion herausgebildet, für die im Verlauf des 16. Jahrhunderts ein “monoflexivisches Prinzip” usuell wird (vgl. Brinkmann 1964: 96; Solms & Wegera 1991: § 143): Die morphosyntaktischen Kategorien des Substantivs werden tendenziell nur an einer Stelle der gesamten Substantivgruppe ausdrucksseitig signalisiert (*der großer Mut* > *der groß-e Mut*; *groß-es Mutes fähig sein* > *groß-en Mutes*); eine ähnliche Entwicklung zeigt sich im Frühneuhochdeutschen vereinzelt auch beim Pronomen: *meinen Wissens*. Parallel der flexivischen Entwicklung wird die Flexionsmöglichkeit bis zum 16. Jahrhundert auf nurmehr die attributive/substantivierte Verwendung begrenzt; die bis in die erste Hälfte des Frühneuhochdeutschen noch mögliche Flexion auch des prädikativen Adjektivs (*ein Baum ist größer*) wird aufgegeben; damit werden prädiktives und adverbiales Adjektiv formal ununterscheidbar.

Markanter noch als beim Substantiv wird das Flexivinventar beim Adjektiv vom Althochdeutschen (47 unterscheidbare Flexive) über das Mittelhochdeutsche (24) zum Neuhochdeutschen hin (6) reduziert (vgl. Sonderegger 1979: 246). Die stufenweise zu verfolgende Reduzierung sowie die beobachtbare Verwendung der jeweiligen Flexive erweist mehrere Tendenzen:

- (a) Es wird eine tendenzielle 1:1-Zuordnung (Form:Funktion) auf der Ebene der “morphosyntaktischen Position” (gemeint ist z. B. ‘Nom. Sg. Mask. nach best. Art.’ oder ‘Dat. Sg. Fem. nach Präposition’; vgl. Solms & Wegera 1991: § 12) herausgebildet. Aufgrund der flexivischen Reduzierung stehen nur noch maximal zwei Flexive je morphosyntaktischem Merkmalbündel zur Verfügung, die auf der Ebene der morphosyntaktischen Position dann eindeutig disambiguierbar werden. So sind z. B. für Nom. Sg. Neutr. -(e) und -(e)s möglich, nach dem bestimmten Artikel kann jedoch nur -(e) (*das liebe Kind*), nach dem unbestimmten Artikel nur -(e)s (*ein liebes Kind*) erscheinen.

- (b) In keinem Fall wird im Frühneuhochdeutschen das flexivische Inventar der Singular-Kasus so verändert, daß eine bestehende Genusdifferenzierung aufgehoben würde. Vielmehr wird dort, wo flexivisch eine Genuspolysémie vorliegt, im Frühneuhochdeutschen eine Eindeutigkeit erreicht: “Genusprofilierung” (z. B. Profilierung des -o im Nom./Akk. Sg. zum genusmarkierenden Flexiv des Neutr. bis ins 16. Jahrhundert: Maskulinum *wahr-er Mut*, Femininum *wahr-e Liebe*, Neutr. *wahr Wort*; vgl. Solms & Wegera 1991: § 51–55).
- (c) Die im Nom./Akk. Sg. Neutr. zum späten Frühneuhochdeutschen erfolgte Verdrängung der unflektierten Form zugunsten der Flexion mit pronominalem -es liegt innerhalb des Prozesses der zwischen dem 14. und 16. Jahrhundert erfolgten formalen Ausdifferenzierung der Substantiv- und Verbalgruppe: die flexionslose Form kommt der nicht substantivischen Komponente der Verbalgruppe (als prädiktives Adjektiv), die flektierte Form der nicht substantivischen Komponente der Substantivgruppe zu.
- (d) Teilweise ergibt sich eine Kasusmarkierung im Zusammenhang der Markierungsfunktion des Adjektivs für die am Substantiv nicht markierten morphosyntaktischen Merkmale: Da sieht neben der “Kasusnivellierung” beim Substantiv eine gegensteuernde “systemerhaltende Tendenz” beim Adjektiv: Eine solche Kasus(gruppen)markierung (Nom./Akk. Pl. versus Gen./Dat. Pl.) zeigt z. B. die insbesondere im Oberdeutschen bis in das 18. Jahrhundert hinein übliche determinierende Flexion mit -e im Nom./Akk. trotz vorausgehenden bestimmten Artikels (*die liebe Kinder*) gegenüber Dat./Gen. mit -en (*den/der lieben Kinder(n)*). Eine generelle Tendenz zur Kasusmarkierung liegt jedoch nicht vor, analog zu den Pronomen (*diu, die* > *die*) wird z. B. die formale Kasusunterscheidung Nom./Akk. Sg. Fem. aufgehoben.

## 6. Wortbildung

Die geschichtliche Ausdifferenzierung von Flexion und Wortbildung erweist sich auch in der funktionalen Profilierung der formalen Ausdrucksmittel. Zwar werden bereits im Germanischen **Komposition**, **Derivation** und auch “innere Wortbildung” genutzt, doch sieht Grimm in letzterer, durch Ablaut erfolgender Wortbildung die älteste Möglichkeit, z. B. erkennbar in ahd. *faran* ‘fahren’, *fuora* ‘Fahrt’ (vgl. Grimm 1826: 1). Dieses Prinzip ist im Althochdeutschen schon nicht mehr produktiv, sondern nurmehr “aktiv” (die Wortschatzeinheiten sind morphologisch-relational motiviert); der Lautwechsel ist seitdem primär flexivisch produktiv; ausgebaut werden die Möglichkeiten der Komposition

und affixalen Wortbildung (als explizite Ableitung durch Suffixe oder als Präfigierung). Aufgrund der in den verschiedenen Wortarten unterschiedlichen Nutzung sind die einzelnen Wortbildungsarten schon im Althochdeutschen als wortartcharakteristisch ausgewiesen: Komposition bei den Substantiven, Derivation bei Substantiven und Adjektiven, Präfigierung bei den Verben (eine wortartbestimmte Bevorzugung erweisen mittelhochdeutsche Texte im Verhältnis von Ableitungen zu Komposita mit 3 : 2 bei Substantiven und 3 : 1 bei Adjektiven; vgl. Zutt 1985: 1160). Dabei zeigt sich in der parallelen Existenz einiger Morphe als noch selbständige Lexeme (z. B. starkes Maskulinum/Neutrum *tuom* ‘Urteil, Gericht, Macht, Fähigkeit’), als erster oder zweiter Teil entsprechender Determinativkomposita (*tuomtac* ‘Gerichtstag’, *cheisertuom* ‘Herrschaft des Kaisers’) sowie auch schon nurmehr als Suffix (*altruom* ‘Alter’) die ursprüngliche Herkunft affixaler Elemente: Es handelt sich um Lexeme, die aufgrund häufiger und reihenhaft empfundener Komposition ihre Selbständigkeit verlieren und zum Affix werden (z. B. auch Präfix *be-* aus der Präposition *bei*). Ähnliche Entwicklungen zeigen sich in allen Sprachstufen des Deutschen, so z. B. auch in *-zeug* (*Nährzeug*, *Werkzeug*).

Die Wortschatzerweiterung durch Wortbildung wird in der geschichtlichen Entwicklung des Deutschen zunehmend genutzt, wobei die deutlichste frequentielle Zunahme im Frühneuhochdeutschen und hier insbesondere im 16. und 17. Jahrhundert stattfindet: So liegt z. B. der Anteil derivierter/zusammengesetzter Nomen in exemplarisch ausgewählten mittelhochdeutschen Texten bei unter 10 Prozent, in entsprechenden Texten des Neuhochdeutschen dagegen bei über 50 Prozent (vgl. Zutt 1985: 1160); allein der Anteil von Komposita am Gesamt des jeweiligen Substantivwortschatzes wächst von ca. 7 Prozent in Texten des 12. und 13. Jahrhunderts auf über 18 Prozent in Texten des späten 17. Jahrhunderts. Im Bereich der Derivation wächst der Bestand desubstantivischer Verben zwischen der Früh- und Spätzeit des Frühneuhochdeutschen um ca. 125 Prozent (vgl. Prell & Scheben-Schmidt 1996: 386–388). In solchen Verhältnissen erweist sich die im Frühneuhochdeutschen konstitutive Offenheit der schriftsprachlichen Entwicklung für varietäre Einflüsse, insbesondere auch die fröhneuhochdeutsche Ausbildung und Entwicklung von Fachsprachen.

Das Inventar der Wortbildungsmittel sowie auch die semantisch-funktionalen wie morphologischen Typen der jeweiligen Wortbildungsmuster sind im wesentlichen bereits im Althochdeutschen ausgebildet (vgl. Prell & Scheben-Schmidt 1996: 380) (z. B. nominale Verbableitung des Typs ‘ornatives Verb’ mittels Verbalisierungsmorphs, z. B. *salb-ōn*). Allerdings finden wesentliche Veränderungen statt, die sich derzeit (möglicherweise nur aufgrund der Forschungslage) als Entwicklungen insbesondere der frühneuhochdeutschen Zeit zeigen: Entwicklungen des Lexembestandes erkennbarer Wortbildungsmuster/-typen und auch des Inventars der genutzten Affixe, sowie Veränderungen der morphologischen Struktur und Basisbeziehungen. So handelt es sich bei der Mehrzahl mittelhochdeutscher Präfixverben (zu ca. 70–90 Prozent) um “untrennbare Präfixverben” (z. B. *be-*, *ent-*; “trennbare Präfixverben” z. B. mit *an-*, *aus-*), wohingegen diese im Neuhochdeutschen nurmehr ca. 30 Prozent ausmachen (vgl. Solms 1990: 118; Herbers 2002: 107 f.). Dem veränderten Wortbildungsmuster geht eine Verschiebung der Wortbildungsbasis der vorhandenen (un)trennbaren Präfixverben parallel. Der Anteil der sogenannten “Präfix-Suffix-Ableitungen” (mit nominaler Basis, z. B. *bebilden*) beträgt neuhochdeutsch knapp 40 Prozent; der mittelhochdeutsche Anteil solcher Bildungen liegt bei höchstens 12 Prozent. Der Umschwung zugunsten einer verstärkten Verbableitung auch aus nominalen Basen wird schon im späten Mittelhochdeutschen bemerkbar und setzt vor allem seit dem 16. Jahrhundert ein (vgl. Prell & Scheben-Schmidt 1996: 388).

Die fortschreitende Verbalisierung nominaler Basen zeigt für die Verben eine den Nomen identische Tendenz: die **Univerbierung** umständlicherer Syntagmen in Komposita und Derivaten, nach Wegera ein Moment sprachlicher Ökonomie des Deutschen (Erben<sup>2</sup> 1983: 122; vgl. Wegera & Prell<sup>2</sup> 2000: 1594).

Die Verbpräfigierung zeigt vom Mittel zum Neuhochdeutschen eine ähnlich auch im Englischen entwickelte Tendenz zur **Iso-Morphie** (vgl. Koziol<sup>2</sup> 1972: § 710). Das Mittelhochdeutsche kennt ebenso wie noch das frühe Frühneuhochdeutsche eine “Überproduktion von Präfixverben”, die zu semantischer Äquivalenz von Simplex und Präfixverb führen kann (vgl. Zutt 1985: 1163; Solms 1990: 122; Herbers 2002: 323 f.). Hier zeigt

sich in der Folge eine Informationsverlagerung vom Simplex auf das Präfixverb, auf die eine semantische Verengung oder der Untergang des Simplexverbs folgt. Der Wortschatz als Inventar der formalen Ausdrucksmittel wird erweitert, auf der Inhaltsseite aber ist nichts Neues entstanden. Die Präfigierung leistet hier einerseits eine Auflösung der bei Simplicia vorhandenen Polysemie. Zugleich erscheinen die Präfixverben als ausdrucksseitig deutlicher markierte Wortschatzeinheiten für gleichbleibende Inhalte. Es werden einzelne Sprachinhalte „reihenhaft gleichbleibend (‘iso-morph’) ausgeprägt“ (Erben 1964), das einzelne Präfix wird zu einem deutlichen Ausdrucksträger profiliert. Da hier eine dem Simplex inhärente Bedeutung durch ein zusätzliches, syntagmatisch gebundenes Element zum Ausdruck gebracht wird, steht auch diese Entwicklung im Zusammenhang einer Veränderung synthetischer zugunsten analytischer Ausdrucksmittel (vgl. Tschirch<sup>3</sup> 1989: 134).

## 7. Zitierte Literatur

- Augst, Gerhard (1975), *Untersuchungen zum Morpheminventar der deutschen Gegenwartssprache*. Tübingen: Narr (Forschungsberichte des Instituts für deutsche Sprache 25)
- Behaghel, Otto (1923), *Deutsche Syntax: Eine geschichtliche Darstellung. Bd. I: Die Wortklassen und Wortformen. A. Nomen. Pronomen*. Heidelberg: Winter (Germanische Bibliothek I, 10)
- Behaghel, Otto (1928), *Geschichte der deutschen Sprache*. Berlin, Leipzig: de Gruyter (Grundriss der Germanischen Philologie 3)
- Bergmann, Rolf & Tiefenbach, Heinrich & Voetz, Lothar (1987, Hrsg.), *Althochdeutsch, Bd I: Grammatik, Glossen und Texte*. Heidelberg: Winter (Germanische Bibliothek, Neue Folge 3. Reihe: Untersuchungen)
- Besch, Werner (1980), „Frühneuhochdeutsch“. In: Althaus, Hans Peter & Henne, Helmut & Wiegand, Herbert Ernst (Hrsg.), *Lexikon der Germanistischen Linguistik*. Tübingen: Niemeyer, 588–597
- Besch, Werner (2003), „Die Entstehung und Ausformung der neuhochdeutschen Schriftsprache/Standardsprache“. In: Besch et al. (Hrsg.), 2252–2296
- Besch, Werner & Betten, Anne & Reichmann, Oskar & Sonderegger, Stefan (1998/2000/2003/2004, Hrsg.), *Sprachgeschichte, Bd. 1–4* [2., vollst. neu bearb. und erw. Aufl.; <sup>1</sup>1984/1989] Berlin usw.: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft 2)
- Braune, Wilhelm (1987), *Althochdeutsche Grammatik*, bearb. v. Hans Eggers. Tübingen: Niemeyer (Sammlung kurzer Grammatiken Germanischer Dialekte A, 5) [<sup>1</sup>1886]
- Brinkmann, Hennig (1964), „Das deutsche Adjektiv in synchronischer und diachronischer Sicht“. *Wirkendes Wort* 14, 94–104
- Durrell, Martin (1980), „Morphologische Regelmäßigkeiten im deutschen Ablautsystem“. In: Rupp, Hans & Roloff, Hans-Gert (Hrsg.), *Akten des VI. Internationalen Germanisten-Kongresses Basel 1980, Teil 2*. Bern usw.: Lang, 19–28 (Jahrbuch für internationale Germanistik A, 8.2)
- Ebert, Robert Peter & Reichmann, Oskar & Solms, Hans-Joachim & Wegera, Klaus-Peter (1993), *Frühneuhochdeutsche Grammatik*. Tübingen: Niemeyer (Sammlung kurzer Grammatiken Germanischer Dialekte A, 12)
- Eichinger, Ludwig M. (1984), „Die großen historischen Grammatiken des Deutschen“. In: Besch et al. (Hrsg.), 433–440
- Erben, Johannes (1964), „Deutsche Wortbildung in synchronischer und diachronischer Sicht“. *Wirkendes Wort* 14, 83–93
- Erben, Johannes (1983), *Einführung in die deutsche Wortbildungslehre*. Berlin: Schmidt (Grundlagen der Germanistik 17) [<sup>1</sup>1975]
- Grimm, Jacob (1819–1837), *Deutsche Grammatik, 4 Teile*. Göttingen
- Habermann, Mechthild (1994), *Verbale Wortbildung um 1500: Eine historisch-synchrone Untersuchung anhand von Texten Albrecht Dürers, Heinrich Deichslers und Veit Dietrichs*. Berlin usw.: de Gruyter (Wortbildung des Nürnberger Frühneuhochdeutsch 2)
- Hempen, Ute (1988), *Die starken Verben im Deutschen und Niederländischen: Diachrone Morphologie*. Tübingen: Niemeyer (Linguistische Arbeiten 214)
- Henzen, Walter (1965), *Deutsche Wortbildung*. Tübingen: Niemeyer (Sammlung kurzer Grammatiken germanischer Dialekte B, 5) [<sup>1</sup>1947]
- Herbes, Birgit (2002), *Verbale Präfigierung im Mittelhochdeutschen: eine semantisch-funktionale Korpusanalyse*. Tübingen: Niemeyer (Studien zur Mittelhochdeutschen Grammatik 1)
- Hirt, Hermann (1932), *Handbuch des Urgermanischen, Teil II: Stammbildungs- und Flexionslehre*. Heidelberg: Winter (Indogermanische Bibliothek I, 21)

- Hoffmann, Walter & Solms, Hans-Joachim (1987), "Zur Flexion der starken und schwachen Verben". In: Besch, Werner & Klaus-Peter Wegner (Hrsg.), *Frühneuhochdeutsch: Zum Stand der sprachwissenschaftlichen Forschung*. Zeitschrift für deutsche Philologie 106, Sonderheft, 37–59
- Hotzenköcherle, Rudolf (1962), "Entwicklungs geschichtliche Grundzüge des Neuhochdeutschen". *Wirkendes Wort* 12, 321–331
- Hotzenköcherle, Rudolf (1968), "Gegenwartsprobleme im deutschen Adjektivsystem". *Neuphilologische Mitteilungen* 69, 1–28
- Joeres, Rolf (1995), *Wortbildungen mit -macher im Althochdeutschen, Mittelhochdeutschen und Neuhochdeutschen*. Heidelberg: Winter (Germanische Bibliothek: Reihe 3, Untersuchungen, N.F. 21)
- Kern, Peter Chr. & Zutt, Herta (1977), *Geschichte des deutschen Flexionssystems*. Tübingen: Niemeyer (Germanistische Arbeitshefte 22)
- Kienle, Richard von (2<sup>1969</sup>), *Historische Laut- und Formenlehre des Deutschen*. Tübingen: Niemeyer (Sammlung kurzer Grammatiken Germanischer Dialekte A, Hauptreihe 11)
- Klein, Thomas (1987), "Zur althochdeutschen Flexionsmorphologie in synchroner Sicht". In: Bergmann et al. (Hrsg.), 146–168
- Kluge, Friedrich (2<sup>1925</sup>), *Abriß der deutschen Wortbildungsllehre*. Halle/S.: Niemeyer
- Koenraads, Willy Henri August (1953), *Studien über sprachökonomische Entwicklungen im Deutschen*. Amsterdam: Rapoldi
- Kohrt, Manfred (1984), "Historische Phonetik und Phonologie". In: Besch et al. (Hrsg.), 388–399
- Koziol, Herbert (2<sup>1972</sup>), *Handbuch der englischen Wortbildungsllehre*. Heidelberg: Winter (Germanische Bibliothek, Reihe 1) [1937]
- Lüdtke, Helmut (1988), "Grammatischer Wandel". In: Ammon, Ulrich & Dittmar, Norbert & Mattheier, Klaus J. (Hrsg.), *Soziolinguistik*, Bd. 1–2. Berlin usw.: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft 3.), 1632–1642
- Mattheier, Klaus (2<sup>1998</sup>), "Allgemeine Aspekte einer Theorie des Sprachwandels". In: Besch et al. (Hrsg.), 824–836
- Michels, Victor (2<sup>1979</sup>), *Mittelhochdeutsche Grammatik*. Heidelberg: Winter
- Moser, Virgil (1929), *Frühneuhochdeutsche Grammatik*. Bd. I: *Lautlehre. 1. Hälfte: Orthographie, Betonung, Stammsilbenvokale*. Heidelberg: Winter (Germanische Bibliothek I, 17)
- Müller, Peter O. (1993), *Substantiv-Derivation in den Schriften Albrecht Dürers*. Berlin usw.: de

## XVIII. Morphologischer Wandel II: Fallstudien

- Gruyter (Wortbildung des Nürnberger Frühneuhochdeutschen 1)
- Paul, Hermann (1916–1920), *Deutsche Grammatik*, Bd. 1–5. Tübingen: Niemeyer
- Paul, Hermann (2<sup>1920</sup>), *Prinzipien der Sprachgeschichte*. Halle/S.: Niemeyer [8<sup>1968</sup> Tübingen: Niemeyer; 1<sup>1880</sup>]
- Paul, Hermann (1981), "Über die Aufgaben der Wortbildungslehre". In: Lipka, Leonhard & Günther, Hartmut (Hrsg.), *Wortbildung*. Darmstadt, 17–35 (WDF 564)
- Paul, Hermann (2<sup>1998</sup>), *Mittelhochdeutsche Grammatik*, neu bearbeitet von Peter Wiehl & Siegfried Grosse. Tübingen: Niemeyer (Sammlung kurzer Grammatiken Germanischer Dialekte A, 2) [1881]
- Penzl, Herbert (1984), *Frühneuhochdeutsch*. Bern usw.: Lang (Germanistische Lehrbuchsammlung 9)
- Prell, Heinz-Peter & Scheben-Schmidt, Marietheres (1996), *Die Verbalieitung im Frühneuhochdeutschen*. Berlin usw.: de Gruyter (Studia Linguistica Germanica 41)
- Putschke, Wolfgang (2<sup>1998</sup>), "Die Arbeiten der Junggrammatiker und ihr Beitrag zur Sprachgeschichtsforschung". In: Besch et al. (Hrsg.), 474–494
- Reichmann, Oskar (1988), "Zur Abgrenzung des Mittelhochdeutschen vom Frühneuhochdeutschen". In: Bachofer, Wolfgang (Hrsg.), *Mittelhochdeutsches Wörterbuch in der Diskussion*. Tübingen: Niemeyer (Reihe Germanistische Linguistik 84), 119–147
- Roelcke, Thorsten (1995), *Periodisierung der deutschen Sprachgeschichte*. Berlin usw.: de Gruyter (Studia Linguistica Germanica 40)
- Roelcke, Thorsten (2001), *Periodisierung: Die zeitliche Gliederung der deutschen Sprachgeschichte*. Frankfurt a. M.: Lang (Dokumentation germanistischer Forschung 4)
- Ronneberger-Sibold, Elke (1989), *Historische Phonetik und Morphologie des Deutschen: Eine kommentierte Bibliographie zur strukturellen Forschung*. Tübingen: Niemeyer (GA: Ergänzungsreihe, 3)
- Russ, Charles V. J. (1978), *Historical Germanic Phonology and Morphology*. Oxford: Oxford University Press
- Russ, Charles (1987), "Die Verben auf -u in der Mundart von Bosco Gurin und die ahd. -ōn-Verben". In: Bergmann et al. (Hrsg.), 332–344
- Scherer, Wilhelm (2<sup>1878</sup>), *Zur Geschichte der deutschen Sprache*. Berlin: Weidmann
- Schirmunski, Victor M. (1962), *Deutsche Mundartkunde: Vergleichende Laut- und Formenlehre der deutschen Mundarten*. Berlin: Akademie

- Schmidt, Karl Horst (2<sup>1998</sup>), "Versuch einer geschichtlichen Sprachtypologie des Deutschen". In: Besch et al. (Hrsg.), 993–1000
- Schwarz, Hans (1986), *Präfixbildungen im deutschen Abrogans: Analyse und Systematik*. Göppingen: Kümmerle (Göppinger Arbeiten zur Germanistik 458)
- Simmler, Franz (1987), "Morphologische Grundlagen zur Beschreibung althochdeutscher Schreibdialekten". In: Bergmann et al. (Hrsg.), 114–146.
- Sodmann, Timothy (2<sup>2000</sup>), "Der Rückgang des Mittelniederdeutschen als Schreib- und Druckersprache". In: Besch et al. (Hrsg.), 1505–1512
- Solms, Hans-Joachim (1984), *Die morphologischen Veränderungen der Stammvokale der starken Verben im Frühneuhochdeutschen*. Ph.D. Dissertation, Universität Bonn
- Solms, Hans-Joachim (1990), "Das System der Präfixverben in der frühesten Überlieferung des Hartmannschen 'Gregorius'". In: Besch, Werner (Hrsg.), *Deutsche Sprachgeschichte: Grundlagen, Methoden, Perspektiven. Festschrift für Johannes Erben zum 65. Geburtstag*. Frankfurt a. M. usw.: Lang, 115–128
- Solms, Hans-Joachim (1991), "Die Wortbildung der Verben in Hartmann von Aues 'Iwein' (Hs.B) und 'Gregorius' (Hs.A): Das Präfix ge- im System der verbalen Präfigierung." In: Wegera, Klaus-Peter (Hrsg.), *Mittelhochdeutsche Grammatik als Aufgabe*. Zeitschrift für deutsche Philologie 110, Sonderheft, 110–140
- Solms, Hans-Joachim (1993), "Ein verspäteter Abalautausgleich im Kontext des regionalen und konfessionellen Meinungsstreits im 18. Jahrhundert. Die starken Verben der Klasse IIIa". In: Mattheier, Klaus J. & Wegera, Klaus-Peter & Hoffmann, Walter & Macha, Jürgen & Solms, Hans-Joachim (Hrsg.), *Vielfalt des Deutschen: Festschrift für Werner Besch*. Frankfurt a. M. usw.: Lang, 331–351
- Solms, Hans-Joachim (2<sup>1998</sup>), "Historische Wortbildung". In: Besch et al. (Hrsg.), 596–610
- Solms, Hans-Joachim & Wegera, Klaus-Peter (1991), *Flexion der Adjektive*. Heidelberg: Winter (Grammatik des Frühneuhochdeutschen: Beiträge zur Laut- und Formenlehre, Band VI)
- Sonderegger, Stefan (1979), *Grundzüge deutscher Sprachgeschichte, Bd. I: Einführung-Genealogie-Konstanten*. Berlin usw.: de Gruyter
- Sonderegger, Stefan (2<sup>2000</sup>), "Morphologie des Althochdeutschen". In: Besch et al. (Hrsg.), 1171–1196
- Stärck, John (1912), *Studien zur Geschichte des Rückumlauts*. Uppsala: Appelbergs Boktryckeri
- Stopp, Hugo (1974), "Veränderungen im System der Substantivdeklination vom Althochdeutschen bis zum Neuhochdeutschen". In: Besch, Werner (Hrsg.), *Studien zur deutschen Literatur und Sprache des Mittelalters. Festschrift für Hugo Moser zum 65. Geburtstag*. Berlin: Schmidt, 324–344
- Tschirch, Fritz (3<sup>1989</sup>), *Geschichte der deutschen Sprache, zweiter Teil*, bearbeitet von Werner Besch. Berlin: Schmidt
- Walch, Maria & Häckel, Susanne (1988), *Flexion der Pronomina und Numeralia*. Heidelberg: Winter (Grammatik des Frühneuhochdeutschen: Beiträge zur Laut- und Formenlehre, Band VII)
- Wegera, Klaus-Peter & Prell, Heinz-Peter (2<sup>2000</sup>), "Wortbildung des Frühneuhochdeutschen". In: Besch et al. (Hrsg.), 1594–1604
- Wegera, Klaus-Peter (1987), *Flexion der Substantive*. Heidelberg: Winter (Grammatik des Frühneuhochdeutschen: Beiträge zur Laut- und Formenlehre, Band III)
- Weinhold, Karl (2<sup>1883</sup>), *Mittelhochdeutsche Grammatik*. Breslau
- Wells, Christopher J. (1990), *Deutsch: eine Sprachgeschichte bis 1945*. Tübingen: Niemeyer (RGL 93)
- Werner, Otmar (2<sup>1998</sup>), "Historische Morphologie". In: Besch et al. (Hrsg.), 572–596
- Wilmanns, Wilhelm (1906–1909), *Deutsche Grammatik: Gotisch, Alt-, Mittel- und Neuhochdeutsch*. Straßburg: Trübner
- Winkler, Gertraud (1995), *Die Wortbildung mit -lich im Alt-, Mittel- und Frühneuhochdeutschen*. Heidelberg: Winter (Sprache – Literatur und Geschichte 11)
- Wolf, Norbert Richard (1981), *Althochdeutsch – Mittelhochdeutsch*. Heidelberg: Quelle & Meyer
- Zutt, Herta (1985), "Wortbildung des Mittelhochdeutschen". In: Besch et al. (Hrsg.), 1159–1164

*Hans-Joachim Solms, Halle/S. (Deutschland)*

## 157. From Latin to French

1. Introduction
2. Inflection
3. Word-formation
4. References

### 1. Introduction

There are few languages whose history has been studied in such depth as French. But though there are many excellent studies on single special issues, coverage is by no means complete and, most disappointingly, we still lack a thorough modern synthesis of the evolution of both inflection and word-formation from Latin to Modern French. It goes without saying that this article cannot fill such a gap. For word-formation, it still had to heavily rely on Meyer-Lübke's elegant monograph of 1921, whose reedition in 1966 by Piel has rightly been the object of severe criticism (see Bork 1968; Höfler 1967; Jänicke 1969). For inflection, the bibliographic situation is more satisfactory, since there exists a series of recent studies that have extended our knowledge of the data or at least reinterpreted well-known data in the light of modern historical linguistics. But we still lack satisfactory explanations for many changes, though there is no dearth of proposals for any single case.

As is well-known, the starting point for the development of the Romance languages is not Classical Latin, but "Vulgar", i.e. spoken Latin. Nevertheless in this article the Vulgar Latin forms are only used where the difference with respect to Classical Latin is significant for the problem under discussion. The length of Latin vowels is also indicated only where this is crucial for the argument.

### 2. Inflection

Ever since Schlegel (1818), the passage from Latin to French has been considered as a paradigmatic case for the evolution from an essentially synthetic language to a predominantly analytic one (cf. Art. 10; Schwesler 1990). **Synthetic** here means that grammatical categories such as gender, number, case, tense, mood, etc. are encoded in the words themselves *via* inflections, while **analytic** refers to situations where these grammatical categories are encoded in separate words or clitics. As we will see below, the analytical tendency has been particularly strong in the

nominal domain, while French verbs have remained closer to the Latin synthetic type, even though the French conjugation is also more analytic than that of other, more conservative Romance languages.

#### 2.1. Nouns and adjectives

Latin nouns and adjectives pertained to one of three genders (masculine, feminine, or neuter) and had to be inflected for number (singular, plural) and case (traditionally six: nominative, genitive, dative, accusative, vocative, and ablative). The categories gender, number, and case were represented by one single exponent (portmanteau-morph): thus, e.g., the ending *-i* of *muri solidi*, from *murus* 'wall' and *solidus* 'solid', encoded at the same time the categories masculine, singular, and genitive (or, alternatively, masculine, plural, and nominative). This single exponent, furthermore, varied according to the declension class the word pertained to. This Latin nominal system underwent dramatic restructuring at the level of morphological categories on its way to French: while the number categories remained unchanged, the genders were reduced to two (masculine, feminine), and so were the cases in Old French (rectus, oblique), before being totally abandoned during the Middle French period (roughly, the 14<sup>th</sup> and 15<sup>th</sup> centuries). The formal expression of the remaining categories has also been heavily changed.

Latin, as we have just seen, had three **genders**, masculine, feminine, and neuter, of which only the first two survived in French. Neuter nouns have joined either the masculine (article: *le*) or the feminine (article: *la*) class: Lat. *vinum* 'wine' > Fr. *le vin*, Lat. *gaudium* 'joy' > Fr. *la joie* (via the Latin plural *gaudia*), etc. We can also observe a number of exchanges between masculine and feminine: Lat. *porticus* (fem.) 'colonnade' > Fr. *le porche*, Lat. *dens*, *-tis* (masc.) 'tooth' > Fr. *la dent*, etc. Many of these changes were intended to make the Latin gender system more coherent by establishing a correlation between the final *-a* and feminine and the final *-us* (*-o* in Vulgar Latin) and masculine, a state of affairs still characteristic of most Romance languages. In French, however, the evolution of final vowels has blurred this Vulgar Latin/Romance correlation: on the one hand, due to the deletion of all final vowels

except *-a*, masculine nouns in *-o* merged with feminine nouns in *-e* (cf. V.Lat. *flore* ‘flower’ > Fr. *la fleur*, etc.), while on the other hand a final *-e* resulting from a Latin final *-a* was also added epenthetically to masculine nouns ending in difficult consonant clusters (cf. V.Lat. *comite* ‘count’ > *conte*, etc.). The gender of a non-derived noun is thus highly arbitrary in Modern French – except for animates, where it tends to coincide with sex – and only becomes apparent through anaphoric pronouns or agreement with determiners and adjectives.

Since **number** and **case** were intimately linked both in Latin and in Old French, it is convenient to deal with them together. The six cases of Classical Latin had already been reduced to two, a rectus and an oblique, in the popular speech of most parts of the Roman Empire, among them Gaul. The rectus formally and functionally corresponded to the classical nominative and also served as vocative, while the oblique, which formally goes back to the accusative, took over most functions of the other cases, either in isolation or in combination with a preposition. Prepositional phrases and cases, by the way, coexisted from the earliest times, but the former grew more and more important with the decline in the case system. The cases that resisted best were the most frequent ones, nominative and accusative (cf. Martínez Moreno 1993: 112–116). Number, on the contrary, was not paralleled by analytic constructions and encoded by inflections throughout the history of Latin.

This is essentially the situation that we still find in Old French, where most nouns obey one of the two patterns displayed in Tab. 157.1:

	singular	plural
	rectus oblique	rectus oblique
m. ‘wall’	<i>murs</i> <i>mur</i>	<i>mur</i> <i>murs</i>
f. ‘door’	<i>porte</i> <i>porte</i>	<i>portes</i> <i>portes</i>

Tab. 157.1: Case and number in Old French

As one can see, in the most frequent declension classes there was no one-to-one correspondence between form and function, so that listeners had to rely on syntactic context and encyclopedic knowledge in order to correctly determine the syntactic function of a noun. In Middle French, this shaky two-case-

system was abandoned, leading to a system where the final *-s* was interpreted as a plural marker, along the lines of the feminine: sg. *mur*/pl. *murs*, sg. *porte*/pl. *portes*. The break-down of the two-case-system has often been attributed to the deletion of final [s] in the Middle French period, but careful investigations (cf. Schöslar 1984) have shown that the break-down occurred slightly earlier, so that there can be no such causal link.

The deletion of the final [s], however, had important consequences for the expression of number in spoken Modern French. While the written language continues to signal plural by attaching *-s*, most nouns and adjectives have the same form for singular and plural in spoken French: *mur* and *murs* are both pronounced [myr], *porte* and *portes* both [pɔrt]. The only exception to this rule is a closed set of highly frequent irregular nouns and an open set of adjectives in *-al*, where plural is expressed by alternations – due to regular sound-change – as in *cheval* ‘horse’ [ʃval]/pl. *chevaux* [ʃvo], *travail* ‘work’ [travaj]/pl. *travaux* [travo], *national* ‘national’/pl. *nationaux*, etc. Apart from these cases, plural in spoken Modern French is indicated at the phrasal level, mainly by determiners (cf. *la porte* [la pɔrt]/pl. *les portes* [le pɔrt]), but also by the number of an agreeing verb (cf. *leur fils dort* ‘their son is sleeping’ [lœr fis dɔr] vs. *leurs fils dorment* ‘their sons are sleeping’ [lœr fis dɔrm]) or by “liaison” (cf. *leur enfant* ‘their child’ [lœr afɑ̃] vs. *leurs enfants* ‘their children’ [lœrz afɑ̃]).

The evolution of case and number in adjectives was essentially the same as in nouns. As far as gender is concerned, we also see the reduction to masculine/feminine: Lat. *bonus* ‘good’ > O.Fr. *bons*/Lat. *bona* > O.Fr. *bone*, etc. In Old French, though, the neuter was still preserved in predicative position after an impersonal pronoun (Lat. *bonum* > O.Fr. *bon*). In the main class of adjectives, feminine was signalled by a final *-e*, and in some cases co-signalled further by a consonantal alternation: V.Lat. *siccō* > O.Fr. *sec*/V.Lat. *sicca* > O.Fr. *seche*, etc. Furthermore, there were two classes without a gender difference in Old French: the class *aigre* ‘sour’ (from V.Lat. *acrolacra*, with an epenthetic *-e* in the masculine) and the class *grant* ‘big’ (from V.Lat. *grande*). During the Old and Middle French period, many members of the third class joined the first by taking *-e* as a feminine marker (masc./fem. *grant* > masc. *grand*/fem. *grande*, etc.), but at the same time some

members of the first class also joined the second (masc. *larc*, fem. *large* > masc./fem. *large*, etc.). All things considered, however, gender inflection has become more complicated in Old, Middle, and early Modern French, contrary to the general trend towards analyticity. At the written level, Modern French still presents an invariant class and a class characterized by the addition of *-e* in the feminine. In spoken Modern French, however, the situation has changed dramatically as a consequence of the deletion of the final schwa (-*e*) and many final consonants: masc. *petit* ‘small’ [pøti]/fem. *petite* [pøtit], masc. *gris* ‘grey’ [grι]/fem. *grise* [grιz], masc. *blanc* ‘white’ [blã]/fem. *blanche* [blãʃ], etc. This situation has led some linguists to consider the feminine form as the base and derive the masculine forms by deleting the final consonant, a solution that creates more problems than it solves. Semantically, the solution is awkward since the unmarked gender in French is clearly masculine. And formally it makes incorrect predictions with respect to the large group of consonant-final invariable adjectives (cf. masc. *clair* ‘clear’ [kler]/fem. *claire* [kler]) and the smaller group of adjectives where the gender distinction is expressed by a consonantal alternation (cf. masc. *vif* ‘lively’ [vif]/fem. *vive* [viv], etc.). Popular neologisms like *bleuse* [bløz] as a feminine form corresponding to invariable *bleu* ‘blue’ [blø] after the model of masc. *creux* ‘hollow’ [krø]/fem. *creuse* [krøz] seem to indicate that speakers take masculine as the unmarked category and consider the feminine to be formed by the addition of a consonant, where there is such a formal difference. Some generative analyses, on the other hand, consider spoken Modern French to be more or less like Middle French at the underlying level, deriving the surface forms by deletion of final schwa and some final consonants. The orthographic final consonant of *petit* and similar words, by the way, is pronounced not only in the feminine form *petite*, but also if the following word is vowel-initial (cf. *petit enfant* ‘small child’ [pøtit ãfã]). This phenomenon, called “liaison” in French grammar, is dependent in a highly intricate manner on context and register (cf. Encrevé 1988).

## 2.2. Determiners

In the determiner-system, the most drastic difference between Classical Latin and French is constituted by the introduction of the category (in)definiteness expressed with articles (Heinz 1982).

As in most other Romance languages, the French **definitive articles** go back to different forms of the Latin demonstrative pronoun *ille*. Old and Middle French still distinguished rectus and oblique forms in the masculine series (sg. *il/le*, pl. *il/les*), while Modern French has generalized the oblique forms (sg. *le*, pl. *les*). The feminine counterparts have always been *la* and *les*, and before vowel-initial words the singular article has always been simply *l'*. As far as the function of the definite articles is concerned, it should be noted that in Old French it mainly served to refer to somebody or something previously introduced or given by context or encyclopedic knowledge, while its generic use, its use with proper nouns, appositions and in several other contexts constitute secondary developments.

The **indefinite articles** go back to different forms of the Latin numeral *unus* ‘one’. Here too, Old French distinguished rectus (*uns* < Lat. *unus*) and oblique (*un* < V.Lat. *uno*) in the masculine series, of which the oblique form was generalized in Modern French. The feminine has always been *une* (< Lat. *una*). In early Old French the indefinite article was restricted to specific reference, but already around 1200 cases of unspecific reference may be found in certain constructions. Before some kinds of nouns, Old and Middle French allowed the plural indefinite articles *uns/unnes* (e.g., *uns guanz* ‘gloves’), which, however, eventually were replaced by *des* (cf. Mod.Fr. *des gants*). This latter form originally had a partitive meaning: *des guanz*, e.g., meant ‘some of the gloves mentioned’ before it simply came to mean ‘(some) gloves’. Such partitive constructions consisting of the preposition *de* ‘from’ and the definite article also existed in the singular: *manger du pain* in Old French meant ‘to eat from the bread in front of us or mentioned before’, while only later, during the Middle French period, it simply came to mean ‘to eat bread’, thus becoming a variant of the indefinite article before mass nouns (*du* is an amalgamated form of *de* and *le*). This so-called “partitive article” is an innovation that sets French apart from most other Romance languages.

The development of articles in Romance is often presented as proof of the general trend towards analyticity. This is awkward, since (in)definiteness was not expressed at all in Latin. Furthermore, one should not forget that a Modern French article like *la* is highly synthetic in itself, since it expresses at the

same time the categories singular, feminine, and definite.

The Latin system of **demonstratives** was characterized by three degrees of distance: near the speaker (*hic*), near the listener (*iste*), and further away (*ille*). While this system is preserved in several Romance languages, in Old French it has been reduced to a two-degree system: *cist* (< V.Lat. *ecce iste*) for proximity and *cil* (< V.Lat. *ecce ille*) for distance (*ecce* was an interjection meaning roughly 'look!'). Both forms could be used as adjectives and pronouns and had special forms for the expression of oblique and plural. The further evolution (cf. Dees 1971) was characterized by a reduction to a one-degree system and by a specialisation of the two formal series in adjectival (*cist*, Mod.Fr. *ce*) and pronominal (*cil*, Mod.Fr. *celui*) use. In the Middle French period, proximity and distance began to be marked by joining the adverbs *ci* 'here' and *là* 'there', which have become obligatory in Modern French with non-modified pronouns (*celui-ci* 'this one' vs. *celui-là* 'that one', etc.), while with demonstrative adjectives these adverbs are still facultative (*ce livre* 'this book', *ce livre-ci*, *ce livre-là*).

Latin had the following **possessives** (only the masculine singular nominative forms are given): 1. *meus* 'my', 2. *tuis* 'your', 3. *suus* 'his' (reflexive), *eius* 'his' (non-reflexive), 4. *noster* 'our', 5. *vester* 'your', 6. *suus* 'their' (reflexive), *eorum* 'their' (non-reflexive). In Vulgar Latin, *suus* was extended to cover the non-reflexive possessive corresponding to the third person singular and in some languages also to the third person plural, which in the Vulgar Latin of Gaul however was marked with *illorum* (> Fr. *leur*). *Vester* furthermore was changed analogically to *voster* (cf. Fr. *votre*, *vôtre*) after *noster* (cf. Fr. *notre*, *nôtre*). All in all, there were few functional changes on the way to Romance. The most important change from Latin to Old French was a formal one, viz. the rise of a stressed and an unstressed series of possessives: unstressed Lat. *meus/meum* (obl.) became O.Fr. *mes/mon*, while under stress it ended up in O.Fr. *miens/mien*, and similarly for the second and third person singular. In Modern French, the series arising from stressed possessives is only used with the article: *mon/ton/son livre* 'my/your/his book' vs. *le mien/tien/sien* 'mine/yours/his', etc.

### 2.3. Personal, relative, and interrogative pronouns

As we have seen in 2.1, with nouns and adjectives the six Latin cases were reduced to two in Old French and have completely disap-

peared in Modern French. **Personal pronouns**, on the contrary, have preserved three cases – nominative, dative, and accusative – up to this day. For the third person feminine, for example, Modern French distinguishes *elle* (subject, V.Lat. *illa*), *lui* (indirect object, < V.Lat. *illui*), and *la* (direct object, < V.Lat. *la*, Cl.Lat. *illam*). This conservative behaviour is probably due to the high frequency of the personal pronouns that have survived, and shows that for highly frequent units inflexion, especially in the form of suppletion, is the most economical way of achieving brevity (cf. Hunnius 1990: 64). Another conservative trait that distinguishes personal pronouns from nouns and adjectives is the conservation of the neuter in the Old French 3. sg. nom. form *el* (< V.Lat. *illu*, Cl.Lat. *illud*), later on replaced by masculine *il* (cf. *il pleut* 'it is raining'). But the pronoun system also shows some innovative aspects. A French innovation is the indefinite personal pronoun *on* 'one' (cf. *on dit* 'one says'), which has still other functions in colloquial French. *On* goes back to unstressed Lat. *homo* 'man', whose accusative form *hominem* is the ancestor of French *homme* 'man'.

Just like in the case of the possessives dealt with at the end of the last paragraph, Latin personal pronouns have also developed differently in Romance in stressed and unstressed positions (cf. Wanner 1987). So, for example, the Lat. 1. sg. acc. pronoun *me* [me:] remained *me* [mə] in the unstressed pre-verbal position (cf. *vous me regardez* 'You look at me'), while in the stressed final position it was diphthongized to *moi* (cf. *regardez-moi!* 'look at me!'). What distinguishes Modern French – together with some Rhoto-Romance and Northern Italian dialects – from the rest of Romance languages is the fact that it also has a stressed and an unstressed series for subject pronouns. In the other Romance languages – just like in Latin, and in Old French – the categories person and number in the verb are expressed by inflectional suffixes, as we are going to see in 2.4, and personal subject pronouns reserved for the expression of emphasis, contrast or disambiguation. But when subject pronouns became obligatory after the Old French period, the stressed direct object pronouns were extended to serve as stressed subject pronouns too. So, O.Fr. *chant* 'I sing' vs. *je chant* 'I sing' corresponds to Mod.Fr. *je chante* vs. *moi, je chante*.

**Relative and interrogative pronouns** also conserved a dative form up to the Old French period (Lat. *cui* > O.Fr. *cui*), later fused with the nominative form *qui*. The rectus/oblique distinction, on the other hand, has been conserved up to the present day in relative pronouns: Lat. *qui* > Fr. *qui*, Lat. *quem* > Fr. *que*. The Latin neuter interrogative pronoun *quid* became *que* in the unstressed position, while under stress it developed into *quoi*.

#### 2.4. Verbs

The Latin conjugation is much better preserved in Romance than the Latin declension (cf. Geckeler 1996: 213–220). So its four conjugation classes are all preserved in French, though many verbs changed class in Vulgar Latin or later on. The most spectacular innovation undoubtedly consisted of the replacement of many Latin synthetic categories, especially the less frequent ones, by analytical ones. Synthetic and analytic forms coexisted during many centuries, and if the latter eventually evinced the former this might be due to their greater simplicity, and hence learnability, a big advantage in the Roman Empire, where millions of non-Latin speakers had to learn the official language in precarious conditions (cf. Wuest 1997).

Latin verbs, in their finite forms, were inflected for person/number, tense/aspect, mood, and voice. In the following discussion, the fate of these categories in the passage from Latin to French will be outlined.

**Person** and **number** are realized by a portmanteau-suffix located at the end of the verb in Latin as well as in Romance (see Tab. 157.2). Written Modern French, which is by and large identical to Old French except for the obligatory presence of the personal pronoun, is still very similar to the Latin/Pan-Romance type, while spoken Modern French has person/number-endings only for the first and second person plural (4 and 5).

Latin	Modern French	
	written	spoken
1 <i>cant-o</i>	<i>je chant-e</i>	[ʒɔ̃ʃãt]
2 <i>canta-s</i>	<i>tu chant-es</i>	[tyʃãt]
3 <i>canta-t</i>	<i>il chant-e</i>	[i(l)ʃãt]
4 <i>canta-mus</i>	<i>nous chant-ons</i>	[nuʃãtõ]
5 <i>canta-tis</i>	<i>vous chant-ez</i>	[vuʃãtɛ]
6 <i>canta-nt</i>	<i>ils chant-ent</i>	[i(l)ʃãt]

Tab. 157.2: Person and number in Latin and French verbs

From a functional point of view, it might seem that the obligatory – if there is no subject noun – presence of the subject pronoun is due to the ambiguity deriving from the phonetic erosion of the person/number-endings in spoken French, but historians of the French language assure that chronology forbids postulating a relationship of cause and effect between the two. Another interpretation of the facts in Tab. 157.2, very popular in the sixties, also has to be rejected, namely the view that in French the Latin person/number-suffixes have been replaced by person/number-prefixes: in reality, *je*, *tu*, etc. do not have the behaviour of prefixes but of clitics, since they may be separated from the stem by other pronouns, as in *je vous la chante* ‘I sing it to you’, and have conserved a certain syntactic mobility, as in *je dis* ‘I say’ vs. *dis-je* ‘say I’ (cf. Hunnius 1977; Koch 1993).

The three-way distinction in past-present-future of the Latin **tense/aspect**-system has been preserved in French. There have been no changes in form (beyond, of course, the effects of regular sound-laws and occasional analogies) or function in the **present indicative** (Lat. *cantabat* > Fr. *il chantait*, etc.). As far as the **perfect indicative** (called “passé simple” in French grammar) is concerned, the most noteworthy development was its total abandonment in the spoken language. There is no agreement among specialists as to when and why this important change took place. In the spoken language, the **passé simple** has been totally replaced by the **passé composé**, a Vulgar-Latin and hence pan-Romance innovation. This periphrastic tense has arisen from Latin resultative constructions of the type *habeo litteras scriptas* ‘I have written letters’, where *scriptas* ‘written’ originally had an adjectival function before it was reanalysed as part of the verb phrase, with *habeo* ‘I have’ becoming an auxiliary verb. As a consequence, agreement in gender and number between the original adjective converted into a participle and the following noun was abandoned: *j'ai écrit* (and not: *écrîtes*) *des lettres* (cf. Loporcaro 1998). Of the Latin **pluperfect indicative** of the type *cantaveram* some traces can be found in Old French, but it has since been completely replaced by a periphrastic construction of the type *j'avais chanté*, composed of the auxiliary verb *avoir* ‘to have’ in the imperfect tense and the past participle. Immediate antRIORITY in the past is expressed

in literary French by the periphrastic **passé antérieur**, a Romance innovation consisting of the passé simple of the auxiliary and the past participle: *j'eus chanté*. If the auxiliary verb is put in the periphrastic past, one gets the **passé surcomposé**: *Je suis parti quand j'ai eu terminé* 'lit. I have gone away when I have had finished' (cf. Holtus 1984). The Latin **future** tense of the type *cantabo* had already been replaced by a periphrastic construction in Vulgar-Latin times, consisting of the infinitive followed by the present tense of *habere* 'to have' (Old French, though, still conserved synthetic future forms for the high-frequency verb 'to be': *(i)er*, *(i)ers*, etc.). This periphrastic construction was then again amalgamated to a synthetic conjugation, a process already completed in Old French (cf. Kuen 1952): V.Lat. *cantare habeo* 'lit. I have to sing' > V.Lat. *cantaraio* > O.Fr. *chanterai*. In Modern French, this synthetic future is again challenged by a new periphrastic future, consisting of the auxiliary verb *aller* 'to go' and the infinitive (*je vais chanter* 'lit. I go to sing'). The Latin **futurum exactum** (past in the future) of the type *cantavero* has also been abandoned and replaced by a periphrastic construction consisting of the future tense of *avoir* and the past participle (*j'aurai chanté* 'I will have sung'). A Romance innovation closely linked to the periphrastic future is the **conditional**, originally a kind of future of the past formed with the infinitive followed by the imperfect tense of *habere*, but later on amalgamated to a new synthetic conjugation along the lines of the future: V.Lat. *cantare habebat* > Fr. *il chanterait* (cf. *il a dit qu'il chanterait* 'He said that he was going to sing'). It is probably no coincidence that the only periphrastic forms which amalgamated were the ones where the auxiliary was postponed to the verb. The conditional later also developed a modal use unknown to Latin: *il chanterait* 'he would sing (if ...)'.

As far as **mood** is concerned, the four Latin subjunctives have been preserved in French as morphological categories, but only the **present subjunctive** still conserves the Latin forms: Lat. *cantem* > (*que*) *je chante* (O.Fr. *chant*, the *-e* is analogical), etc. The **imperfect subjunctive** in Modern French only survives in very stilted literary language and formally goes back to the Latin pluperfect subjunctive: Lat. *cantavissem* > Fr. (*que*) *je chantasse*, etc. The Latin **perfect subjunctive** of the type *cantaverim* has been replaced by a periphrasis composed of the present subjunctive of the

auxiliary and the past participle: (*que*) *j'aie chanté*, etc. The **pluperfect subjunctive**, whose forms, as we have seen, are used as imperfect subjunctive in Romance, is also expressed peripherastically in French, combining the imperfect subjunctive form of the auxiliary with the past participle: (*que*) *j'eusse chanté*, etc. It is as marginal in Modern French as the imperfect subjunctive.

Of the different Latin **imperatives**, only the imperative of the singular has conserved the Latin forms: Lat. *canta* > Fr. *chante*, etc. The imperative of the second person plural has been replaced by the corresponding indicative ending (cf. Lat. *cantate* vs. Fr. *chantez*, etc.), and the negative imperatives are formed with the negative particle *ne ... pas*: Lat. *ne cantaveris*, *ne cantes*, *noli cantare* vs. Fr. *ne chante pas*, etc. The Latin future imperative of the type *cantato* has disappeared completely.

Changes have been equally drastic in the domain of infinite verb forms. Latin had a wealth of **infinitives**: present, perfect, and future, all of them in an active and a passive variant. Of these, only the present active infinitive is conserved both in function and form: Lat. *cantare* > Fr. *chanter*, etc. The present passive infinitive has been replaced by a periphrastic form (Lat. *cantari* vs. Fr. *être chanté*), and so was the perfect active infinitive (Lat. *cantavisse* vs. Fr. *avoir chanté*). The perfect passive infinitive (*cantatus/atum esse* 'to have been sung') was already periphrastic in Latin, but has been adapted in French as *avoir été chanté(e)*, since the formal equivalent *être chanté(e)* means 'to be sung'. The future infinitives both disappeared. Of the three Latin **participles**, past, present, and future, the latter has disappeared completely. The past participle has been conserved and has even strongly extended its use due to the many periphrastic tenses: Lat. *cantatus/la* > Fr. *chanté(e)*, etc. The Latin present participle was replaced in Late Latin by the ablative of the gerund, which in French later on was used to translate Latin present participles and so got a participial character (cf. Lausberg 1962: 197): *ses mains tremblant de peur* 'his hands trembling with fear'. The Latin gerund is also conserved in the form V.Lat. *in cantando* > Fr. *en chantant*. The Latin future participle has disappeared completely, and so have the two forms of the Latin **supine**.

In the category **voice** Latin distinguished active and passive. **Passive** was expressed by special inflectional forms in the imperfect,

present, and future tenses, as well as in the infinitive present and perfect, as we have already seen. For the future infinitive as well as the perfect and pluperfect tense, passive was realized peripherastically. In French, as generally in Romance, synthetic passives have disappeared completely, maybe as a consequence of the phonetic erosion of some of the endings (Togeby 1980: 136). Thus, there are also no longer any deponent verbs in French, i.e. active verbs with passive inflection, typical of Latin. Passive now is formed peripherastically with *être* ‘to be’ and the past participle, the tense of the auxiliary being identical to the tense of the periphrastic construction: *la chanson est/la été/sera/etc. chantée* ‘the song is/was/will be/etc. sung’, etc. Alternatively, in certain contexts, the reflexive construction can be used instead of the periphrastic passive: *cette chanson se chante beaucoup* ‘lit. this song sings itself very much’, etc.

### 3. Word-formation

Word-formation has also undergone deep changes from Latin to French, but it cannot be said that word-formation is more analytic in French than in Latin. Though some patterns have remained more or less unchanged over the centuries, many others have been abandoned or have undergone semantic change or reanalysis, and some French patterns may be considered as innovations (see also Art. 121). From the Middle French period onwards, however, French and Latin word-formation have again become more similar as a consequence of massive borrowing of Latin words on the part of French. Besides Classical Latin, also Neo-Latin, English and, to a lesser extent, other European languages contributed to the renovation of French vocabulary, creating in this way many correspondences with word-formation in Latin and other European languages, to the extent that someone has even spoken of the emergence of a “euromorphology” (Schmitt 1996; Rainer 2002). Another general characteristic that sets Old and Middle French apart from Modern French is the wealth of synonymous derivations in the former stages of the language; this obviously is due to the fact that the language had not yet gained complete geographic homogeneity in those times and that normative pressure was much lower.

#### 3.1. Compounding

Latin had relatively few **compounds** in comparison to Greek or Germanic languages (cf. Lindner 2002). Only very few compounds, furthermore, were passed on to French directly and those that were often lost their transparency on the way, like Lat. *undecim* ‘eleven’ (*unus* ‘one’ + *decem* ‘ten’) > Fr. *onze*, etc. Modern French compound types are thus in general later innovations, due in many cases to loan-translations. Only the more important types can be mentioned here.

One of the most hotly debated questions in Romance word-formation is the origin of verb-noun compounds of the type *essuie-glace* ‘windshield-wiper’, lit. ‘wipe-windshield’ (cf. Bierbach 1982). Traditional wisdom attributes its origin to a reanalysis, in early Romance, of imperative sentences as proper names, whence the type is supposed to have been extended to common nouns. This developmental sequence, however, is purely speculative, since the first examples, going back to the 11<sup>th</sup> and 12<sup>th</sup> centuries in French, already show a great variety of semantic types. It has thus been recently proposed (cf. Bork 1990) to see in this Romance compound type the descendant of the Latin compound type *poscinummius* ‘money-asking’ (from *poscere* ‘to ask’ and *nummus* ‘coin’), which was influenced by the corresponding Greek type. Only sixteen, mostly marginal Latin compound words of this type, however, are documented, and none of them has directly survived in Romance.

Noun-noun compounds – where the first noun is the head, contrary to English – became more important in French only in the 19<sup>th</sup> century, but are quite frequent nowadays in the written language. They do not constitute one homogeneous type, but rather a conglomerate of subtypes with varying semantics and productivity. This synchronic heterogeneity reflects the different origin of the subtypes. *Hôtel-dieu* ‘main hospital’, e.g., goes back to an asyndetic Old-French genitive construction meaning ‘lodging (*hôtel*) of god (*dieu*)’. At least from the 16<sup>th</sup> and 17<sup>th</sup> centuries onwards we find copulative compounds of the type *imprimeur libraire* ‘lit. printer librarian’ or *comédie-ballet*. The Romantic poets liked compounds with a metaphor as second element: *cité-cadavre* ‘a town (*cité*) that is like a corpse (*cadavre*)’, etc. But the great vogue of noun-noun compounds and the diversification of their semantics in the 19<sup>th</sup> and 20<sup>th</sup> centuries is attributable to the influ-

ence of English and German, manifest in many loan-translations like *éetalon-or* ‘gold (or) standard (*éetalon*)’, etc. A thorough investigation of the history of this compound type is badly lacking.

The most frequent nominal compounds in French are the ones formed by noun-preposition-noun or adjective-noun/noun-adjective, like *pomme de terre* ‘potato, lit. apple-of-earth’, *moulin à vent* ‘windmill, lit. mill-by-wind’, *casque bleu* ‘blue helmet, lit. helmet blue’, etc. It seems better, however, to treat such sequences as essentially syntactical constructions.

Some subordinative adjective-adjective (or adverb-adjective) compounds may already be found in Old-French (*nouveau-né* ‘new-born’, *clair-voyant* ‘clear-sighted’, etc.), but the pattern has remained marginal. The copulative adjective-adjective compound type *aigre-doux* ‘sour-sweet’ was adapted directly from Greek by the Pléiade poets in the sixteenth century, but never did become very productive in standard French (Hatcher 1951: 11 f.). The copulative type *éthico-moral*, on the contrary, is due to adaptations of Neo-Latin compounds such as *ethico-moralis* from the 17<sup>th</sup> century onwards (Hatcher 1951: 135–137). As a consequence of its Neo-Latin origin, the first element of this type of compound is often a learned allomorph, like *franco-* for *français*, etc.

The Vulgar Latin noun-verb compound *\*manutenerē* ‘to hold (*tenere*) with the hand (*manu*)’ survived in French as *maintenir*, and so did some other compound verbs of this type (cf. Klingebiel 1989). In Old and Middle French, a few neologisms based on this model like *saupoudrer* (from *sau* (Mod.Fr. *sel*) ‘salt’ and *poudrer* ‘to strew over’) are attested, but the type never became really productive.

Many Latin compounds were borrowed by French in the late Middle Ages and even more Neo-Latin compounds during the Modern Times. Due to the changes in the vernacular language, most such compounds however have remained more or less opaque to the average speaker. While Latin speakers could easily analyse *agricola* ‘farmer’ into *ager* ‘field’ + linking vowel *-i-* + *colere* ‘to till’ + ending *-a*, Fr. *agricole* ‘agricultural’ is much more opaque for speakers of French, where ‘field’ now is *champ* and ‘to till’ *labourer*. Where many (Neo-)Latin compounds of the same make-up were borrowed, however, these sometimes have given rise to a

productive pattern in Modern French. Such was the case, e.g., with compounds with a second element *mania* (Höfler 1972), which entered the French language in loan-translations like *démonomanie*, *bibliomanie*, etc. and then came to be attached even to unequivocally French bases as in *bureaumanie*, *peinturomanie* (from *peinture* ‘painting’), etc. Elements like the *-omanie* of *peinturomanie*, however, should better be analysed as suffixes from a synchronic point of view (cf. also Rainer 2003 on the integration of the Latin type **aurifer**).

### 3.2. Derivation

#### 3.2.1. Prefixation

Latin had a rich system of **prefixes**, many of them with a clear link to corresponding prepositions or adverbs, which diachronically constitute the most important source for prefixes. It is interesting to note, however, that the opposite evolution, from a prefix to an adverb, is also possible. Lat. *trans-* ‘across’, e.g., has become the French adverb *très* ‘very’, and the same process of adverbialisation of a prefix may be observed in Old French *par* (Lat. *per-*, cf. *permagnus* ‘very big’) and *re* (Lat. *re-* ‘again, back’, cf. *revenire* ‘to come back’): *par est granz* ‘lit. very it-is big’, *r-est venuz* ‘lit. back he-has come’ (Meyer-Lübke<sup>2</sup> 1966: 139 f.). Due to this close relationship between prefixes on the one side and prepositions and adverbs on the other, it is not always easy to draw a neat borderline between prefixation and compounding.

Only relatively few deverbal verb-forming prefixes were passed on directly to French. The most productive ones are *re-* (< Lat. *re-*; *refaire* ‘to do again’, etc.) and reversative *dé(s)-* (< Lat. *dis-*; *défaire* ‘to undo’, etc.). The others are much rarer: Lat. *ad-* > Fr. *a-* (*apporter* ‘to bring; lit. here-carry’, etc.), *contra-* > *contre-* (*contredire* ‘to contradict’, etc.), *ex-* > *é-* (*s'écouler* ‘to flow off’, etc.), *in-* > *en-* (*enfermer* ‘to enclose’, etc.), *inter-* > *entre-* (*s'entreuer* ‘to kill each other’, etc.), *pro-* > *pour-* (*poursuivre* ‘to pursue’, from *suivre* ‘to follow’, etc.), *super-* > *sur-* (*surcharger* ‘to overload’, etc.). Lat. *sub-* is only present in some lexicalized forms like (*se*) *souvenir* ‘to remember’ (< Lat. *subvenire*), while in productive formations it has been replaced by *sous-* (*sous-estimer* ‘to underestimate’, etc.), the follower of *subtus*, or by Latinate *sub-* (*subdiviser* ‘to subdivide’, etc.). Other Lat-

nate deverbal prefixes are *post-* (*postposer* ‘to postpone’, etc.), *pré-* (*prévoir* ‘to foresee’, etc.), *super-* (*superposer* ‘to super(im)pose’, etc.), *trans-* (*transporter* ‘to transport’, – the popular outcome *tres-* is still present in lexicalized *trépasser* ‘to die’), and some others. For prefixes in parasynthetic verbs see 3.2.2.

Some of the Latinate prefixes just mentioned also occur before nouns and adjectives. Other Latinate denominal prefixes are *bis-* (*bisaïeu* ‘forefather’), *co-* (*copilote* ‘co-pilot’, etc.), *ex-* (*ex-mari* ‘ex-husband’, etc.), *non-* (*non-intervention* ‘non-intervention’, etc.), *vice-* (*vice-chancelier* ‘vice-chancellor’, etc.), etc. As far as the deadjectival prefixes are concerned, one must mention, among others, the negative prefixes *in-* (*incroyable* ‘incredible’, etc.) and *anti-* (*anti-atomique* ‘anti-atomic’, etc.), as well as the intensive prefixes *archi-*, *extra-*, *super-*, and *ultra-*. Many such “Latinate” formations, of course, were not borrowed directly from Latin but from other European languages, mainly English. Popular transmission, here too, is much rarer: *minus-* > *mé(s)-* (*mécontent* ‘discontent’, etc.), and some others. Some French prepositions and adverbs have also developed prefixal uses: *après-midi* ‘afternoon’, *arrière-garde* ‘rear guard’, *avant-garde* ‘vanguard’, *contre-sens* ‘nonsense’, *entretemps* ‘entremets; lit. between-dishes’, *sous-officier* ‘non-commissioned officer’. As one can see, such formations may be exocentric (the *après-midi* ‘afternoon’, for example, is not a *midi* ‘midday’). There is even one adjective that has been converted into a prefix: *meidius* > *mi-* (*la mi-février* ‘mid-February’, etc.).

All in all, one can see that French prefixation closely resembles that of Latin and other European languages due to intensive borrowing from Latin and other European languages.

### 3.2.2. Parasynthesis

**Parasynthesis** is a term used in Romance word-formation for what seems to be the simultaneous application of two word-formation rules to a base. *Enrichir* ‘to enrich’, e.g., where neither *\*enrich* nor *\*richir* are independently existing words of the language, is generally believed to be formed by the simultaneous adjunction of *en-* and *-ir* to the base *rīche* ‘rich’. Since *-ir* marks the infinitive and hence is an inflectional ending, some prefer to define this kind of formation as formed through the simultaneous application of *en-* prefixation and conversion, while others hold

that the whole derivational process is located in the – then category-changing – prefix. Under this last hypothesis, parasynthesis is reduced to simple prefixation.

Parasynthesis is most prominent in verbs. The type is already present in Old Latin and became particularly frequent in Late Latin (Crocco Galéas & Iacobini 1993): *clarus* ‘clear’ > *acclarare* ‘to clarify’, *uncus* ‘hook’ > *inuncare* ‘to hook’, etc. Originally the prefix of such verbs was a preposition, so that the whole could be viewed as a derivation on a phrasal base: *per noctem* ‘over night’ > *pernoctare* ‘to stay overnight’, etc. In the course of time, however, such formations were reanalyzed as consisting of nominal or adjectival bases “circumfixed” by a combination of a generally desemanticized prefix and conversion. In French, e.g., the prepositional phrase corresponding to *cadre* ‘frame’ > *encadrer* ‘to frame’ is (*mettre*) *dans un cadre*, not *en cadre*, that of *lune* ‘moon’ > *alunir* ‘to land on the moon’ is (*aborder*) *sur la Lune*, not *à la Lune*, and in *courage* ‘courage’ > *décourager* ‘to discourage’ or in deadjectival formations such as *long* ‘long’ > *allonger* ‘to make longer’ no plausible prepositional phrase is in sight anyway. Parasynthetic verbs have been very productive throughout the history of French.

### 3.2.3. Suffixation

French suffixation is so rich that it will only be possible to mention the most important categories and suffixes here.

Nominal derivatives are by far the largest group. In a first subgroup, the base is also a noun and the derivative only implies a slight modification.

**Collective nouns** did not form a well-established category in Latin. French (Baldinger 1950) inherited some Latin suffixes like *-alia* (Fr. *-aille*; *pierraille* ‘crushed stone’) or *-eta* (Fr. *-aie*; *hétraie* ‘beech forest’), but most formations in French arose through semantic extensions or reanalysis of other derivational patterns. Such was the case, among others, of *-age* (*feuillage* ‘leafwork’, etc.), which goes back to neuter nominalized adjectives in *-atium* and of extensions from deadjectival and deverbal abstracts (*argenterie* ‘silverware’, *gouvernement* ‘government’, etc.).

The characteristic Latin feminine ending *-a* (alternating with masculine *-us*) is reflected in French *-e*: Lat. *ursus* ‘bear’/*ursa* ‘she-bear’ > Fr. *ours/ourse*, etc. A suffix that gained great momentum in Old French was *-esse* (< Lat.

*-issa): comtesse* ‘countess’, etc. In the 16<sup>th</sup> century, however, *-eur* took over the feminine ending *-euse* from *-eux* instead of traditional *-eresse* (Lindemann 1977). For loan-words, the Latin feminine ending *-trix* was adapted as *-trice*. Many feminine forms can be expressed only in a rather roundabout way in French, such as *Madame le ministre*; feminists, however, insist that the corresponding feminine forms should be imposed.

One of the most salient traits of Romance word-formation, inherited from Latin, is its wealth of **diminutive suffixes** (cf. also Art. 99). Old French fits into this picture, while in Modern French only the diminutive suffix *-ette* is marginally productive in journalistic and commercial language, less so in spontaneous conversation. The abandonment of diminutive suffixes – augmentatives have never been very productive in French – and their substitution with the adjective *petit* ‘small’ was a process that began in Middle French and came to an end in the 17<sup>th</sup> century, after a short Renaissance in the poetry of the Pléiade (cf. Eckert 1986: 284–318). The reasons for the agony of diminutives in Middle and early Modern French are difficult to pin down and have been the subject of much debate: homonymy of *-et* with several other suffixes, after final consonants having become mute, may have been one decisive factor. It should not be forgotten, however, that French has created a new reduplicative diminutive type (see 3.2.5), and that popular French and slang have several “evaluative” suffixes of dialectal origin like *-uche* (Baldinger 1997), which do not have any denotative meaning but only serve to give the word a particular, often slang connotation.

Another group of denominal nouns is that of **status suffixes**, i.e. suffixes that express the status of the person referred to by the base. Latin had quite a number of suffixes with that function (*consulatus* ‘consulate’, etc.), but only very few were directly passed on to French, among them *-ia* (*mairie* ‘office of mayor’, etc.). French however created new status suffixes like *-age* (*baronnage* ‘barony’, etc.) or later on reintroduced some of the Latin suffixes (*professorat* ‘professorship’, *préfecture* ‘charge of a préfet’, etc.). Many such formations have locative, temporal or collective extensions; so the *préfecture*, nowadays, is more commonly the place where the *préfet* has his office.

**Personal nouns** also constitute an important class of denominal nouns. Among the

names for traditional social roles we find several suffixes that were passed on directly from Latin to French: *-ain* (< Lat. *-anus*; *châtelain* ‘lord of a castle’, etc.), *-on* (< Lat. *-o*, *-onis*; *charron* ‘cartwright’, etc.), *-eron* (from *-ier* + *-on*; *forgeron* ‘smith’, etc.), and, most important, *-ier* (cf. Roché 1998), which arose through an ellipsis in collocations such as *faber ferrarius* ‘smith’ (Lat. *carbonarius* > Fr. *charbonnier* ‘charcoal (*charbon*) burner’, etc.). In more modern professional spheres, *-ier* was ousted later on by learned *-iste* (< Lat. *-ista*): *tracteuriste* ‘tractor driver’, etc. Equally learned are *-ien* (< Lat. *-ianus*; *mathématicien* ‘mathematician’, etc.) and *-aire* (< Lat. *-arius*; *fonctionnaire* ‘civil servant’, etc.). The suffix *-iste* also commonly refers to the follower of a person or doctrine: *marxiste*, etc. The same holds for the more colloquial *-ard* (from Germanic *-hard*; *dreyfusard* ‘follower of Dreyfus’, etc.), in one of its several uses.

The last important group of denominal nouns is that of **instrumental** and **locative nouns**. The borderline between these two categories is not always very neat, especially with nouns referring to containers. We thus sometimes find the same suffix for both subgroups, as with *-ier* (masc.) and *-ière* (fem.), going back to Lat. *-arius* and *-aria*: *cendrier* ‘ashtray’, *cafetièr* ‘coffee-pot’, etc. The gender of the derivative is in most cases the opposite one of that of the base. Purely locative are Old Fr. *-oi* (< Lat. *-etum*; *sapinoi* ‘fir wood’, etc.) and synonymous *-aie*, which has already been mentioned under the collective suffixes. The most important locative suffix is *-ie*, originally an extension of the homonymous status suffix: *boulangerie* ‘bakery’, etc. Through the reanalysis of words like *tuilerie* ‘tilery’, which could not only be referred plausibly to *tuilier* ‘tile-maker’ but also to *tile* ‘tile’, a new suffix *-erie* was created.

In the class of **deverbal nouns** we again find many of the suffixes and categories just treated. The category of **agent nouns** is dominated by *-(at)eur* (< Lat. *-(at)or*): *vainqueur* ‘winner’, *exportateur* ‘exporter’, etc. The Latinate allomorph in *-ateur* is chosen if the corresponding action noun ends in *-ation*. From the 18<sup>th</sup> century onwards this suffix has also become by far the most important deverbal **instrumental suffix** (cf. Spence 1990): *condensateur* ‘condenser’, etc. The traditional instrumental suffixes like *-ail* (< Lat. *-aculum*; *épouventail* ‘scarecrow’, etc.), or *-oir* (< Lat. *-orium*; *rasoir* ‘razor’, etc.) have lost ground.

Here too we find the place/instrument ambivalence noted with respect to the denominal formations: *abrevoir* ‘watering place’, *baignoire* ‘bathtub’, etc.

The huge category of **action nouns** has undergone severe changes from Latin to French. The suffix *-(at)io*, *-onis* (applied to the past participle stem) survived as *-aison*, but was later on pushed aside by the learned *-ation* (cf. Schmitt 1988). Lat. *-ura* survives in Fr. *-ure* (*ouverture* ‘opening’, etc.). Latin action nouns in *-us* (attached to the past participle stem) survive in the feminine form (due to a reinterpretation as feminine singular nouns of neuter plurals; Alsdorf-Bollée 1970): *mise* ‘putting’, *entrée* ‘going in’, *sortie* ‘going out’, etc. French formations like *mise* are interesting examples of form-function-anisomorphism: they are action nouns, but have the form of feminine perfect participles. The most productive action noun suffix during the history of French has been *-ment*, whose Latin antecedent *-mentum* had been relatively rare in this function until Late Latin. French *-ance* (*naissance* ‘birth’, etc.) goes back to another Late Latin creation, *-ntia*, which was due to a coalescence of the present participle in *-nt-* + *-ia*. Old Fr. *-ëiz* (cf. Malkiel 1986), which survives in some Modern French derivatives like *clapotis* ‘ripple’, and Fr. *-age* (cf. Fleischman 1977) eventually go back to the Latin adjectival suffixes *-atīcius* and *-aticus*. Action nouns in *-erie* like *rêverie* ‘dreaming’ arose when speakers reanalysed denominal formations like *rêveur* + *-ie* as deverbal (*rêv(er)* + *-erie*). Another important means of forming action nouns is conversion (see 3.2.4).

Latin had a lot of suffixes for forming **deadjectival abstract nouns**. Unstressed *-ia* was subject to phonetic erosion and only survives in few words like *force* (from V.Lat. *fortia*) as well as with bases going back to present participles (*importance* ‘importance’, etc.). It has been replaced by the stressed *-ia*, which underlies Fr. *-ie*: *jalousie* ‘jealousy’, etc. Deverbal *-or*, which had already been reanalysed as deadjectival in Latin, was passed on to French as *-eur*: *hauteur* ‘height’, etc. Lat. *-tas* is present both in popular derivatives like *bonté* and in learned ones like *latinité* ‘latinity’. The same is true for *-itia*, which we find as *-esse* (*richesse* ‘richness’, etc.), *-ise* (*franchise* ‘exemption’, etc.), and *-ice* (*justice* ‘justice’). Lat. *-tudo* was changed to *-tume*, as in *amertume* ‘bitterness’. The extremely polysemous suffix *-erie* is also found with deadjecti-

### XVIII. Morphologischer Wandel II: Fallstudien

val abstracts like *pruderie* ‘prudery’. Last but not least, *-isme* (< Lat. *-ismus*) may also be used to form abstract nouns (*conservatisme* ‘conservatism’, etc.), besides its many other uses.

**Adjectival derivatives** (cf. Gawełko 1977) may be classified in deadjectival, deverbal, and denominal both in Latin and French. The latter group may again be subdivided into qualitative and relational adjectives, though the suffixes are partly the same.

Among the **denominal adjectives** the most important category is that of **relational adjectives**. Of the numerous Latin adjectives of this type only relatively few survived as such in Old French: Lat. *carnalis* > Fr. *charnel*, Lat. *aquaticus* > O.Fr. *evage*, Lat. *marinus* > Fr. *marin*, etc. Their decline was probably due, in part at least, to the fact that many adjectives had been converted into nouns. In the course of the centuries, and especially from the 18<sup>th</sup> century onwards, however, relational adjectives were borrowed massively from Latin, Neo-Latin, and other European languages, so that in present-day, particularly written French they again constitute a very prominent category. The same is true, essentially, for the subcategories of adjectives derived from place names (cf. Wolf 1964) and personal names (cf. Schweickard 1992). Non-relational denominal adjectives express resemblance (*-esque*, etc.), possession (*-u*, etc.), and support (*-iste*, etc.), among others.

**Deverbal adjectives** have witnessed great changes. Of Lat. *-abilis/-ibilis* only the former was passed on directly to Old French, while the latter was reintroduced later on through loan-words. The active sense of the suffix (*durable* ‘durable’, etc.) was more prominent in the old language, while later on it specialized in the passive meaning (*faisable* ‘feasible’, etc.), evincing completely the competing Old French suffix *-(e)is*. In Old French, denominal *-able* was also relatively prominent; this had arisen through the reanalysis of ambiguous formations like *taillable* ‘obliged to pay the *taille* (a tax)’, originally derived from the verb *tailler* ‘to subject to the *taille*’. Another prominent active deverbal suffix in Old French was *-eux* (Lat. *-osus*), later on reduced to a more marginal position by *-(at)eur* (Lat. *-(at)or*), which is the most important suffix with an active sense in Modern French, together with *-ant* (*important* ‘important’, etc.), which goes back to the Latin present participle. With respect to the originally deverbal *-(at)if* and *-(at)oire* it is noteworthy that

many formations were reanalysed as denominal, like *circulatoire*, which does not mean ‘that circulates’ but ‘related to circulation’. Among the suffixes of Germanic origin, *-ard* was the one that gained the greatest importance.

**Deadjectival adjectives** are rare in French. The Latin comparative suffix *-ior* has disappeared and has survived only in several residual forms like *majeur* (< Lat. *maior*), while the Latin superlative suffix *-issimus* has disappeared completely. Both have been replaced by analytic constructions: Lat. *grandior* ‘bigger’ ≈ Fr. *plus grand*, Lat. *grandissimus* ‘the biggest’ ≈ Fr. *le plus grand*, etc. The intensive use of Lat. *-issimus*, which was so firmly reestablished in Italian or Spanish, never became very productive in French (*richissime* ‘very rich’, etc.). Equally low is the productivity of approximative suffixes like *-ot* (*viellot* ‘a bit old’, etc.) or *-âtre* (< Lat. *-aster*; *blanchâtre* ‘whitish’, etc.).

Latin had four **adverbial suffixes** (cf. Karlsson 1981) and several adverbial phrases consisting of adjective + noun. None of the suffixes has remained transparent in French, but the adverbial phrase consisting of adjective (in the ablative) + *mente* ‘with an adj. mind’ gave rise to the very productive adverbial suffix *-ment*. In Old French, the possibility of omitting *-ment* after the first adjectival base of two coordinated adverbs (*humble et doucement* ‘lit. humb- and softly’) still betrayed the syntactic origin of the suffix.

Of the Latin **verbal suffixes**, *-icare* and *-iare* had to succumb to the erosive force of the sound-laws: Lat. *caballicare* ‘to ride’ > Fr. *chevaucher*, Lat. *acutiare* ‘to sharpen’ > Fr. *aiguiser*, etc. The intensive type *cantare* (Fr. *chanter*) also became opaque when the base *canere* ‘to sing’ went out of use. A verbal suffix, on the contrary, destined to a great future was *-idiarel-izare*, borrowed from Greek. In the popular transmission it became *-oyer*, a suffix quite productive in Old French (*guerre* ‘war’ > *guerroyer* ‘to make war’, etc.). It has gained much more importance in the modern language since its reintroduction as Latinate *-iser* ‘-ize’. The suffix *-ifier* ‘-ify’ has also been reintroduced through latinisms. Besides these denominal and deadjectival suffixes, French also has a small series of deverbal suffixes that modify the *Aktionsart* of the base: *sauter* ‘to jump’ > *sautiller* ‘to make small and repeated jumps’, *crier* ‘to shout’ > *criailler* ‘to shout continuously’, *tousser* ‘to cough’ > *toussoter* ‘to cough slightly’, etc.

### 3.2.4. Conversion

**Conversion**, i.e. the transposition of a word from one part of speech to another without formal change, was not unknown to Latin, but became even more important in French.

Most conversions yield a noun. Many Latin adjective-noun conversions – often of elliptical origin – survive in French as simple nouns, since the corresponding adjective has disappeared: Lat. *alba* ‘dawn’ (related to *albus* ‘white’) > Fr. *aube*, etc. In other cases, the conversion of relational adjectives into nouns has given rise to new nominal suffixes: Lat. *-arius* > Fr. *-ier*, Lat. *-aticus* > Fr. *-age*, etc. Adjective-noun conversions (cf. Malkiel 1938), of the type *le rouge* ‘lit. the red (colour)’, *l'essentiel* ‘lit. the essential (thing)’, *le ridicule de sa conduite* ‘lit. the ridiculous(ness) of his conduct’ or *un curieux* ‘a curious (person)’ have gained importance in Modern French, partly as a consequence of the imitation of Latin usage on the part of the Pléiade poets. As far as verb-noun-conversions are concerned, in Old French the infinitive could freely be used as a noun, while in Modern French the process has become unproductive and only few lexicalized conversions of this type like *un baiser* ‘a kiss’ have survived. Action nouns formed by conversion were already present in Latin (*pugnare* ‘to fight’ > *pugna* ‘a fight’, etc.), but became much more important in French (cf. Lené 1899); they can be masculine (*soutenir* ‘to support’ > *le soutien* ‘support’, etc.) or feminine (*chasser* ‘to hunt’ > *la chasse* ‘hunting’, etc.). Agent and instrument nouns like *le guide* ‘guide’ (from *guider* ‘to guide’) or *la bêche* ‘rake’ (from *bêcher* ‘to rake’) are probably off-shoots of the action nouns.

Noun-adjective conversion, which is also highly characteristic of French, is of Latin origin but has been greatly extended in French: *une adultère* ‘an adulteress’/*une femme adultère* ‘an adulterous woman’, *une orange* ‘an orange’/*une robe orange* ‘an orange dress’, etc. It is not always easy to distinguish such conversions from copulative noun-noun compounds.

There is also a conspicuous number of noun-verb and adjective-verb conversions (*-er* and *-ir* are infinitive endings) in French: *sel* ‘salt’ > *saler* ‘to salt’, *sale* ‘dirty’ > *salir* ‘to make dirty’, etc. Both types are of Latin origin and have given rise to many new formations during the history of French, but they are subject to heavy restrictions. Only

the first conjugation (-er) has remained productive.

Adjective-adverb conversions of the type *parler haut* ‘to speak loudly’ may be traced back to Latin adverbs in -e which came to coincide with adjectives when the final [e] became mute.

### 3.2.5. Reduplication

Latin used reduplication for the formation of the perfect tense of some verbs, but had no reduplicative word-formation rule. Reduplication is also unknown to Romance word-formation, with the exception of Portuguese and especially French, where a diminutive rule using partial reduplication has developed (cf. Rainer 1998). This rule, whose earliest traces may be found at the end of the Middle Ages or the beginning of Modern Times, developed through the imitation of child language in baby talk, but has since spread to colloquial language. It consists of the repetition of the onset and the nucleus of the base, which must always be a monosyllable or be shortened to a monosyllable: *patte* [pat] ‘paw’ > *papatte* [papat], *Pompidou* > *Pompom*, etc.

## 4. References

- Alsdorf-Bollée, Annegret (1970), *Die lateinischen Verbalabstrakta der u-Deklination und ihre Umbildungen im Romanischen*. Ph.D. Dissertation, Universität Bonn
- Baldinger, Kurt (1950), *Kollektivsuffixe und Kollektivbegriff: Ein Beitrag zur Bedeutungslehre im Französischen mit Berücksichtigung der Mundarten*. Berlin: Akademie-Verlag
- Baldinger, Kurt (1997), “Das Suffix -uche in der Volkssprache und im Argot”. In: Holtus et al. (eds.), 115–147
- Bierbach, Mechthild (1982), *Die Verbindung von Verbal- und Nominalelement im Französischen*. Tübingen: Narr
- Bork, Hans Dieter (1968), Review of Meyer-Lübke (1966). *Romanische Forschungen* 80, 421–439
- Bork, Hans Dieter (1990), *Die lateinisch-romantischen Zusammensetzungen Nomen + Verb und der Ursprung der romanischen Verb-Ergänzungskomposita*. Bonn: Romanistischer Verlag
- Crocco Galées, Grazia & Iacobini, Claudio (1993), “Parasintesi e doppio stadio derivativo nella formazione verbale del latino”. *Archivio Glottologico Italiano* 78, 167–199
- Dees, Anthonij (1971), *Étude sur l'évolution des démonstratifs en ancien et en moyen français*. Groningen: Wolters-Noordhoff
- Eckert, Gabriele (1986), *Sprachtypus und Geschichte: Untersuchungen zum typologischen Wandel des Französischen*. Tübingen: Narr
- Encrevéd, Pierre (1988), *La liaison avec et sans enchaînement*. Paris: Seuil
- Fleischman, Susanne (1977), *Cultural and Linguistic Factors in Word Formation: An Integrated Approach to the Development of the Suffix -age*. Berkeley: University of California Press
- Gawelko, Marek (1977), *Évolution des suffixes adjectivaux en français*. Breslau: Polskiej Akademii Nauk
- Geckeler, Horst (1996), “Gemeinromanische Tendenzen II: Flexionslehre”. In: Holtus et al. (eds.), 199–222
- Hatcher, Anna Granville (1951), *Modern English Word-Formation and Neo-Latin: A Study of the Origins of English (French, Italian, German) Copulative Compounds*. Baltimore: The Johns Hopkins Press
- Heinz, Sieglinde (1982), *Determination und Repräsentation im Altfranzösischen*. München: Fink
- Höfler, Manfred (1967), Review of Meyer-Lübke (1966). *Zeitschrift für romanische Philologie* 83, 104–114
- Höfler, Manfred (1972), *Zur Integration der neutestamentlichen Kompositionsweise im Französischen*. Tübingen: Niemeyer
- Holtus, Günter (1984), “L'emploi des formes sur-composées dans les variétés linguistiques du français et l'attitude des grammairiens”. *Französisch heute* 15, 312–329
- Holtus, Günter & Kramer, Johannes & Schweickard, Wolfgang (1997, eds.), *Italica et Romania. Festschrift für Max Pfister zum 65. Geburtstag, Vol. II*. Tübingen: Niemeyer
- Holtus, Günter & Metzeltin, Michael & Schmitt, Christian (1996, eds.), *Lexikon der Romanistischen Linguistik, Vol. II.1: Latein und Romanisch: Historisch-vergleichende Grammatik der romanischen Sprachen*. Tübingen: Niemeyer
- Hunnius, Klaus (1977), “Frz. *je*: ein präfigiertes Konjugationsmorphem? Ein Forschungsbericht zur Frage der Prädetermination”. *Archiv für das Studium der neueren Sprachen und Literaturen* 224/129, 37–48
- Hunnius, Klaus (1990), “Französisch: Flexionslehre”. In: Holtus, Günter & Metzeltin, Michael & Schmitt, Christian (eds.), *Lexikon der Romanis-*

- tischen Linguistik, Vol. V.1: Französisch. Tübingen: Niemeyer, 59–71
- Jänicke, Otto (1969), Review of Meyer-Lübke (²1966). *Vox Romanica* 28, 329–339
- Karlsson, Keith E. (1981), *Syntax and Affixation: The Evolution of MENTE in Latin and Romance*. Tübingen: Niemeyer
- Klingebiel, Kathryn (1989), *Noun + verb Compounding in Western Romance*. Berkeley etc.: University of California Press
- Koch, Peter (1993), “Le ‘chinook’ roman face à l’empirie: Y a-t-il une conjugaison objective en français, en italien et en espagnol et une conjugaison subjective prédéterminante en français?”. In: Hilty, Gerold (ed.), *Actes du XX<sup>e</sup> Congrès International de Linguistique et Philologie Romanes, Vol. III*. Tübingen, Basel: Francke, 169–190
- Kuen, Heinrich (1952), “Rückläufige Bewegungen in der Entwicklung der romanischen Sprachen zum analytischen Typus: *dabis – dare habes – das*”. In: *Festgabe Ernst Gamillscheg zu seinem 65. Geburtstag*. Tübingen: Niemeyer, 140–163 [reprint: Kuen, Heinrich (1970), *Romanistische Aufsätze*. Nürnberg: Carl, 72–92]
- Lausberg, Heinrich (1962), *Romanische Sprachwissenschaft, Vol. III: Formenlehre*. Berlin: de Gruyter
- Lené, Gustav (1899), *Les substantifs postverbaux dans la langue française*. Uppsala: Almqvist & Wiksell
- Lindemann, Margarete (1977), *Zum Suffixwechsel von -eresse zu -euse und -trice im Französischen*. Tübingen: Narr
- Lindner, Thomas (2002), *Lateinische Komposita*. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck
- Loporcaro, Michele (1998), *Sintassi comparata dell’ accordo participiale romanzo*. Torino: Rosenberg & Sellier
- Malkiel, Jacques (1938), *Das substantivierte Adjektiv im Französischen*. Berlin: Speer & Schmidt
- Malkiel, Yakov (1986), “The Old French Verbal Abstracts in *-eiz*”. *Zeitschrift für romanische Philologie* 102, 1–39
- Martínez Moreno, Annette (1993), *Sprachwandel und Irregularität: Morphosyntaktische Veränderungen im Bereich französischer Nominalkategorien*. Berlin: Schmidt
- Meyer-Lübke, Wilhelm (²1966), *Historische Grammatik der französischen Sprache, Zweiter Teil: Wortbildungslehre* [zweite, durchgesehene und ergänzte Auflage von J. M. Piel]. Heidelberg: Winter
- Rainer, Franz (1998), “La reduplication française du type *fille* d’un point de vue diachronique”. In: Ruffino, Giovanni (ed.), *Atti del XXI Congresso di Linguistica e Filologia Romanza* (Università di Palermo 18–24 settembre 1995). Tübingen: Niemeyer
- Rainer, Franz (2002), “Convergencia y divergencia en la formación de palabras de las lenguas románicas”. In: García-Medall, Joaquín (ed.), *Aspectos de morfología derivativa del español*. Lugo: Tris Tram, 103–133
- Rainer, Franz (2003), “L’intégration des composés latins du type *aurifer* en français”. In: Fradin, Bernard & Dal, Georgette & Hathout, Nabil & Kerleroux, Françoise & Plénat, Marc & Roché, Michel (eds.), *Les unités morphologiques*. Villeneuve d’Ascq: SILEX, 151–168
- Roché, Michel (1998), “Deux études sur la dérivation en *-ier(e)*”. *Carnets de grammaire*, Rapport n° 2. Toulouse: Université de Toulouse-Le Mirail
- Schlegel, August Wilhelm von (1818), *Observations sur la langue et la littérature provençales*. Paris: Librairie grecque-latine-allemande [reprint: (1971), Tübingen: Narr]
- Schmitt, Christian (1988), “Funktionale Variation und Sprachwandel: Zum Verhältnis von ererbter und gelehrter Wortbildung im Spanischen und Französischen”. In: Albrecht, Jörn & Lüdtke, Jens & Thun, Harald (eds.), *Energeia und Ergon: Sprachliche Variation, Sprachgeschichte, Sprachtypologie, Vol. II*. Tübingen: Narr, 183–203
- Schmitt, Christian (1996), “Euromorphologie: Perspektiven einer neuen romanistischen Teildisziplin”. In: Dahmen, Wolfgang & Holtus, Günter & Kramer, Johannes & Metzeltin, Michael & Schweickard, Wolfgang & Winkelmann, Otto (eds.), *Die Bedeutung der romanischen Sprachen im Europa der Zukunft*. Tübingen: Narr, 119–146
- Schöslar, Lene (1984), *La déclinaison bicasuelle de l’ancien français*. Odense: University Press
- Schwegler, Armin (1990), *Analyticity and Syntheticity: A Diachronic Perspective with Special Reference to Romance Languages*. Berlin: Mouton de Gruyter
- Schweickard, Wolfgang (1992), “Deonomastik”: Ableitungen auf der Basis von Eigennamen im Französischen: Unter vergleichender Berücksichtigung des Italienischen, Rumänischen und Spanischen. Tübingen: Niemeyer
- Spence, N. C. W. (1990), “Instrumental *-eur/-aeurl-eusel-abice*”. *The Modern language Review* 85, 29–35

- Togeby, Knud (1980), “Romance Historical Morphology”. In: Posner, Rebecca & Green, John N. (eds.), *Trends in Romance Linguistics and Philology, Vol. I: Romance Comparative and Historical Linguistics*. The Hague, Paris, New York: Mouton, 105–155
- Wanner, Dieter (1987), *The Development of Romance Clitic Pronouns: From Latin to Old Romance*. Berlin etc.: Mouton de Gruyter

## XVIII. Morphologischer Wandel II: Fallstudien

- Wolf, Heinz Jürgen (1964), *Die Bildung der französischen Ethnica (Bewohnernamen)*. Geneva: Droz; Paris: Minard
- Wuest, Jakob (1997), “Zur Entstehung der analytischen Verbformen in den romanischen Sprachen”. In: Holtus et al. (eds.), 31–44

Franz Rainer, Vienna (Austria)

## 158. From Vedic to modern Indic languages

1. Introduction
2. Old Indo-Aryan
3. Middle Indo-Aryan
4. New Indo-Aryan
5. Transliteration and transcription
6. References

### 1. Introduction

Indo-Aryan, among the most richly documented of Indo-European language groups, is represented by Old, Middle and New Indo-Aryan languages (see Bloch 1965, Masica 1991, Cardona & Jain 2003 for historical surveys and Cardona & Jain 2003, eds. for sketches of individual languages). The oldest documents – composed centuries earlier than they are attested in writing – are the Vedas: the R̥gveda, the most ancient, certainly composed before the end of the second millennium B. C. and possibly in the third millennium; the Atharvaveda; the Sāmaveda; and the Yajurveda, this last in two major groupings, called *śukla* ('white') and *kṛṣṇa* ('black'), and various recensions. A variety of Old Indo-Aryan systematically close to a late Vedic stage is described by the grammarian Pāṇini (see Art. 5), who distinguishes between the spoken language (*bhāṣā*) and usage proper to Vedic (*chandas*). These and other early Indo-Aryan dialects are included in a language termed *sāṁskṛta* (Sanskrit, abbr. Skt.), the pure or polished language, opposed to other vernaculars, which were considered corruptions (*apabhrañśa*) of Sanskrit. These vernaculars make up the group of Prakrits (Skt. *prākṛta*- 'natural'), Middle Indo-Aryan languages and dialects (see Pischel 1965; Hinüber 2001). The earliest attested Prakrit (abbr. Pkt.) documents are the inscriptions of the emperor Aśoka (3rd c. B. C.), which are

especially valuable because they appear in the different dialects spoken in areas of Aśoka's empire (see Oberlies 2003). The major texts of Theravāda Buddhism are represented in the mixed language called Pāli; Jaina texts also are composed in Prakrit (see Geiger 1965; Hinüber 2001; Oberlies 2001). The exemplar of literary Prakrits, used in both Sanskrit dramas and independent works, is called Māhārāṣṭrī. The latest Middle Indo-Aryan stage is represented by various dialectal forms of Apabhrañśa (see Bubenik 2003). New Indo-Aryan documents date from approximately the twelfth century, and these languages cover the largest part of the Indian subcontinent, from Assamese, Bangla, and Oriya in the east, to Gujarati, Marathi, and Konkani in the west. Hindi, spoken in the midlands and used as a lingua franca, is one of the official languages of the Republic of India.

Indo-Aryan languages not only have an extremely rich store of literary sources from various eras, they have also been expertly described by grammarians of Sanskrit (see Art. 5) and of Middle Indo-Aryan dialects. The following sketch is meant to give an idea of the major developments in Indo-Aryan. Obviously, a great many details have been omitted, which are covered in grammars and studies of individual languages (for bibliography, see Masica 1991: 486–510 and the references given in individual chapters of Cardona & Jain 2003, eds.).

### 2. Old Indo-Aryan

Old Indo-Aryan morphology is fairly typical of early Indo-European inflectional languages. It is possible also to trace developments from early Vedic to later stages (see Cardona 2003).

## 2.1. Inflectional morphology

### 2.1.1. Nominal morphology

The Old Indo-Aryan **nominal system** has three **numbers** (singular, dual, plural), three **genders** (masculine, feminine, neuter), and up to seven **case** forms plus a vocative. Pronouns and pronominally inflected adjectives have special characteristics. The dual shows the least number of contrasts, with three distinct forms: ‘NOM/ACC’, ‘INSTR/DAT/ABL’, ‘GEN/LOC’. Nominative and accusative are identical for neuters; e.g. sg. *phalam*, du. *phale*, pl. *phalāni* ‘fruit’. Ablative and genitive singular are formally identical except for *a*-stems, and dative and ablative plural are identical for all stems.

In addition to a variety of endings, nominals can have vowel alternation (ablaut); there are also nouns with heterogeneous stems. The greatest number of formal contrasts occurs in *a*-stems (e.g., *deva-* ‘god’, *phala-* ‘fruit’). The forms of *deva-* and the consonant-stem *vāc-* are shown in Art. 5 (section 4.2.2, (9)–(10)). *senā-* ‘army’, *agni-* ‘fire’ (m. *i*-stem), *vāyu-* ‘wind’ (m. *u*-stem), *kṛti-* ‘deed’, (f. *i*-stem), *dhenu-* ‘cow’, (f. *u*-stem), *nadi-* ‘river’ (f. *ī*-stem), *vadhū-* (f. *ū*-stem) ‘bride’, *pitṛ-* ‘father’, *mātṛ-* ‘mother’, *kartr-* ‘doer, agent’, *rājan-* ‘king’, and *asthi-/asthan-* ‘bone’ (n. *i-/n*-stem) will illustrate further, as in Tab. 158.1.

In accordance with Pāṇini’s system, case forms are labeled as follows: 1. nominative, 1a. vocative, 2. accusative, 3. instrumental, 4. dative, 5. ablative, 6. genitive, 7. locative (cf. Art. 5, section 4.2.2).

Stems such as *agni-/agne-* (prevocalic *agnay-*), *vāyu-/vāyo-* (*vāyav-*), *pitṛ-* (*pitr-*) *lpitar-*, *kartr-/kartar-/kartār-*, *rājan-/rājān-/rājñ-/rāja-* show vocalic alternation.

The **pronominal system** includes personal pronouns (see Art. 5: 4.2.2, (29)–(30)), interrogative, relative, and demonstrative pronouns. Pronominals in *-a*, of the type *sarva-* ‘all’, *ka-* (interr.), *ya-* (rel.) have a particular set of forms:

- (1) (a) M. sg.: 4. *sarvasmai*, 5. *sarvasmāt*, 7. *sarvasmīn*; pl. 1. *sarve*, 6. *sarveṣām* (the rest like *deva-*, *phala-*);
- (b) F. sg.: 4. *sarvasyai*, 5–6. *sarvasyās*, 7. *sarvasyām*; pl. 6. *sarvāsām* (the rest like *senā-*).

A subset of pronominals in *-a* have a nominative-accusative singular neuter with *-at* instead of *-am*: *anyat* ‘other’, *anyatarat* ‘either

of two’, *itarat* ‘other’, *katarat* ‘which of two?’, *katamat* ‘which of many?’.

**Demonstrative pronouns** have degrees of deixis: *idam* ‘this’, *etat* ‘this, that’, *tat* ‘that’, *adas* ‘that (yonder)’ (all neuter nominative/accusative singular). There are alternating stem forms: *eṣa-*, *eṣā-*, *sa-*, *sā-* in the nominative singular non-neuter (m. *eṣas*, *sas* [*eṣa*, *sa* before consonant], f. *eṣā*, *sā*), *eta-*, *ta-* elsewhere, according to the model shown for *sarva-*. Forms for the proximate and distant demonstratives noted are as follows:

- (2) (a) M. sg. 1. *ayam* ‘this’ *asau* ‘that’, 2. *imam amum*, 3. *anena amunā*, 4. *asmai amusmai*, 5. *asmāt amusmāt*, 6. *asya amusya*, 7. *asmin amusmin*; du. 1–2. *imau amū*, 3–5. *ābhyām amūbhyaṁ*, 6–7. *anayoḥ amuyos*; pl. 1. *ime amī*, 2. *inān amūn*, 3. *ebhis amībhis*, 4–5. *ebhyas amībhayas*, 6. *eṣām amīṣām*, 7. *eṣu amīṣu*;
- (b) F. sg. 1. *iyam asau*, 2. *imām amūm*, 3. *anayā amuyā*, 4. *asyai amusyai*, 5–6. *asyās amusyāh*, 7. *asyām amusyām*; du. 1–2. *ime amū*, 3–5. *ābhyām amūbhyaṁ*, 6–7. *anayos amuyos*; pl. 1–2. *imāh amūh*, 3. *ābhih amūbhīh*, 4–5. *ābhyah amūbhyaḥ*, 6. *āsām amūṣām*, 7. *āsu amūṣu*;
- (c) N. 1–2 sg. *idam adas*; du. *ime amū*; pl. *imāni amūni*.

**Indefinite pronouns** are formed with elements postposed to formal interrogatives, as in *kaścit* (nom. sg. m.) ‘some one’, *kiñcit* ‘some thing’ (nom./acc. sg. n.), *ko’pi* ‘any one’.

Moreover, some pronouns have orthotonic and enclitic alternant forms (see Art. 5: 4.2.2, (29)–(30)). Anaphoric pronominal forms such as *anena* (3rd sg. instr. m./n.), *enam* (acc. sg. m.) ‘this one’ are also atonic.

### 2.1.2. Verbal morphology

The Old Indo-Aryan verb system distinguishes the following tense and modal forms: present, future, aorist, imperfect, perfect; indicative, optative, subjunctive, conditional. There is a subtype of optative used in wishes (termed precative, benedictive) and a variety of future that originally was periphrastic. Forms called injunctive – equivalent to unaugmented imperfects and aorists – were also used. Singular, dual, and plural are distinguished in all persons. There is also a contrast between active and medio-passive. Some verbs are activa tantum or media tantum; e.g., *as* ‘be’ and *ās* ‘be seated’ have

## (a) Singular

1	1a	2	3	4	5–6	7
<i>senā</i>	<i>sene</i>	<i>senām</i>	<i>senayā</i>	<i>senāyai</i>	<i>senāyās</i>	<i>senāyām</i>
<i>agnis</i>	<i>agne</i>	<i>agnim</i>	<i>agninā</i>	<i>agnaye</i>	<i>agnes</i>	<i>agnau</i>
<i>vāyus</i>	<i>vāyo</i>	<i>vāyum</i>	<i>vāyunā</i>	<i>vāyave</i>	<i>vāyos</i>	<i>vāyau</i>
<i>kṛtis</i>	<i>kṛte</i>	<i>kṛtim</i>	<i>kṛtyā</i>	<i>kṛtyail/kṛtaye</i>	<i>kṛtyās/kṛtes</i>	<i>kṛtyām/kṛtau</i>
<i>dhenus</i>	<i>dheno</i>	<i>dhenum</i>	<i>dhenvā</i>	<i>dhenvai/dhenave</i>	<i>dhenvās/dhenos</i>	<i>dhenvām/dhenau</i>
<i>nadī</i>	<i>nadi</i>	<i>nadīm</i>	<i>nadyā</i>	<i>nadyai</i>	<i>nadyās</i>	<i>nadyām</i>
<i>vadhūs</i>	<i>vadhu</i>	<i>vadhūm</i>	<i>vadhvā</i>	<i>vadhvai</i>	<i>vadhvās</i>	<i>vadhvām</i>
<i>pitā</i>	<i>pitar</i>	<i>pitaram</i>	<i>pitrā</i>	<i>pitre</i>	<i>pitur</i>	<i>pitari</i>
<i>mātā</i>	<i>mātar</i>	<i>mātaram</i>	<i>mātrā</i>	<i>mātre</i>	<i>mātur</i>	<i>mātari</i>
<i>kartā</i>	<i>kartar</i>	<i>kartāram</i>	<i>kartrā</i>	<i>kartre</i>	<i>kartur</i>	<i>kartari</i>
<i>rājā</i>	<i>rājan</i>	<i>rājānam</i>	<i>rājñā</i>	<i>rājñe</i>	<i>rājñas</i>	<i>rājñi</i>
<i>asthi</i>		<i>asthi</i>	<i>asthnā</i>	<i>asthne</i>	<i>asthnas</i>	<i>asthni</i>

## (b) Dual

1–2	3–5	6–7
<i>sene</i>	<i>senābhymām</i>	<i>senayos</i>
<i>agnī</i>	<i>agnibhymām</i>	<i>agnyos</i>
<i>vāyū</i>	<i>vāyubhymām</i>	<i>vāyvos</i>
<i>kṛtī</i>	<i>kṛtibhymām</i>	<i>kṛtyos</i>
<i>dhenū</i>	<i>dhenubhymām</i>	<i>dhenvos</i>
<i>nadyau</i>	<i>nadībhymām</i>	<i>nadyos</i>
<i>vadhvau</i>	<i>vadhūbhymām</i>	<i>vadhvos</i>
<i>pitaraū</i>	<i>pitrbhymām</i>	<i>pitros</i>
<i>mātarau</i>	<i>mātrbhymām</i>	<i>mātros</i>
<i>kartārau</i>	<i>kartrbhymām</i>	<i>kartros</i>
<i>rājānau</i>	<i>rājabhymām</i>	<i>rājños</i>
<i>asthīnī</i>	<i>asthibhymām</i>	<i>asthnos</i>

## (c) Plural

1–1a	2	3	4–5	6	7
<i>senās</i>	<i>senās</i>	<i>senābhīs</i>	<i>senābhīyas</i>	<i>senānām</i>	<i>senāsu</i>
<i>agnayas</i>	<i>agnīn</i>	<i>agnibhīs</i>	<i>agnibhīyas</i>	<i>agnīnām</i>	<i>agniṣu</i>
<i>vāyavas</i>	<i>vāyūn</i>	<i>vāyubhīs</i>	<i>vāyubhīyas</i>	<i>vāyūnām</i>	<i>vāyuṣu</i>
<i>kṛtayas</i>	<i>kṛtīs</i>	<i>kṛtibhīs</i>	<i>kṛtibhīyas</i>	<i>kṛtīnām</i>	<i>kṛtiṣu</i>
<i>dhenavas</i>	<i>dhenūs</i>	<i>dhenubhīs</i>	<i>dhenubhīyas</i>	<i>dhenūnām</i>	<i>dhenuṣu</i>
<i>nadyas</i>	<i>nadīs</i>	<i>nadībhīs</i>	<i>nadībhīyas</i>	<i>nadīnām</i>	<i>nadīṣu</i>
<i>vadhvas</i>	<i>vadhūs</i>	<i>vadhūbhīs</i>	<i>vadhūbhīyas</i>	<i>vadhūnām</i>	<i>vadhūṣu</i>
<i>pitaras</i>	<i>pitrīn</i>	<i>pitrbhīs</i>	<i>pitrbhīyas</i>	<i>pitrīnām</i>	<i>pitṛṣu</i>
<i>mātaras</i>	<i>mātrīs</i>	<i>mātrbhīs</i>	<i>mātrbhīyas</i>	<i>mātrīnām</i>	<i>mātrṣu</i>
<i>kartāras</i>	<i>kartīn</i>	<i>kartrbhīs</i>	<i>kartrbhīyas</i>	<i>kartīnām</i>	<i>kartṛṣu</i>
<i>rājānas</i>	<i>rājñas</i>	<i>rājabhīs</i>	<i>rājabhīyas</i>	<i>rājñām</i>	<i>rājasu</i>
<i>asthīni</i>		<i>asthibhīs</i>	<i>asthibhīyas</i>	<i>asthīnām</i>	<i>asthiṣu</i>

Tab. 158.1: Nominal system in Old Indo-Aryan

active and middle inflexion, respectively: 3rd sg. pres. indic. *as-ti*, *ās-te*; 3rd pl. *s-anti*, *ās-ate*. Other verbs can have both inflexions, with a semantic difference; e.g., *yaj-a-ti* ‘ven-

erates, performs a sacrificial rite on another’s behalf’; *yaj-a-te* ‘arranges for a sacrificial rite on his own behalf’, *kar-o-ti* ‘makes for someone else’, *kur-u-te* ‘makes for himself’ (see

Art. 5: 4.2.2, (11)). In addition, stems can have ablaut variants; e.g., *as/əs*.

**Tense stems** consist of bases alone or bases with affixes, either suffixed or infixated. In addition to the root type *as-ti s-tas s-anti* (impf. *āsīt* [archaic *ās*] *āstām āsan*) there are nine types of present-imperfect stems formed with different affixes, as follows (3rd sg., pl.):

- (3) (a) *bhav-a-ti bhav-anti* (*abhavat abhavan*) ‘be, become’: full-grade root with unaccented suffix *-a*.
- (b) *juho-ti juhv-ati* (*ajuhot ajuhavus*) ‘offer oblations’, *jihīt-e jih-ate* ‘flee’: reduplicated stem with ablaut.
- (c) *pus-ya-ti pus-yanti* (*apusyat apusyan*) ‘thrive’: suffix *ya*.
- (d) *su-no-ti su-nv-anti* (*asunot asunvan*) *su-nu-te su-nv-ate* (*asunuta asunvata*) ‘press juice out of’: suffix *-no-l-nu* follows zero-grade root.
- (e) *tud-a-ti tud-anti* (*atudat atudan*) *tud-a-te tud-ante* (*atudata atudanta*) ‘goad, wound’: zero-grade root with accented suffix *-a*.
- (f) *yu-na-k-ti yu-ñ-j-anti* (*ayunak ayuñjan*) *yu-ñ-k-te yu-ñ-j-ate* (*ayuñkta ayuñjata*) ‘join, connect, yoke’: zero-grade root with infixated *-na-/n-*.
- (g) *tan-o-ti tan-v-anti* ‘stretch’, *kar-o-ti*: suffix *-o-/u-* (see Art. 5: 4.2.2 (11)).
- (h) *kri-ñā-ti kri-ñ-anti* (*akrīñāt akrīñan*) *kri-ñī-te kri-ñ-ate* (*akrīñīta akrīñata*) ‘buy’: zero-grade root with suffix *-ñā-/ñī-/ñ-*.
- (i) *cor-ay-a-ti cor-ay-anti* ‘steal’, *kath-ay-a-ti kath-ay-anti* ‘tell relate’: suffix *-i-/e* (-*ay-* before vowels).

Similarly, there are several types of aorist stems, as in the following examples:

- (4) (a) *adā-t ad-us* (*dāld* ‘give’), *abhū-t ab-hūv-an* (*bhū* ‘be, become, come into being’): root aorist.
- (b) *apus-a-t apus-an* (*pus* ‘thrive’), *agam-a-t agam-an* (*gam* ‘go’): root followed by stem vowel *-a*.
- (c) *acūcur-a-t acūcur-an* (*cur* ‘steal’), *ajījan-a-t ajījan-an* ‘brought into being’ (*jan* ‘be born’): reduplicated aorist with stem vowel *-a*, the regular aorist corresponding to presents with *-i* of the type (3i) and causatives (*jan-i*: *jan-ay-a-ti* ‘brings into being’).
- (d) *anai-ş-īt anai-ş-us*, mid. *ane-ş-ṭa ane-ş-ata* (*nī* ‘lead’); *akār-ş-īt akār-ş-us* *akr-ta* (1st sg. *akr-ş-i*) *akr-ş-ata* (*kr*

‘do, make’), *abhait-s-īt abhait-s-us abhit-ta abhit-s-ata* (*bhid* ‘split’): *s-aorist*.

- (e) *anam-s-īt anam-siş-us* (*nam* ‘bow, bend’), *ayā-s-īt ayā-siş-us* (*yā* ‘go’): stem with *-siş-*, regular for bases in *-ā*.
- (f) *adhuk-şa-t adhuk-şa-n* (*duh* ‘milk’): stem in *-sa-*, regular for *aniṣ* verbs with final consonants that in combination with a following *-s-* give a cluster *-kṣ-*, as in *-dhukṣa-*; this verb originally had an *s-aorist* (3rd sg. mid. *adugdha*, 1st sg. mid. *adhukṣi*) which was the analogic source of this formation.

The general future stem is formed with a suffix *-sya-*, as in *dā-sya-ti* ‘will give’, *kar-iṣya-ti* ‘will do, make’. There is also a complex future, with an additional suffix *-tās-*, of the type:

- (5) 3rd sg. *kartā kartārau kartā* (act./mid.);  
2nd sg. *kartāsi kartāsthas kartāsthā* (act.), *kartāse kartāsāthe kartādvhe* (mid.);  
1st sg. *kartāsmi kartāsvas kartāsmas* (act.), *kartāhe kartāsve kartāsmahe* (mid.).

The source of this type is a periphrastic formation consisting of an agent noun in *-tr* and forms of the verb *as*, except for the third person forms. But this came to have the structure Pāṇini ascribes to it: a suffix *-tās-* is added.

In the language Pāṇini describes, the **aorist** (e.g., *avadhīt* ‘slew, has slain’) is used for a general past, including the day of reference; the **imperfect** (e.g., *ahan* ‘slew’) and **perfect** (e.g., *jaghāna* ‘slew’) are used for a past excluding the day of reference, with the perfect further distinguished by being used if the act in question was not witnessed by the speaker. These contrasts are also found in certain Vedic texts. There is a similar contrast between the future types *karisyāmi* ‘I shall do, make’ and *kartāsmi* ‘I shall do, make (at some time after today)’.

The **conditional** has an augmented stem with a suffix *-sya-* followed by secondary endings, of the type *adā-sya-* (3rd sg. *adāsyat* ‘would give, would have given’), *akar-iṣya-* (*akarisyat* ‘would do, would have done’). This is used in contrary-to-fact conditional sentences.

**Perfect stems** are generally characterized by reduplication, ablaut, and a particular set

of endings. For example, the perfect of *kṛ* ‘do, make’ is:

- (6) 3rd sg. *cakāra cakratus cakrus* (act.), *cakre cakrāte cakrire* (mid.);  
 2nd sg. *cakṛtha cakrathus cakra* (act.), *cakṛṣe cakrāthe cakṛdhve* (mid.);  
 1st sg. *cakāral/cakara cakṛva cakṛma* (act.), *cakre cakṛvahē cakṛmahe* (mid.).

Not all perfects have reduplicated stems. The most notable exception is the perfect of *vid* ‘know’: *veda vidatus vidus*. A periphrastic perfect of the type *gamayām cakāra* ‘caused to go’ is the norm for derived verbs such as causatives.

There are also **suppletive stems**; e.g., *ab-hū-t*, *avadh-īt* are the aorists corresponding to the presents *as-ti* ‘is’, *han-ti* ‘kills’.

**Imperative** forms have particular endings in second and third person forms. To indicatives with *-ti*, *-antil-ati*, *-e* (see (3)) correspond imperatives with *-u* instead of *-i* and *-ām* instead of *-e* in active and middle forms respectively. For example: *as-tu*, *s-antu*, *juho-tu*, *juhv-atu*; *sunu-tām*, *sunv-ātām*, *sunv-ātām*, *sunv-āthām*, *yaj-antām*. Second person singular forms have *-hi*, *-dhi*, *-o*, and the corresponding middle ending is *-sva*. For example: *i-hi* ‘go’, *juhu-dhi* ‘offer oblation’, *bhind-dhi* ‘split’, *paca* ‘cook’, *sunu* ‘press juice’, *kuru* ‘do, make’; *paca-sva*, *sunu-sva*, *kuru-sva*. The second plural middle has *-dhvam*; e.g., *sunu-dhvam*. Imperatives of the type *kṛ-ṇī-hi* ‘buy’ (see (3h)) are regular for vocalic bases; consonantal bases have corresponding imperative forms of the type *gr̥h-āṇa* (*gr̥hṇāti* ‘grasps’). First person imperative forms contain *-ā* followed by *-ni* (1st sg. act.), *-vahai*, *-mahai* (1st st. du., pl. mid.), with *-ai* instead of *-e*; similarly, the first singular middle has *-ai*; e.g., *bhav-āṇi* ‘I would be’, *sunav-ai*, *sunav-āvahai*, *sunav-āmahai*. Other imperative forms have secondary endings, as in *bhavatām*, (3rd du.), *bhavathām* (2nd du.), *bhavata* (2nd pl.), *bhavā* (1st du.), *bhavāma* (1st pl.).

**Optative** forms to stems in *-a* are characterized by *-e*, and optatives from other stems have active forms with *-yā-*, middle forms with *-ī(y)-*. Third plural forms contain endings *-us* (act.), *-ran* (mid.), and *-a* is the first singular middle ending; otherwise, secondary endings follow the stems. E.g., *pac-e-t pac-ey-us*, *pac-e-ta pac-e-ran pac-ey-a*, *sunu-yāt sunu-yus*, *sunv-ī-ta sunv-ī-ran sunv-īy-a*. **Precative** forms are radical and have endings like those of the optative, but the stem is characterized by *-s-* in the middle and active endings

## XVIII. Morphologischer Wandel II: Fallstudien

have *-s-* before nonfinal sounds; e.g., *sū-yā-t sū-yās-tām* *sū-yās-us*, *so-śīs-ṭa so-śīy-āstām* *so-śī-ran* (*su* ‘press juice’).

### 2.2. Derivational morphology

Early Indo-Aryan has a rich system of deriving bases from nominals and verbs.

#### 2.2.1. Nominal derivation

**Derived nominals** are denominative or deverbative. Nouns formed directly from verb roots include derivates signifying actions, agents, objects, instruments, and loci; e.g., *kṛ-ti-* ‘doing, deed’, *kar-tr-* ‘doer, agent’, *kar-man-* ‘action, object’, *kar-anā-* ‘doing, means, instrument’ (base *kṛ*), *-dhānt-* ‘receptacle’ (base *dhā-* ‘put’). Action nouns in *-tu* supply what are traditionally called infinitives. What is etymologically an accusative singular in *-tum* (e.g., *kar-tum*) is the norm in the language Pāṇini describes, but earlier Vedic also had other case forms of such derivates (dat. *-tave*, abl./gen. *-tos*). Moreover, forms in *-tum* alternate with datives of action nouns; e.g., *bhok-tum*, *bhojanāya* (dat. sg. of *bhoj-ana-*) ‘to eat’. There are also participles etymologically related to derivates with *-tu-*, of the type *kar-tavya-* ‘to be done, which can be done’, which alternate with derivates in *-anīya-* (e.g., *kar-anīya-*) and other suffixes. In addition, there are participles with tense values; e.g., *-ta-* (*kṛ-ta-* ‘done, made’, *ga-ta-* ‘gone’), present and future participles in *-at-l-ant-* (act.) and *-āna-*, *-māna-* (med./pass.), such as *kur-v-at-* (m. nom. sg. *kurvan*, instr. sg. *kurvatā*) ‘doing, making’, *kurvāṇa-* ‘id.’, *kariṣyat-* (*kariṣyan*, *kariṣyatā* ‘going to do’), and perfect participles in *-vas-l-us-* (act.), *-āna-* (mid.), such as *cakryas-* (*cakrvān* [nom. sg. m.], *cakruṣā* [instr. sg. m./n.]). There are, in addition, indeclinables with *-tvā* or *-ya* – the latter in nonnegative compounds – which are used with reference to an action that is performed prior to another; e.g., *kṛ-tvā* ‘after making, doing’, *ni-śad-ya* ‘after sitting’.

Several kinds of nominals are derived with secondary derivative affixes (called *taddhita*). A series of affixes occurs in derivates signifying the property of being what a given nominal base denotes; for example, *puruṣa-tva-*, *puruṣa-tā-* signify the property of being a man (*puruṣa-*), and *prath-imā-* designates the property of being broad (*prthu-*). There is a large group of derivates that correspond to phrases of the type *X-E Y*, with which they alternate, where the values of *X-E* are case

forms of nominals and *Y* stands for a nominal whose meaning is attributed to the derivational affix. For example, there are patronymics like *gārgī-* ‘son of Garga’, *aupagava-* ‘descendant of Upagu’: a case form of *gārgī-* or *aupagava-* corresponds to and alternates with a phrase containing a genitive of *garga-* or *upagu-* and a form of *putra-* ‘son’, *apatyā-* ‘offspring, descendant’, or a synonym of such a noun.

Some **taddhita derivates** are formed from a more restricted group of nominals. Thus, *tāvat-* (nom. sg. m. *tāvān*) ‘this much’, *yāvat-* ‘as much’, *kiyat-* ‘how much?’ are formed from a subset of pronouns. Some such derivates correspond to case forms. For example, *tatas* ‘from that, thence’, *tatra* ‘in, at that, there’ correspond to and alternate with ablative and locative forms of *tad-* ‘that’. Other adverbial terms derived from pronouns are *tathā* ‘thus’, *yathā* ‘as’, *katham* ‘how?’.

A particularly wide-spread taddhita affix is *ka*, which serves to form derivates with diminutive and other semantic shadings (e.g., *putra-ka-* ‘poor little child’, *ásva-ka-* ‘wretched horse’) and also occurs redundantly, as in *avi-ka*, synonymous with *avi-* ‘sheep’.

Sanskrit is like other Indo-European languages in that it forms what are traditionally called **comparatives** and **superlatives** of adjectives by means of suffixes. For example, *madhumat-tara-* ‘quite full of honey’, *priyatama-* ‘most dear’, *gar-īyas-* ‘weightier, more prolix (*guru-* ‘weighty, prolix’), *gar-iṣṭha-* ‘most weighty, prolix.’ As can be seen, derivates with *īyas*, *iṣṭha* are formed from roots, not to adjective stems, and in Vedic the type *yaj-iṣṭha-* ‘who best performs a rite’ is widespread. Moreover, while eliminating the use of derivates with these suffixes in agentive derivates formed from productive verb roots, Sanskrit developed the type *pacati-tarām* (= *pacati-tara-ām*) ‘cooks very well’, *pacati-tamām* ‘cooks extremely well’, with the suffixes *-tara*, *-tama* attached to finite verb forms.

**Compounds** are of four general kinds: *tatpuruṣa* (determinative), *dvandva* (copulative), *bahuṛihi* (exocentric), and a type that is usually invariant (*avyayībhāva*). The first member of a *tatpuruṣa* is regularly equivalent to a case form other than a nominative; e.g. *tatpuruṣas* (nom. sg.), *grāma-gatas* ‘gone to the village’ are equivalent to *tasya puruṣaḥ* ‘his (gen. sg. *tasya*) man, servant’ and *grāmar̥i gatas*. A subtype of *tatpuruṣa*, called *karmadhāraya*, has a first member coreferen-

tial with the second, which it modifies, as in *nīlotpalam* ‘blue (*nīla-*) lotus’, equivalent to *nīlam utpalam*. Like *tatpuruṣa-*, the compound *bahuṛihi-* is an instance of the class: *bahuṛihi* (nom. sg.) is equivalent to *bahuṛihi asya*, used of someone who has (*asya* ‘of this one’) much (*bahu-*) rice (*vrīhi-*). Dvandvas are equivalent to phrases with *ca* ‘and’; e.g., *mātāpitaraū* (nom./acc. du.) is the equivalent of *mātā ca pitā ca* ‘mother and father’. *avyayībhāva* compounds are generally though not always invariant; e.g., *upāgni* ‘near the fire’.

*upāgni-* is an obligatory compound in the sense that the constituents *upa* and *agnes*, are not used together in a syntagma that alternates with the compound. Similarly, there are *tatpuruṣa* compounds such as *kumbhakāra-* ‘pot maker’: the derivate *-kāra-* does not occur outside of a compound. Certain nominals also have variants that occur only in compounds. For example, *rājan-* is an *n*-stem, but there is an *a*-stem *-rāja-* found in compounds, as in *mahārāja-* ‘great king’.

Number words designating units also form compounds with terms for decades.

- (7) '11' *ekā-daśa* (*eka-* ‘one’)
- '12' *dvā-daśa* (*dvi-* ‘2’ [archaic nom./acc. du. *dvāl*])
- '13' *trayo-daśa* (< *trayas-daśa*) (*tri-* ‘3’ [nom. pl. m. *trayas*])
- '14' *catur-daśa*
- '15' *pañca-daśa*
- '16' *śo-daśa* (*śaṣ-* ‘6’)
- '17' *sapta-daśa*
- '18' *aṣṭā-daśa* (*aṣṭān* ‘8’ [archaic nom./acc. pl. *aṣṭāl*])
- '19' *nava-daśa* (also *ekona-viṁśatis*; *ūna-viṁśatis* [*‘20 less one’*])
- '20' *viṁśatis*
- '50' *pañcāśat*
- '51' *eka-pañcāśat*
- '52' *dvi-pañcāśat*
- '53' *tri-pañcāśat*
- '54' *catuh-pañcāśat*
- '55' *pañca-pañcāśat*
- '56' *śat-pañcāśat*
- '57' *sapta-pañcāśat*
- '58' *aṣṭa-pañcāśat*
- '59' *nava-pañcāśat* (*ekona-śaṣṭis*)

Sanskrit also has iterative sequences like *damē dame* ‘in each house, in house after house’. The second element does not have a high-pitched vowel. In *pada-pāṭhas* (analyzed versions) of Vedic texts, such sequences

are treated as compounds, but Pāṇini and other grammarians treat them as sequences that involve repetition of a word.

### 2.2.2. Verb derivation

Derived verbal bases are either deverbal or denominative. To the former group belong causatives, desideratives, and intensives. **Causatives** are formed with a suffix *-i-*, which generally conditions substitution by a vowel *ā ai au* in a preceding base. For example, the causatives of *yaj* ‘venerate, perform a rite’, *kṛ* ‘do, make’ are *yāj-i-* (3rd sg. pres. ind. *yāj-ay-a-ti*), *kār-i-* (*kār-ay-a-ti*). In addition, verbs in *-ā* and a small group of other verbs show a final *-p* preceding the causative suffix, as in *sthāp-i-* ‘set up, found’ (*sthā* ‘be in place’), *arp-i* ‘cause to reach’ (*r* ‘go to’). **Desideratives** have the suffix *-sa-*, usually with reduplication; e.g., *vi-vak-ṣa-* (*vivakṣati*) ‘wish to speak’. **Intensives** are generally formed to consonant-initial monosyllabic bases, with the suffix *-ya-* and reduplication; such bases are inflected medially; e.g., *yā-yaj-ya-* (*yāyajyate*) ‘repeatedly perform a rite’. There are also intensives of the type *roravīti* ‘roars’, which involve reduplication without a suffix.

**Denominatives** are regularly formed with the suffix *-ya-*, as in *putrī-ya-* (3rd sg. pres. indic. *putrīyati*) ‘wish for a son, treat as a son’, *vrsā-ya-* (*vrsāyate*) ‘act like a bull’, *apsarā-ya-* (*apsarāyate*) ‘act like an Apsaras’. In addition, verb endings could directly follow nominal bases, without a derivational suffix, a type found mainly in Vedic; e.g., *bhiṣaj-* ‘healer’, *bhiṣak-ti* ‘heals’.

### 2.3. Vedic features

Early Vedic had variants that were eliminated in dialects that make up later Sanskrit. For example, Vedic dialects have *a*-stem forms of the types *vīryā* ‘heroic power’ (instr. sg.), *ubhā* ‘both’ (nom./acc. du. m.), *devāsā* ‘gods’ (nom. pl.), *āyūdhā* ‘weapons’ (nom./acc. pl. n.), *devām* (gen. pl.). In early Vedic, a distinction is also made between derivative stems of the types *vrkti-* ‘female wolf’ and *devī-* ‘goddess’. The first type has *-s* in the nominative singular (*vrktis*) and inflects like a consonant stem (e.g., acc. sg. *vrkyām*); the second lacks *-s* in the nominative singular (*devī*) and has alternative stems of the types *devī-*, *devyā-*. Forms of the two types were later collapsed into a single type, illustrated by *nadī-* in Tab. 158.1.

Sanskrit has particles such as *ā*, which cooccur with particular case forms, as in *ā*

*pāṭaliputrāt* (abl. sg.) ‘up to Pāṭaliputra’. In Vedic, one also has *ā* redundantly postposed to a case form, as in *diva ā* ‘from heaven’ (abl. sg. *divas*), *ghra āñ* ‘in the cloud’ (loc. sg. *ghbre*).

Vedic shows traces of an aspectual contrast (perfective : imperfective), most clearly seen in negative imperatives of the types *mā vādhīs* ‘do not slay’ (perfective: 2nd sg. aor. *avadhīts*), meant to keep someone from initiating an act, versus *mā dīvyas* ‘stop gambling’ (imperfective: 2nd sg. impfv. *adīvyas*), meant to have someone cease what he or she is already doing. From earliest attested texts, however, the contrasts are primarily among temporal stems. Some perfect forms in early Vedic reflect the Indo-European use as stative (e.g., *bī-bhāy-ā* ‘is afraid’ [3rd sg. pf.]), but from quite early on the perfect too is incorporated into a tense system.

The subjunctive (e.g., *as-ā-t*, 3rd sg. subj. of *as* ‘be’) gradually goes out of use in Sanskrit, being supplanted by forms of the optative and present and future indicative. In Pāṇini’s description, subjunctive forms are marked as particular to Vedic usage. A remnant of this type persists in first person imperative forms of the type *karavāni* (1st sg.) ‘I ... do, make’ (e.g., *kim karavāni* ‘what shall I do, what may I do?’). Similarly, the earlier injunctive remains in the negative imperative types *mā vādhīs*, *mā dīvyas* (see above).

In some instances, there are different stems which are historically related. For example, in Vedic *kṛ* ‘make, do’ has a root aorist of the type *akār* (3rd sg. act. < \*ákart [1st sg. *akāram*]), *akrān* (3rd pl. < \*ákr-ant), in which the stem is simply the base (with possible augment), with either full grade (*kar*) or zero-grade (*kr*). There is also a type *akār-a-t* (3rd sg.), developed analogically from the athematic root aorist, in which the stem is invariant: *akar-a-*, with the stem vowel *-a-*. Both these stems are found in earliest Vedic texts. Later, however, the productive aorist is sigmatic: *akār-s-ūt*, with the stem suffix *-s-* and lengthened grade of the base (*kār-*). Forms of the earlier root aorist persist in the medio-passive; e.g., 3rd sg. *akr-ta* ‘did, made (for himself)’ forms part of a paradigm that includes 1st sg. *akr-s-i* and 3rd pl. *akr-s-ata*. In early Vedic also, middle forms like *stave* ‘is praised’ could function as passives and coexisted with marked passives of the type *stū-ya-te* ‘is praised’. Later, the second type prevailed as the norm for passives.

A particularly noteworthy contrast between early Vedic and the later language which Pāṇini describes concerns imperative and optative forms. The indicative (3rd sg.), imperative (2nd sg.) and optative (3rd sg.) forms in (8) represent the norm at Pāṇini's time:

- (8) (a) *sad* 'sit down': *s̄dati*, *s̄da*, *s̄det*  
      (b) *śru* 'hear, listen': *śr̄noti*, *śr̄nu*, *śr̄nuyāt*

Such forms occur in earlier Vedic, which, however, also has nonindicative forms like 2nd sg. imper. *sada*, *śr̄udhi*, from the aorist stems. Early Vedic had comparable modal forms also to perfect stems (e.g., *sasadyāt* [*sad* 'sit']).

This abundance is also to be found in other categories. Thus, there is a large number of accusative, dative, ablative, and genitive forms used as infinitives that Pāṇini explicitly recognizes as proper to Vedic usage only, and he also mentions Vedic absolutives in *-tvāya*, *-tvī*, *-tvīnam* in addition to those in *-tvā*.

Some formations that were eliminated in what came to be viewed as standard Sanskrit are reflected in Middle Indo-Aryan.

### 3. Middle Indo-Aryan

#### 3.1. Inflectional morphology

##### 3.1.1. General

Middle Indo-Aryan morphology differs from that of Old Indo-Aryan with respect to the number of contrasts observed. The dual is eliminated as a distinct category, and the use of ablaut variants is severely restricted both in the nominal and verbal systems. The number of nominal cases is diminished. In the verb system also the number of contrasts gradually lessens. The contrast between active and medio-passive endings is gradually eliminated, and the large number of stems within categories is reduced.

##### 3.1.2. Nominal morphology

The major differences between Old and Middle Indo-Aryan **nominal morphology** consists in the reduction of contrasts in the latter. The number of cases formally distinguished is reduced. In general, the dative is replaced by the genitive, except in expressions of purpose. The instrumental and ablative plural merge in the Pāli type *purisehi* 'men', which has the ending *-hi* found in other instrumentals (cf. Old Indo-Aryan *-bhīs*). In addition, the con-

trast between nominal and pronominal endings of the types *-āt*, *-e*, *-āyās*; *-smāt*, *-smin*, *-asyās* in Old Indo-Aryan (see 2.1.1) tends to be levelled, so that Pāli has nominal forms of the type *purisamhā*, *purisamhi* (*purisa-* 'man', cf. also Ásokan *vijitamhi* 'in the empire'), and pronominal *tāya*, comparable to *tamhā*, *tamhi*, and *kaññāya* (*kaññā-* 'girl'). Moreover, ablaut alternation of the type *agni-lagne-*, *vāyu-lvāyo-* is gradually eliminated.

The loss of final consonants, along with analogic remodelings, results in the replacement of consonant-stems by vowel-stems, with alternants found in early stages. For example, Pāli *gacchanto* 'going', *mahanto* 'great:NOM.SG.M', *manassa* 'mind:GEN.SG', are inflected like Skt. *devas*, *devasya* (see Art. 5: 4.2.2, (10)) as opposed to Skt. *mahān*, *manasas*. These developments continued in other dialects and later times. As a consequence, there are paradigms of the type *devo* (see Art. 5: 5.2, (38)). According to the descriptions of Prākṛit grammarians and literary sources, the following are among Prākṛit forms equivalent to the forms of Sanskrit *agni-* 'fire', *vāyu-* 'wind', *pitr-* 'father', *bhartr-* 'husband' (see Tab. 158.1; Pischel 1965: 266–76).

Analogical extensions have been carried out, so that stems like *piara-*, *bhattāra-* have been abstracted. The obliteration of formal contrasts continues to develop in later Middle Indo-Aryan, so that, for example, in Apabhrāmī nominative and accusative forms merged (e.g., *dahamuhu* 'ten.faced:NOM.SG.M', *caūmuhu* 'four.faced:ACC.SG.M'), as did the ablative and genitive (e.g., *mahu* 'of, from me'), and etymologically instrumental forms could function as locatives, as in *dāhiṇābhāēm* 'in the southern section', *ṇiyaniyathā-nehim* 'in their proper places'.

One additional point that is worth mentioning concerns the accumulation of affixes in some case forms. For example, *-himto*, *-suṁto* (see Tab. 158.2(b)) obviously consist of *-him*, *-su* and an appended element *-to*. In Old Indo-Aryan, *-tas* was indeed used in forms that alternated with ablatives; e.g., *grāmatas* = *grāmāt* 'from the village'. Moreover, the R̄gveda already shows *-tas* appended to an actual case form: *patsyatas* 'at the feet', with *-tas* following a locative plural. The extension of such a procedure has been carried further in Middle Indo-Aryan.

Already in Old Indo-Aryan, the affix *-ka* was used not only in derivates with particular semantic shadings such as diminutive (e.g., *putra-ka-* 'little child, poor child') but also

## (a) Singular

1	2	3	5	6	7
<i>aggī</i>	<i>aggīn</i>	<i>aggīnā</i>	<i>aggīo/aggīnō/aggīhil aggīhiṁto</i>	<i>aggīnō/agissa</i>	<i>aggimmi</i>
<i>vāū</i>	<i>vāūn</i>	<i>vāūnā</i>	<i>vāūo/vāūnō/vāūhiṁto</i>	<i>vāūnō/vāūssa</i>	<i>vāūmni</i>
<i>piā</i>	<i>piaram</i>	<i>piuṇāl/piareṇā</i>		<i>piuṇol/piussa</i>	
<i>bhattā</i>	<i>bhattāram</i>	<i>bhattāreṇā</i>		<i>bhattuṇol/bhattassa</i>	

## (b) Plural

1	2	3	5	6	7
<i>aggī/aggīo/</i>	<i>aggī/aggīnō</i>	<i>aggīhi(m)</i>	<i>aggīhiṁto/</i> <i>aggīsuṁto</i>	<i>aggīnā(m)</i>	<i>aggīsu(m)</i>
<i>aggīnō</i>			<i>vāūhiṁtol</i>	<i>vāūnā(m)</i>	<i>vāūsu(m)</i>
<i>vāūl/vāūo/</i>	<i>vāūl/vāūnō</i>	<i>vāūhi(m)</i>	<i>vāūsuṁto</i>		
<i>vāūnō</i>			<i>piarāhiṁtol</i>	<i>piuṇam</i>	<i>piaresu(m)</i>
<i>piarol/piuṇo</i>	<i>piarel/piuṇo</i>	<i>piarehi(m)/</i> <i>piūhi(m)</i>	<i>piarehiṁto</i>	<i>piaraṇa(m)</i>	
<i>bhattārāl</i>	<i>bhattārel</i>	<i>bhattārehi(m)/</i>	<i>bhattārāhiṁtol</i>	<i>bhattārāṇa(m)</i>	<i>bhattāresu(m)</i>
<i>bhattānō</i>	<i>bhattārāl</i>	<i>bhattāūhi(m)</i>	<i>bhattārehiṁto</i>		
	<i>bhattāūnō</i>				

Tab. 158.2: Nominal system in Middle Indo-Aryan

with no difference from the base to which the suffix was added; e.g., *avi-*, *avi-ka-* ‘sheep’. This redundant affix and its variant *-aka-* became wide spread in Middle Indo-Aryan. For example, Aśokan *hakam* ‘I’ is equivalent to *aham*; as a result of phonological developments, in Apabhraṃśa, the form is *haiṁ*. Similarly, Apabhraṃśa nominative-accusative forms in *-ai* (m.) and *-aüṁ* (n.) are developed from such enlarged bases, and feminines in *-ī* developed from comparable enlarged stems (Old Indo-Aryan *-ikā*). These are the sources of New Indo-Aryan *-o*, *-ā*, *-u* *ṁ*, *-ī* in variable nouns and adjectives (see 4.1.1).

The Middle Indo-Aryan **pronominal system** is systematically like that of Old Indo-Aryan (see 2.1.1). The major differences between the two have to do with dialectal variations, phonological changes, and analogical developments. Interrogative and relative pronouns are characterized by *k-* and *j-*, the latter due to a phonological change of *y-* – still found in early Middle Indo-Aryan (Pāli *yo* (nom. sg. m.) etc.) – to *j-*. Personal pronoun forms like Pāli *amhehi* (instr./abl. pl.), *amhesu* (loc. pl.) ‘us’ (Pkt. *amhehiṁ*, *amhesu*) have *amhe-* as in the nom./acc. pl. *amhe*. Similarly, *tumhehi*, *tumhesu* ‘you’ (Pkt. *tumhehiṁ*, *tumhesu*) have *tumhe-*, as in the nom./acc. *tumhe*; and

forms with *yu-* have been eliminated in favor of forms with *tu-*. Apabhraṃśa again shows a considerable reduction in forms:

- (9) (a) Sg. 1. *haiṁ*, 2–3–7. *maiṁ*, 5–6. *mahu/majju*; pl. 1–2. *amhelamhaiṁ*, 3–7. *amhehim* 5–6. *amhaham*, 7. *amhāsu*;
- (b) Sg. 1. *tuhum*, 2–3–7. *taüṁlpaiṁ* 5–6. *taul/tujha/tudhra*; pl. 1–2. *tumhel* *tumhaṁ*, 3/7. *tumhehim*, 6. *tumhaham* 7. *tumhāsu*.

Middle Indo-Aryan also has demonstrative pronouns comparable to those of Old Indo-Aryan: *sa-/ta-* ‘that’ (e.g., Pāli *so* (nom. sg. m.), *sā* (f.), *tam* (acc. sg. m./f.), *te* (nom./acc. m.), *tā* (nom./acc. f.), and so on. Again various phonological and analogical changes account for differences between Old and Middle Indo-Aryan. For example, since final stops are lost, the Old Indo-Aryan contrast between the types Skt. *sarvam* ‘all’ and *anyat* ‘other’ is absent; thus Pāli *añña-* ‘other’ has an inflection like that of *sabba-* ‘all’. Pāli *imam* ‘this:ACC.SG.F’ as opposed to Skt. *imām* shows a regular shortening of *-ā* in final syllable, but forms like *iminā*, *imamhā*, *imassa*, *imamhi* (m./n. sg. 3, 5, 6, 7), *imehi*, *imesam*, *imesu* (m./n. pl. 3/5, 6, 7) and *imayā*, *imāya*, *imissā*, *imissam* (f. sg. 3, 5, 6, 7), *imāhi*, *imā-*

*sam*, *imāsu* (f. pl. 3/5, 6, 7) show a generalization of *ima*-*.* Similarly, Pāli forms like *amūhi*, *amūsam*, *amūsu*, used for all genders, show a generalization of *amū-* at the expense of *amī-*.

In late Middle Indo-Aryan, Apabhraṃśa has *tahāṁ*, *jahāṁ*, *kahāṁ*, *tahim*, *jahim*, *kahim* equivalent to Old Indo-Aryan ablative and locatives *tasmāt* ‘from that ...’, *yasmat* ‘from which ...’, *kasmāt* ‘from which ...?’, *tasmin*, *yasmin*, *kasmin*.

### 3.1.3. Verb morphology

In its course of development, Middle Indo-Aryan also eliminated many contrasts found in the Old Indo-Aryan verb system. One now has a system with present, future, and preterit stems. The perfect and the periphrastic future are done away with. In addition, forms with active endings gradually replace forms with middle endings, as in presents like Pāli *maññati* ‘thinks’, *vuccati* ‘is said’; contrast Skt. *man-ya-te*, *uc-ya-te*, with the middle ending *-te*.

Ablaut variations are gradually eliminated in favor of single stems; e.g. Pāli *e-ti* (3rd sg.) ‘go:3.SG (goes)’, *e-nti* ‘go:3.PL (go)’ as opposed to Skt. *e-ti*, *y-anti*. In addition, the present stem tends to function as the basis for deriving stems of other verb forms; e.g., Pāli *gacch-a-ti* ‘goes’, fut. *gacch-issa-ti*, aor. *agacch-i* (*agañch-i*), past participle *gacch-ita* as opposed to Skt. *gacch-a-ti*, *gam-isy-a-ti*, *agam-a-t*, *ga-ta-*, which also have counterparts in Middle Indo-Aryan. In later Middle Indo-Aryan, moreover, the present stems themselves are generally reduced to two prevalent types, in *-a-* and *-e-*; e.g., *kar-a-i* ‘does’, *kah-e-i* ‘says, tells’ (Pāli *katheti*, Skt. *kathayati*). Further, these two types could alternate, as in *sun-a-/su-e* (*sun-a-i*, *suṇe-i*) ‘hear, listen’.

In early Middle Indo-Aryan, the prevalent preterital forms were sigmatic (e.g. Pāli *agacch-is-arṇ* ‘I went’). However, in the course of history, Middle Indo-Aryan tended to eliminate verbally inflected preterital forms in favor of past participles of the type *gaya-*, *gaa-* ‘gone’, *kaya-*, *kaa-* ‘done, made’ (Skt. *ga-ta-*, *kṛ-ta-*). If such participles derive from transitive verbs, they agree in number and gender with nominals denoting objects, but derivates from intransitives agree with subjects. By the late Middle Indo-Aryan period, this was the norm.

Future stems of the Pāli type *karissati* ‘do:FUT:3.SG (will do)’ represent phonological outcomes of Old Indo-Aryan *-sya-* (Skt. *kari-*

*syati*). Due to a phonological change, moreover, later Prakrit has not only a future with *-ssa-* but also one with *-ha-*, as in *pucchihām* ‘I will ask’.

Suppletion of course occurs also in Middle Indo-Aryan. For example, the past participle corresponding to the Prakrit present *jā-* (3rd sg. *jā-i*) ‘go’ is *gaya-/gaa-*. Moreover, as a result of phonologic and analogic developments, there are other instances where suppletion can justifiably be invoked although etymologically a single root is involved. For example: *ne-i* ‘leads’ (phonologically <*naya-ti*>) : *nī-ya-* (<*nī-ta-*>), *suṇa-i* ‘hears, listens’ (generalized stem *suṇa-* <*śr-ṇv-a-* [cf. Skr. *śrṇvanti*]) : *su-ya-*. Etymologically, the present and participle forms contain the same base.

Middle Indo-Aryan maintains not only indicative but also imperative and optative forms, with varying developments at particular stages in different dialects. For example, Pāli *gaccha*, *gacchāhi* ‘go:IMP.2.SG’, *gacchatha* ‘go:IMP.2.PL’, *karotu* ‘go:IMP.3.SG (should go)’, *karontu* ‘go:IMP.3.PL’, *gaccheya* ‘go:OPT.3.SG’.

Middle Indo-Aryan also has passives, again with various dialect forms. Passives like Pāli *vuccati* represent developments of passives with the suffix *-ya-*. In addition, there are passives with a suffix *-iya-/īya-*, as in Pāli *har-iya-tilhar-īya-ti* ‘is taken away’, Pkt. *pucch-ijja-i* ‘is asked’ (*pucch-a-ti*, *pucch-a-i* ‘asks’). Moreover, new passive aorists based on present stems were developed, as in Pāli *chijjīnsu* ‘were cut down’ (pres. *chijjanti* ‘are cut down’).

## 3.2. Derivational morphology

### 3.2.1. Nominal derivation

Nominal forms derived from verbs are typologically like those of Old Indo-Aryan, with variation: **action nouns** such as Pāli/Pkt. *karaṇa-* ‘doing’; **agent nouns** such as Pkt. *kattār-* (nom. sg. *kattā*, acc. sg. *kattāram*, etc.; see Tab. 158.2); present and past participles (e.g., *gacchanta-*; cf. 3.1.1), *gacchita-*, *gaya-*, *gaa-*); gerundives (e.g., Pāli *kātabba-* ‘to be done’, *dassanīya-* ‘worthy of being seen’; Skt. *kartavya-*, *darśanīya-*); absolutives (e.g., Pāli *nikkham-itvā* ‘after leaving’); infinitives (e.g., Pāli *pappotum* ‘obtain’; Pkt. *soum* ‘hear, listen’). As can be seen, the rules of distribution differ from those of Old Indo-Aryan: *nikkhamitvā* has *-tvā* after a complex with a preverb and *pappo-tum* has *tum* added to the present stem, but Sanskrit has *niṣkramya*,

with *-ya*, and *prāptum*, in which the derivational suffix follows the verb root. In addition, Middle Indo-Aryan has some derivates not found in standard Sanskrit; for example, Prakrit absolutives with *-ānal-iṭṇa*, as in *dāṭṇa* ‘after giving’, *genh-iṭṇa* ‘after taking’. Apabhrāṁśa has a series of different gerundives, absolutives, and infinitives. Hemacandra (4.438–441; cf. Vaidya 1958) enumerates the following: gerundives in *-ievvaiūṁ* (e.g., *kariievvaiūṁ* = *kartvayam*), *-evvaiūṁ* (*sahевваиūṁ* = *sodhavyam* [*sah* ‘bear’]), *-evā* (*soevā* = *svapitavyam* [*svap* ‘sleep’]); absolutives in *-i* (*māri* = *mārayitvā* ‘after first killing’), *-iu* (*adohiu* ‘without plunging into’), *-ivi* (*cumbivi* = *cumbitvā* ‘kissing’), *-avi* (*vichodavi* ‘letting go of’), *-eppi* (*jeppi* = *jītvā* ‘conquering’), *-eppiṇu* (*deppiṇu* = *dattvā* ‘giving’), *-evi* (*levi* ‘taking’), *-eviṇu* (*jhādeviṇu* = *dhyātvā* ‘meditating on’); infinitives in *-evaṁ* (*devaṁ* = *dātum* ‘give’), *-anya* (*karana* = *kartum* ‘do’), *-anahāṁ*, *-anahim* (*bhuñjaṇahāṁ*, *bhuñjaṇahim* = *bhoktum* ‘enjoy’), *-eppi* (*jeppi*), *-eppiṇu* (*caeppiṇu* ‘abandon’), *-evi* (*pālevi* = *pālayitum* ‘maintain’), *-eviṇu* (*leviṇu*). As can be seen, some absolute and infinitive forms are identical. Moreover, the infinitives in question are etymologically absolutives in origin. For the use of absolutives as infinitives, compare also *pievae laggā* ‘they began to drink’, with the locative singular of a gerundive of *pi* ‘drink’ (Bhayani 1953, ed.: 70).

The formation of derivates signifying degrees of qualities by means of suffixes continues in Middle Indo-Aryan. For example: Pāli *piya-tara-* ‘dearer’, *sat-tama-* ‘best’, Pkt. *daḍha-yara-* ‘firmer’, *pia-ama-* ‘dearest’.

As in Old Indo-Aryan, so too in Middle Indo-Aryan there are derivates formed from pronouns. For example: Pāli *yāvat-* ‘as much’, *tāvat-* ‘that much’, *yathā* ‘as’, *tathā* ‘thus’, *kathā* ‘how?’ and their Prakrit equivalents *jāvaṭa-*, *tāvaṭa-*, *jahāljaḥ*, *tahāltaha*, *kahāmī*.

Middle Indo-Aryan also has compounds comparable to those of Old Indo-Aryan. E.g. *tatpuruṣa*: Pāli *aggi-sikhā-* ‘flame of fire’, *suvanṇa-kāra-* ‘goldsmith’, Pkt. *ahinda-* (*ahi* + *inda*) ‘king of snakes’, *kumbha-āra-* ‘pot maker’; *bahuvrīhi*: Pāli *rattakkha-* (*ratta* + *akkha*) ‘red-eyed’, Pkt. *candarīnsu-* (*candā* + *arīnsu*) ‘hot-ray, i.e. the sun’; *dvandva*: Pāli *ajelakā-* (*aja* + *elaka*) ‘goats and sheep’, Pkt. *ammā-piyara-* ‘parents’; *avyayībhāva*: Pāli *abhi-dosam* ‘at evening time’, Pkt. *upa-gaṇ-garī* ‘near the Gangā’.

One difference between Old and Middle Indo-Aryan compounds is due to phonologi-

cal behavior (see Davane 1956: 144–151). Such phonological differences are evident also in other instances. For example, the Apabhrāṁśa equivalents of Skt. *yādṛṣa-* ‘of what sort’, *tādṛṣa-* ‘of that sort’ (Pāli *yādisa-*, *tādisa-*, Pkt. *jārisa-*, *tārisa-*) are *jaṭsa-*, *taṭsa-*, and corresponding to the Old Indo-Aryan number words for ‘11’ to ‘19’ (see 2.2.1, (7)), Apabhrāṁśa has:

- (10) ‘11’ *eāraha*
- ‘12’ *bāraha*
- ‘13’ *teraha*
- ‘14’ *caīuddaha*
- ‘15’ *panṇaraha*
- ‘16’ *solaha*
- ‘17’ *dahasatta*
- ‘18’ *aṭṭhāraha*
- ‘19’ *eguṇavīnsā*

The first vowels of the taddhita derivates Pāli *nemittika-*, Pkt. *ṇemittia-* ‘one who knows the omens’, Pāli/Pkt. *komāra-* ‘juvenile’ are *e*, *o*, as opposed to *ai*, *au* of the corresponding Sanskrit derivates *naimittika-*, *kaumāra-*. This is due to the change of *ai au* to *e o*.

### 3.2.2. Verb derivation

Causatives generally contain *-e-* (< *-aya-*) as in Pāli *kār-e-ti*, Pkt. *kār-e-i* ‘has ... do’. As there are alternant noncausative present stems with *-al-e* (Pkt. *suna-i/sune-i* ‘hears, listens’), so are there causatives with *-a-* and *-e-*; e.g., *pāde-ilpāda-i* ‘causes to fall’. Further, due to sound changes, consonant alternation came to be linked with causative derivation, as in *toda-i* ‘breaks (trans.)’ versus *tutṭa-i* ‘breaks (intr.)’; contrast Skt. *trutyati* : *troṭayati*. A type with *-āp-ay-a-* was gaining in productivity from early Middle Indo-Aryan on. Thus, Aśoka uses *lekh-āpi-ta-* ‘caused to be inscribed’, and causatives with this extended affix are formed to present stems, as in Pkt. *pucch-āve-* ‘have ... ask’. In addition, there are alternative causatives in *-āve-l-āva-*, e.g., *kar-āve-l kar-āva-* ‘have ... do, make’. Apabhrāṁśa also has causatives such as *bham-ād-a-* ‘cause to turn’. Desideratives and intensives of the Pāli types *bubhukka-* (3rd sg. pres. *bubhukkhati*) ‘wish to eat’, Pkt. *susūssa-* (*susussaī*) ‘attend to, serve’ and Pāli *caṅkama-* (*caṅkamati*), Pkt. *cakkamma-* (*cakkammaī*) ‘stride’ reflect the Old Indo-Aryan types *bubhukṣa-* (*bubhukṣate*), *śuśrūṣa-* (*śuśrūṣate*), *caṅkramya-* (*caṅkramyate*).

Denominatives are formed, as in Old Indo-Aryan, with a suffix *-ya-*, e.g., Pāli *sukhāyati*, Pkt. *suhāadi* ‘is happy’. In addi-

tion, there are denominatives such that endings follow a nominal directly; e.g., *missaī* ‘mixes’.

#### 4. New Indo-Aryan

New Indo-Aryan grammatical systems represent results of developments whose beginnings appear in earlier stages, but the outcome makes this stage considerably different from Old Indo-Aryan.

##### 4.1. Nominal morphology

###### 4.1.1. Nominal system

The old system of inflection, whereby a single inflected form signalled both case and number, has been greatly reduced, with an increasing use of postpositions. In a large part of the Indo-Aryan area, there is an opposition between **direct** forms, used independently, and **oblique** forms. The latter occur with postpositions and connectives – postposed elements that function in the manner of genitive endings in the earlier stages – followed by postpositions. The following examples from Gujarati and Hindi will illustrate:

###### (11) Gujarati

- (a) *chokro* ‘boy:DIR.SG’  
*chokrī* ‘girl:DIR.SG’  
*chokrā(o)* ‘boy:DIR.PL’  
*chokrī(o)* ‘girl:DIR.PL’
- (b) *chokrā* ‘boy:OBL.SG’  
*chokrī* ‘girl:OBL.SG’  
*chokrā(o)* ‘boy:OBL.PL’  
*chokrī(o)* ‘girl:OBL.PL’
- (c) *chokrā-ne* (obj.) ‘the boy, to the boy’  
*chokrā-thī* ‘from, by the boy’  
*chokrā-ne-māṭe* ‘for the boy’  
*chokrā-nī-pāchal* ‘behind the boy’  
*chokrā-nī-sāthe* ‘with the boy’  
*chokrī-ne* ‘the girl, to the girl’  
*chokrī-thī* ‘from, by the girl’  
*chokrī-ne-māṭe* ‘for the girl’  
*chokrā-nī-pāchal* ‘behind the girl’  
*chokrī-nī-sāthe* ‘with the girl’
- (d) *chokrā(o)-ne*  
*chokrā(o)-thī*  
*chokrā(o)-ne-māṭe*  
*chokrā(o)-nī-pāchal*  
*chokrā(o)-nī-sāthe*  
*chokrī(o)-ne*  
*chokrī(o)-thī*  
*chokrī(o)-ne-māṭe*  
*chokrī(o)-nī-pāchal*  
*chokrī(o)-nī-sāthe*

###### (12) Hindi

- (a) *laḍkā* ‘boy:DIR.SG’  
*laḍkī* ‘girl:DIR.SG’  
*laḍke* ‘boy:DIR.PL’  
*laḍkiyāṁ* ‘girl:DIR.PL’
- (b) *laḍke* ‘boy:OBL.SG’  
*laḍkī* ‘girl:OBL.SG’  
*laḍkom* ‘boy:OBL.PL’  
*laḍkiyom* ‘girl:OBL.PL’
- (c) *laḍke-ko* (obj.) ‘the boy, to the boy’  
*laḍke-se* ‘from, by the boy’  
*laḍke-ke-liye* ‘for the boy’  
*laḍke-ke-pīche* ‘behind the boy’  
*laḍke-ke-sāth* ‘with the boy’  
*laḍkī-ko* ‘the girl, to the girl’  
*laḍkī-se* ‘from, by the girl’  
*laḍkī-ke-liye* ‘for the girl’  
*laḍkī-ke-pīche* ‘behind the girl’  
*laḍkī-ke-sāth* ‘with the girl’
- (d) *ladkom-ko*  
*ladkom-se*  
*ladkom-ke-liye*  
*ladkom-ke-pīche*  
*ladkom-ke-sāth*  
*ladkiyom-ko*  
*ladkiyom-se*  
*ladkiyom-ke-liye*  
*ladkiyom-ke-pīche*  
*ladkiyom-ke-sāth*

There are also nominals that do not have different direct and oblique forms throughout. For example, Hindi *ādmī* ‘man’ has an oblique plural *ādmiyom*, but *ādmī* occurs otherwise; a feminine noun such as *mez* ‘table’ has a direct plural *mezem* and an oblique plural *mezom*. Gujarati *koṭ* (m.) ‘coat’, *vāṭ* (f.) ‘matter, story’, *kām* (n.) ‘work’ show no variation between direct and oblique forms. The contrast between direct and oblique forms is not always observed elsewhere also. For example, eastern languages like Bangla lack the contrast in the singular. On the other hand, even in such instances, the contrast appears for certain forms. For example, Bangla *chele-ra* ‘boys’ (nom. pl.), *chele-d-er* ‘of the boys’, *chele-d-er-ke* ‘(to) the boys’ (colloquial *cheleder*).

As can be seen, New Indo-Aryan languages agree in the general opposition of direct versus oblique forms and the use of postpositions but differ with respect to the exact forms of the nominals and postpositions. Moreover even where postpositions are the norm, inflectional forms occur. For example, Hindi has a vocative plural of the type *laḍko* ‘boys’, *ladkiyo* ‘girls’, and Panjabi has abla-

tive singular (e.g., *khetom* ‘from the field’) and locative plural (e.g., *khetīn* ‘in the fields’) forms for subsets of nouns. In addition, Gujarati has a plural marker *-o-*, which can be omitted if plurality is otherwise shown (e.g., by use of a number word) in a context. Further, the connective *-nā-/ne-/nī* can be omitted in Gujarati in certain instances. Comparable differences are seen in other languages.

The status of postpositions is like that of the English genitive *s* in that they regularly follow phrases, as in Gujarati *pelā māṇas-na dīkrā-ne*, Hindi *us ādmī-ke bētē-ko* ‘that man’s son’, where the postpositions *-ne* and *-ko* follow phrases with oblique forms. In addition, more than one postposition can be attached to a nominal. For example, in Gujarati *ophis-e-thī* ‘from the office’, Hindi *ghar-mem-se* ‘from the house’, the postpositions *-thī*, *-se* follow complexes that include the locative postpositions *-e*, *-mem*. Such usage has parallels in earlier Indo-Aryan (see 3.1.2). Indeed, the use of postpositions goes back to earliest Indo-Aryan (see 2.3), and Rgvedic *samiddhasya ... purastāt* ‘... in front of the lighted (fire)’ (see Bloch 1965: 159) is equivalent to the Hindi type *āg-ke sāmne* ‘before the fire’, in which *sāmne* ‘before’ with the connective *-ke* follows *āg* ‘fire’. What sets the New Indo-Aryan system apart is that it involves direct forms, which occur independently, in opposition to oblique forms that occur with postpositions.

In the western area of Indo-Aryan, languages like Gujarati, Marathi, and Konkani have three formal genders; e.g., Gujarati *chokro* ‘child:DIR.SG.M (boy)’, *chokrī* ‘child:DIR.SG.F (girl)’, *chokrum* ‘child:DIR.SG.N’ (*chokrām* ‘child:DIR.PL.N’); adjectives *nāno*, *nānt*, *nānum* ‘small’. In other areas, two-genders contrast, as in Hindi *ladkā* versus *ladkī* ‘girl’, *chotā*, *chotī* ‘small’. In the east, on the other hand, languages like Bangla lack such formal gender distinctions.

#### 4.1.2. Pronominal system

The New Indo-Aryan pronominal system has reflexes of the earlier types: personal, demonstrative, interrogative, indefinite, relative. E.g., Hindi *maiṁ* ‘I’ (*mujhe* ‘me’, *mujh-se* ‘from me, by me’, etc.), *ham* ‘we’ (*hamer* ‘us’, *ham-se* etc.); *kaun* ‘who?’, *kyā* ‘what?’ (obl. sg. *kis-*, pl. *kin-*, pl. obj. *kinhem*), *koi* ‘some one’ (obl. *kisī-*), *jo* ‘who, which’ (obl. sg. *jis-*, pl. *jin-*, pl. obj. *jinhem*). In some areas, reflexive pronouns occur obligatorily in certain

syntactic configurations; e.g., Hindi *apnā* (obl. pl. *apne*), *apnī*.

In Indo-Aryan, as in many other areas of Indo-European, personal pronouns retain morphological complexities that are eliminated elsewhere. Thus, Hindi has *maiṁ*, also used in the agentive *maiṁ-ne*, opposed to *mujh-*, and the derivate equivalent to a noun with *-kā* is *merā* (f. *merī*): *merā bētā* ‘my son’, *merī bētī* ‘my daughter’, *mere liye* ‘for me’. Correspondingly, *tum* ‘you’ pairs with *tumhārā* (*tumhārī*) ‘your’. Gujarati *hum* ‘I’, *mē* (agentive), *māruṁ* and *tame* ‘you’ (also agentive), *tamāruṁ* has comparable variants. New Indo-Aryan languages also show the effects of regularization. For example, Old Gujarati had forms such as the accusative singulars *mujha* ‘me’, *tujha* ‘you’, which have been replaced by *mane*, *tane* in modern Gujarati.

New Indo-Aryan languages differ with respect to degrees of **deixis**; e.g., Hindi *yah* (obl. sg. *is-*, pl. *in-*) ‘this’, *vah* (obl. sg. *us-*, pl. *un-*) ‘that’; Bangla *e* ‘this one’ (obj. sg. *e-ke*, gen. sg. *e-r*), nom. pl. *e-rā*, obj./gen. pl. *e-der*), *se* [ʃe] ‘he, she it’ (obj. sg. *tā-ke*, gen. sg. *tār*, nom. pl. *tā-rā*, obj./gen. pl. *tā-der*), *o* ‘that one’ (*o-ke*, *o-r*, *o-rā*, obj./gen. pl. *o-der*); Gujarati *ā* ‘this’, *e* (agentive sg. *ene*, pl. *enē*), *te* (agentive sg. *tene*, pl. *tenē*), ‘that’, *pelum* (m. *pelo*, f. *pelā*, etc.) ‘yonder one’. Languages of the west also distinguish inclusive and exclusive first person plural pronouns, as in Gujarati *ame* ‘1.PE’ versus *āpne* ‘1.PI’ ‘we’; Marathi *āmhi*, *āpan*.

In general, modern Indo-Aryan observes a contrast of degrees of respect or distancing conveyed by the use of different second person pronouns: **honorific**, **ordinary**, **intimate**. Thus, Hindi *tum* is a non-distancing pronoun used towards close friends, one’s wife and such; *āp* is the more generally used, respectful pronoun; and *tū* is intimate. Similarly, Bangla has three degrees of formality represented in the use of *tui*, *tumi*, *āpni*, but these are singular pronouns with plural counterparts (*torā tomrā āpnāra*). In Gujarati, on the other hand, *tame* is the pronoun generally used, *tum* is used as the intimate pronoun, and *āp* is not generally used. If someone is addressed with this, it implies either extreme respect, even fawning, or that one is making fun of the person. The contrast between ordinary and honorific can be shown by formal contrasts for the third person also, as in Bangla: neutral *se* [ʃe] versus formal *tini* (pl. *tāra*, *tāmrā*). Even in those languages that do not have comparable formal distinctions among

third person pronouns, there are strategies to signal the difference: plural forms are used – including the neuter plural for respectful reference to a woman (as in Gujarati) – along with plural agreement in verb forms that take gender-number affixes, and an honorific second person pronoun can be used, as in Hindi (*āp*).

#### 4.2. Verb morphology

The sources of the verbal systems of New Indo-Aryan languages are of course found in late Middle Indo-Aryan (*Apabhraṃśa*) and early stages of New Indo-Aryan. By the late Middle Indo-Aryan period, as noted earlier (see 3.1.3), verbally inflected preterit forms had been eliminated in favor of nominal forms. Nevertheless, present indicative forms of the types *karaī* ‘does’, *kahei* ‘says’ continued in use, right down into the early New Indo-Aryan period, and such forms are still used in some languages as presents (e.g., see (18)). In the course of development, however, these too came not to be used as present indicative forms. For example, Gujarati and Hindi have the type *kar-ε*, *kar-e* ‘might do, may do’ (Old Gujarati, Old Hindi *kar-a-i*), but this does not function as a present indicative. In effect, the system represented in the greater part of New Indo-Aryan languages is based on a contrast between **imperfective** and **perfective**, along with temporal auxiliary forms. The following Gujarati examples will illustrate:

- (13) (a) *mohan śum̥ kar-ε*  
Mohan what do-3  
‘what might Mohan do?’
- (b) *mohan śum̥ kar-ε ch-ε*  
Mohan what do-3 AUX-3  
‘what is Mohan doing?/what does Mohan do?’
- (14) (a) *mohan śum̥ kar-t-o*  
Mohan what do-IPFV-M.SG  
*ha-t-o*  
AUX-IPFV-M.SG  
‘what was Mohan doing?’
- (b) *mohan-e śum̥ kar-y-uṁ*  
Mohan-AGT what do-PFV-N.SG  
‘what did Mohan do?’

The imperfective stem *kar-t-* is followed by a gender-number element, as is the temporal auxiliary stem *ha-t-*. The imperfective stem *kar-t-* appears also in negative sentences of the type:

- (15) *mohan kaśum̥ nathī kar-t-o*  
Mohan something NEG do-IPFV-M.SG  
‘Mohan is doing nothing at all’

but in positive sentences of the type (13 b), one has what I call the neutral form *kar-ε* with the temporal auxiliary form *ch-ε*. The ergative type (14 b) has the perfective stem *kar-y-* along with a gender-number element. The imperfective and perfective markers appear also in *jam-t-* and *kar-y-* of examples such as in (16).

- (16) (a) *jam-t-ī wakhat-e*  
eat-IPFV-F time-LOC  
‘at the time to eat, at dining time’
- (b) *kām kar-y-ā pachī*  
work do-PFV-OBL after  
‘after doing the work’

Other languages have comparable system, with variations. Thus, Hindi distinguishes a **continuous** from a **non-continuous**, as in (17).

- (17) (a) *mohan kyā kām kar-t-ā*  
Mohan what work do-IPFV-M.SG  
*h-ai*  
AUX-3.SG  
‘what work does Mohan do?’
- (b) *mohan kyā kar rah-ā*  
Mohan what do remain-PFV.M.SG  
*h-ai*  
AUX-3  
‘what is Mohan doing?’

Although modern Hindi has the type *kar-e*, it does not use this in conjunction with a temporal auxiliary to form present forms, as in Gujarati.

**Habitual** and **continutive** presents contrast elsewhere also, in different ways. For example, in Bangla personal endings follow the base or the affix *-ch-*:

- (18) (a) *tumi ki kar-o [koro]*  
you what do-2.NEUTRAL  
‘what do you do?’
- (b) *tumi ki kar-ch-o [korcho]*  
you what do-IPFV-2.NEUTRAL  
‘What are you doing?’

The ergative type (14 b) also shows variations in different languages with respect to **agreement** phenomena. For example, contrast the following examples from Gujarati (19) and Hindi (20):

- (19) *mē tamār-ī dīkr-ī-ne jo-o-ī*  
1.SG.AGT your-F child-F-OBJ see-PFV-F  
‘I saw your daughter’

- (20) *maiñ-ne āp-kī bet-ī-ko*  
 1.SG-AGT you-POSS.F child-F-OBJ  
*dekh-ā*  
 see-PFV.M.SG  
 'I saw your daughter'

In (19), the perfective *jo-ī* ( $\leftarrow$  *jo-y-ī*) is feminine, agreeing with *dikr-ī-ne*, but in (20) the verb agreement is suspended, since *bet-ī-ko* has the postposition *-ko* in the object phrase. The agreement appears in an example like (21), where the object phrase *kitāb* 'a book' does not have *-ko*.

- (21) *maiñ-ne kitāb padh-ī*  
 1.SG-AGT book read-F  
 'I read a book'

In languages of the east, on the other hand, perfective forms show person number agreement of non-ergative systems. The following illustrate the difference (Gujarati, Hindi, Bangla):

- (22) (a) *hum̩ gayol/gaī* 'I went'  
*ame gayāl/gaī* 'we went'  
*tame gayāl/gaī* 'you went'  
*e gayol/gaī* 'he/she went'  
*e gayāl/gaī* 'they went'  
 (b) *maiñ gayāl/gaī* 'I went'  
*ham gael/gaīñ āp gael/gaīñ* 'we went'  
*vah gayāl/gaī* 'he/she went'  
*ve gael/gaīñ* 'they went'  
 (c) *āmil/āmrā gelum (gelām)* 'I/we went'  
*tumil/tomrā gele* 'you went'  
*se [ʃe]/tarā gela [gelo]* 'he/she/they went'

The perfectives *gayo*, *gai*, *gaya*, *gaī*, and so on in (22 a-b) agree in gender and number with the subjects, but in (22 c) the verb forms differ according to person; similarly, Bangla *āmi karlum* [korlum] and so on from the base *kar* 'do, make'.

In addition, New Indo-Aryan languages have **future stems** with various formants. For example, Gujarati has a future with *-ś-* that shows person-number differences: *kar-iś* 'I shall do', *kar-ś-e* 'he/she/they will do', *kar-iś-ūñ* 'we will do', *kar-ś-o* 'you will do'. Bangla *kar-b-o* [korbo] (1st person) *kar-b-e* [korbe] (ordinary 2/3 person) 'will do' has person markers following the future suffix *-b-*. Hindi, on the other hand, has a future with *-g-* that shows not only person-number but also gender differences: sg. *kar-uñ-gā-ī* 'I shall do', *kar-e-gā-ī* 'he/she shall do', pl. *kar-eñ-gē-ī*.

Forms of the type Hindi *kar-e*, Gujarati *kar-e* have modal functions. They are used,

for example, in the protasis of a **conditional sentence** as well as in sentences like (23) from Hindi (23 a) and Gujarati (23 b).

- (23) (a) *maiñ kyā kar-uñ*  
 1.SG.AGT what do-1.SG  
 'what am I to do?'  
 (b) *hum̩ sūñ kar-uñ*  
 1.SG.AGT what do-1.SG  
 'what am I to do?'

There are also forms used in contrary-to-fact conditional sentences, as in the following examples from Gujarati (24 a) and Hindi (24 b):

- (24) (a) *mārī-pāse paisā ho-t to*  
 me-by money be-COND then  
*tamārī-sāthe jā-t*  
 you-with go-COND  
 (b) *agar mere-pās paise ho-t-e*  
 if by me money be-COND-M.PL  
*to āp-ke sāth jā-t-ā*  
 then you-OBL with go-COND-M.SG  
 'If I had money I would go with you.'

In Gujarati, forms of the type *hot*, *jāt* have *-t* following a verbal base; in Hindi, *-t-* is followed by a gender-number affix: m. sg. *hotā*, pl. *hote*. The forms in question, however, can be neutral with respect to time reference, as is the Sanskrit conditional of the type *abhavīyat*. Historically, the forms with *-t-* reflect the use of present participles in conditional sentences, a usage that goes back to late Middle Indo-Aryan.

In addition, New Indo-Aryan languages have formally distinct **imperatives**. For example, the imperatives of the bases *baith*, *bas* [bɔʃ/bɔ], *bes* [bes] 'sit' linked to the use of the pronouns Hindi *tū tum āp*, Bangla *tui tumi āpni*, Gujarati *tum̩ tame* are:

- (25) (a) Hindi: *baith*, *baith-o*, *baith-iyē*  
 (b) Bangla: *bas*, *bas-o*, *bas-ūñ*  
 (c) Gujarati: *bes*, *bes-o*

The first consists simply of the base; others have particular affixes, as shown. In addition, Hindi allows the use of imperatives in *-o* with the pronoun *āp*: *āp āo*. Various other forms function as imperatives, with semantic shadings. For example, Hindi has imperatives with *-iegā*, as in *ā-iegā*, which are used to attenuate the feeling of command, and Gujarati has forms with *-je*, *-jo* (the latter more deferential), as in *āv-je*, *āv-jo*; the last cited form is also the common expression for 'goodbye'.

In addition, there are third person imperative forms, e.g., Bangla *kar-uk*, Gujarati *kar-o* ‘... should do’. Moreover, some languages distinguish tenses for imperatives; thus, the Bangla forms in (25 b) are present imperatives, and the corresponding future imperatives are: *kar-is*, *kar-o*, *kar-ben*. The very polite imperative Gujarati type *kar-s-o-ji* ‘(you will) please do’ contains a formal future.

As in earlier stages, there is **suppletion** in New Indo-Aryan, and the situation is comparable to that of Middle Indo-Aryan (see 3.1.3). Thus, the Hindi preterit forms *thā the thī thīm* (m. sg., pl., f. sg., pl.) of the temporal auxiliary correspond to present forms with *h-*: *hūm* (1st sg.), *hai* (2nd sg. [t̪e], 3rd sg.), *ho* (2nd [tum]), *haiṁ* (2nd [āp], pl.); in Gujarati, the preterit forms corresponding to the present *ch-* (*hum̄ chum̄*, *tum̄ che*, *e che*, *amelāpne chie*, *tame cho*, *e che*) are *hato*, *hatā*, *hatī*, *hatum̄*. The following examples of infinitives and perfectives will serve further to illustrate:

- (26) (a) Hindi  
*de-nā : di-yā* ‘give’  
*jā-nā : ga-yā* ‘go’  
(b) Gujarati  
*āp-vum̄ : dī-dhum̄* ‘give’  
*ja-vum̄ : ga-yum̄* ‘go’

As expected, levelling has also taken place; for example, in standard Gujarati the perfective of *jo-vuṁ* ‘see, look’ is *jo-yuṁ*, though *diṭhum̄* occurs in dialects.

New Indo-Aryan languages have **passive** formations, but these occupy a special status, for two reasons. Passive sentences that correspond to actives usually involve obligatory omission of an agent expression. Passive sentences with agent expressions have particular semantics: that one is or is not capable of doing something, that one has done something accidentally, and are most commonly negative. Moreover, passive formations are not limited to transitive verbs. Passives are formed both with particular affixes and by periphrasis. For example, Gujarati has a passive suffix *-ā-*, as in *kar-ā* ‘be done’, *av-ā-* (passive to *āv-* ‘come’). In Hindi, on the other hand, a passive consists of a perfective form of the verb followed by a form of *jā* ‘go’, as in *kīyā jā-*, *calā jā-* (*cal-nā* ‘move’).

#### 4.3. Derivation

##### 4.3.1. Nominal derivation

New Indo-Aryan languages have a large variety of nominals derived from verbs, such as Hindi *lad-āt* ‘fighting’ (*lad-nā* ‘fight-INF’), *bah-āv* ‘flow’ (*bah-nā* ‘flow-INF’), *ruk-āvat*

‘obstruction’ (*rok-nā* ‘stop-INF’). Of particular interest are derivates that are generally classed as **infinitives** and **absolutives**. The examples from Hindi (27 a) and Gujarati (27 b) both of which mean ‘Allow him to go’, have the oblique forms *jāne*, *javā* in construction with an imperative of *de-* ‘give’.

- (27) (a) *us-ko jā-ne do*  
this-OBJ go-INF give:IMP  
‘allow him to go’  
(b) *ε-ne ja-vā dō*  
this-OBJ go-INF give:IMP  
‘allow him to go’

Action noun forms in *-nā*, *-vum̄* (e.g., *jā-nā*, *ja-vum̄*) serve as lexical entries for verbs in dictionaries of Hindi and Gujarati. Oblique forms of such derivates appear also with postpositions, as in Hindi *jāne-ke liye*, Gujarati *javā māte* ‘in order to go’. Hindi (28) and the equivalent Gujarati sentences (29) contain masculine *kar-nā*, feminine *padh-nī*, and neuter *kar-vum̄*, feminine *wāṁc-vī* in concord with the nouns *kām*, *kitāb* and *kām*, *copdī*.

- (28) (a) *mujhe kām kar-nā*  
to.me work do-ACNNR.M  
*pad-ā*  
fall-PFV.M.SG  
‘I had to work’  
(b) *āp-ko yah kitāb padh-nī*  
you-to this book read-ADJR-F  
*pad-eg-ī*  
fall-FUT-F  
‘you will have to read this book’  
(29) (a) *māre kām kar-vum̄*  
to.me work do-ACNNR.N  
*pad-y-um̄*  
fall-PFV-M.SG  
‘I had to work’  
(b) *tamāre ā copdī wāṁc-vī*  
you.to this book read-ACNNR.F  
*pad-ś-e*  
fall-FUT-F  
‘you will have to read this book’

In these examples, the verbal derivates function as adjectival modifiers, in the manner that gerundives function in earlier Indo-Aryan. Similarly, Gujarati *pī-vā-nuṁ pāṇī* ‘drinking water’ has an adjectival derivate *pīvānuṁ* (n. sg.) construed with *pāṇī* ‘water’. In Hindi, an oblique stem of the type *jā-ne* is followed by the suffix *vālā* to form agentive nouns such as *jānevālā* ‘one who is going, is about to go’, *rahnevālā* ‘one who dwells (in an area, resident of a place)’.

Another type of infinitive appears in *kar-ī śak* ‘be able to do’ of Gujarati:

- (30) *ā kām ām kar-ī śak-ā-y*  
 this work thus do-INF can-PASS-3  
 ‘this work can be done this way’

The equivalent modern Hindi infinitive has the simple base (*kar*): *kar sak-* ‘be able to do’.

A derivate in *-ī* occurs also as an absolute in Gujarati:

- (31) *kām kar-ī-ne gher jā-o*  
 work do-ABS-to home go-2  
 ‘do your work before you go home’

Other New Indo-Aryan languages have comparable constructions, with differences in how the absolutives are formed. For example, in Hindi, *kar* follows a verb base (e.g., *jā-kar*) other than *kar*; this takes *-ke* (*kar-ke*), which can be used with other verbs also.

**Attributive participles** corresponding to perfectives are formed in different manners. For example, Hindi has the construction type *padhīt hut kitab* ‘the book that has been read’, in which a perfective is followed by a form of the perfective of *ho-* (*huā, hue, huī, huīṁ*) agreeing in number and gender with the modified noun. In Gujarati, on the other hand, the equivalent construction has a suffix *-el-* after the perfective stem, as in *wāṁcelī cōpdī* ‘a book that has been read’, *pidhelo māṇas* ‘drunken man’. The formation with *-l-* occurs elsewhere and is doubtless also incorporated in verb forms of the type given in (22c).

Denominative formations involve affixation and composition. There is a large number of suffixes used in forming substantives and adjectives. Some examples from Hindi will suffice to illustrate: *lamb-āt* ‘length’ (*lambā* ‘long’); *bac-pan* ‘childhood’ (*baccā* ‘child’); *dukān-dār* ‘shopkeeper’ (*dukān* ‘shop’); *taiksi-vālā* ‘taxi driver’; *dhārmik* ‘religious’ (*dharma* ‘religion’). There are also prefixes, though less numerous. For example: Hindi *anu-kram* ‘sequence’; *gair-hāzir* ‘absent’ (*hāzir* ‘present’); *be-kār* ‘unemployed, useless’. Many such derivates involve elements and processes borrowed from Sanskrit and Perso-Arabic sources. Thus, *dharma* is a Sanskrit borrowing and *dhārmik* involves a Sanskrit vowel replacement along with the affix *-ik* (Skt. *-ika*).

As in earlier stages of Indo-Aryan, there are particular derivates from pronouns, illustrated here by the following examples

from Hindi (32a) and Gujarati (32b), with equivalent meanings:

(32) (a)	<i>aisā</i>	‘of this sort’
	<i>vaisā</i>	‘of that sort’
	<i>jaisā</i> (rel.)	‘of which sort’
	<i>kaisā</i>	‘of what sort?’
	<i>itnā</i>	‘this much’
	<i>utnā</i>	‘that much’
	(b) <i>āvūṁ</i>	‘of this sort’
	<i>evūṁ</i>	‘of that sort’
	<i>jevūṁ</i> (rel.)	‘of which sort’
	<i>kevūṁ</i>	‘of what sort?’
	<i>ātlūṁ</i>	‘this much’
	<i>eṭlūṁ</i>	‘that much’

There are also adverbial counterparts: Hindi *aise* ‘in this way’, *vaise* ‘in that way’, *kaise* ‘how?’, *yahāṁ* ‘here’, *vahāṁ* ‘there’, *jahāṁ* ‘where’, *kahāṁ* ‘where?’, *jab* ‘when’, *kab* ‘when?’. The equivalent Gujarati derivates are *ām*, *em*, *jem*, *kem*, *ahīṁ*, *tyāṁ*, *jyāṁ*, *kyāṁ*, *jyāre*, *kyāre*. Synchronously, these involve affixation, but historically a form like Hindi *jaisā* reflects an earlier compound, and Hindi *yahāṁ* and similar as well as Gujarati *em*, *jem* reflect earlier case forms (cf. 3.1.2).

**Compounds** such as Hindi, Gujarati *tār-ghar* ‘telegraph office’ and Hindi, Gujarati *mā-bāp* ‘parents’ conform to the tatpuruṣa and dvandva compound types. Compounds of the bahuvrīhi type are commonly Sanskrit (e.g., Hindi *pragati-śīl* ‘one with progressive character’). The productivity of composition and of different compound types varies across languages.

There are also compound-like sequences of closely connected units. Particularly noteworthy because it is not limited to Indo-Aryan is the “echo formation” of the type Hindi *pustak vustak*, Gujarati *cōpdī bōpdī*, Bangla *boi ̄toi* ‘book and such’: A nominal is repeated with a change of its first sound to form a sequence meaning ‘x and related matter’.

#### 4.3.2. Verb derivation

New Indo-Aryan languages also continue the use of morphologically derived **causatives**, reflecting both the earlier type Old Indo-Aryan *troṭayati*, Middle Indo-Aryan *todaī* ‘breaks’ and the more prevalent type with a morpheme that reflects the Middle Indo-Aryan generalization of *-āp-ay-* (Middle Indo-Aryan *-āp-e-, -āv-e-*). For example:

- (33) (a) *khul-* : *khol-* (Hindi)  
*khul-* : *khol-* (Gujarati)  
 ‘open (intr.)’ ‘open (trans.)’

<i>tūt-</i>	:	<i>tod-</i> (Hindi)
<i>tūt-</i>	:	<i>tod-</i> (Gurajati)
‘break’		‘break’
(b) <i>kar-</i>	:	<i>kar-ā-</i> (Hindi)
<i>kar-</i>	:	<i>kar-āv-</i> (Gurajati)
‘do, make’		‘have ... do, make’
<i>cal-</i>	:	<i>cal-ā-</i> (Hindi)
<i>cāl-</i>	:	<i>cal-āv-</i> (Gurajati)
‘move, go’		‘cause to move, to go’
<i>padh-</i>	:	<i>padh-ā-</i> (Hindi)
‘read, study’		‘teach’
<i>vāñc-</i>	:	<i>vañc-āv-</i> (Gurajati)
‘read’		‘have ... read’

A noteworthy characteristic of modern languages of India is the use of complexes usually referred to as compound verbs, involving particular verbs, which have been labelled **vector verbs**. The complex commonly involves particular semantic features. For example, Hindi *mār-* means ‘beat’ but *mār dāl-* means ‘kill’; cf. also: *bol-* ‘speak’ : *bol uth-* ‘blurt out’. There are also oppositions of the type *dēkh de-* ‘let someone have a look’ : *dēkh le-* ‘have a look’. In addition, aspectual features and restrictions concerning possible negation are linked to the use of compound verb complexes.

## 5. Transliteration and transcription

I generally follow the standard system of transliteration (see Art. 5), but where *ai*, *āi*, and so on are sequences of vowels – as in some Middle Indo-Aryan forms – I transliterate *ai* and so on. In addition, *ṁ* following a vowel indicates nasalization. For modern Indo-Aryan, I use a combination of transliteration and transcription. *ɖ* represents the flap commonly represented by *r*. Since vocalic [l] does not occur in Middle and New Indo-Aryan languages, *l* represents retroflex [l]. Although in some languages, such as Bangla, Gujarati, and Marathi, high vowels do not contrast by length, I still write long vowels in accordance with spelling conventions. If the script is ambiguous and an unambiguous representation is required, a transcription is given. For example, in Hindi the graphemes *ai au* represent [ɛ] [ɔ], so that a transliteration is used, but in Gujarati the same symbols (*e o*) are used to represent [e] [o] and [ɛ] [ɔ], so that I use IPA symbols for the lower vowels.

## 6. References

- Bhayani, Harivallabh Chunnilal (1953, ed.), *Pau-macariu (A Purānic Epic in Apabhrāṃśa)*, Composed by Kavirāja Svayambhūdeva. Bombay: Bharatiya Vidya Bhavan (Singhi Jain Series 34)
- Bloch, Jules (1965), *Indo-Aryan from the Vedas to Modern Times* [English edition, largely revised by the author and translated by Alfred Master; Original French: Paris, 1934]. Paris: Adrien Maisonneuve
- Bubenik, Vit (2003), “Prākrits and Apabhrāṃśa”. In: Cardona & Jain (eds.), 204–249
- Cardona, George (2003), “Sanskrit”. In: Cardona & Jain (eds.), 104–160
- Cardona, George & Jain, Dhanesh (2003), “General Introduction”. In: Cardona & Jain (eds.), 1–45
- Cardona, George & Jain, Dhanesh K. (2003, eds.), *The Indo-Aryan Languages*. London: Routledge
- Davane, G. V. (1956), *Nominal Composition in Middle Indo-Aryan*. Poona: Deccan College (Deccan College Dissertation Series 11)
- Geiger, Wilhelm (1965), *Pāli Language and Literature*, authorised English translation by Batakrishna Ghosh. Calcutta: University of Calcutta [Original German: Strassburg, 1916]
- Hinüber, Oscar von (2001), *Das ältere Mittelindisch im Überblick*. Wien: Verlag der Österreichischen Akademie der Wissenschaften (Österreichische Akademie der Wissenschaften, Philosophisch-historische Klasse, Sitzungsberichte 467, Veröffentlichungen zu den Sprachen und Kulturen Südasiens, Heft 20)
- Masica, Colin (1991), *The Indo-Aryan Languages*. Cambridge: Cambridge University Press
- Oberlies, Thomas (2001), *Pāli, A Grammar of the Language of the Theravaada Tipiṭaka with a Concordance to Pischel's Grammatik der Prakrit-Sprachen*. Berlin, New York: Walter de Gruyter (Indian Philology and South Asian Studies 3)
- Oberlies, Thomas (2003), “Āśokan Prakrit and Paali”. In: Cardona & Jain (eds.), 161–203
- Pischel, Richard (1965), *Comparative Grammar of the Prākrit Languages*, translated from the German by Subhadra Jhā. Delhi: Motilal Banarsiādass. [Original: Strassburg, 1900]
- Vaidya, P. L. (1958), *Prākrit Grammar of Hemacandra, Being the Eighth adhyāya of his Siddha-Hema-Śabdānuśāsana*. Poona: Bhandarkar Oriental Research Institute. (Bombay Sanskrit and Prakrti Series No. LX)

George Cardona, Philadelphia (U.S.A.)

## 159. From Archaic Chinese to Mandarin

1. Introduction
2. Affixation in Old Chinese
3. Prosodic alternation – Type A and B syllables
4. Syllabic prefixes and suffixes in Old Chinese
5. Reduplication
6. Compounding
7. New flectional suffixes
8. New derivational affixes
9. References

### 1. Introduction

The non-phonetic nature of the Chinese script places peculiar difficulties in the way of the historian of the language. The native tradition of rhyme dictionaries and rhyme tables, supplemented by comparison of modern dialects and other indirect evidence, makes it possible to reconstruct the phonological system of Late Middle Chinese (LMC), the *koine* of the Tang dynasty (A. D. 618–906), and Early Middle Chinese (EMC), the standard common to the educated classes of North and South during the period of division from the fall of Western Jin in A. D. 316 to the reunification by Sui in 589, with a fair degree of confidence. The first reconstruction of Early Middle Chinese on a modern scientific basis (under the name Ancient Chinese) was achieved by Bernhard Karlgren (1915–1926) and is still widely used. The newer reconstructions of Late and Early Middle Chinese cited in this article are those of Pulleyblank (1984; 1991a).

For Old Chinese, the language of the Chinese classics and other early texts, the evidence is much less satisfactory. The reconstruction (under the name Archaic Chinese) by Karlgren (1940; 1957) based on the rhyming categories of the earliest poetic anthology, the *Shijing* (*terminus ante quem* 600 B. C.) and the evidence of phonetic loans in the script, is commonly cited but has been much criticized (e.g. Li 1971; 1974–1975; Pulleyblank 1962; 1977–1978; 1991c). Tentative reconstructions cited here are preceded by an asterisk. References to Modern Chinese in this article are to the current standard called variously Mandarin, National Language (*Guoyu*) or Common Language (*Putonghua*) based on the vernacular of the capital, Beijing.

The traditional characterization of Chinese as monosyllabic is not far off the mark.

As one distinguished linguist put it, “The so-called ‘monosyllabic myth’ is in fact one of the truest myths in Chinese mythology” (Chao 1968: 139). Though much of the vocabulary of Modern Chinese consists of disyllabic or polysyllabic lexemes these are mostly compounds of monosyllables that have the status of morphemes. Exceptions are some foreign loanwords, words like *méigui* ‘rose’, and *chóuchú* ‘hesitate’ that are reduplicative in form though there are no simpler morphemes to which they can be referred, and a small residue of unanalyzable disyllabic or polysyllabic forms like *wúgong* ‘caterpillar’ and *hàshimǎ* ‘Chinese forest frog’ (cf. *hámá* ‘frog’). The traditional term “isolating” is also not inappropriate. Grammatical relationships are mostly conveyed by rules of word order and particles. In the modern language there are a few noun-forming suffixes, the plural or collective suffix *-men* for personal pronouns and animate nouns, and a handful of aspect suffixes for verbs, but even these affixes often do not have the obligatory character that such elements commonly have in inflected languages.

It was not always so. Isolated traces of ancient morphological processes survive to the present. Thus *cháng* ‘long’ and *zhǎng* ‘grow; elder; chief’ are written with the same graph and must be from the same root. *Zhāng* ‘stretch’, *zhàng* ‘swell’ and *zhàng* ‘curtain (i.e. something stretched)’, which are written with the graph for *cháng* ‘long’ with additional semantic determinatives, must also be related. In other cases obsolete variant readings of the same graph associated with differences in meaning are recorded in old dictionaries and commentaries, or sound similarities suggest etymological relationships between words even when there is no graphic connection.

Early attempts to analyze morphological processes in Old Chinese (Conrad 1896) could make little progress before phonological reconstruction had been put on a sound basis. Serious analysis of the morphology of grammatical particles also began with Karlgren (1920), who observed that the first person pronouns, *wú* and *wǒ*, had different distribution in texts of the classical period and suggested that *wú* had originally been used for the nominative and genitive and *wǒ* for the accusative. While the interpretation of the distinction as one of case on the Indo-Euro-

pean model is unsatisfactory (Kennedy 1956), the general conclusion that these and other pronouns and particles were subject to morphological alternations associated with changes in grammatical function is well established. Karlgren also assembled sets of words that appeared to be related in both sound and meaning which he called “word families” and pointed out a number of recurrent patterns (Karlgren 1933; 1949: 70–101; 1956). He concluded that Proto-Chinese must have had formal word classes and regular systems of inflection and word derivation but was unable to find clearcut associations between patterns of phonetic alternation in Old Chinese as he reconstructed it and specific semantic or grammatical functions and concluded that by the time of the earliest texts the derivational processes had been too much obscured by phonetic change to yield to analysis. The study of morphological processes can, however, suggest alternative solutions to problems of Old Chinese phonology. Comparisons with other Sino-Tibetan languages, especially Tibetan, which in its classical form preserves a rich (though imperfectly understood) system of affixation in verbal morphology as well as word formation, can also suggest hypotheses to explain patterns of alternation in Chinese. The proposals that are put forward below are based on a thoroughly revised reconstruction of Middle Chinese (Pulleyblank 1984; 1991a) and on new ideas about the reconstruction of Old Chinese that are still being tested and revised.

## 2. Affixation in Old Chinese

### 2.1. Departing Tone as derivational – the suffix \*-s

In Middle Chinese **syllables** were divided into four categories, labelled Level, Rising, Departing, and Entering, that correspond to the **tones** of modern spoken forms of Chinese. The Entering Tone in particular, which consisted of syllables ending in a stop consonant, was not “tonal” in a phonetic sense, and it is now thought that the Rising and Departing Tones had also developed between Old and Middle Chinese out of segmental features at the end of the syllable, a final glottal stop in the case of the Rising Tone and a suffix \*-s in the case of the Departing Tone (Haudricourt 1954 a; 1954 b; Pulleyblank 1962; 1973 b; 1978; Mei 1970). During the Late Middle Chinese period a further split occurred into

Upper and Lower registers conditioned by the devoicing of initial obstruents, giving a schematic system of eight tone classes which have been simplified or further elaborated in various ways in different dialects. The four tones of Modern Mandarin correspond roughly to the Middle Chinese categories as follows: (1) *mā* ‘mother’ Upper Level, (2) *má* ‘hemp’ Lower Level, (3) *mǎ* ‘horse’ Rising, (4) *mà* ‘curse’ Departing. Neutral tone on particles and the second morphemes of some compounds is unmarked.

Even in the modern language there are many examples of words originally in Level, Rising or Entering Tone with cognate forms in the Departing Tone. *Mó* ‘rub, grind’ and *mò* ‘mill’, written with the same character, show such an alternation between Level and Departing. Similar alternations between Rising and Departing and between Entering and Departing are *hǎo* ‘good’ and *hào* ‘love’, *è* ‘bad’ and *wù* ‘hate’. In the last example both words now have the Departing Tone but *è* ‘bad’, Cantonese *ɔ:k*, Sino-Japanese *aku*, ended in -k in Middle Chinese. Though the semantic relationships associated with this pattern are quite various, it seems clear that, in general, forms in Departing Tone must be derivatives of forms in the other categories. It has sometimes been thought that the conditioning factor was always a matter of tone (e.g. Wang 1958). In the case of the Entering Tone, however, this would imply that variation in pitch could lead to the loss of a final stop consonant, a kind of phonetic change for which it is difficult to find a parallel. Comparison with cognate Austroasiatic languages led to the discovery that the Vietnamese tone corresponding to the Middle Chinese Departing Tone originated in the loss of final \*-h, from an earlier \*-s, and to the suggestion that Chinese might have had a parallel history (Haudricourt 1954 a; 1954 b). A parallel was also drawn with the -s suffix that figures prominently in Tibetan verbal morphology (Forrest 1960) and there is much evidence that a final sibilant survived in certain classes of Departing Tone syllables throughout the Han period (206 B. C. – A. D. 220) and as late as ca. A. D. 500 in the southern form of Early Middle Chinese (Pulleyblank 1962; 1973 b; 1979). The semantic relationships corresponding to Departing Tone derivation are quite varied and difficult to classify (Downer 1959; compare also Chou 1962: 50–92). It is possible that more than one original suffix is involved. Apart from its

use as a marker of the **Perfect** in certain classes of verbs, the suffix *-s* in Tibetan appears in other forms where its meaning is hard to determine and it is nearly in complementary distribution with another suffix *-d*.

## 2.2. Voicing of initial obstruents – \**ã* as a prefix

There are many alternations involving a contrast in Early Middle Chinese between voiceless and voiced **initial obstruents** (Karlgren 1949: 90–92; Downer 1959: 263; Chou 1962: 50–92). Typical examples are *jiān* EMC *ken<sup>h</sup>* ‘see’, *xiān* EMC *yen<sup>h</sup>* ‘appear’, *zhū* EMC *teuawk* ‘attach’, *shū* EMC *dzuawk* ‘be attached, belong’. It has been proposed (Pulleyblank 1973 a: 114–116; 1989) that the voiced alternants in such pairs reflect a prefix *ã-*, a non-syllabic, pharyngeal glide that caused voicing, or in some dialects, voiced aspiration of a following obstruent, cognate to the Tibetan prefix, *ha-čhūn*, and to the prefix *?ã-* of Written Burmese. It typically had the effect of changing transitive verbs into intransitives or to adjectives describing a state. In Classical Tibetan the prefix *h-* is associated especially with the so-called “present” stem in verbs, which has an aspectual rather than temporal meaning. It expresses what is universally the case, is going on at the present time, or is treated as going on in a narrative referring to the past (Hahn 1974: 61). The *?ã-* of Burmese is used especially to derive nouns from verbs, e.g. *wak* ‘to halve’, *?awak* ‘half’. It seems to be the same morpheme *\*a*, with “introvert” meaning, that was involved in the *ə/a* ablaut (2.3) and it could, perhaps, also occur as a suffix (2.4).

## 2.3. Ablaut – *\*a* as an infix

As a corollary to the hypothesis that, like that of modern Mandarin (Hartman 1944; Hockett 1947; 1950; 1955: 88; Martin 1957; Chao 1968; Cheng 1973; Chen 1976; Pulleyblank 1981; 1984), the vowel system of Old Chinese was based on a two-way ‘vertical’ contrast in rhyme vowels, /ə/ and /a/, with features of frontness and roundedness encoded underlyingly as glides, /j/, /w/ and /ɥ/, it has been proposed that there was a morphological system of ablaut between the vowels /ə/ and /a/, associated with an extrovert/introvert contrast in meaning. This will account for alternations such as *tán* EMC *dəm* ‘talk about (trans.)’, *tān* EMC *dam* ‘talk (intrans.); conversation’; *sì* EMC *zi'*, < \*-ጀ? ‘resemble’, *xiàng* EMC *zian'*, < \*-ጀ? ‘imi-

tate, represent; image’; *qiū* EMC *kʰuw* < \*-xʷጀy ‘hill’, *qū* EMC *kʰiā* < \*-xʷጀy ‘large hill, mound, abandoned city’ (Pulleyblank 1963; 1965; 1973 a; 1989). In spite of the example of Mandarin, vertical vowel systems of this kind have been regarded as unnatural (Ting 1975: 32, citing Jakobson & Halle 1971) but the development of nonlinear phonology in the seventies has made such an analysis more acceptable (Anderson 1978). A further point that has become a commonplace of recent phonological theory is that shwa vowels are often best analyzed as epenthetic, inserted by rules of syllabification, rather than underlyingly present in the lexicon (cf. Anderson 1978; 1982). If Old Chinese is analyzed in this way, the *\*ə/a* ablaut can be interpreted as a process of affixation, i.e. as infixation of *\*a*, rather than alternation between two ends of a polar opposition and is easier to understand as a normal linguistic process.

Much work remains to be done on identifying and analyzing the role of the *\*ə/a* ablaut. It plays a role in function words as well as families of content words. For example, the prohibitory negative *wú* EMC *muā* ‘do not’ and *wú* EMC *muā* ‘not have, there is no ...’ were already homonyms by the fourth or third century B. C. but were originally written with different graphs, implying that the former had the vowel *\*ə* and the latter had the vowel *\*a*. The particles *zhū* EMC *teiā* < \**tāy* ‘the collectivity of, those of a certain class’, and *zhē* EMC *teia'*, < \**tāy?* (?), pronoun substitute for the head of a noun phrase, must be related by way of the *\*ə/a* ablaut to *zhī* EMC *tei* < \**tə(y)*, the particle that functions both as the mark of subordination before the head of a noun phrase and as a general third person object pronoun.

## 2.4. Rising Tone – *\*a* as suffix?

Oppositions within word families between the Middle Chinese Level and Rising Tones are less frequent than those involving the Departing Tone but are by no means rare. *Zhǎng* EMC *trian'* ‘grow; elder’, contrasting with *zhāng* EMC *triaŋ* ‘stretch’ is one such case. As noted above, there is good evidence that the Rising Tone reflects an earlier final glottal stop. Though glottal stop *?-* was also an initial consonant in Middle Chinese, there is reason to think that it was not a lexically distinctive phoneme in Old Chinese but rather an automatic accompaniment of vowel onset (Pulleyblank 1991c). In this case, the source of final glottalization also becomes

problematical. Parallelism suggests that it may have been an automatic feature after an open vowel, e.g. *\*-aŋ* (> Level Tone) vs. *\*-aya?* > *\*-aŋ?* (> Rising Tone). In this case the affix that changed ‘stretch’ into ‘grow’ would be a suffix *\*-a*, presumably the same morpheme that, as a prefix, was responsible for voicing the initial in *cháng* EMC *driŋ* ‘long’. In both cases the derived form had a less active, more introvert, meaning, though in the case of the prefix the result was an adjective, while in the case of the suffix the result was an intransitive verb. Another example of Rising Tone corresponding to a clearly defined difference in meaning is *zhéng* EMC *tɕiaŋj* ‘in good order; put in order; whole’, which is related both graphically and phonetically to *zhéng* EMC *tɕiaŋj* ‘to set right, regulate’, now found only in the sense of *zhéng yuè* ‘regulating moon, the month that starts the year’ and, with an additional semantic indicator in the graph, in *zhéng* ‘mount a (punitive) expedition against, attack; tax’, as well as in the more common derivatives in Departing Tone, *zhèng* EMC *tɕiaŋj<sup>h</sup>* ‘correct, right’, and *zhèng* EMC *tɕiaŋj<sup>h</sup>* ‘government’.

## 2.5. Prefix *\*s*

In Tibetan *s-* clearly has a **causative** function and this is also found in other Tibeto-Burman languages, for example, Burmese, where *\*s-* has given rise to aspiration of the following consonant (Benedict 1972: 105 ff.). Traces of the same prefix can, perhaps, be seen in Old Chinese in a case such as *sì* EMC *z<sup>i</sup>ŋ<sup>h</sup>* < *\*sləks* ‘feed; food, fodder’, from *shí* EMC *zik* < *\*lək* ‘eat’ (Pulleyblank 1973 a:117), but parallel examples are few and the phonology of Old Chinese *\*s-* clusters is still quite uncertain.

## 2.6. Initial aspiration

Though Middle Chinese had a three-way manner contrast for obstruents, the voiced stops were of secondary origin, partly derived from the effect of the prefix *\*ü-* and partly from the ‘hardening’ of voiced continuants in Type A syllables (*\*l-* > *d-*, *\*w-* > *yw-*), leaving only a two-way contrast between plain voiceless and voiceless aspirates in Old Chinese. Classical Tibetan also has a three-way contrast in obstruents between plain voiceless, voiceless aspirates and voiced, but the two voiceless series are in complementary distribution after prefixes and only the aspirates occur freely without a prefix as syllable onsets. Plain voiceless stops in this position are

mostly confined to special categories of words, including reduplicative forms, foreign loanwords and, in some cases, doublets of forms with prefixes such as *kog-pa* ~ *skog-pa* ‘shell, rind, bark’ or *pags-pa* ~ *lpags* (in compounds) ‘skin, hide’ (Li 1933; Benedict 1972: 20). Benedict concluded that only the distinction of voice was pertinent and aspiration of the voiceless initials was allophonic in Tibeto-Burman but the comparative evidence will equally well support the hypothesis that the original contrast was between plain voiceless and voiceless aspirates and that voicing of the non-aspirates in Classical Tibetan was secondary. In that case, alternations that we find in Tibetan verbal morphology between voiced stops in Present and Future and the corresponding plain voiceless or voiceless aspirate in Perfect and Imperative, as in *hgebs-pa*, *bkab*, *dgab*, *khob* ‘cover’ imply that an original plain voiceless stop became aspirated as a result of some morphological process. The source of this aspiration cannot be the prefix *s-*, which survives in Tibetan, but there could have been another prefix, such as *s*, *x*, or *h*, that did not survive in Classical Tibetan (Shafer 1950: 721 f.). An alternative suggestion, also by Shafer, is that aspiration could have resulted from the coming together of two identical plain unaspirated stops: *p-p-* > *p<sup>h</sup>-*, *k-k-* > *k<sup>h</sup>-*, etc., a process that is alleged to occur in the morphophonemics of Tsangla, that is, from a kind of initial reduplication. In Chinese aspiration also plays a role in word families. In a few cases, such as *pàn* EMC *p<sup>h</sup>an<sup>h</sup>* ‘divide’ vs. *bàn* EMC *pan<sup>h</sup>* ‘half’ or *kuò* EMC *k<sup>h</sup>wak* ‘expand, stretch’, vs. *guǎng* EMC *kwaŋ* ‘wide’ (the variation in final consonants is also to be noted) it appears to be associated with causative formations and, in the absence of clear evidence as to the phonology of the prefix *\*s* in Chinese, one might conjecture that this was the source of aspiration. In other cases the semantic relationship between the unaspirated and the aspirated form is much less clear, as in *qī* EMC *k<sup>h</sup>iā* ‘one-footed’, cf. *jī* EMC *kiā* ‘odd number’, *qǐ* EMC *giā* ‘strange’; *cǐ* EMC *ts<sup>h</sup>iā* ‘this’, cf. *zī* EMC *tsi* ‘this’; *chē* EMC *te<sup>h</sup>ia* ‘carriage’, also *jū* EMC *kiā* ‘carriage’ (variant pronunciations of the same character). Comparison with Tibetan suggests that aspiration probably arose independently of prefixed *\*s-* in Chinese also.

## 2.7. Prefix (or infix?) *\*r*

The Tibetan prefix *r-* with parallels in Kachin, Bodo-Garo and Mikir (Benedict 1972: 109 f.) may have an Old Chinese paral-

lel in cases like *zhì* EMC *tri<sup>h</sup>* ‘cause to arrive, bring’, derived from *zhì* EMC *tsei<sup>h</sup>* ‘arrive at, reach’, and *chù* EMC *tr<sup>h</sup>wit* ‘expel’, derived from *chū* EMC *te<sup>h</sup>wit* ‘go out’ (Pulleyblank 1973 a:118). Though the Middle Chinese form might seem to imply a \*Cr- cluster, in which the derivation would have to be interpreted as infixation, a prefixed \*r- could have given the same result.

#### 2.8. Presyllables \*kə-, \*tə-, \*pə-, \*mə-

Assuming that Tibetan orthography was based on the conventions of the Indian alphabet on which it was modelled, the Tibetan prefixes *g*, *d*, *b*, *m*, which are written on the line before the root initial rather than on top of it, were probably originally followed by a shwa vowel and formed **enclitic presyllables**. Isolated parallels to this are found in modern Chinese dialects, e.g. colloquial Pekingese *gēbei* or *gēbo* ‘arm’, which must be related to Classical *bì* EMC *pjiā<sup>h</sup>* ‘arm’ (Yang 1977–1978: 292–294). Another trace of this may be the so-called “softened” initials in North-western Min dialects. Thus, Kienyang sometimes has zero or *v*- instead of *p*-, corresponding to Middle Chinese *p*- and *b*-, *l*- corresponding to Middle Chinese *t*-, *d*-, *ts*-, *dz*-, and zero or *h*- corresponding to Middle Chinese *k*- and *g*- (or *y*-). The common Min word for ‘cockroach’ has a presyllable in Fuzhou *ka<sup>6</sup> sak<sup>8</sup>*, Amoy *ka<sup>1</sup> tsua<sup>7</sup>*, but a “softened” initial in Kienyang *hue<sup>8</sup>*, Chungan *huai<sup>8</sup>*, suggesting that in at least some cases the “softening” of the initials may have resulted from weakening in intervocalic position, with subsequent loss of the presyllable (Norman 1986: 383 f.). A similar “softening” of initial obstruents in Vietnamese corresponds to presyllables in other Mon-Khmer languages, e.g. *gà* ‘chicken’ Ruc *rəka*; *vái* ‘cotton’ Ruc *kupal* (Haudricourt 1965; Thompson 1976: 1131–1133) and such “softened” initials also occur in early Vietnamese loans from Chinese (Pulleyblank 1981: 279). In particular, Vietnamese has *vai* ‘shoulder’ corresponding to Chinese *bì* EMC *pjiā<sup>h</sup>*, colloquial *gēbo*, *gēbei* ‘arm’. In a number of cases the Min softened stops correspond to voiced stops in Yao loanwords which, in turn, correspond to prenasalized stops in Miao, which has led to the suggestion that prenasalization was also the general source of the phenomenon in Chinese (Norman 1986: 381–383). Prenasalization in Miao-Yao could, however, simply be the effect of a preceding unaccented shwa vowel derived from a lost presyllable. The *ha-čhün*

prefix in Tibetan, written in Tibetan script as a presyllable, had the effect of prenasalizing obstruents in that language.

#### 2.9. Dental suffixes \*-n, \*-t, \*-l

Besides \*-s, Tibeto-Burman had the dental suffixes \*-n and \*-t whose precise functions are difficult to analyze (Benedict 1972: 98–103). There is also internal Chinese evidence for \*-n and \*-t as suffixes. In a number of sets of particles in the preclassical and early classical language \*-n seems to be associated with **durative** or **continuative** aspect, while \*-t is associated with **punctual** or **perfective** aspect (Graham 1983) and derivatives in \*-n or \*-t can be recognized among ordinary content words as well (Pulleyblank 1991b: 29–33). It seems likely that \*-l could also function as a suffix in word formation.

#### 2.10. Miscellaneous affixation

A **distributive** suffix \*-k appears in a number of pronominal forms: *huò* EMC *ywək* < \*wók ‘(there is) some one, something’ from *yǒu* EMC *wu'* < \*wó? ‘have; there is/are’; *mò* EMC *mak* < \*mák ‘(there is) no one, nothing’ from *wú* EMC *muă* < \*máy ‘not have, there is/are not’; *shú* EMC *dzuwk* ‘which’ from *shuí* EMC *dzwi* ‘who’; *gè* EMC *kak* ‘each’, cf. *jǐ* EMC *kiǎ* < \*kày? ‘all’, *jiē* EMC *kəj* ‘all’, *jǐ* EMC *kuă* ‘both’, etc. These particles are placed between the subject and the verb, or in front of the verb without an expressed subject, and have the effect of selecting among possible subjects: ‘some of X’, ‘none of X’, ‘which of X?’, ‘each of X’. The same suffix can be recognized in the early classical demonstrative pronoun *shí* EMC *dzik* < \*-ək ‘this’, with a dialectal variant showing palatalization of the final consonant *shí* EMC *zit* (for *dzit*), which is related to the common demonstrative *shi* EMC *dziă* ‘this’ and has the function of recapitulating a preceding subject and giving it contrastive emphasis (Pulleyblank 1960).

A type of alternation of which there are several examples is between Old Chinese \*-ŋ and either the corresponding voiced continuants \*-ɣ or zero: *děng* EMC *təŋ?* < \*təŋ? ‘step, degree’, also read EMC *təj?* < \*təɣ?, *wáng* EMC *muaj* < \*máy ‘there is no; disappear, die; lose’, *wú* EMC *muă* < \*má(y) ‘there is no’; *wǎng* EMC *wuaj?* < \*wáŋ? ‘go’, *yú* EMC *wuă* < \*wá(y) ‘go to; to’; *fāng* EMC *puaj* < \*páŋ ‘just now’, *fǔ* EMC *puă?* < pà(y)? ‘begin, for the first time’ (Pulleyblank 1962: 232 f.).

### 3. Prosodic alternation – Type A and B syllables

A phonological alternation that was probably not the result of affixation is the contrast between what have been called Type A and Type B syllables. In Karlgren's reconstructions this phonological contrast is represented by the absence (Type A) or presence (Type B) of a medial *yod*, *-j-*. Though the contrast between syllable types that it marks is real and important, the interpretation in terms of a medial segmental element is a basic error in his system (Pulleyblank 1984; 1992 a). In EMC Type A syllables had a mid or low nuclear vowel, while Type B syllables had a high vowel, *i*, *i*, or *u*, either as the sole element in the nucleus or as the first mora of a diphthong, *ia*, *ia*, or *ua*. Since one can in general derive corresponding A and B pairs from every Old Chinese rhyme category, it seems likely that the Middle Chinese contrast reflects an earlier **prosodic contrast** in Old Chinese which put prominence either on the second mora of the syllable (Type A) or the first (Type B). There is a parallel in Sizang (Siyin) Chin, a Tibeto-Burman language of Burma, which distinguishes within sequences ending in -VV, -VC, and -VVC between a syllabic peaking that falls upon either the first or the second mora, e.g.: *pai* 'to lose' vs. *pai* 'to go'; *bel* 'to take shelter under' vs. *bel* '(superlative particle)'; *tam* 'jungle' vs. *tam* 'to number, many' (Stern 1963; Pulleyblank 1992 b). The process of diphthongization affecting Type B syllables that gave rise to the Early Middle Chinese contrast seems to have occurred around the beginning of the present era and there is comparative evidence that it spread as an areal feature to Vietnamese and other languages that were within the Chinese sphere of influence at that time (Pulleyblank 1992 b). Type A and B syllables are marked by an acute or a grave accent respectively in Old Chinese reconstructions, e.g. \*-án > EMC -an (Type A), \*-àn > EMC -ian or -ian, depending on the initial (Type B).

There are alternations between the two syllable types both in grammatical particles: *ān* EMC ?an < \*?án 'where? how?' (Type A), *yān* EMC ?ian < \*?án 'where? how?' (Type B) (with no apparent difference in meaning); *nuò* EMC nak < \*nák 'agree, say yes', *ruò* EMC niak < \*nák 'thus; conform to, accord with' – and content words: *biān* EMC pen < \*pján and EMC pjian < \*pján 'plait; compile'. Compare also *māo* EMC maiw <

\*mráw 'cat', also read EMC miaw < \*mràw with corresponding forms in some modern dialects, e.g. Changsha miau, as well as *mau*. There is a more marked semantic difference between Type A and Type B in *nāi* EMC nɔj' < \*nóy? 'then, thereupon' and *ér* EMC ni < \*nóy 'then, and'. The stronger form, *nāi*, is a connective between independent clauses or sentences and can also serve to emphasize a following noun predication, while *ér* is a colourless connective between serial verbs.

### 4. Syllabic prefixes and suffixes in Old Chinese

Besides processes that changed the forms of monosyllables, there are certain constructions in the preclassical language that have been described as 'prefixation' that involve syllables represented by separate graphs. *Yōu*, which ordinarily means 'have; there is ...' sometimes appears in front of nouns, especially proper names, where it seems to be part of the noun phrase though what, if anything, it adds to the meaning is not clear (Wang 1958: 219 f.). There are also certain prefixed syllables that mark distinctions of aspect in a following verb (Chou 1962: 248 ff.; Graham 1983: 56; Pulleyblank 1986: 3). They may be compared to the aspect suffixes of Mandarin (see below). A formative element that was used in Classical Chinese and is also found in the modern language is -rán '-like, -ly', a derivative in -n of *rú* 'is like', used as a suffix forming adverbs of manner, e.g. *huírán* 'kindly' from *huí* 'to favour; kindness'. Unlike purely colloquial suffixes, it retains its full tone in the modern language, a mark of its literary origin.

### 5. Reduplication

A process that was already present in the ancient language and has continued to be active until the present is complete or partial **reduplication** (cf. Art. 57). Examples of reduplication used in onomatopoeia, expressive adjectives and adverbs, and names of plants and small animals are *guān'guān*, 'cry of a bird', *xiāoyá* 'free and unfettered, carefree', *zhīzhū* 'spider'. Of more interest from the point of view of word formation are expressions such as: *xūyú* EMC suā juā 'a moment', from *xū* EMC suā 'wait', and *cōngróng* EMC tsʰuawŋ 'juawŋ (note the aspiration) 'at leisure, unhurriedly', from the root found in *zòng* EMC

*tsuawŋ<sup>h</sup>* ‘release; loose’ and *cóng* EMC *dzuawŋ* ‘follow’; *pífú* EMC *bəbək* < \**bábək* ‘crawl’. In a few cases the analysis of such disyllabic forms as reduplicative breaks down, e.g. *húdié* EMC *yɔ dep* ‘butterfly’ (Kennedy 1955), but the second syllable does occur separately and is probably cognate to Tibetan *leb* in *phye-ma-leb* ‘butterfly’ (where *phye-ma* means ‘flour’ and refers to the flour-like scales on a butterfly’s wing), and written Burmese *lip-pra* ‘butterfly’.

Certain productive categories of full reduplication that are normal in the modern language – familiar terms of address like *gēge* ‘elder brother’, distributives like *rénrén* ‘person:RDP (everybody)’, *tiāntian* ‘every day’, and repetition of a verb to express tentative aspect as in *kānkān* ‘have a look’ – are not found in the classical language.

## 6. Compounding

Compounding, by which word sequences arising out of regular syntactic processes of the language are lexicalized with special meanings (Lyons 1977: 535), which is so prominent in modern Chinese, was already in progress in the classical language. There are many disyllabic expressions whose meaning cannot be inferred transparently from that of the separate elements, e.g.: *báixìng* ‘common people’ (literally: ‘hundred surnames’), *guárén* ‘I (term of self-reference for the ruler of a state)’ (literally: ‘solitary man’), *jūnzǐ* ‘gentleman’ (literally: ‘son of a lord’), *shùjí* ‘almost, nearly, it is to be hoped that’ (literally: ‘many-few’). Unfortunately there is no way to tell whether there were any phonological correlates. In modern Mandarin the second member of a compound is frequently toneless, which binds the two syllables together into a single tone-bearing unit. Since there is no clue to it in the writing system, it is difficult to know when this appeared historically. It is not found in more conservative dialects such as Cantonese.

## 7. New flectional suffixes

### 7.1. The plural suffix *-men*

In modern spoken Chinese *-men* is used to mark the **plural** of personal pronouns: *wǒmen* ‘we’, *zánmen* ‘we (inclusive)’, *nǐmen* ‘you’, *tāmen* ‘they’ (only when referring to persons, otherwise simply *tā*). In this sense it clearly fits the definition of an inflectional suffix. Its use with nouns is not strictly comparable. It

has been called a mark of the **indefinite plural** (Kupfer 1980: 56), but according to another interpretation, it forms **collective nouns** (cf. Art. 101), which would make it a derivational suffix (Chao 1968: 244 f.). With nouns it cannot be used when a definite number is specified: *liángge qiángdao* ‘two bandits’, not \**liǎngge qiángdaomen*, while with pronouns it can, though it is also possible to omit it: *wǒmen liǎngge* or *wǒ liǎngge*, *wǒ liǎ*, ‘we two’. The etymology of *-men* is obscure. In the literary language certain words like *chái*, *shǔ*, *cáo*, *děng*, *bèi* meaning ‘class, category, group’ or the like are used from time to time with both nouns and pronouns to indicate a collectivity but there is nothing approaching a regular grammatical process of plural marking. The suffix *-men*, written in a variety of ways that seem to indicate that it was an innovation in the vernacular without a recognized antecedent, begins to appear in colloquial texts from about the tenth and eleventh centuries (Lü 1955; Wang 1958: 272 f.). From an early period it appears to have had the allomorphs *-me* and *-m*, still found in present-day usage. Out of the latter have developed polite forms of the second and third person pronouns used without reference to number, *nǐn* < *nǐm* and *tān* < *tām*, by the regular shift of syllable final *-m* to *-n* that occurred around the end of the fifteenth century (Lü 1955: 213). In epistolary style, though not in speech, one can even create new plural forms, *nǐnmen* and *tānmen* (Chao 1968: 244 f., 640 f.).

### 7.2. The verbal suffixes *-le*, *-zhe*, *-guo*

The **perfective** suffix *-le* and its homonym, the sentence final particle *le*, are derived from *liǎo* ‘end, finish’, which is still current as a monosyllabic verb in certain expressions and is also found in such combinations as *liǎojíé* ‘finish, settle’, *liǎojú* ‘end, solution, settlement’, *zhōngliǎo* ‘end (of a period)’, etc. The full form *liǎo* is also found as what Chao calls a ‘phase complement’ to other verbs, that is, a complement that expresses the phase of an action of the preceding verb, rather than some result in the action or goal like an ordinary resultative complement (Chao 1968: 449). This may have been the earliest development in its progress towards **grammaticalization**. Its use as a perfective suffix seems to have become established in the vernacular by the tenth and eleventh centuries, by which time it had also come into use as a sentence final particle, taking over the functions of the classical final particle *yǐ* (Wang 1958: 304–307, 447 f.). Unfortunately

textual evidence does not enable us to trace the steps by which the full form *liǎo* became reduced phonetically to *le* in the suffix and sentence final particle. The suffix *-zhe* for **progressive** aspect is similarly related both graphically and etymologically to *zháo* EMC *driak* ‘come in contact with, touch’, which also occurs both in the form *zháole*, with the perfective suffix *-le*, as an ordinary resultative complement ‘hit the mark, touched the essential point’, and, with optional neutral tone, as a phase complement, translated by Chao as ‘touched, got at, successful after an attempt’ (Chao 1968: 447). As an aspect suffix it immediately follows the verb and cannot be separated from it by an object (Chao 1968: 248–250). The suffix *zhe* was a little later in appearing than *le* but there is evidence for it already in the eleventh century (Wang 1958: 309–311). The verb *guò* ‘pass’ can also be used, with optional neutral tone, as a complement, in which case it can be followed by the perfective suffix *-le*, as in *chīguòlē fàn* ‘have had one’s meal’, but it can also be used as a suffix in the neutral tone to mean ‘completed action at an indefinite time in the past’, as in *méi chīguó yáng* ‘has never been abroad’ (Chao 1968: 252f.). This usage can also be found by the eleventh century (Wang 1958: 312). Two compound suffixes, *-qilai*, from *qilai* ‘rise, get up’, and *-xiaqu*, from *xiàqu* ‘go down, descend’, which one scholar would prefer to relegate to the domain of word formation or syntax (Kupfer 1980: 57), are used as suffixes for **inchoative** and **successive** aspect respectively (Chao 1968: 251f.).

## 8. New derivational affixes

### 8.1. Noun prefixes

*A-*, as in *āgē* ‘elder brother’, is found widely in modern Chinese dialects (though not in Peking Mandarin) as a prefix with proper names and words of relationship, conveying a feeling of endearment or familiarity. Though it is not found in texts of the classical period, it occurs quite early in postclassical times (Wang 1958: 222) and is probably an inheritance outside the literary tradition from Sino-Tibetan. Benedict (1972) argued that, as a prefix of familiarity in Tibeto-Burman, *ʔa-* was a variant of the *ha-čhun* prefix of Tibetan that has been discussed in 2.2 and this seems very likely. The difference would be that, in the one case, \**a-* acquired a glottal stop onset and remained a distinct syllable, while in the other, it was a non-syllabic prefix, that in

Chinese voiced a following obstruent initial and then disappeared but in Tibetan was still pronounced as an onglide at the time when the script was invented and afterwards realized as prenasalization (Pulleyblank 1986). Although it does not appear as a separate syllable in the classical language, the prefix of familiarity has left its trace in the Middle Chinese voiced initial of standard Chinese *fù* “father” EMC *bua’* < \**pà?*. The unprefixed form with voiceless initial is *fǔ* EMC *pua’* < \**pà?*, written with the same character, which was used as an honorific form of address outside the family context in Old Chinese (Benedict 1972: 121ff.; Pulleyblank 2000).

In the present standard language the function of *a-* as a prefix of familiarity has been taken over by *lǎo* ‘old’, which is found in this usage as early as the Tang period (7th to 9th centuries), e.g. *lǎo xiōng* ‘elder brother’ in a collection of conversations by a Chan master or *lǎo Yuán* ‘old Yuan’ in a poem addressed by Bai Juyi to his friend Yuan Zhen. The prefix *lǎo* has become attached to some names of animals with no implication of age, e.g. *lǎoshǔ* ‘rat’, *lǎohǔ* or *lǎohu* ‘tiger’. This usage, too, begins from Tang times (Wang 1958: 223–225).

### 8.2. Noun-forming suffixes, *-zi*, *-(ə)r*, *-tou*, *-zhē*, *-de*

*Zi* ‘child’ already occurs in some combinations in Classical Chinese that superficially appear to anticipate its later use as a noun-forming suffix but they can all be derived from semantic extensions that are also found for the independent morpheme. Compare the modern compound *shízi* ‘cobblestone’ made up of *shí* ‘stone’ plus the same morpheme *zǐ* in the sense of ‘small, hard object’. It is a reasonable conjecture that its use as a suffix at first had a **diminutive** force but by the sixth century, when we first have clearly documented examples, it had become fully established in its modern usage, with no discernible diminutive force. We find such examples as *núzǐ* ‘slave’, *rızǐ* ‘day’, *gézǐ* ‘clam’ and examples multiply in the following centuries (Wang 1958: 226f.). In modern usage *-zi* as a suffix is distinguished from *zǐ* as the second member of a compound by its neutral tone but it is not known when this distinction appeared diachronically. In the modern language the suffix *-zi* is no longer productive and nouns which require it have become fully lexicalized.

The similar use of *ér* ‘child’ as a noun suffix is later in appearing. By the eleventh cen-

tury, however, there are unambiguous examples in which *ér* is attached to nouns referring to inanimate objects, apparently with a diminutive sense. Unlike *zǐ*, which lost any diminutive force that it may have once had at an early date, *ér* has retained this connotation, combined with an implication of endearment or familiarity (Wang 1958: 227–229; Chao 1968: 230). It is still a productive suffix in the modern language. Except for some speakers of Manchu background who still pronounce it as a separate, toneless syllable *-er* (Chao 1968: 140), the phonological reduction has gone farther for this suffix than for other flexional and derivational suffixes of modern Mandarin, so that it has become simply the consonant *-r* which is attached to the preceding syllable with various morphophonemic changes. Etymologically, suffixal *-r* also has two other sources in Mandarin, a reduced form of *rì* ‘day’ in *jīnr* ‘today’, *zuó* ‘yesterday’, etc., and a reduced form of *lì* ‘inside’ in the place words *zhèr* ‘there’, *nàr* ‘there’ and *nār* ‘where’ (Chao 1968: 229).

A noun-forming suffix of less frequent occurrence, for which there are examples from the fifth century onwards, is *-tou*, derived from *tóu* ‘head’, as in *shítou* ‘stone’, *gǔtou* ‘bone’ (Wang 1958: 230 f.). The suffix *-ba*, apparently derived from the verb *bā* ‘cling’, is found in a few nouns such as *wéiba* ‘tail’, *níba* ‘mud, clay’ (Chao 1968: 244).

The classical particle *zhě* which functions as a pronominal substitute for the head of a noun phrase has been borrowed into the modern spoken language as suffix for **agent nouns**, as in (*xīnwén*) *jízhě* ‘(newspaper) reporter’, *láodòngzhě* ‘worker’. It shows its literary origin, however, in that it retains its tone and cannot be combined with purely non-literary verbs, such as *chī* ‘eat’ or *páo* ‘run’. The true modern equivalent of *zhě* is *de*, which, like *zhě* in the classical language, is not only a suffix for agent nouns, as in *zhǎngsháorde* ‘chef’, but, more generally, a **nominalizer** of a noun-modifier phrase with the head omitted or without a specified head, e.g. *Tā shi cóng Zhōngguó lái de* ‘He is one who comes from China’ instead of *Tā shi cóng Zhōngguó lái de rén* ‘He is a man who comes from China’. In the modern language the particle of **subordination** in *de rén* is homonymous with the nominalizing particle *de* that is used when *réni* is omitted. As a particle of subordination *de* is a variant of the classical particle *zhī* that has developed separately in the colloquial language. Nominalizing *de* can similarly be regarded as the colloquial

development of *zhě*. For the original morphological connection between *zhī* and *zhě* see 2.3.

### 8.3. Other derivational suffixes

*Shi*, a reduced form of the **copula** *shì* ‘is’, is a suffix in a number of particles: *yaòshí* ‘if’, *ruòshí* ‘if’, *dōushí* ‘all, in all cases’, *kěshí* ‘but’, etc. (Chao 1968: 722 f.). Chao notes that in the case of *ruòshí* the two components can still be separated in *ruòbushi* ‘if (is) not’, showing that the meaning of the suffix as a copula is not completely attenuated.

**Prepositions**, or “coverbs”, in Chinese are a special class of transitive verbs that form phrases that modify a main verb rather than standing on their own. They normally do not take aspect particles since they take their aspect from the main verb. In some cases, however, we do find them with the aspect suffixes *-zhe* or *-le* with no apparent change in meaning or with some specialized meaning that cannot be directly related to the usual force of the particle. Thus, *wèi* ‘for; because of; for the sake of’ also has the forms *wèizhe* and *wèile*. Authorities differ as to the precise shades of meaning that these forms convey.

## 9. References

- Anderson, Stephen R. (1978), “Syllables, Segments and the Northwest Caucasian Languages”. In: Bell, Alan & Hooper, Joan Bybee (eds.), *Syllables and Segments*. Amsterdam, New York: North Holland Publishing Company, 47–58
- Anderson, Stephen R. (1982), “The Analysis of French Shwa”. *Language* 58, 534–573
- Benedict, Paul K. (1972), *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge University Press
- Chao, Yuen Ren (1968), *A Grammar of Spoken Chinese*. Berkeley, Los Angeles: University of California Press
- Chen, Matthew Y. (1976), “From Middle Chinese to Modern Peking”. *Journal of Chinese Linguistics* 4, 113–277
- Cheng, Chin-chuan (1973), *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton
- Chou, Fa-kao (1962), *Zhongguo gudai yufa, Gouci bian* (A Historical Grammar of Ancient Chinese, Part II: Morphology). Taipei: Academia Sinica
- Conrad, August (1896), *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten*. Leipzig: Harrassowitz
- Downer, Gordon, B. (1959), “Derivation by Tone Change in Classical Chinese”. *Bulletin of the School of Oriental and African Studies* 22, 258–290

- Forrest, R. A. D. (1960), "Les occlusives finales en Chinois archaïque". *Bulletin de la Société de Linguistique de Paris* 55, 228–239
- Graham, A. C. (1983), "Yün and Yüeh as Verbs and as Particles". *Acta Orientalia Hafniensia* 44, 33–71
- Hahn, Michael (1974), *Lehrbuch der Klassischen Tibetischen Schriftsprache*. Bonn: Indologisches Seminar
- Hartman, Lawton M. (1944), "The Segmental Phonemes of the Peiping Dialect". *Language* 20, 28–42
- Haudricourt, André G. (1954 a), "Comment reconstruire le chinois archaïque". *Word* 10, 351–364
- Haudricourt, André G. (1954 b), "De l'origine des tons en Viêtnamien". *Journal Asiatique* 242, 69–82
- Haudricourt, André G. (1965), "Les mutations consonantiques des occlusives initiales". *Bulletin de la Société de Linguistique de Paris* 60, 160–172
- Hockett, Charles F. (1947), "Peiping Phonology". *Journal of the American Oriental Society* 67, 253–267
- Hockett, Charles F. (1950), "Peiping Morphophonemics". *Language* 26, 63–85
- Hockett, Charles F. (1955), *A Manual of Phonology*. Baltimore: Waverly Press (Memoir 11 of the *International Journal of American Linguistics*)
- Jakobson, Roman & Halle, Morris (1971), *Fundamentals of Language*. The Hague: Mouton
- Karlgren, Bernhard (1920), "Le Proto-Chinois langue flexionnelle". *Journal Asiatique* 15, 205–233
- Karlgren, Bernhard (1933), "Word Families in Chinese". *Bulletin of the Museum of Far Eastern Antiquities* 5, 1–120
- Karlgren, Bernhard (1940), *Grammata Serica*. Taipei: Ch'eng-Wen [reprinted from: *Bulletin of the Museum of Far Eastern Antiquities* 12, 1–471]
- Karlgren, Bernhard (1949), *The Chinese Language*. New York: Roland Press
- Karlgren, Bernhard (1956), "Cognate Words in the Chinese Phonetic Series". *Bulletin of the Museum of Far Eastern Antiquities* 28, 1–18
- Karlgren, Bernhard (1957), *Grammata Serica Recensa*. Stockholm: Museum of Far Eastern Antiquities [reprinted from: *Bulletin of Far Eastern Antiquities* 29, 1–332]
- Kennedy, George A. (1955), "The Butterfly Case". *Wennti* 8 [reprinted in: Li, Tien-yi (1964, ed.), *Selected Works of George A. Kennedy*. New Haven/CT: Far Eastern Publications Yale University, 274–322]
- Kennedy, George A. (1956), "Zai lun wu wo" (The Classical Pronoun Forms *ngo* and *nga*). *Bulletin of the Institute of History and Philology, Academia Sinica* 28, 273–281 [reprinted in: Li, Tien-yi (1964, ed.), *Selected Works of George A. Kennedy*. New Haven/CT: Far Eastern Publications Yale University, 434–442]
- Kupfer, Peter (1980), "Morphemklassen und Wortstrukturen im modernen Chinesischen". *Sprachwissenschaft* 5, 53–72
- Li, Fang-kuei (1933), "Certain Phonetic Influences of the Tibetan Prefixes upon the Root Initials". *Bulletin of the Institute of History and Philology, Academia Sinica* 4, 135–157
- Li, Fang-kuei (1971), "Shang-ku yin yen-chiu". *Tsinghua Hsuehpao* n.s. 9, 1–61
- Li, Fang-kuei (1974–1975), "Studies on Archaic Chinese". *Monumenta Serica* 31, 219–287
- Lü, Shuxiang (1955), "Shuo men" (A Study of the Plurality Suffix *mén*, with English Summary). In: Lü, Shuxiang, *Hanyu yufa lunwenji*. Beijing: Kexue chubanshe, 145–168, 212–213
- Lyons, John (1977), *Semantics*, Vol. 1–2. Cambridge: Cambridge University Press
- Martin, Samuel E. (1957), "Problems of Hierarchy and Indeterminacy in Mandarin Phonology". *Bulletin of the Institute of History and Philology, Academia Sinica* 29, 219–229
- Mei, Tsu-lin (1970), "Tones and Prosody in Middle Chinese and the Origin of the Rising Tone". *Harvard Journal of Asiatic Studies* 30, 86–110
- Norman, Jerry (1986), "The Origin of the Proto-Min Softened Stops". In: McCoy, John & Light, Timothy (eds.), *Contributions to Sino-Tibetan Studies*. Leiden: E. J. Brill, 375–384
- Pulleyblank, Edwin G. (1960), "Studies in Early Chinese Grammar, Part I". *Asia Major* 8, 36–67
- Pulleyblank, Edwin G. (1962), "The Consonantal System of Old Chinese". *Asia Major* 9, 58–144, 206–265
- Pulleyblank, Edwin G. (1963), "An Interpretation of the Vowel Systems of Old Chinese and Written Burmese". *Asia Major* 10, 200–221
- Pulleyblank, Edwin G. (1965), "Close/Open Ablaut in Sino-Tibetan". In: Milner, G. B. & Henderson, E. J. A. (eds.), *Indo-Pacific Linguistic Studies* [= *Lingua* 14], 230–240
- Pulleyblank, Edwin G. (1973 a), "Some New Hypotheses Concerning Word Families in Chinese". *Journal of Chinese Linguistics* 1, 111–125
- Pulleyblank, Edwin G. (1973 b), "Some Further Evidence Regarding Old Chinese -s and Its Time of Disappearance". *Bulletin of the School of Oriental and African Studies* 36, 368–373

- Pulleyblank, Edwin G. (1977–1978), “The Final Consonants of Old Chinese”. *Monumenta Serica* 33, 180–206
- Pulleyblank, Edwin G. (1978), “The Nature of the Middle Chinese Tones and Their Development to Early Mandarin”. *Journal of Chinese Linguistics* 6, 173–203
- Pulleyblank, Edwin G. (1979), “Some Examples of Colloquial Pronunciation from the Southern Liang Dynasty”. In: Bauer, Wolfgang (ed.), *Studia Sino-Mongolica: Festschrift für Herbert Franke*. Wiesbaden: Franz Steiner Verlag, 315–328
- Pulleyblank, Edwin G. (1981), “Some Notes on Chinese Historical Phonology”. *Bulletin de l'Ecole Française d'Extrême-Orient* 69, 277–288
- Pulleyblank, Edwin G. (1984), *Middle Chinese: A Study in Historical Phonology*. Vancouver: UBC Press
- Pulleyblank, Edwin G. (1986), “The Locative Particles *yü*, *yü*, and *hu*”. *Journal of the American Oriental Society* 106, 1–12
- Pulleyblank, Edwin G. (1989), “Ablaut and Initial Voicing in Old Chinese Morphology: \*a as an Infix and Prefix”. *Proceedings of the Second International Conference on Sinology: Section on Linguistics and Paleography*. Taipei: Academia Sinica, 1–21
- Pulleyblank, Edwin G. (1991a), *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*. Vancouver: UBC Press
- Pulleyblank, Edwin G. (1991b), “Some Notes on Morphology and Syntax in Classical Chinese”. In: Rosemont, Henry Jr. (ed.), *Chinese Texts and Philosophical Contexts: Essays Dedicated to Angus C. Graham*. La Salle, Illinois: Open Court, 21–45
- Pulleyblank, Edwin G. (1991c), “The *ganzhi* as Phonograms and Their Application to the Calendar”. *Early China* 16, 39–80
- Pulleyblank, Edwin G. (1992 a), “European Studies on Chinese Phonology: the First Phase”. Paper for the International Symposium on the History of European Sinology, Taipei, April 1992
- Pulleyblank, Edwin G. (1992 b), “Type A and B Syllables in Middle Chinese and Their Origin in Old Chinese”. Paper for the 25th International Conference on Sino-Tibetan Languages and Linguistics, Berkeley, 14–18 October 1992
- Pulleyblank, Edwin G. (2000), “Morphology in Old Chinese”. *Journal of Chinese Linguistics* 28, 26–51
- Shafer, Robert (1950–1951), “Studies in the Morphology of Bodic Verbs”. *Bulletin of the School of Oriental and African Studies* 13, 702–724, 1017–1031
- Stern, Theodore (1963), “A Provisional Sketch of Sizang (Siyin) Chin”. *Asia Major* 10, 222–278
- Thompson, Laurence C. (1976), “Proto-Viet-Muong Phonology”. In: Jenner, Philip N. & Thompson, Laurence C. & Starosta, Stanley (eds.), *Austroasiatic Studies*. Honolulu: University Press of Hawaii, 1131–1203
- Ting, Pang-hsin (1975), *Chinese Phonology of the Wei-Chin Period*. Taipei: Academia Sinica [Institute of History and Philology, Academia Sinica, Special Publications, no. 65]
- Wang, Li (1958), *Hanyu shigao*, Vol. 1–3. Beijing: Kexue chubanshe
- Yang, Paul Fu-mien S. J. (1977–1978), “Prefix *Kə* in Modern Chinese Dialects and Proto-Chinese”. *Monumenta Serica* 33, 286–299

*Edwin G. Pulleyblank, Vancouver (Canada)*

## 160. From Classical Arabic to the modern Arabic vernaculars

1. History of the Arabic language
2. Morphology of the verb
3. Morphology of the noun
4. Demonstratives and interrogatives
5. Pidginization and creolization in Arabic
6. References

### 1. History of the Arabic language

For the purposes of this article the term “Classical Arabic” will refer to the type of

Arabic that was used in the Arabian peninsula during the pre-Islamic period in poetry and the *Qur’ān* and that was to remain the standard language throughout the entire Arabo-Islamic history. Its modern form is usually called “Modern Standard Arabic”; this variety is distinguished from Classical Arabic mainly by its lexicon and phraseology, as well as minor syntactic adaptations. According to most Western Arabists, in the pre-Islamic

period Classical Arabic was an intertribal and supratribal language, based on the dialects of the Bedouin in the central and eastern parts of the peninsula (cf. Rabin 1951). In this view, the “poetico-Qur’ānic **koine**”, as it is sometimes called (cf. Zwettler 1978: 97–188), was never used in everyday communication. The vernacular of the Arab tribes is assumed to have differed from the Classical type of Arabic in lexicon and phonology, but chiefly in the absence of declensional and modal markers (*i'rāb*). The proponents of this view point to the evidence found in the inscriptions of early forms of Arabic, such as Lihyanite (cf. Müller 1982; Robin 1992). According to Corriente (1976) the *i'rāb*-less dialects were mainly found in the periphery of the peninsula, namely in those areas where contact with other languages took place.

The Arabs themselves view the history of their language in a different way. According to them, there was only one Arabic language in the pre-Islamic period, the *Arabiyya*. It is true that the Arab grammarians mention a considerable amount of differences between tribal dialects (*lughāt*), but these concern almost without exception lexical and phonetic features. In this account, the Arabic language was “corrupted” by the contact with the inhabitants of the conquered territories after the conquests. This resulted in a drive for the preservation of the language by the grammarians, thanks to whom the Classical language remained the standard language for cultural, political, and above all, religious purposes. The Arabs emphasize the vulgarity and decadence of the language of the sedentary population as against the purity of the language of the Bedouin. Some Western Arabists have followed this view of the development of the language (cf. Nöldeke 1904; Blau 1977), and it has been re-interpreted in terms of the **acquisition process** of the language, which may be compared with processes of pidginization and creolization (cf. Versteegh 1984; 1997: 102–112; but see Holes 1995: 19–24).

In this article, the type of Arabic that became the colloquial language after the period of the conquests will be called “New Arabic” or “Arabic dialects”. In the contemporary Arab world this colloquial language is usually called the “common, vulgar language” (*āmmiyya*) as against the “correct, Classical language” (*fushā*). Both varieties serve as the extremes in a continuum of speech variation, which was described as one of the classic cases of **diglossia** by Charles Ferguson

(1959a). Even according to those Arabists who maintain that some of the features of New Arabic were already present in the posited pre-Islamic colloquial language, the bulk of the linguistic changes were introduced during the period of the conquests. Explanations for the development of the colloquial language are complicated by the occurrence in the dialects of common differences that set them apart from the Classical language, as well as differences among them. The common features of New Arabic are explained either by a theory of **monogenetic origin** (cf. Ferguson 1959b), or by a **general drift**, which is compared to similar trends in other Semitic languages, such as Neo-Aramaic and Amharic. Others draw the attention to subsequent processes of **convergence**, which led to greater similarity between the dialects (cf. Cohen 1970; Diem 1978). The differences between the various dialect areas are sometimes explained by **substratal influence** (cf. Diem 1979; Versteegh 1997: 104–107).

In some instances, the dialects exhibit analogous processes of change with varying lexical realization, such as the development of an analytical possessive construction and a system of aspectual markers in the verb: each dialect has its own characteristic genitive exponent and aspectual particles. In view of this, the most reasonable account for the emergence of the dialects seems to be an explanation in terms of a polygenetic origin. Accordingly, the similarity of the development they went through must be the result of the similarity in acquisition process of the language by the inhabitants of the conquered territories. This may explain some of the common features. Other features, especially in the lexicon, may have spread by subsequent convergence.

An important common factor in the development of New Arabic is the influence of the Classical language, which was pervasive in all Arabophone areas, except in the so-called **language islands** (Malta, Cyprus, Anatolia, Uzbekistan, Afghanistan). The development in the latter dialects is characterized by the fact that from a certain point in time onwards they no longer underwent the influence of the Classical standard, so that elements of the original structure could be maintained. Changes in the structure of the dialects of the language islands did occur under the influence of the **adstratal languages** (e.g., Italian in the case of Maltese, Uzbek in the case of Uzbekistan Arabic, Turkish and Kurdish in the case of Anatolian Arabic).

Within the Arabic dialects the dialects of the Bedouin occupy a special position (cf. Rosenhouse 1984). The Arab grammarians asserted that only the Bedouin spoke a pure Arabic, with all the inflectional features preserved. According to them, the Bedouin only gradually succumbed to the corruption of sedentary speech. This constitutes an important argument for the thesis according to which the *Arabiyya* was, indeed, of the Classical type, with only minor regional differences. The main criteria that distinguish Bedouin from sedentary dialects are: the realization of Classical /q/ (Bedouin voiced [g], [g<sup>y</sup>] etc.; sedentary voiceless [q], [χ] etc.); the realization of the Classical interdentals (Bedouin [t], [d], [d̪]; sedentary [t̪], [d̪], [d̪̫]); the realization of some morphological features in Bedouin dialects (e.g., the ending *-in*, *-an* as a marker of **indefiniteness**; gender distinctions in the plural of the verb and the pronoun; internal **passive**; **causative**). The characterization of a dialect as Bedouin or sedentary is a linguistic one, since there have been many secondary processes of bedouinization and sedentarization, for instance in the dialects of Baghdad, where the Muslim dialect was bedouinized, whereas the Christians and the Jews continued to speak their sedentary dialects (cf. Blanc 1964).

Geographically, the contemporary dialects are divided into the Western and the Eastern dialects. The main isogloss between these two areas is that of the 1st person singular and plural of the imperfect verb (cf. 2.1). The Western dialects include the dialects of North Africa: the Ḥassāniyya dialect of Mauretania, Moroccan (for the dialect of Rabat, cf. Harréll 1962; Caubet 1993), Algerian, Tunisian, and Libyan. The Eastern dialects are usually divided into four major groups: Egyptian Arabic (cf. Behnstedt & Woidich 1982–1999; Woidich 1990), including Sudanese and Chad Arabic, with outliers stretching into Nigeria (cf. Owens 1993); Syro-Lebanese dialects (for the dialect of Damascus, cf. Grotzfeld 1965); Mesopotamian dialects (cf. Jastrow 1978; for the dialect of Baghdad, cf. Erwin 1963), including the dialects of language islands in Anatolia and Uzbekistan; and the dialects of the Arabian peninsula: Yemeni (cf. Behnstedt 1985/1992; Watson 1993), dialects of Saudi-Arabia (cf. Prochazka 1988; Ingham 1982; 1994), and Eastern Arabic Gulf dialects (cf. Holes 1990). Within each area, there is a division into Bedouin and sedentary dialects.

The description of the Arabic dialects is complicated by the diglotic situation. On the one hand, the existing continuum between dialect and standard and the constant interference between the extremes make it very difficult to elicit ‘pure’ dialect forms. Especially in the presence of a researcher speakers tend to use the prestigious variants more often than they normally do. As a result, the description of dialects often reflects Classical structures, and researchers usually interpret New Arabic structures in terms of Classical Arabic. On the other hand, speakers actually use the prestigious variants much more frequently nowadays because of increased education and exposure to the language of the media. A case in point is the analytical possessive construction, which for most speakers competes with the Classical construction even in informal speech. In this connection, the dialects of the language islands may serve as a point of reference: to some degree, they have escaped the influence of the Classical language and are, therefore closer to the original dialect structure.

In general, scholars maintain that the development of the Arabic dialects (apart from those in the language islands and from the attested cases of pidginization/creolization in the Arab world, cf. 5) is not very spectacular compared to the development in, for instance, the Neo-Aramaic dialects (cf. Diem 1978). To some extent, this is correct because of the influence of the Classical language, but on the whole, such a view tends to obscure the most salient traits of New Arabic, especially in the field of syntax.

Most scholars would agree that the development of the dialects is characterized by a process of **reduction** and a tendency towards greater **analyticity** vis-à-vis the Classical language. Partly because of changes in stress and accent (cf. Janssens 1972) New Arabic has lost the final short vowels, some of which served as the core of the inflectional endings. This has led some scholars to assume that the disappearance of the inflectional system may be explained phonologically. On the whole, however, this is not acceptable (cf. Diem 1991) since the inflectional system has broken down in its entirety, and the changes involved are clearly structural ones.

## 2. Morphology of the verb

Classical Arabic has a verbal system with distinctions for number (singular, dual, plural), gender (masculine, feminine in the 2nd and

3rd person), person (1st, 2nd, 3rd). Most verbs consist of three radicals that are usually indicated by the three consonants *f*-*l*. Traditionally, a perfect (*fa'ala*) and an imperfect (*yaf'alilu*) verb are distinguished, along with an imperative (*if'al*). Morphologically the perfect, which denotes past actions with a punctual, perfective or stative aspect, is marked by suffixes, and the imperfect, which denotes actions with a durative or habitual aspect, is marked by prefixes and suffixes. The imperfect verb has a three-way distinction between indicative, subjunctive (mainly after conjunctions) and apocopate or jussive (mainly for adhortative and prohibitive functions), which are marked by a change in endings, e.g. *yaf'al-u* (indicative), *yaf'al-a* (subjunctive), *yaf'al* (jussive).

### 2.1. The sound verb

The paradigm for **sound verbs** (i.e., those not containing *w* or *y* as one of their radicals) is as follows (cf. Tab. 160.1).

In all or most dialects the following developments have taken place, usually as a result of reduction or analogous change:

- the category of the **dual** has disappeared without leaving any traces;
- in most sedentary dialects the **gender** distinction has disappeared in the 2nd and 3rd person plural, in Western dialects even in the singular;
- the **modal** distinctions in the imperfect verb have disappeared through the loss of the short vowel endings and the ending *-na*;
- in most dialects the personal endings *-tu/-ta* have merged in *-t*; for the feminine ending of the 2nd person singular a new

ending *-tī* (> *-ti*) has emerged; in the Western dialects *-ta/-ti* have merged in *-tī* (> *-ti*), whereas the 1st person has become *-t*;

- the personal prefixes in the imperfect verb have the vowel *-i-* instead of Classical *-a-* in almost all dialects; this phenomenon (called *taltala*) was already known in some of the pre-Islamic dialects (cf. Rabin 1951: 61);
- the imperfect verb has developed a system of **aspectual markers** (cf. 2.4);
- some dialects have achieved greater regularity for the perfect plural endings *-tum/-ū*, e.g., in some Egyptian dialects *katabtum*, *katabum*; other dialects have *-tul-u*; Western dialects have *-tum/-tīw*;
- all Western dialects are characterized by a 1st person singular prefix of the imperfect verb *ni-*, plural *ni-..-u*; the existence of a mixed paradigm in the Egyptian delta (*aktib/niktibu*) shows that the earlier explanation of this development as a contraction from *'ana 'aktubu* 'I write' > *naktub* is probably wrong (cf. Behnstedt 1978: 69);
- in Classical Arabic verbal bases of the form *fa'ala* alternate with verbal bases of the pattern *fa'ila* (for non-permanent states, e.g., *gadiba* 'to be angry') and *fa'ula* (for permanent states, e.g., *hasuna* 'to be good'); in the dialects only two bases have been preserved without semantic distinction, namely *fa'al* and *fi'il*; in Moroccan and Algerian Arabic there is only one base, *f'āl*;
- in Classical Arabic the perfect bases correlate with imperfect patterns: *fa'ala* verbs have imperfect *yaf'lu* or *yaf'ilu* (in

'to write'		singular	dual	plural
perfect	1st	<i>katabtu</i>		<i>katabnā</i>
	2nd masc.	<i>katabta</i>	<i>katabtumā</i>	<i>katabtum</i>
	2nd fem.	<i>katabti</i>		<i>katabtunna</i>
	3rd masc.	<i>kataba</i>	<i>katabā</i>	<i>katabū</i>
	3rd fem.	<i>katabat</i>	<i>katabatā</i>	<i>katabna</i>
imperfect	1st	<i>'aktubu</i>		<i>naktubu</i>
	2nd masc.	<i>taktubu</i>	<i>taktubāni</i>	<i>taktubūna</i>
	2nd fem.	<i>taktubīna</i>		<i>taktubna</i>
	3rd masc.	<i>yaktubu</i>	<i>yaktubāni</i>	<i>yaktubūna</i>
	3rd fem.	<i>taktubu</i>	<i>taktubāni</i>	<i>yaktubna</i>

Tab. 160.1: The sound verb in Classical Arabic

Classical	Egyptian (Cairo)	'Iraqi (Baghdad)
<i>kataba/yaktubu</i>	<i>katablyiktib</i>	<i>kitablyiktib</i>
<i>darabalyadribu</i>	<i>darablyidrib</i>	<i>durablyudrub</i>
<i>harabalyahrabu</i>	<i>harablyihrab</i>	<i>hirablyuhrub</i>
<i>'amila/ya'malu</i>	<i>'amallyi'mil</i>	<i>'imallyi'mal</i>
<i>labisalyalbasu</i>	<i>libislyilbis</i>	<i>libaslyilbas</i>
<i>šaribalyašrabu</i>	<i>šriblyišrab</i>	<i>širablyišrab</i>

Tab. 160.2: Perfect/imperfect patterns in Classical Arabic and two dialects

some phonological contexts *yaf'alu*); *fa'ila* verbs regularly have *yaf'alu*; and *fa'ula* verbs regularly have *yaf'ulu*; in the dialects reflexes of these three patterns occur, but often with a different distribution (cf. Tab. 160.2).

As an example of one of the most conservative dialects, in which only the modal endings have disappeared, we quote here the verbal paradigm of a Saudi-Arabian Bedouin dialect, Gāmid (cf. Prochazka 1988:27) (cf. Tab. 160.3):

'to write'		singular	plural
perfect	1st	<i>katabt</i>	<i>katabna</i>
	2nd masc.	<i>katabt</i>	<i>katabtum</i>
	2nd fem.	<i>katabti</i>	<i>katabtunnah</i>
	3rd masc.	<i>katab</i>	<i>katabaw</i>
	3rd fem.	<i>katabat</i>	<i>katabnah</i>
imperfect	1st	<i>aktub</i>	<i>niktub</i>
	2nd masc.	<i>tiktub</i>	<i>tiktubūn</i>
	2nd fem.	<i>tiktubīn</i>	<i>tiktubnah</i>
	3rd masc.	<i>yiktub</i>	<i>yiktubūn</i>
	3rd fem.	<i>tiktub</i>	<i>yiktubnah</i>

Tab. 160.3: Verbal paradigm of Gāmid (Saudi-Arabia)

As an example of an innovative dialect with complete disappearance of gender distinction, except in the 3rd person singular we quote the verbal paradigm of sedentary Moroccan (Rabat) (cf. Tab. 160.4):

## 2.2. The weak verb

In Classical Arabic the conjugation of the **weak verbs** (i.e., those containing a *w* or *y* as one of their radicals) differs from that of the sound verbs as the result of various morpho-phonological rules. The following types may be distinguished (I, II, III indicating the three radicals of the verbal stem) (cf. Tab. 160.5):

		singular	plural
perfect	1st	<i>ktəbt</i>	<i>ktəbna</i>
	2nd masc.	<i>ktəbti</i>	<i>ktktəbtiw</i>
	2nd fem.	<i>ktəb</i>	
	3rd masc.	<i>ktəb</i>	<i>kətbu</i>
	3rd fem.	<i>kətbət</i>	
imperfect	1st	<i>nəktəb</i>	<i>nkətbu</i>
	2nd masc.	<i>təktəb</i>	<i>tkətbu</i>
	2nd fem.	<i>tkətbī</i>	
	3rd masc.	<i>yəktəb</i>	<i>ykətbu</i>
	3rd fem.	<i>təktəb</i>	

Tab. 160.4: Verbal paradigm of sedentary Moroccan (Rabat)

In New Arabic the following developments in the conjugation of the weak verbs may be noted; they all tend towards greater regularity:

- in most dialects the type Iw has become sound, e.g., in Egyptian the verbs *wazan/iywzin* 'to weigh', *wa'adlyiw'id* 'to promise', *wi'i/yiw'a'* 'to fall down';
- the two types IIIw and IIIy have merged in all dialects, e.g., in Egyptian *mišil/mišeṭ/yimši* 'to walk' (IIIy, ē < ay) and *da'ala/ēṭ/yid'i* 'to invite' (IIIw); the same development took place in the perfect of the *laqiya* group, e.g., Egyptian *la'allā/ēṭ/yil'a* 'to meet';
- the type II = III has merged with the II measure of the IIIy verbs, e.g., in Egyptian *raddal/raddēt* 'to return' like the II measure of *maša*: *mašša/maššet*.

A tendency may be noted in the dialects towards greater similarity between the endings of the weak and those of the sound verbs, especially in the case of the type IIIy. In many cases this similarity was achieved by

	underlying	3rd person perfect	3rd person imperfect	1st person imperfect
Iw	wa`adal	<i>wa`ada</i>	<i>ya`idu</i>	<i>wa`adtu</i>
IIw	qawama	<i>qāma</i>	<i>yaqūmu</i>	<i>qumtu</i>
IIy	sayara	<i>sāra</i>	<i>yasīru</i>	<i>sirtu</i>
	hawifa	<i>hāfa</i>	<i>yahāfu</i>	<i>hiftu</i>
IIIw	da`awal	<i>da`ā</i>	<i>yad`ū</i>	<i>da`awtu</i>
IIIy	mašaya	<i>mašā</i>	<i>yamšī</i>	<i>mašaytu</i>
	laqiyā	<i>laqiyā</i>	<i>yalqā</i>	<i>laqītu</i>
II = III	radada	<i>radda</i>	<i>yaruddu</i>	<i>radadtu</i>

Tab. 160.5: Types of weak verbs in Classical Arabic

assigning the sound endings to the weak verb, but alternatively, there are dialects in which the weak endings are transferred to the sound verb, either in the perfect (e.g., Muslim dialect of Baghdad) or in the imperfect (e.g., Jewish dialect of Baghdad), as shown in Tab. 160.6 (where  $\bar{e} < ay$ ,  $\bar{o} < aw$ ).

Classical	Egyptian (Cairo)	Muslim Baghdad	Jewish Baghdad
<i>katabū</i>	<i>katabu</i>	<i>kitbaw</i>	<i>katbu</i>
3rd plural			
<i>mašaw</i>	<i>mišu</i>	<i>mašaw</i>	<i>mašaw</i>
3rd plural			
<i>yaktubūna</i>	<i>yiktibu</i>	<i>yikitbūn</i>	<i>ykətbōn</i>
3rd plural			
masc.			
<i>yansayna</i>	<i>yinsu</i>	<i>yinsūn</i>	<i>yənsōn</i>
3rd plural			
masc.			
<i>taktabūna</i>	<i>tiktibi</i>	<i>tikitbīn</i>	<i>tkətbēn</i>
2nd singular			
fem.			
<i>tansayna</i>	<i>tinsi</i>	<i>tinsīn</i>	<i>tənsēn</i>
2nd singular			
fem.			

Tab. 160.6: Weak and sound endings in Arabic dialects

The development of the weak endings was important in another respect as well. In Maltese Arabic the pattern of the weak verbs has proliferated to such an extent that foreign loans, especially from Italian, which often ended in a vowel, could be integrated easily (cf. Mifsud 1995), e.g., *iddependaljidddependi* ‘to depend’ (< Italian *dipendere*), *importaljimporta* ‘to matter’ (< Italian *importare*).

Similar developments have taken place in French loans in Moroccan Arabic (cf. Heath 1989: 108–112).

### 2.3. The verbal measures

In Classical Arabic the verb system may be enlarged to form semanto-syntactic categories, the so-called **verbal measures**. Usually the following measures are recognized as productive stem formations (the main function indicated as far as possible, although in many cases the meaning is fixed lexically):

II	<i>kattaba</i>	intensive, causative, denominative
III	<i>kātaba</i>	transitive to direct object
IV	<i>'aktaba</i>	causative
V	<i>takattaba</i>	reflexive-intransitive to II
VI	<i>takātaba</i>	reciprocal to III
VII	<i>inkataba</i>	reflexive-passive
VIII	<i>iktataba</i>	reflexive-intransitive to I
IX	<i>(il)marra</i>	colours, bodily defects
X	<i>istaktaba</i>	request, opinion of I

Tab. 160.7: The verbal measures in Classical Arabic

In most dialects the II and III measures have been preserved, although only the II measure is productive (as intensive, denominative or causative); the III measure is lexically fixed and functions just as an alternative vowel pattern, e.g., Moroccan *qatəl* ‘to kill’/*qatəl* ‘to fight’. The IV measure has disappeared from almost all dialects and is replaced by the II measure, e.g., Moroccan *ktəb* ‘to write’/*kattəb* ‘to make so. write’ (Classical Arabic *'aktaba*), or by a periphrastic expression. Only in some Bedouin dialects in the East a IV measure is still productive, e.g., Dēr iz-Zōr (Syria) *haḍar/yihḍur* ‘to be present’, *ihḍar/yihḍir* ‘to

produce' (Fischer & Jastrow 1980, eds.: 46; cf. Ingham 1994: 77 for Nağdī Arabic). In Classical Arabic, the IX measure was strictly limited to verbs denoting colours and bodily defects; in the dialects this form has often been replaced by the II measure, e.g., in the Gulf dialects *bayyad* 'to make white' or 'to be white' (Fischer & Jastrow 1980, eds.: 71); other dialects have preserved a restricted IX measure, e.g., Egyptian *iħmarra* 'to be red'. In Maghrebine dialects a new form has developed, e.g., Moroccan *ħmar* 'to be red'. The X measure exists only in isolated lexical items, often loans from Classical Arabic, e.g., Moroccan *stħsen* 'to approve', *stəd'a* 'to summon'.

Each verbal measure has a passive in Classical Arabic, marked by a characteristic **vowel pattern** (so-called 'internal passive'), e.g., I measure *katabalyaktubu*, passive *kutibalyuktabu*; VIII measure *iktatabalyaktatibu*, passive *uktutibalyuktatabu*. In most dialects the **internal passive** has disappeared (Retsö 1983), but some Bedouin dialects have preserved a passive formation, e.g., in Saudi-Arabia (Ristāq) from *gatal* 'to kill' *gtill/yugtal* (Fischer & Jastrow 1980, eds.: 117, 262; cf. Ingham 1994: 26–28). In the Hassāniyya dialect of Mauretania a new internal passive has developed for derived stems, e.g., from *sahħah* (II) 'to correct' *uṣahħah/yusahħah* (Zavadovskij 1981: 35 f.).

In the other dialects some of the verbal measures have become specialized for the passive (cf. Retsö 1983). In some dialects the VII measure is used for this purpose, e.g., in Hassāniyya from *šrab* 'to drink' *nšrab* 'to be drunk' (Zavadovskij 1981: 34). In other dialects the Classical V and VI measures with their intransitive meaning are used for the passive, building on the Classical pattern of verbs like *gayyara* II 'to change st.', *tagayyara* V 'to be changed, to change [intrans.]'. In these dialects the marker *ta-* has become a passive marker which may even be used with the base stem (measure I). Examples are Moroccan *ktab* (I) 'to write'/*ttaktab* (tI, with secondary initial gemination) 'to be written'; *səmma* (II) 'to call'/*tsəmma* (tII) 'to be called'; Egyptian *misik* 'to catch'/*itmasak* 'to be caught'. In the tIII stems the Classical reciprocal meaning has usually been preserved, e.g., Moroccan *tqatel* 'to fight each other'.

The vowel patterns of the derived measures show a tendency towards a uniform vocalization. In Egyptian, for instance, the per-

fect and the imperfect of the II measure have acquired identical vowels: Classical Arabic *kallamalyukallimu* 'to talk', Egyptian *kalliml yikallim*. In the tII and tIII measures the vocalization is identical: Classical Arabic *ta'allaħħalayataħħamu* 'to learn', *taqāħbalalayataqāħbalu* 'to be met, to meet each other', Egyptian *it'ħallim/yit'ħallim*; *it'ħabbilyit'ħabil* (Jastrow 1982: 136).

#### 2.4. The aspectual markers

The most radical departure from the structure of Classical Arabic is the marking of verbal **aspect**. The main function of the two verbal forms, perfect and imperfect, in Classical Arabic seems to have been aspectual (Denz 1982: 71 f.; but cf. Eisele 1999). They may be combined with the auxiliary verb *kāna* to indicate past **perfectivity** and past **durativity** or **habituality**, respectively, e.g., *kāna qad kataba* 'he had written', *kāna yaktubu* 'he was writing, he used to write'. There are two particles that are connected with aspect/tense: *qad* indicates perfectivity and *sawfa* (*sa-*) indicates **future tense**. Neither particle has any reflex in New Arabic. In the dialects a new form of aspectual marking is used. The old **imperfect**, stripped of its modal endings, has become a modal form itself, e.g., Egyptian *tiħrab* 'ahwa 'would you like to have coffee?'. Durative and future/volitional aspect are indicated by clitic particles, whose etymology is not always clear, but seems to be connected with the notion of 'being at, sitting' for the durative and 'going' or 'desiring' for the future. The particles are highly characteristic for each dialect (Czapkiewicz 1975) (cf. Tab. 160.8).

	durative	future
Egyptian Arabic	<i>bi-</i>	<i>ha-</i>
Moroccan Arabic	<i>ka-</i>	<i>ġa-</i>
Syrian Arabic	<i>'am-</i>	<i>bi-</i>
Baghdad (Muslim)	<i>da-</i>	<i>rah-</i>

Tab. 160.8: Aspectual markers in Arabic dialects

Apart from the geographical variation the markers vary with regard to their semantic domain: the Syrian marker *bi-*, for instance, is used for a **volitional** aspect (probably <*yibgi* 'to desire'), whereas certain future actions are indicated in this dialect with a marker *rah-* (<*rāyih* 'going'). Another example is the indication of habitual or general statements. In some dialects, for instance

Egyptian and Moroccan, the durative marker is used for habitual actions, whereas in Syrian the *bi-* marker is used for that purpose. In 'Iraqi Arabic (cf. Erwin 1963: 335–344) the durative particle is restricted to a certain period of time, whereas the modal imperfect is used for general and habitual statements. The dialects also differ with regard to the combinability of the aspectual markers. In most dialects the durative marker may be combined with the past marker *kān* to indicate past durativity, e.g., Egyptian *kān b-yiktib* 'he was writing'; the combination *kān* + future marker to indicate **irreality** is not available in all dialects: in Egyptian *kān hā-yiktib* means 'he was going to write (but he didn't)', but in Moroccan the combination *kān ga-* is impossible.

In Classical Arabic the active **participle** *fā ilun* may be used to indicate durativity: *huwa kātibun* 'he is writing, he is going to write' (Denz 1982: 72 f.). In many dialects the participle is used to indicate a **resultative** or **perfective** aspect, e.g., Egyptian *ana nāsi* 'I have forgotten', *ana wākil* 'I have eaten'. In Uzbekistan Arabic the participle has become an integral part of the verbal system: it is used instead of the perfect for punctual actions in the past, and it takes both subject and object suffixes, e.g., *zōrib* 'he hit', *zorbinī* 'I hit', *zorbinik* 'I hit you' (Fischer 1961: 255).

## 2.5. The verbal complex

In Classical Arabic **pronominal suffixes** referring to the object may be added to the verb, e.g., *katabtu-hu* 'I wrote it'. There are no restrictions with regard to the person; although fairly unusual, combinations of suffixes may occur with ditransitive verbs, e.g., '*a'lamtukahu*' 'I let you know it' (the preferred form is with the particle *iyyā*: '*a'lamtukahu iyyā*'). The **indirect pronominal object** is indicated with the preposition *li-*, e.g., *katabtu-hu la-hā* 'I wrote it to her'. The indirect object is not incorporated in the verbal form, and it can be separated from it, e.g., *qultu dālikā la-ka* 'I told you that'.

In general, the modern dialects have preserved the form of the pronominal suffixes. Noteworthy is the disappearance of the *-h* of the 3rd person masc. sing. in virtually all dialects, e.g., Egyptian *katabt-o* 'I wrote it'. In some dialects this leads to a secondary stress contrast with suffixless forms, e.g., Syrian *kátabu* 'they wrote' vs. *katabú* 'they wrote it'. The normal morphophonemic rules lead to epenthesis and vocalic switch, e.g., Syrian

*yaktubu-hu* > *yiktbo* > *yəkətbo*, or Egyptian (with stressed epenthetic vowel) *yiktibō*.

An innovation is the development of the indirect object as a clitic to the verbal form, e.g., Egyptian *katab-līna* 'he wrote to us'. In most dialects a direct pronominal object of the 3rd person may be combined with an indirect object suffix, usually without constraints as to the person of the indirect object, e.g., Egyptian *iddet-hā-lo* 'I gave her to him', *bi-ywarri-ho-lhom* 'he shows him to them'. In Anatolian Arabic (Daragözü) the indirect object precedes the direct object, which may only be 3rd person, e.g., *wəddī-lī-a* 'bring her to me!', *wəddī-lā-an* 'bring them to her!' (Jastrow 1973: 68). For the development of subject suffixes with the participle in Uzbekistan Arabic see 2.4.

As we have seen in 2.4 the aspectual markers are added clitically as prefixes to the verbal form. The negative **circumfix** *mā ... š* may be added to the verbal form as well, so that morphologically rather complex forms may result. As an example we may quote Moroccan *ma-ğā-nəktəb-o-lək-š* 'NEG-FUT-I.write-it-to.you-NEG (I won't write it to you)' or Egyptian *ma-bi-tgib-hā-lhom-š* 'NEG-DUR-you.bring-her-to.them-NEG (you are not bringing her to him)'.

## 3. Morphology of the noun

In Classical Arabic substantives and adjectives are inflected for **case** (nominative, genitive, accusative), **number** (singular, dual, plural) and **(in)definiteness**. There is a two-way distinction between masculine and feminine nouns, the feminine nouns being characterized by the ending *-at* (pausal form *-ah*), and for some categories of nouns, by *-ā*, *-ā'*. Indefiniteness is marked by the ending *-n* (the so-called **nunation**), definiteness by the definite article *al-* and the absence of nunation.

### 3.1. Case

The singular declension is as follows (cf. Tab. 160.9):

	masculine 'man'	feminine 'city'
nominative	<i>rağul-u-n</i>	<i>madīn-at-u-n</i>
genitive	<i>rağul-i-n</i>	<i>madīn-at-i-n</i>
accusative	<i>rağul-a-n</i>	<i>madīn-at-an</i>

Tab. 160.9: Singular declension of the noun in Classical Arabic

In New Arabic the **feminine** ending has become *-a* in sedentary dialects; the form with *-at* (> *-it*) appears before pronominal suffixes and in possessive constructions (Egyptian *‘ōda* ‘room’, *‘ōdtak* ‘your room’, *‘ōdit il-bēt* ‘the room of the house’). Bedouin dialects retain the pausal form *-ah*, *-ih*. In some dialects a redistribution of the feminine ending has taken place, depending on the preceding consonant, e.g., Syrian *mnīha* ‘good:F’ vs. *kbīra* ‘big:F’. The alternative endings *-ā*, *-ā'* have disappeared, e.g., Egyptian *hamra* ‘red:F’ vs. Classical *hamrā’u*.

Since the short vowel endings of Classical Arabic have all disappeared in New Arabic, there is no inflection for case in the singular. The fact that in the sound plural, too, only one case ending has been preserved (cf. 3.2) and the fact that the nunciation as marker of indefiniteness has disappeared completely (apart from a few traces in Bedouin dialects), clearly show that the development is a morphological one (cf. Diem 1991). The syntactic function of the cases has been transferred to syntactic devices: the word order in the dialects has become more or less fixed to SVO; the adverbial and adnominal functions of the accusative have been transferred to other constructions, and the possessive function of the genitive has been taken over by the genitive exponent.

The emergence of the **genitive** exponent constitutes one of the most obvious changes in the structure of the language towards greater analyticity, e.g., Egyptian *il-bēt bitā’* ‘*il-wizīr* ‘the house of the minister’ instead of Classical Arabic *bayt-u l-wazīr-i*. The etymological origin of the genitive exponent is usually a noun meaning ‘possession’, ‘belonging to’, ‘thing’. Just as with the aspectual markers of the verb (cf. 2.4) the choice of the exponent is highly characteristic for each dialect area (cf. Harning 1980) (cf. Tab. 160.10):

Syrian Arabic	<i>taba’</i>
Egyptian Arabic	<i>bitā’</i>
Moroccan Arabic	<i>d-</i> , <i>dyl</i>
Baghdad Arabic (Muslims)	<i>māl</i>
Sudanese Arabic	<i>haqq</i>
Cypriot Arabic	<i>šāyt</i>
Sa‘īdī Egyptian	<i>šugl</i>
Chad Arabic	<i>hana</i>
Anatolian Arabic (Daragözü)	<i>līl</i>
Maltese	<i>ta</i>

Tab. 160.10: Genitive exponents in the Arabic dialects

Although not strictly relevant for the morphological development, two remarks should be added here. In most dialects, the analytical possessive construction coexists with the Classical construction. The choice for the latter is in some cases obligatory, due to semantic constraints. In Moroccan Arabic, for instance, the analytical construction is not used with terms of kinship, e.g., *mnī-i* ‘my mother’ (not *\*l-umm dyl-i*). In most dialects there is a distinction between **alienable** and **inalienable** possession (cf. Art. 103), e.g., Egyptian *il-lahm ibtā’i* ‘my piece of meat’ vs. *lahm-i* ‘my flesh’. Apart from these semantic distinctions there is a high degree of sociolinguistic variation in the use of the possessive construction. The current diglossia has (re-)introduced the Classical construction as an alternative in the dialect speech community. Not surprisingly, the language islands do not exhibit such variation. In Maltese Arabic, for instance, the analytic construction with the genitive exponent *ta* is obligatory, except for a few isolated expressions and terms of kinship. Conversely, in some of the more conservative Bedouin dialects the analytic genitive is hardly ever used.

In some dialects the disappearance of the formal marker for indefiniteness has led to the development of an **indefinite article**. Examples are ‘Iraqi *fat* (< Classical Arabic *fard* ‘individual’), also in Uzbekistan Arabic (which has lost the definite article), Moroccan Arabic *wahd* (< Classical Arabic *wāhid* ‘one’), Cypriot Arabic *ehen/lehte* (< Classical Arabic *’ahad* ‘one’; cf. Borg 1985: 145).

In a few Bedouin dialects the nunciation has not disappeared completely since there is a generalized nunate ending, acting as a connector between noun and attribute and indicating indefiniteness without marking the syntactic function, e.g., in Nağdī Arabic (Saudi-Arabia) the indefinite marker *-in* (cf. Ingham 1982: 53–56), *kalmit-in gālo-hā-li* ‘a word which they said to me’, *giz-in min-h* ‘a part of it’; in Uzbekistan Arabic the same marker *-in* is used, e.g., *fat haġart-in kabīra* ‘a big stone’ (cf. Fischer 1961: 244, who derives this marker from Classical Arabic *’ayna* ‘where?’). Andalusian Arabic seems to have known the ending *-an* as a connective element (cf. Corriente 1977: 121). In Yemeni Arabic a contrast exists between *bayt-u* ‘a house’ and *al-bayt* ‘the house’ (cf. Fischer & Jastrow 1980, eds.: 120).

### 3.2. Number

Classical Arabic has three numbers, **singular**, **dual**, and **plural**. The dual is marked by the endings *-āni* (nom.)/-*ayni* (gen./acc.) for masculine, and *-atānīl*-*atayni* for feminine nouns. One of the characteristics of the modern dialects is that they have lost the dual in the verb and the pronoun. In most dialects the formation of dual nouns is not productive either, except for a limited category of nouns. In Egyptian and Syrian, for instance, the dual is current with terms of kinship, e.g., *bintēn* ‘two daughters’, but it may be used occasionally with other nouns as well. In Moroccan the dual is only used with expressions indicating a period of time, e.g., *yumayn* ‘two days’, *sa'tayn* ‘two hours’. In some dialects the ending *-ayn* is also used for paired body parts, the so-called **pseudo-dual** (cf. Blanc 1970), e.g. Syrian *'enēn* ‘eyes’, *idēn* ‘hands’. The difference between this pseudo-dual and the original dual is obvious in the contrast between *'enēn*, which functions as a plural, and *'entēn* ‘two eyes’, which functions semantically as a dual. In combination with pronominal suffixes the pseudo-dual loses the ending *-n*, as in Classical Arabic – compare *īdēk* ‘your hands’ with *bintēnak* ‘your two daughters’.

In Classical Arabic there are two kinds of plural. The so-called **sound plural** is formed with the help of suffixes: *-ūna* (nom.)/-*īna* (gen./acc.) for masculine, and *-ātunl*-*ātin* for feminine nouns. The sound plural is used almost exclusively with some adjectives and all active participles, e.g., *gāmīlūnā/gāmīlātūn* ‘beautiful’; *muslīmūnā/muslīmātūn* ‘Muslims [act. part. IV measure]’. The second type of plural (the so-called **broken plural**) consists in a rearrangement of the vowel pattern of the singular. More than thirty different patterns are used, some of them connected with specific singular patterns, but by no means exclusively. A few examples: *rağulun/riğālun* ‘man’; *kitābun/kutubun* ‘book’; *wazīrun/wu-zarā'u* ‘minister’; *ṣadīqun/'aṣdiqā'u* ‘friend’; *ṭalibun/tullābun* ‘student’. The broken plurals are declined as singular nouns, except for the ones ending in *-u* (diptotic plurals) which have their gen./acc. in *-a* when they are indefinite.

As we have seen in 3.1, New Arabic has lost its declensional system: the sound plural is marked by the oblique ending *-īn* for masculine and *-āt* for feminine nouns. In the iso-

lated dialects there is a tendency to generalize these endings. In Uzbekistan Arabic (Fischer 1961: 243), for instance, the ending *-īn* is reserved for male persons, e.g., *uhwīn* ‘brothers’ (Classical Arabic *ihwatun*), *zuḡīn* ‘husbands’ (Classical Arabic *'azwāġun*), whereas the ending *-āt* is used for female persons and all other plurals, e.g., *rāsāt* ‘heads’ (Classical Arabic *ru'ūsun*), *uhāt* ‘sisters’ (Classical Arabic *'ahawāt*). In Cypriot Arabic (cf. Borg 1985: 180) *-āt* is the general plural ending for a large number of inanimate nouns, even for those which etymologically are broken plurals, e.g., *humāt* ‘kinds of meat’ (Classical Arabic *lahm*, plural *luhūm*).

The usual plural ending for adjectives in the dialects is *-īnl*-*āt*. For adjectives of the type *fa'īl*, e.g., *kabīr* ‘big’, all dialects have developed a new form, *fi'īl*; the plural for this type has become *fu'āl* (Classical Arabic *fi'āl*), e.g., Egyptian *kubār*, whereas substantives with the same plural ending in Classical Arabic retain *fi'āl*, e.g., *gamall/gimāl* ‘camel’.

Most dialects have retained a number of broken plural types, although with less variation than in Classical Arabic. In some cases the dialects have expanded the domain of certain types of plurals. The two patterns *fa'ālīl* and *fa'ālīl* for quadrilateral nouns have merged in most dialects into *fa'ālīl*, e.g., Egyptian *šābabīk* (Classical Arabic *šābabīku*) from *šubbāk* ‘window’. In Western dialects the resulting pattern is *f'ālīl*, e.g., Moroccan *šabbak/šbabak*. The plural pattern *'af'āl* has become identical in many dialects with the pattern *fi'āl*, due to the disappearance of the pretonic vowel, e.g., Moroccan *qləm/qlam* ‘pen’ (Classical Arabic *qalamun/qaqlamun*). The plural pattern *fa'ālt* > *fa'āli*, which in Classical Arabic was reserved for the singular types *fa'lā*, *fa'lā'u*, has become much more general, e.g., Egyptian *hāral/hawāri* ‘alley’ (Classical Arabic *hāra/hārāt*). In one category of nouns, expressing professional activities, the ending *-āt* has been replaced with the feminine singular (and **collective**) ending *-a*, e.g., Syrian *laḥḥām/laḥḥāme*. Weak nouns of this type have the ending *-iyya*, e.g., Egyptian *ḥarāmil/ḥarāmiyya* ‘thief’. The latter ending became generalized in foreign loans, e.g., Syrian *sōfēr/sōfērīye* ‘driver’.

In most dialects new patterns for the plural have developed. In Yemeni Arabic, for instance, the plural *af'ūl* is found, e.g. *'amm/ā'mūm* ‘uncle’ (cf. Fischer & Jastrow 1980, eds.: 91). Another new plural pattern in Ye-

meni Arabic is *fi'wal*, *fu'wal* for singulars of the type *fa'il*, e.g., *tariq/turwaq* 'street'. This pattern has sometimes been associated with the plural *qetwōl*, *qetyōl* in Modern South Arabian, but according to others, it is a case of common Semitic stock rather than direct substratal influence (cf. Diem 1979: 64–75).

### 3.3. Nominal patterns

As we have seen in the case of the broken plurals in the modern dialects, New Arabic has preserved the triradical system of Classical Arabic. Alongside the inflectional patterns Classical Arabic used a large number of derivational patterns, which were partly productive, partly fixed. In the case of singular nouns, some types appear to constitute semantically constrained categories, e.g., the type *fa'il*, which is prevalent in adjectives. Clear examples of derivational patterns are the following: *'af'alu* (colours), *fa-'āl* (occupations), *mif'āl* (instruments), *maf'ala*, *maf'ala* (places of action). In New Arabic most of these patterns have been preserved, and in some cases they have regained productivity through the efforts of the Arab Academies to create a pure Arabic lexicon with the help of derivational patterns (cf. Ali 1987). This development stimulated the creation of new word patterns in the dialects, or the expansion of old patterns. A case in point is the pattern *fu'āl*, which has become the normal pattern for illnesses, e.g., *sufār* 'jaundice' (from *asfar* 'yellow').

In a few cases the interference from other languages led to the formation of completely new patterns; an example is the pattern *tafsa-'ālat* indicating professions in Moroccan Arabic, which was borrowed from Berber, e.g., *tanəżżarət* 'carpentry'.

The most serious threat to the triradical structure of Arabic was the enormous influx of **foreign loans**. In Classical and Modern Standard Arabic most loans were integrated more or less fully into the system, e.g., *faylasūf/falāsifa* 'philosopher' (< Greek *philó-sophos*); *film/filmāṣlām*. In the modern dialects of North Africa the borrowing of French loans as a rule did not cause any major deviation from the triradical structure since most of these could be integrated. As examples we may quote from Algerian Arabic (cf. Hadj-Sadok 1955) *foršēṭal/frāṣāṭ* (< French *fourchette*), *bordill/bṛāḍel* (< French *bordel*). In cases where an internal plural is impossible

most loans receive the plural ending *-āt* or *-iyye*. The existence of a large number of quadrilateral consonants in Classical Arabic facilitated this incorporation of foreign loans.

In Moroccan Arabic a similar situation obtains. A detailed analysis (Heath 1989) has shown the amount of segmental conversion and adaptation that was needed in order to incorporate the nouns borrowed from French into the Arabic system. Noteworthy is the re-analysis of final vowels as the feminine ending *-a*: *anṭirna* (< French *l'internat*) is feminine in Moroccan Arabic, but *duš* (< French *douche*) is masculine. The former example also shows the re-analysis of the French article as part of the noun; other examples of this are *liba* (< French *les bas*), plural *libāt*, and *luṭuṣṭup* (< French *l'autostop*).

In Maltese, the incorporation of Italian loans has led to a completely new system of nominal patterns and to a proliferation of patterns based on Arabic weak stems (cf. Mifsud 1995). Due to the structure of Italian words, the morphology of the nominal patterns had to be restructured, and in many words the pattern has come to control only part of the word, e.g., *umbrella/umbrelel* (< Italian *umbrella*), *gverra/gverer* 'war' (< Italian *guerra*).

The reduction in variation also applies to the patterns of the verbal nouns. In Classical Arabic the **verbal noun** (the so-called **masdar**) of the base stem occurs in a large variety of forms, e.g., *darb* 'hitting', *hurūg* 'going out', *dahāb* 'going away'; some verbs may have more than one masdar. In the derived measures, the verbal nouns have a fixed form. In the dialects there is a general tendency to reduce the number of masdar types (Jastrow 1982: 139) and some of the more isolated dialects have preserved only one type. In Uzbekistan Arabic all verbal nouns end in *-ān*, e.g., *gatalān* 'killing' (Classical Arabic *qatl*), *dahālān* 'entering' (Classical Arabic *duhūl*). The verbal nouns of the derived measures in this dialect use the same ending attached to the participle, e.g., *inkasar* (VII), verbal noun *minkisriyān* (Fischer 1961: 257). Likewise, in Daragözü Arabic *f'il* is the pattern for all verbal nouns, e.g., *fith* 'opening', even in the derived measures, e.g., *tammam* (II), verbal noun *təmmim* 'completing' (Jastrow 1973: 53). In the Maghrebine dialects, which have preserved a much larger number of patterns for the verbal noun, the pattern *f'il* < *fa'il* is also much more frequent than in Classical Arabic.

#### 4. Demonstratives and interrogatives

##### 4.1. Demonstratives

A good example of the enormous variation between the Arabic dialects, within an identical frame of development, is provided by the deictic system. Classical Arabic has two series of demonstrative pronouns, one for near, and one for far deixis. The canonical forms for **near deixis** are *hādā*, fem. *hādihi*, pl. *hā'ulā'i*; for **far deixis** the forms are *dāl-ika*, *tilka*, *ūlā'ika*. These forms are indeclinable, except for the dual, which has two cases (masc. nom. *hādāni*; gen./acc. *hādayni*; fem. *hātānil/hātayni*). The deictic element *-t-* in the feminine *tilka* is also found in an alternative form of the demonstrative pronoun for the near deixis, *tā*, *tī*, which is, however, seldom used.

All modern dialects have at least two series of demonstrative pronouns (cf. Fischer 1959). The near demonstrative pronouns either continue the Classical form or they lose the element *hā-*. The forms for the far deixis are all new developments and the Classical forms are not found in any dialect. The demonstratives for the near deixis are sometimes found in two forms, a short one for attributive use, and a long one for substantivized use, e.g., Algerian Arabic (Djidjelli) *had* vs. *hada*, *hadi*, *hadu*. With regard to the form of the near deixis demonstratives the dialects may be divided into three groups: those which use only forms without *hā-* (e.g., Egyptian, Eastern Sudanese); those which use only forms with *hā-* (e.g., Tunisian, Palestinian, Lebanese); and those which use both (e.g., Moroccan, Algerian, Saudi-Arabian, Andalusian Arabic). In the third group the forms with *hā-* are usually less emphatic than the ones without.

In general the number of different forms in the dialects is staggering. As an example we quote here the demonstrative series from a number of dialects that may be regarded as typical. They all exhibit in varying combinations the deictic elements *-h-*, *-d-*, *-k-*, *-l-* (cf. Tab. 160.11).

The pronouns for the far deixis in Egyptian Arabic are fairly uncommon and used almost exclusively in situations of contrasting deixis. In Cairene there is an alternative set of forms for the near deixis consisting of a combination of demonstrative and personal pronoun (*dawwat*, *diyyat*, *dōlat*); these forms have become slightly stigmatized. In some dialects there is a third series for **middle deixis**, consisting of a combination between the de-

	near deixis	far deixis
Saudi Arabian (Mecca)	<i>hādā</i> , <i>dā</i>	<i>(ha)-</i>
	<i>hādi</i> , <i>dī</i>	<i>dāk(a)</i>
	<i>(ha)dōl(a)</i>	<i>(ha)dīk(a)</i>
Lebanese Arabic	<i>hā</i> , <i>hāda</i>	<i>hadāk(i)</i>
	<i>hay</i> , <i>haydi</i>	<i>haydīk(i)</i>
	<i>haw</i> , <i>hawdi</i>	<i>hawdīk(i)</i>
Egyptian Arabic	<i>da</i>	<i>dukha</i>
	<i>di</i>	<i>dikha</i>
	<i>dōl(a)</i>	<i>dukham(a)</i>
Moroccan Arabic	<i>hada</i>	<i>hadak</i>
	<i>hadi</i>	<i>hadik</i>
	<i>hadu</i>	<i>haduk</i>
Daragözü Arabic	<i>āza</i>	<i>āk</i>
	<i>ayzi</i>	<i>ayk(i)</i>
	<i>ōzi</i>	<i>ōk</i>
Cypriot Arabic	<i>ada</i>	<i>adak</i>
	<i>adi</i>	<i>adik</i>
	<i>alli</i>	<i>allik</i>

Tab. 160.11: Demonstratives in the Arabic dialects

monstratives of the near and the far deixis, for instance in the dialect of Ġiblah (Yemen) *hādkāh*, *hādkih*, *hādkūh*; in the dialect of Bəh-zānī (Iraq) *hākā*, *haykā*, *hawkā* (cf. Fischer & Jastrow 1980, eds.: 116, 151); and in the dialect of Daragözü *ukkan*, *ukkanī*, *ukkanō* (cf. Jastrow 1973: 41).

##### 4.2. Interrogatives

The interrogatives of Classical Arabic *man* ‘who?’ and *mā* ‘what?’ have not been preserved in the majority of Arabic dialects. For ‘who?’ many dialects have a form that goes back to a non-attested form *mīn* (except for Moroccan which has *škun*, probably going back to a combination *'ayyu šay'in yakūn* ‘what thing is it?’); for ‘what?’ almost all dialects have opted for a reflex of *'ayyu šay'in (huwa)* ‘which thing?’, e.g., Moroccan *aš*, *šnu*, Egyptian *ē*, Chad Arabic *šinu*, Syrian *šū*, *ēš*.

The development of a new interrogative ‘what?’ illustrates a process of periphrasis that is particularly obvious in the interrogative adverbs. For ‘how much?’, for instance, many dialects have a form that goes back to a substantive meaning ‘form’ or ‘measure’, e.g., Syrian *addēš* (< \**qadr* *'ayyi šay'in* ‘which measure?’), Cypriot Arabic *áška* (< \**'ayyu šay'in qadr*). Likewise, in the interrogative adverb ‘how?’ we find Syrian *šlōn* (< \**'ayyu šay'in lawn* ‘which colour?’), Egyptian

*izzayy* (< \*'ayyu šay'in zayy 'which appearance?') and Cypriot Arabic *áššik* (< \*'ayyu šay'in šakl 'which form?').

These different realizations of the interrogative adverbs seem to point at a polygenetic origin in different contact situations. They all represent the product of an attempt at communication in which salient and semantically transparent forms are preferred over opaque ones such as Classical *kam* 'how much?' and *kayfa* 'how?'.

### 5. Pidginization and creolization in Arabic

There is one documented case of **pidginization** and subsequent **creolization** in Arabic (cf. also Art. 153). The structure of this dialect, (Ki-)Nubi, is an interesting example of a non-Indo-European based creole. Moreover, this dialect plays an important role in the discussions about the developmental processes in the central Arabic dialects (cf. Heine 1982; Versteegh 1984: 113–128; Prokosch 1986; Owens 1990). The origin of Nubi lies in the 19th century when a pidginized variety arose in the Anglo-Egyptian army camps in Upper Egypt as the principal means of communication between the Arabic-speaking subalterns and the multilingual recruits from the Southern Sudan, most of whom were speakers of Nubian. It is possible that in the formation of this pidgin elements from an earlier Arabic trade language in the region were used. Some of the soldiers followed the English army to Uganda and Kenya, where they settled and married indigenous women. The children who were born in these mixed marriages creolized the pidginized variety. Nowadays there are ±80,000 speakers of Nubi, all of them living in Uganda and Kenya. In the Southern Sudan the pidginized variety has remained in use as a **lingua franca** under the name of 'Juba Arabic'. In the cities it tends to become creolized by children from bilingual marriages (cf. Versteegh 1993).

Nubi exhibits the classic features of other creole languages (Wellens 2003):

- as in all creole languages there is a complete breakdown of the inflectional system of the original language; in the pidginized variety nouns had only one form for singular and plural, in Nubi there seems to be a new plural marker -á, which is, however, frequently omitted (cf. Heine

### XVIII. Morphologischer Wandel II: Fallstudien

- 1982: 29); there are a few fossilized broken plurals, e.g., *kááfír/kufará* 'infidel';
- adjectives have no inflection for gender or number, e.g., *mářia úe-dé kebír* 'this woman is big'; there seems to be a plural marker -ín, but this marker, too, is often omitted (Heine 1982: 33);
  - there is an analytical **collective** form, e.g., *nas-yalá* '(group of) children' (< *nás* 'people');
  - there are many compound nouns, e.g., *rásu-lakáta* 'tree-top' (lit. 'head + tree');
  - there is no reflex of the Arabic **article**, but the deictic form *dé* (< Egyptian *da*) has taken over some of the functions of the article, e.g., *lakáta dé kebír* 'the [above mentioned] tree is big'; a new series of demonstratives has developed: *úe-dé*, plural *dól-dé* for the near deixis; *náa-dé* for the far deixis;
  - the verb has only one form, diachronically connected with the 3rd plural of the perfect or with the imperative, e.g., *áalim* 'to teach' (< imperative *'allim?*), *ásurubu* 'to drink' (< imperative *iśrab?*); the ending -*u* seems to function as a **transitivity marker**, possibly derived from the 3rd person singular object suffix -*hu* (cf. Versteegh 1984: 124);
  - from verbs a verbal noun is derived by a change in stress (or possibly tone) pattern, e.g., *kárabú* 'to spoil', verbal noun *karáb*; this is probably an innovation that arose at the creolization stage;
  - there is a system of aspectual markers that is remarkably similar to that of other creole languages; at the same time it is not completely different from the system in central Arabic dialects (cf. Tab. 160.12).

---

#### 'to come'

---

<i>rúa</i>	neutral form
<i>bi rúa</i>	future
<i>kalà(s) rúa</i>	perfective
<i>káán rúa</i>	pluperfect
<i>gí rúa</i>	durative
<i>bi gí rúa</i>	future durative
<i>káán gí rúa</i>	past durative
<i>káán bi rúa</i>	hypothetical, unrealis

---

Tab. 160.12: Aspectual markers in Nubi (Heine 1982: 38)

In Juba Arabic exposure to the prestige variety of Khartoum Arabic has led to a change

in the verbal system. The personal prefixes of the imperfect verb in Khartoum Arabic are first re-interpreted as alternative aspectual markers; with continued contact they are assigned the function of indicating person (cf. Mahmud 1979). In this way, the dichotomy perfect/imperfect of the central Arabic dialects is re-instated in Juba Arabic.

## 6. References

- Ali, Abdul Sahib Mehdi (1987), *A Linguistic Study of the Development of Scientific Vocabulary in Standard Arabic*. London & New York: Kegan Paul International
- Behnstedt, Peter (1978), "Zur Dialektgeographie des Nildeltas". *Zeitschrift für arabische Linguistik* 1, 64–92
- Behnstedt, Peter (1985/1992), *Die nordjemenitischen Dialekte*, Vol. I: *Atlas*; Vol. II: *Glossar*. Wiesbaden: L. Reichert
- Behnstedt, Peter & Woidich, Manfred (1982–1999), *Die ägyptisch-arabischen Dialekte*, Vol. 1–5. Wiesbaden: L. Reichert
- Blanc, Haim (1964), *Communal Dialects in Baghdad*. Cambridge/MA: Harvard University Press
- Blanc, Haim (1970), "Dual and Pseudo-dual in the Arabic Dialects". *Language* 46, 42–57
- Blau, Joshua (1977), *The Beginnings of Diglossia: A Study of the Origin of Neo-Arabic*. Malibu: Undena
- Borg, Alexander (1985), *Cypriot Arabic*. Wiesbaden: F. Steiner
- Caubet, Dominique (1993), *L'arabe marocain*. Vol. 1–2. Paris, Louvain: Peters
- Cohen, David (1970), "Koinè, langues communes et dialectes arabes". In: Cohen, David (ed.), *Etudes de linguistique sémitique et arabe*. The Hague, Paris: Mouton, 105–125
- Corriente, Federico C. (1976), "From Old-Arabic to Classical Arabic through the Pre-Islamic Koine: Some Notes on the Native Grammarians' Sources, Attitudes and Goals". *Journal of Semitic Studies* 1, 62–98
- Corriente, Federico C. (1977), *A Grammatical Sketch of the Spanish Arabic Dialect Bundle*. Madrid: IHAC
- Czapkiewicz, Andrzej (1975), *The Verb in Modern Arabic Dialects as an Exponent of the Development Processes Occurring in Them*. Wrocław: Wydawnictwo Polskiej Akademii Nauk
- Denz, Adolf (1982), "Die Struktur des klassischen Arabisch". In: Fischer (ed.), 58–82
- Diem, Werner (1978), "Divergenz und Konvergenz im Arabischen". *Arabica* 25, 128–147
- Diem, Werner (1979), "Studien zur Frage des Substrats im Arabischen". *Der Islam* 56, 12–80
- Diem, Werner (1991), "Vom Altarabischen zum Neuarabischen: Ein neuer Ansatz". In: Kaye, Alan S. (ed.), *Semitic Studies in Honor of Wolf Leslau*, Vol. I. Wiesbaden: O. Harrassowitz, 297–308
- Eisele, John C. (1999), *Arabic Verbs in Time: Tense and Aspect in Cairene Arabic*. Wiesbaden: O. Harrassowitz
- Erwin, Wallace M. (1963), *A Short Reference Grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press
- Ferguson, Charles A. (1959 a), "Diglossia". *Word* 15, 325–340
- Ferguson, Charles A. (1959 b), "The Arabic Koine". *Language* 25, 616–630
- Fischer, Wolfdietrich (1959), *Die demonstrativen Bildungen der neuarabischen Dialekte: Ein Beitrag zur historischen Grammatik des Arabischen*. The Hague: Mouton
- Fischer, Wolfdietrich (1961), "Die Sprache der arabischen Sprachinsel in Uzbekistan". *Der Islam* 36, 232–263
- Fischer, Wolfdietrich (1982, ed.), *Grundriss der arabischen Philologie*, Vol. I: *Sprachwissenschaft*. Wiesbaden: L. Reichert
- Fischer, Wolfdietrich & Otto Jastrow (1980, eds.), *Handbuch der arabischen Dialekte*. Wiesbaden: O. Harrassowitz
- Grotfeld, Heinz (1965), *Syrisch-arabische Grammatik (Dialekt von Damaskus)*. Wiesbaden: O. Harrassowitz
- Hadj-Sadok, Mohammed (1955), "Dialectes arabes et francisation linguistique de l'Algérie". *Annales de l'Institut des Etudes Orientales d'Alger* 13, 61–97
- Harning, Kerstin E. (1980), *The Analytical Genitive in the Modern Arabic Dialects*. Göteborg: Acta Universitatis Gothoburgensis
- Harrell, Richard S. (1962), *A Short Reference Grammar of Moroccan Arabic*. Washington: Georgetown University Press
- Heath, Jeffrey (1989), *From Code-Switching to Borrowing: A Case Study of Moroccan Arabic*. London, New York: Kegan Paul International
- Heine, Bernd (1982), *The Nubi-Language of Kibera: An Arabic Creole*. Berlin: D. Reimer

- Holes, Clive (1990), *Gulf Arabic*. London, New York: Routledge
- Holes, Clive (1995), *Modern Arabic: Structures, Functions and Varieties*. London, New York: Longman
- Ingham, Bruce (1982), *North East Arabian Dialects*. London, New York: Kegan Paul International
- Ingham, Bruce (1994), *Najdi Arabic: Central Arabian*. Amsterdam, Philadelphia: J. Benjamins
- Janssens, Gerard (1972), *Stress in Arabic and Word Structure in the Modern Dialects*. Leuven: Peeters
- Jastrow, Otto (1973), *Daragözü: Eine arabische Mundart der Kozluk-Sason-Gruppe (Südostanatolien)*. Nürnberg: H. Carl
- Jastrow, Otto (1978), *Die Mesopotamisch-Arabischen Qəltu-Dialekte, Vol. I: Phonologie und Morphologie*. Wiesbaden: F. Steiner
- Jastrow, Otto (1982), "Die Struktur des Neu-Arabischen". In: Fischer (ed.), 128–141
- Mahmud, Ushari Ahmad (1979), *Variation and Change in the Aspectual System of Juba Arabic*. Ph.D. Dissertation, Georgetown University
- Mifsud, Manwel (1995), *Loan Verbs in Maltese: A Descriptive and Comparative Study*. Leiden: E. J. Brill
- Müller, Walter M. (1982), "Das Altarabische und das klassische Arabisch", "Das Altarabische der Inschriften aus vorislamischer Zeit". In: Fischer (ed.), 17–36
- Nöldeke, Theodor (1904), "Das klassische Arabisch und die arabischen Dialekte". In: Nöldeke, Theodor (ed.), *Beiträge zur semitischen Sprachwissenschaft*. Strassburg: K. J. Trübner, 1–14
- Owens, Jonathan (1990), "East African Nubi: Bio-program vs. Inheritance". *Diachronica* 7, 217–250
- Owens, Jonathan (1993), *A Grammar of Nigerian Arabic*. Wiesbaden: O. Harrassowitz
- Prochazka, Theodore (1988), *Saudi Arabian Dialects*. London, New York: Kegan Paul International
- Prokosch, Erich (1986), *Arabische Kontaktsprachen (Pidgin- und Kreolsprachen) in Afrika*. Graz: Institut der Sprachwissenschaft der Universität Graz
- Rabin, Chaim (1951), *Ancient West-Arabian*. London: Taylor's Foreign Press
- Retsö, Jan (1983), *The Finite Passive Voice in Modern Arabic Dialects*. Göteborg: Acta Universitatis Gothoburgensis
- Robin, Christian (1992), *L'Arabie antique de Karib'il à Mahomet: Nouvelles données sur l'histoire des Arabes grâce aux inscriptions*. Aix-en-Provence: Editions Edisud
- Rosenhouse, Judith (1984), *The Bedouin Arabic Dialects*. Wiesbaden: O. Harrassowitz
- Versteegh, Kees (1984), *Pidginization and Creolization: The Case of Arabic*. Amsterdam: J. Benjamins
- Versteegh, Kees (1993), "Leveling in the Sudan: From Arabic Creole to Arabic Dialect". *International Journal of the Sociology of Language* 99, 65–97
- Versteegh, Kees (1997), *The Arabic Language*. Edinburgh: Edinburgh University Press [2nd, revised edition, 2001]
- Watson, Janet C. E. (1993), *A Syntax of San'ānī Arabic*. Wiesbaden: O. Harrassowitz
- Wellens, Ineke (2003), *An Arabic Creole: The Nubi Language of Uganda*. Ph.D. diss., University of Nijmegen
- Woidich, Manfred (1990), *Ahlan wa sahan: Eine Einführung in die Kairoer Umgangssprache*. Wiesbaden: L. Reichert
- Zavadovskij, Jurij N. (1981), *Mavritanskij dialekt arabskogo jazyka (Chassaniia)*. Moscow: Nauka
- Zwettler, Michael (1978), *The Oral Tradition of Classical Arabic Poetry: Its Character and Implications*. Columbus: Ohio State University Press
- Kees Versteegh, Nijmegen (*The Netherlands*)

## 161. Tok Pisin

1. Pidgins and development
2. Historical background
3. History of research
4. Documentation
5. Internal history of Tok Pisin word formation
6. Theoretical consequences
7. References

### 1. Pidgins and development

Pidgin languages, because of their relatively short history and rapid development, are ideal test cases for general linguistics. Tok Pisin (from English *talk pidgin* also referred to as Neo-Melanesian or New Guinea Pidgin) is a particularly important example because

- (a) it is one of the very few languages whose entire linguistic development can be documented in some detail;
- (b) it illustrates two types of developments: that from simple jargon to complex creole and that of a conservative rural language to an anglicised urban one;
- (c) it illustrates both the spontaneous development of new morphology and the influence of deliberate acts of standardisation and planning.

While the development of inflectional morphology in Tok Pisin has been widely used (Sankoff 1977; 1980) in the debate about substratum influence and universals of linguistic development as well as the testing of other theoretical proposals, data from Tok Pisin derivational morphology have rarely been employed.

### 2. Historical background

Like the other two big Melanesian pidgins, Solomon Islands Pijin and Vanuatu Bislama, the history of Tok Pisin began outside its present area of currency. Between 1879 and 1914, thousands of New Guineans were recruited to work on the plantations of Samoa, as well as a much smaller number for a much shorter time (1880–1885) on the Queensland sugar plantations. In Samoa, Papua New Guineans picked up the plantation jargon that had developed among the multilingual workforce from Kiribati, Vanuatu and elsewhere and, from about 1890 onward, they also influenced this plantation language lexi-

cally and grammatically. Whilst it was a reasonably stable pidgin, Samoan Plantation Pidgin English remained a rather rudimentary language with no inflectional morphology, circumlocution and borrowing being the main sources of lexical expansion. As the Melanesian workers returned from Samoa to New Guinea and as a local plantation economy developed, the number of Tok Pisin speakers in former German New Guinea (which comprised of Northeastern New Guinea and the Bismarck Archipelago but not Papua) rapidly increased. From a few hundred in 1880, the numbers had grown to about 15,000 in 1914. Today there are about 2 million speakers of Tok Pisin, i. e. more than 50 percent of the population of independent Papua New Guinea. With the increase in speakers, Tok Pisin became more useful and was employed in a growing number of new situational contexts. It also began to acquire first language speakers. The main new contexts were (a) the gold industry from 1930 onwards, (b) Christianisation and missionisation, also from about 1930 onwards, (c) extension of Australian administrative control from 1920 onwards and, most important of all, (d) increasing use of Tok Pisin as the language of inter-tribal communication, replacing previously established local pidgins, bilingualism or silent barter.

During the Second World War Tok Pisin was used extensively in the information campaigns of the allied forces and the Japanese. After 1945 a Tok Pisin-medium press was set up and it also became a medium of education (until its condemnation by the United Nations in 1953). In spite of negative official policies Tok Pisin developed into the *de facto* language of village councils, new political parties in pre-independence days and the language of political self determination after independence. In the first seventy years of its existence intertribal marriages were rare, but, after 1945, new non-traditional settlements developed (mission stations, inter-tribal cargo cult centres, boarding schools and towns), significantly increasing the number of intertribal marriages. **Creolisation** of Tok Pisin began around 1950 and this process has accelerated greatly since. In most large urban areas, children have grown up with Tok Pisin as their first language.

### 3. History of research

The study of the morphology of Tok Pisin is a very recent phenomenon, mainly because until very recently there was little morphology to study. Early missionary grammars (surveyed by Mühlhäusler 1985 a) contain occasional references to apparent patterns of morphology, and in Hall (1943) a more detailed statement of zero derivation is given. The absence of productive reduplication is noted, an observation that was confirmed subsequently in Laycock (1970) and Wurm (1971), both of whom identified some other productive word formation devices. It was only when investigators turned to data from speakers who had acquired Tok Pisin after the Second World War that larger areas of morphological regularities became evident. Thus, the beginnings of verbal inflections are noted in Sankoff (1977), and in Mühlhäusler (1979) a dramatic increase in derivational productivity in the early 1970s is documented. Coinciding with political independence, there was a rash of activity in the area of lexical planning (surveyed by Wurm et al. 1977) and some detailed studies of recent developments in Tok Pisin morphology have been prepared by Romaine (1990). Smith (2002) contains valuable information on recent trends in creolised Tok Pisin and the continuing uncertainty about its social position and grammatical standards.

### 4. Documentation

The emergence of inflectional categories is relatively easy to document because of their high token frequency. There are, however, some problems for those who rely on the written text, including:

- (a) the failure of recorders to pick up phonologically reduced forms;
- (b) the lack of clear indications of what is an affix and what is an independent form (e.g. should *wanfela botol* be analyzed as 'one fellow bottle' or not);
- (c) the emergence of norms for writing, (e.g. *mipela* 'we') can conceal new developments, such as the morphologically reduced *mitla*.

The development of derivational morphology is much more difficult to document, as many lexical items are of a very low token frequency. However, research on Tok Pisin is in the fortunate position of being able to draw

on extensive lexicographical work by German missionaries. In my own studies I have used these to compile a very large corpus, supplemented by observational data and questionnaires.

An important consideration in all this is the locus of lexical regularities, as with a rapidly developing language such as Tok Pisin one cannot assume that a complex morphological form is the result of the application of any rule at all. For instance, for some speakers the compound adjective *bolhet* 'bald' forms part of the pattern:

- (1) *bikhet* 'conceited'  
*guthet* 'intelligent'

i.e. a pattern *someone i gat* adjective noun > adjective noun. For others, it is interpreted as *het i olsem bol* 'the head is like a ball', just like adjectives of the type *sakman* 'the man is like a shark (sharp customer)'. For yet others *bolhet* may simply be an unanalysed simple lexical entry.

It has been found that speakers who learnt Tok Pisin before World War II do not have the same number of productive patterns as younger speakers. It has also been observed that they will produce fewer morphologically complex forms and, in some instances, Melanesian-type forms produced by older speakers are being replaced by more English derived ones among younger speakers, as in *kot ren* 'rain coat' being replaced by *renkot*. To what extent such clear differences in grammar affect the communication is not understood. However, one suspects that the postulation of a common code is not a satisfactory model for a dynamic system such as Tok Pisin.

### 5. Internal history of Tok Pisin word formation

The internal history of Tok Pisin is most conveniently portrayed by referring to a number of idealised stages. It should be kept in mind that the processes are at least partly continuous and that at most times speakers of different developmental stages can interact. The stages are:

	first found
(a) Jargon	1870
(b) Stabilisation	1890
(c) Expansion	1925
(d) Creolisation	1950
(e) Post pidgin/post creole	1960

### 5.1. The Jargon stage

**Jargons** are distinguished from proper **pidgin** languages by their excessive instability, extremely impoverished expressive power and high context dependence. As they are individual rather than social solutions to the problem of crosslinguistic communication, it would seem to make little sense to postulate ideal speaker/hearers. Rather, one is to assume that interaction is characterised by different communication strategies and resulting misunderstandings and failures to communicate. At this stage, speakers of the **lexifier language** English (i. e. the language which provided most of the lexical roots) continue to use their own inflectional morphology, though they also make unsystematic simplifications characteristic of foreigner talk. Local speakers tend not to have any productive inflectional patterns, though fossilised inflections are in evidence, such as English plural -s in:

- (2) *masis* ‘match’  
*anis* ‘ant’

or the genitive affix -s in

- (3) *boskru* ‘boat’s crew’

the German plural -en in

- (4) *binen* ‘bee’

and the German infinitive -en in

- (5) *beten* ‘pray’  
*singen* ‘sing’

Some of these endings have since found their way into the grammar of Tok Pisin.

The derivational morphology of English is similarly eroded in the jargon stage, and morphologically complex forms typically end up with simple roots, as in the case of

- (6) *bilinat* ‘betelnut’  
*trausel* ‘tortoise shell’  
*katuana* ‘guard of honour’

Differences between indigenes’ and expatriates’ perceptions of morphological relatedness of words persist for long periods of time. As late as 1970 one could note differences such as:

expatriate	indigenous	gloss
<i>go – go ap</i>	<i>go – kwap</i>	‘go-climb’
<i>daun-daun</i>	<i>daun-tambilō</i>	‘down-hold’
<i>bilo</i>	–	‘of a ship’

Tab. 161.1: Expatriates’ and indigenes’ words

The jargon phase is characterised both by a very reduced and mostly nominal lexicon, and considerable pressure on speakers to say more than the language can express adequately.

In fully developed languages there tends to be constructional **iconicity** (cf. Art. 30) of the type:

- (a) all central concepts to be expressed by lexical stems;
- (b) all complex concepts to be expressed by complex or derived forms;
- (c) all percepts and other singularities to be expressed by phrases, sentences or other syntactic means.

In early **jargon** texts, one encounters many examples of speakers/hearers resorting to either non-verbal means of expression (e. g. by pointing to the chest to indicate ‘chest’, an item for which there is no word, but one that was used frequently as chest infections were common in the plantation context), or else, the formation of *ad hoc* circumlocutions, though this device was limited by the absence of syntactic regularities and lexical items useful in defining other items. Some early examples include:

- (7) *kas bilong pulimap wara*  
‘cask for filling with water (barrel)’  
*bokis bilong man i dai*  
‘box of person who died (coffin)’  
*wara bilong maus*  
‘water of mouth (saliva)’  
*lata bilong haus*  
‘ladder of house (stairs)’

Pointing and circumlocution were individual *ad hoc* solutions and most of these were probably unrecorded.

### 5.2. The stabilisation stage

**Stabilisation** of a **pidgin** language refers to its becoming a linguistic system separate from its lexifier languages, and the first language of its users. Stabilisation in the lexical component is manifested (a) in the development of socially sanctioned conventions about lexical information, (b) the reappearance of lexical structure, and (c) the use of stable syntax as a means of deriving new lexical items. Although additions to the lexical inventory do occur, they filter slowly into the language without causing severe structural disruption or major changes in lexical organisation. The main function of lexical additions is to con-

solidate the pidgin as a means of communication in a relatively narrow field of interlingual contacts.

The stabilisation of Tok Pisin was not a sudden event but a continuing process, starting with the stabilisation of Samoan Plantation Pidgin on the Samoan plantations and culminating in the crystallisation of a flexible and stable **lingua franca** spoken throughout German New Guinea. The lexicon which developed during this period derived its separate identity from the impact of Tolai and German and, to a lesser degree, Malay and Samoan, which replaced English as Tok Pisin's principal **lexifier languages** between 1880 and 1920.

The effect of stabilisation on inflectional morphology was relatively small. It manifested itself mainly in word class marking and number marking in pronouns.

The form *-pela* which occurred in a number of jargons variably as a kind of adjective marker or noun classifier became a marker of attributive, but not predicative monosyllabic adjectives in stable Tok Pisin as in:

- (8) *gutpela man* 'a good fellow, a good man'  
*smolpela pik* 'a small pig'  
*bikpela tais* 'a big swamp'

Similarly, *-im* (from English *him*) became a marker of transitive verbs as in:

- (9) *lukim* 'to see'  
*tanim* 'to turn'  
*ritim* 'to read'

An incipient marking of intransitives by means of **reduplication** as in Tolai, the language spoken around Rabaul, the former capital, remained unsuccessful and is now found only in a handful of items including:

- (10) *waswas* 'bathe'  
*wasim* 'wash someone'  
*singsing* 'dance'  
*singim* 'sing a song'  
*lukluk* 'watch'  
*hukim* 'see someone'

Pronominal number marking in Tok Pisin is the result of an encounter of two solutions, the Samoan Plantation Pidgin one using *ol* after pronouns as in:

- (11) *me ol* 'we (excl. & incl.)'  
*you ol* 'you (pl.)'  
*(em)ol* 'they'

and a different tradition employing *-pela* for this purpose.

Most present-day variants of Tok Pisin employ

- (12) *mi pela* 'we (excl.)'  
*you pela* 'you'  
*(em)ol* 'they'

*Ol* rather than *-pela* was subsequently adopted as a nominal plural marker (see Mühlhäusler 1981).

Tense, aspect and modality in stable Tok Pisin continued to be marked by fully stressed adverbials and auxiliaries. Developments of word formation devices were more spectacular, suggesting that the expansion of the lexicon was more important than stylistic flexibility or syntactic refinements.

By the time German control came to an end in 1919, the lexical inventory of Tok Pisin comprised perhaps 750–1,000 items, most of them lexical bases. Throughout the stabilisation stage, borrowing from outside sources remained the most important means of increasing Tok Pisin's referential adequacy. However, in contrast with the lexicon of Jargon English, that of stabilised Samoan Plantation Pidgin and Tok Pisin was no longer an unstructured list. Instead, one can observe the development of lexical structures manifested in (a) the emergence of tightly structured semantic fields; (b) the use of stable syntax in circumlocutions; and (c) the development of a small number of regularities of lexical derivation.

With regard to the last category, most complex forms were still phrase level items as:

(a) names for different kinds of people:

- (13) *man bilong pekato* 'a sinner'  
*man bilong kamda* 'a carpenter'  
*man bilong stia* 'a helmsman'  
*man bilong limlimbur* 'an idler'

(b) names for certain kinds of buildings:

- (14) *haus kuk* 'kitchen'  
*haus boi* 'hut for labourers'  
*haus kot* 'court house'  
*haus pekpek* 'toilet'  
*haus marit* 'married quarters'

(c) distinction of sex in humans and animals:

- (15) *hos man* 'stallion'  
*hos meri* 'mare'  
*pikinini man* 'young boy'  
*pikinini meri* 'young girl'  
*dok man* 'male dog'

<i>dok meri</i>	'bitch'
<i>pik man</i>	'boar'
<i>pik meri</i>	'sow'
<i>kakaruk man</i>	'cock'
<i>karakaruk meri</i>	'hen'

(d) antonyms of adjectives and intransitive verbs:

(16) <i>gut</i>	'good'
<i>nogut</i>	'bad'
<i>strong</i>	'strong'
<i>nostrong</i>	'feeble'
<i>stret</i>	'correct'
<i>nostret</i>	'incorrect, unacceptable'
<i>nap</i>	'sufficient'
<i>nonap</i>	'insufficient'

Another important innovation at this stage are conventions for changing the subcategorial status of words. Thus nouns referring to trees can also refer to the fruit from these trees, thus *bata* 'avocado tree' can also mean 'avocado pear', *saksak* 'sago palm' and 'sago', or *muli* 'orange tree' and 'orange'.

The name of a sound is derived from the name of the instrument with which it is produced. Examples are *tandok* 'horn trumpet' and 'signal for beginning or ending work'; *kundu* 'drum' and 'sound of a drum'; *visil*, *kaviv* 'flute' and 'sound of a flute', and so forth.

Next we can observe a third convention, which was less generally maintained than the first two: a noun referring to a material can at the same time refer to something typically made out of this material, e. g. *purpur* 'flower, purpur bush' and 'grass skirt', *ain* 'metal' and 'flat-iron, anvil', *kapa* 'shiny metal' and 'washing copper', *let* 'leather' and 'belt'.

Finally a very small number of causative verbs were formed by adding the transitivity marker *im* to intransitive verbs. As late as 1940 observers remark on the absence of word formation in Tok Pisin. There are certainly few signs of the sometimes allegedly vigorous role of reduplication in early pidgin development.

### 5.3. The expansion stage

The realisation that pidgins can be classified not only in terms of their social functions, but also in terms of their structural complexity, is a relatively recent one, stemming from the work of Samarin (1971), Mühlhäusler (1974), and Todd (1974). It was found that, as pidgins are used in new functions, their linguistic complexity also increases. Tok

Pisin's structural expansion appears to have taken place at a time when the number of first language speakers was diminutive. Whilst these few first language speakers may have contributed to its structural development, it appears that the bulk of innovations originated among second language speakers. Thus, the view that significant structural expansion of a pidgin only occurs with its creolisation is not borne out.

Tok Pisin's structural expansion after 1920 is defined by two main developments, namely the emergence of sentence embedding and discourse structure and the development of a derivational lexicon. Developments in the area of inflectional morphology remained modest, as phonological reduction and consequent clitisation, or affixation, is not favoured by second language speakers of Tok Pisin. As the age at which Tok Pisin was learnt decreased progressively from 1920 onwards, some of the processes usually associated with creolisation are found among fluent younger speakers of Tok Pisin, particularly those that have grown up in the towns. Derivational morphology on the other hand experienced a dramatic expansion during this phase. This asymmetry of development in the inflectional and derivational areas is an index of the relative unimportance of stylistic flexibility and importance of referential adequacy in a pidgin language.

The following general tendencies in the growth of the derivational lexicon have been observed:

- (a) a continued increase in the number of productive lexical regularities;
- (b) an increase in the distance between lexical structures and syntactic structures, in particular a tendency for lexical items to shift down from higher to lower lexical size levels;
- (c) an increase in the numbers of lexical items accounted for by individual lexical regularities, i. e. increase in productivity.

Thus, instead of the handful of word formation patterns found in 1920, 57 such patterns were identified by Mühlhäusler (1979), involving **conversion**, **compounding** and **reduplication**, enabling speakers to create hundreds of new morphologically complex words. This increase in productivity of the lexicon appears to have been governed by a number of principles, particularly those of gradual downshifting and an increase in derivational complexity.

Downshifting, i. e. the shift from phrase to word level lexical items can be illustrated with examples such as:

1945	present day Tok Pisin
<i>lam wokabaut</i> 'hurricane lantern'	<i>wokabautlam</i>
<i>manki bilong masta</i> 'servant (male)'	<i>mankimasta</i>
<i>mekim hariap</i> 'to speed up someone'	<i>hariapim</i>
<i>hatpela wara</i> 'soup, hot water'	<i>hatwara</i>
<i>mani pepa</i> 'paper money'	<i>pepamani</i>
<i>wara bilong skin</i> 'sweat'	<i>skinwara</i>
<i>man bilong stil</i> 'thief'	<i>stilman</i>
<i>man save</i> 'knowledgeable person expert'	<i>saveman</i>

Tab. 161.2: Downshifting

This trend is continuing today. One consequence is the decreasing reliance of speakers on circumlocutions.

An interesting though insufficiently understood phenomenon is the creation of morphologically complex nonce words (cf. Downing 1977), e.g. *hetman* 'a ghost that was a head only' in a ghost story, rather than the lexicalised item *hetman* 'village chief'.

Increasing derivational complexity is illustrated by a loss of morphological transparency. For example, one can contrast the early emergence of cumulative compounds of the type:

- (17) *papamama* 'parents'  
*manmeri* 'people'

with the late appearance of:

- (18) *wantok* 'someone who speaks the same language'  
*wankaikai* 'mess mate'  
*wansmok* 'someone who shares a smoke'  
*wanmak* 'someone of the same size'  
*wanlain* 'someone working in the same labour line'

In addition, many derivational regularities have become considerably more productive over the years. Thus, no compounds ending

### XVIII. Morphologischer Wandel II: Fallstudien

in the agent suffix *-man* 'person, doer' are found before 1920 though lexical phrases of the form [man bilong Vint] expressing 'someone who usually does what is referred to by the verb' are documented in fair numbers for the mid 1920s, including:

- (19) *man bilong singaut* 'noisy person, beggar'  
*man bilong slip* 'sleepy, lazy person'  
*man bilong stil* 'thief'

The only word-level items at this point are *sutman* 'policeman', and *fulman* 'jester'. Most commonly nouns with an agentive meaning are fully lexicalised. Typical examples from an unpublished 1924 vocabulary by Leo Brenninkmeyer are *kuskus* 'clerk, writer', *ridima* 'redeemer', *kamda* 'carpenter' and *kundar* 'acolyte'.

For the mid-1930s the authors of a dictionary (*Wörterbuch* n. d.: 53) remark that "-man as the suffix of verbs forms agent nouns" (author's translation). However, only a few word-level items are listed:

- (20) *wasman* 'watchman'  
*sikman* 'patient'  
*daiman* 'dead, dying man'  
*stilman* 'thief'

Phrase-level items listed in the *Wörterbuch* (n. d.) include:

- (21) *man bilong toktok* 'talkative person'  
*man bilong save* 'wise, knowledgeable person'  
*man bilong pait* 'warrior, fighter'  
*man bilong pret* 'fearful person'

The only additional items culled from Kutscher (1940) are:

- (22) *rabisman* 'destitute person'  
*lesman* 'loafer'

It is interesting to note that Kutscher does not list 'thief' as *stilman*, as had been done in the *Wörterbuch* (1935) but as *man bilong stil*.

The three trends identified in the derivational morphology of expanded Tok Pisin continue to dominate the development of the next stage, creolisation.

#### 5.4. The creolisation stage

**Creolisation** (i. e. the language becoming adopted as a first language) in Tok Pisin began in the early 1950s but has only become a large scale phenomenon since urbanisation, from the 1970s onward. Today an estimated 50,000 children speak Tok Pisin as their first or main language, mainly in towns such as

Port Moresby, Lae or Madang. The main changes once Tok Pisin becomes a first language appear to lie in the area of inflectional morphology. Many of these changes are the result of the growing tempo at which Tok Pisin is now spoken and the consequent phonological reductions. As yet the underlying morphological structure of complex forms does not appear to have been affected, though quite significant deeper restructuring may occur within a few generations.

Generally speaking one can observe two trends:

- (a) that from lesser to greater morphological synthesis;
- (b) that from lesser to greater fusion.

These combined trends make for a dramatic increase in morphological complexity and result in communication problems between speakers of creolised and non-creolised varieties as well as problems with the existing spelling system.

Typical examples of such problems include possessive pronouns, indirect object pronouns and future/irrealis modality in verbs (cf. Romaine & Wright 1982; Mühlhäusler 1983; 1985 b).

In traditional second language varieties of Tok Pisin the pattern for expressing the possessive pronoun is:

- (23) *bilong mi* ‘my’  
*bilong yu* ‘your’  
*bilong em* ‘his/her’

and so on. In progressive creolised varieties these forms are reduced to *blomi*, *bloyu*, *blem*, and *blusat* (from *bilong husat*) ‘whose’.

Most speakers would produce such eroded forms as well as intermediate forms such as *belomi* and so on.

As regards the dative case of pronouns, in traditional varieties one finds:

- (24) *long mi* ‘to me’  
*long yu* ‘to you’  
*long em* ‘to him/her’

In creolised varieties these appear as *leme*, *leyu* and *lem* respectively. Note that in at least two examples there is a loss in regularity (instead of *bilong* as the possessive marker, we now have *blo* or *blu*, depending on what noun or pronoun follows) as well as a loss of transparency: *lem* is not as transparent and easy to recognise as *long em*.

A final example of the loss of morphological transparency is that of future/irrealis mo-

dality in verbs. Labov’s (1985) demonstration that **pidgins** and **creoles** develop verbal affixes out of free-standing adverbials or auxiliaries is borne out by evidence from Tok Pisin. Most commentators, e.g. Sankoff (1980), have postulated a developmental hierarchy such as:

- (a) *baimbai mi kam* ‘I shall come’, with a fully stressed sentence initial adverbial;
- (b) *bambai mi kam*, exhibiting phonological reduction;
- (c) *bai mi kam*, with a fully stressed monosyllabic adverbial;
- (d) *mi bai kam*, with a shift of the adverbial to a position in front of the main verb;
- (e) *mi bai kam*, with further phonological reduction and finally:
- (f) *mi bakam*, with cliticisation of the future marker.

Whilst examples of all of the above constructions can still be found, most speakers use these forms variably, but not necessarily in an implicity ordered manner. Further research is needed to determine what the role of creolisation is in the development of inflectional morphology.

The effects of **creolisation** on derivational morphology as characterised by Mühlhäusler (1979) are:

- (a) an increase in the productivity of existing patterns of word formation;
- (b) a greater relaxation of earlier restrictions on word formation rules, i.e. a move from context-sensitive to context-free;
- (c) a relaxation of the constraint on the recursive application of word formation rules.

These three combined trends make for a much more powerful derivational lexicon. Note that all of the above trends have already begun with fluent second language speakers of expanded pidgin. We are thus dealing with a quantitative rather than a qualitative difference.

I shall now turn to more specific examples which demonstrate these trends. Increased productivity of derivational morphology in Malabang creolised Tok Pisin can be illustrated with the occurrence of new abstract nominals derived from verb or adjective bases. For example:

- (25) *no ken gat sot bilong wara*  
‘there won’t be a shortage of water’  
*I gat planti harim bilong Tok Pisin*  
‘there are many dialects of Tok Pisin’

- maski mi gat sting nogut*  
 ‘although I have a bad infection’  
*i gat planti kain kolim bilongen*  
 ‘there are many different names for it’  
*rait bilong mi i no klia tumas*  
 ‘my writing is not very clear’  
*karim bilong dispela meri i hat tru*  
 ‘the act of childbirth is very painful for  
 this woman’  
*dilim bilong yu i no strett*  
 ‘your way of dealing (cards) is not cor-  
 rect’  
*em i lukim ron bilong olgeta pis*  
 ‘he looked at the movement of the fish’

A second instance of functional change is the direct derivation of **causative** verbals from verb or adjective bases by means of the affix *-im*, which in Malabang Tok Pisin appears to apply to all adjective or intransitive verb bases and to be fully recognised as a stylistic variant of the periphrastic construction [mekim N i Adj/V]:

- (26) *wara i stingim ol plang*  
 ‘the water makes the planks rot’  
*dispela kaikai i switim maus bilong mi*  
 ‘this food gives my mouth a pleasant  
 taste’  
*meri i bonim pikinini*  
 ‘the woman gave birth to a child’  
*em i wok long raunim diwai*  
 ‘he is busy making a piece of wood  
 round’  
*meri i smatim em yet*  
 ‘the girl dolled herself up’

A further example are verbs derived from nouns referring to instruments. In second language Tok Pisin these can be formed as long as they are not used as instruments for cutting, a contextual constraint that may have been introduced from Tolai. Whilst forms such as *akisim* ‘to cut with an axe’ and *naipim* ‘to cut with a knife, to stick a knife into’ are not documented for non-creole varieties, they occur freely in the Malabang creole of Manus Island.

Next, one finds the relaxation of derivational constraints on the formation of new word-level items. The disappearance of the constraint specifying that words should not consist of more than two morphemes can be seen in the following examples observed in spontaneous conversations:

- (27) *nek-tai-im*  
 ‘tie a necktie around someone’  
*han-kap-im*  
 ‘put handcuffs on’  
*kol-sisel-im*  
 ‘cut with a cold chisel’  
*pinis-taim-man*  
 ‘returned labourer’  
*bik-het-pasin*  
 ‘stubbornness’

More important, and hardly ever observed in non-creolised varieties of Tok Pisin, is the high incidence of **multiple derivation**, i. e. the operation of derivational processes on derived lexical items, as in:

- |                 |                 |                |
|-----------------|-----------------|----------------|
| (28) <i>huk</i> | > <i>huk</i>    | > <i>huk</i>   |
| ‘hook’          | ‘to go fishing’ | ‘fishing’      |
| <i>kuk</i>      | > <i>kukim</i>  | > <i>kukim</i> |
| ‘boil, cook’    | ‘cook           | ‘cooking       |
|                 | something’      | method’        |

### 5.5. The post-pidgin and post-creole stages

Post-pidgin and post-creole development, i. e. the language coming under heavy influence from English, can be dealt with together, since in the towns where these developments are encountered, the difference between first and second language varieties of Tok Pisin is not great, particularly among young speakers. Generally speaking, what one expects to find at this stage are the beginnings of a post-creole continuum. However, in the case of Tok Pisin, one cannot easily identify the operation of certain, generally accepted principles such as “new forms are adopted before new meanings”. Since there is a growing number of speakers who are bilingual in English and Tok Pisin and as access to the acrolect continues to increase with the introduction of new media such as T.V. and video, in at least some cases, contact with the acrolect has led not to a gradual approximation, but to new morphology distinct from that of the systems in contact.

As English is not a highly inflected language, the number of inflectional morphemes taken on by speakers of urban Tok Pisin remains limited. Most notable is the very considerable increase in the use of plural *-s*, replacing or supplementing the conservative *ol* + noun construction. My own data (Mühlhäusler 1981) suggest that the adaptation of the *-s* does not follow any known developmental path, such as the **animacy hierarchy**. However, other researchers have found evidence of such patterns emerging in the speech

of young urban speakers. Possessive *-s* and verbal agreement *-s* are absent in Tok Pisin as is the past tense *-ed*. A number of examples of *-ing* were found, particularly in written texts. They appear to be the result of a reinterpretation of the transitive/causative suffix *-im*, rather than a straight borrowing of the English continuous marker *-ing*: Consider the following examples from private letters:

- (29) *mi no man bilong spiking English*  
       'I do not speak English'  
       *ol meri i ken draiving kar*  
       'the women can drive a car'  
       *dispela toktok i helibing yu*  
       'this talk supports you'  
       *mi laik askim yu long wanpela samtim*  
       'I want to ask you for something'

Most writers would prefer *spikim*, *draivim*, *helipim*, and *samtng* instead of the highlighted forms. The effects of renewed massive contact with its original lexifier language has brought about some important changes in the derivational morphology of Tok Pisin. Generally speaking, the current preferred strategy of many urban speakers is to borrow a word from English rather than make use of existing derivational regularities. In some cases, existing complex items are being replaced by borrowed ones, with the resulting decline in the productivity of derivational morphology and an increase in suppletion.

The following table illustrates how borrowing has replaced previously existing regularities with suppletion as effect:

rural Tok Pisin	urban Tok Pisin	meaning
<i>bekim</i>	<i>bekim</i>	'answer'
<i>bekim</i>	<i>ansa</i>	'the answer'
<i>bungim</i>	<i>kolektim</i>	'gather'
<i>bung</i>	<i>bung</i>	'gathering place'
<i>hariap</i>	<i>hariap</i>	'hurry up'
<i>hariap</i>	<i>spit</i>	'speed'
<i>subim</i>	<i>subim</i>	'shove'
<i>subim</i>	<i>pos</i>	'force'
<i>tingtink</i>	<i>tingting</i>	'think'
<i>tingting</i>	<i>aidia</i>	'idea'
<i>hevi</i>	<i>hevi</i>	'heavy'
<i>hevi</i>	<i>wait</i>	'weight'
<i>peim</i>	<i>peim</i>	'pay'
<i>pe</i>	<i>wes</i>	'pay, wages'

Tab. 161.3: Suppletion as effect of borrowing

A second trend is the reordering of the elements of nominal compounds to make them conform more closely to the English model:

rural Tok Pisin	urban Tok Pisin	meaning
<i>tok ples</i>	<i>plestok</i>	'local language'
<i>sit bet</i>	<i>betsit</i>	'bed sheet'
<i>masta rot</i>	<i>rotmasta</i>	'road constructor'
<i>haus</i>	<i>morata-</i>	'house constructed of
<i>morata</i>	<i>haus</i>	'bush materials'
<i>gris pik</i>	<i>pikgris</i>	'pig grease, lard'

Tab. 161.4: Reordering in nominal compounds

As yet, no productive pattern directly resulting from English has been observed. However, if other creoles are anything to go by, then these will appear in Tok Pisin before too long.

The ultimate fate of Tok Pisin would not seem to be a linguistic matter. It will depend on the degree to which Tok Pisin is needed by its speakers. With the reemergence of regional *lingue franche* in the provinces and English in the towns, the role of Tok Pisin remains unclear.

## 6. Theoretical consequences

The evidence presented here raises a number of questions including:

- (a) the question of substratum influence;
- (b) the question of linguistic universals;
- (c) the question of structure and function.

These questions can only be addressed if proper attention is paid to the temporal dimension. Large scale changes in morphological power and productivity can be observed over a very brief period of time. Comparing arbitrarily selected points of this developmental continuum with the grammar of a particular substratum or superstratum language could be quite misleading because of the numerous changes the language has undergone. There is also the analytical problem of how one can identify morphological material across two or more languages.

As regards inflectional morphology, the sole successful example of **substratum influence** are a few (subsequently lexicalized) examples of nominal plurals being indicated by means of **reduplication**, as in Tolai:

- (30) *sipsip*      'sheep'  
       *meme*      'goat'

The bulk of the innovations in this area are the result of internal developments such as increasing tempo, reassignment of stress, re-interpretation of word boundaries and similar well known phenomena. These developments parallel those found in other creoles such as those discussed by Labov (1985). For derivational morphology, Mosel (1980) has prepared a painstaking comparison between the fifty-seven word formation patterns of Tok Pisin, identified by Mühlhäusler (1979), and those of Tolai. Mosel's findings include:

- (a) in the vast majority of cases (40 out of 57 patterns) there are no close parallels between Tok Pisin and its alleged substratum languages;
- (b) in many cases where there are parallels, other forces such as universals or reanalysis seem to have been at work simultaneously.

Not mentioned by Mosel but highly significant, is that many of the similarities between Tok Pisin and Tolai developed at a time when Tolai's importance as a lexifier language of Tok Pisin had already declined. In the foundation years of Tok Pisin (1880–1910) virtually no word formation was in evidence. There is another temporal or developmental factor that weakens substratum arguments, i.e. the observable fact that the productivity of derivational regularities is very low until well into the expansion phase of the language. This would seem to indicate that substratum languages contribute a few bases for analogical extension rather than productive lexical rules. There are numerous other ways in which derivational morphology can arise. These are discussed by Mühlhäusler (1979: 257–273). Among these, grammatical reanalysis is a particularly important process.

The question of **universals** arises mainly in connection with the sequence in which morphological regularities are added to the grammar of a language such as Tok Pisin. The presence of **substratum** or **superstratum** sources in itself is neither a necessary nor a sufficient condition for the development of new morphology. Rather, it appears that its growth is governed by language independent principles (such as causatives will be derived first from static verbs and all verbs before adjectives or nominals or the emergence of plural affixation will proceed along the accessibility and animacy hierarchies). Insufficient developmental data for other pidgins and

other types of second language acquisition are available to warrant firmer conclusions.

The presence of universal trends in the development of pidgin morphology needs to be regarded as independent of the question of the ontological status and sources of such universals (cf. Art. 118, 153). It certainly does not follow that the postulated hierarchies of morphological development are **biopgrammatic** (Bickerton 1981). Functional explanations for the development of pidgin morphology remain underexplored, though it is widely acknowledged that the structural expansion of these languages is a response to increased functional requirements.

The often heated debate about "**substratum** or **universals**" has failed to ask to what end such processes need to operate. As has been pointed out above, **pidgins**, like other languages, serve to express a number of concepts and generalities as well as percepts or singularities. The latter are expressed by syntactic means, for instance: *dispela naispela meri i sindaun long ret pela sia* 'this nice girl who is sitting in the red chair'.

For concepts, however, one needs lexical items such as:

- (31) *meri* 'woman, girl'

Typically, basic concepts are expressed by word bases. Complex concepts are expressed by morphologically complex lexical units such as in:

- (32) *singelmeri* 'single girl'  
*develmeri* 'spirit woman'  
*pamukmeri* 'prostitute'

It is the growing need for more complex concepts that has brought into being greater lexical power.

At the same time, because **pidgins** are second languages they need to remain learnable, and enriching the lexicon by borrowing more and more lexical bases from **substratum** or **superstratum** languages is therefore costly in terms of learnability. The emergence of lexical regularities of the type described in this article is an answer to this problem. This view is in harmony with the observation that pidgins begin not as simple languages, but as highly complex, heavily lexical and impoverished ones, and that simplicity in the sense of regularity or grammaticality emerges only in the process of their becoming referentially and expressively more adequate.

Many morphological questions still remain unanswered. A corpus of linguistic and func-

tional data at present being compiled in connection with the international project on “Languages of Intercultural Communication in the Pacific” (see Mühlhäusler 1986) is likely to make available information which can illuminate further the issues raised here.

## 7. References

- Bickerton, Derek (1981), *Roots of Language*. Ann Arbor: Karoma
- Downing, Pamela (1977), “On the Creation and Use of English Compound Nouns”. *Language* 53: 810–842
- Hall, Robert A. Jr. (1943), *Melanesian Pidgin English: Grammar, Texts, Vocabulary*. Baltimore: Linguistic Society of America
- Kutscher, P. (1940), *Wörterbuch Deutsch–Pidgin-English*. Vunapope: Rabaul
- Labov, William (1985), *On the Adequacy of Natural Languages*. Trier: Linguistic Agency, University of Trier
- Laycock, Donald C. (1970), *Materials in New Guinea Pidgin (Coastal and Lowlands)*. Canberra: Pacific Linguistics
- Mosel, Ulrike (1980), *Tolai and Tok Pisin: The Influence of the Substratum on the Development of New Guinea Pidgin*. Canberra: Pacific Linguistics
- Mühlhäusler, Peter (1974), *Pidginization and Simplification of Language*. Canberra: Pacific Linguistics
- Mühlhäusler, Peter (1979), *Growth and Structure of the Lexicon of New Guinea Pidgin*. Canberra: Pacific Linguistics
- Mühlhäusler, Peter (1981), “The Development of the Category of Number in Tok Pisin.” In: Muysken, Pieter (ed.) *Generative Studies on Creole Languages*. Dordrecht: Foris, 35–84
- Mühlhäusler, Peter (1983), “The Development of Word Formation in Tok Pisin”. *Folia Linguistica* 17, 463–487
- Mühlhäusler, Peter (1985 a), *History of the Study of Tok Pisin*. In: Wurm & Mühlhäusler (eds.), 15–33
- Mühlhäusler, Peter (1985 b), “Internal development of Tok Pisin”. In: Wurm & Mühlhäusler (eds.), 75–166
- Mühlhäusler, Peter (1986), “Pidgins and Creoles of Australia and the Pacific”. *Australian Journal of Linguistics* 6, 181–199
- Romaine, Suzanne (1990), “The Use of *bai* in Young Children’s Creolized Tok Pisin”. In: Verhaar, John (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: Benjamins, 187–203
- Romaine, Suzanne & Wright, Fiona (1982), “A Short Note on Short Forms in Tok Pisin”. *Journal of Pidgin & Creole Languages* 2, 64–67
- Samarin, William J. (1971), “Salient and Substantive Pidginization”. In: Hymes, Dell (ed.) *Pidginization and Creolization of Language*. Cambridge: Cambridge University Press, 117–140
- Sankoff, Gillian (1977), “Variability and Explanation in Language and Culture: Cliticization in New Guinea Tok Pisin”. In: Saville-Troike, Muriel (ed.), *Linguistics and Anthropology*. Washington, D.C.: Georgetown University Press, 59–73
- Sankoff, Gillian (1980), “Variations, Pidgins & Creoles”. In: Valdman, Albert & Highfield, Arnold (eds.) *Theoretical Orientations in Creole Studies*, New York: Academic Press, 139–164
- Smith, Geoff P. (2002), *Growing up with Tok Pisin*. London: Battlebridge Publications
- Todd, Loreto (1974), *Pidgins and Creoles*. London: Routledge & Kegan Paul (Language and Society Series)
- Wörterbuch mit Redewendungen Pidgin-Englisch–Deutsch*. (Pidgin-English–German dictionary, with idioms). (n. d.) [1935] Mimeo
- Wurm, Stephen A. (1971), *New Guinea Highlands Pidgin: Course Materials*. Canberra: Pacific Linguistics
- Wurm, Stephen A. & Mühlhäusler, Peter (eds. 1985), *Handbook of Tok Pisin*. Canberra: Pacific Linguistics
- Wurm, Stephen A. & Mühlhäusler, Peter & Laycock Donald C. (1977), “Language Planning and Engineering in Papua New Guinea”. *New Guinea Area Languages and Language Study* (Canberra: Pacific Linguistics) C-40, 1151–1177

Peter Mühlhäusler, Adelaide (Australia)

## XIX. Psycholinguistische Perspektiven

### Psycholinguistic perspectives

#### 162. Mentale Repräsentation morphologischer Strukturen

1. Mentale Repräsentationen
2. Das mentale Lexikon und das Lexikon der Grammatik
3. Beobachtungsbereiche und Forschungsmethoden
4. Einige Grundprobleme der Forschung
5. Zitierte Literatur

##### 1. Mentale Repräsentationen

Der Ausdruck "mentale Repräsentationen" bezeichnet ein Kernkonzept für die Frage nach der Struktur, dem Format und der Funktionsweise kognitiver Prozesse. In den **Kognitionswissenschaften** hat der Begriff unterschiedliche Ausformungen erhalten, die schon früh in der Geistesgeschichte nachweisbar sind (Scheerer 1992). Die beiden Grundpositionen des symbolverarbeitenden Ansatzes und des **Konnektionismus** sind dabei beide durch ihren Bezug auf Computerorganisation und Programmiersprachen gekennzeichnet. Im symbolverarbeitenden Ansatz wird die Manipulation von Symbolen als zentrale Aufgabe kognitiver Prozesse angesehen; mentale Repräsentationen sind durch Regelanwendung zustande gekommene syntaktisch wohlartikulierte und semantisch interpretierbare Verkettungen von Symbolen (vgl. u. a. Pylyshyn 1984). Die Grundeinheiten in den konnektionistischen Ansätzen sind dagegen subsymbolischer Natur: Es gibt nur Knoten (*nodes*) und ihre mannigfachen Verknüpfungen (*connections*); mentale Repräsentationen können sich nur als emergente Eigenschaften von Netzwerken insgesamt ergeben, und zwar als multivariate Aktivierungsmuster. Die Diskussion zwischen den beiden Grundpositionen ist derzeit nicht entschieden, vgl. die Reaktionen auf den expositorischen Artikel "On the proper treatment of connectionism" in der Zeitschrift *Behavioural and Brain Sciences* (Smolensky 1988). Dabei spielt sich die Debatte neuerdings vornehmlich auf dem Gebiet der Modellierung sprachverarbeitender

Prozesse ab: Der symbolverarbeitende Ansatz findet ja eine vorbildliche Ausformung in den Regeln und Repräsentationsebenen generativer Grammatiken.

Morphologie ist erst relativ spät Gegenstand psycholinguistischer Analysen geworden. Zunächst modellierte man auf der Basis linguistischer Beschreibungen (Halle 1973) den Worterkennungsvorgang als einen Zerlegungsprozeß, in dem zuerst die Bausteine des Wortes bestimmt werden, weil nur so auf das mentale Lexikon zugegriffen werden kann (s. 2); z. B. engl. *re-tell* besteht aus dem Präfix *re-* und dem Stamm *tell*, beides lexikalische Einträge, die durch eine Regel "Verb → Präfix plus Stamm" zu der komplexen Einheit *retell* führen (Taft & Forster 1975). Die Diskussion dieses in der Folgezeit vielfach modifizierten Ansatzes wird mit wenigen Ausnahmen auf der Grundlage der symbolverarbeitenden Position geführt (vgl. zusammenfassend Taft 1988 und Art. 163).

Erst in den letzten Jahren werden konnektionistische Modelle mit linguistischen Erklärungsansätzen für spezielle Erscheinungen wie **Kasussynkretismus** in Zusammenhang gebracht (Carstairs & Stemberger 1988) oder die Ähnlichkeit des alten Terminus **Analogie** (z. B. Paul<sup>5</sup> 1920; Art. 148) mit dem konnektionistischen Grundschema hervorgehoben (Derwing & Skousen 1989). Neuerdings ist der Erwerb der Flexionsmorphologie zu einem zentralen Gegenstand in der Konnektionismusdiskussion geworden. In Rumelhart & McClelland (1986) war gezeigt worden, daß **Übergeneralisierungen** vom Typ engl. *goed* (statt des suppletiven *went*), die seit langem (z. B. Berko 1958) als prototypische Evidenz für regelbasiertes Verhalten galten, auch in einem konnektionistischen Netzwerk ohne Regeln durch *back propagation* erzeugt werden können. Die Debatte darüber ist nicht abgeschlossen, u. a. weil nicht klar ist, in welcher Weise konnektionistische Systeme den un-

markierten Fall erkennen und als solchen lernen (vgl. u. a. Clahsen & Rothweiler 1992; Hare et al. 1995).

## 2. Das mentale Lexikon und das Lexikon der Grammatik

### 2.1. Das Lexikon in modernen Grammatiktheorien

In der Auffassung der neueren Grammatiktheorien erzeugt die Syntax reguläre abstrakte Strukturen, die durch die Einsetzung lexikalischer Elemente individuiert werden. Es ist das Geschäft der lexikalischen Komponente einer Grammatik, lexikalische Repräsentationen zu erzeugen; diese (und nicht die lexikalischen Minimaleinheiten) sind es, die in syntaktische Strukturen eingesetzt werden. Für diesen Zweck spezifiziert die morphologische Theorie den Begriff "mögliche Wortform" einer Sprache.

Dementsprechend wird von den meisten Linguisten angenommen, daß zumindest reguläre Flexionsformen wie etwa im Paradigma der schwachen Verben des Deutschen nicht jede einzeln im Lexikon der Grammatik aufzuführen sind, sondern durch eine Regel "Verbform = Verbstamm plus Endung" generiert werden. Lexikalische Einheiten sind also nur der Stamm, z. B. *sag-*, sowie die einzelnen Endungen, so daß bei der lexikalischen Einsetzung die grammatische Form "Verb + 2. Person Singular Präteritum Indikativ Aktiv" identifiziert werden kann mit, z. B., *sag- + -test*.

In neueren Wortstrukturtheorien wird ferner angenommen, daß auch die Wortbildung weitgehend analog zur Syntax organisiert ist (vgl. unter vielen anderen Olsen 1986). Deutsche Adjektivableitungen mit dem Suffix *-bar* oder dreiwertige deutsche Verben mit dem Präfix *be-* lassen sich informal so beschreiben:

Sei X ein deutsches transitives Verb. Dann ist X-*bar* ein Adjektiv mit der Bedeutung "kann (ge)-x-t werden", z. B. *hörbar*, *bewohnbar*, *ableitbar*, etc.

Sei X ein deutsches transitives dreiwertiges Verb mit Subjekt, direktem Objekt und direktonaler Präpositionalphrase. Dann ist *be-X-en* ein transitives dreiwertiges Verb gleicher Bedeutung mit Subjekt, Objekt und einer Präpositionalphrase mit der Präposition *mit*, wobei Objekt und Präpositionalphrase bei Basis

und Ableitung vertauscht sind, vgl. *belegen*, *bemalen*, *bepflanzen* etc.

Diese Strukturbeschreibungen treffen sowohl auf eine Anzahl existierender deutscher Wörter zu als auch auf Wörter, die wir gegenwärtig kaum in einem auch umfangreichen Lexikon finden wie z. B. *drückbar*, *zerrbar*, *ionisierbar* oder *bekugeln*, *betüpfeln*, *besprøyen*, wobei die Regeln auch kombiniert werden können wie in *betüpfelbar* oder (mit weiteren Regeln) *Unbetüpfelbarkeitssymptomatik*. Solche **lexikalischen Regeln** für die Generierung von Ableitungen und Zusammensetzungen spezifizieren lexikalische Repräsentationen, d. h. das, was beim Vorgang der lexikalischen Einsetzung in die durch die Syntax generierten Strukturen eingesetzt wird. Unabhängig vom konkreten Format, in dem derlei Regeln zu formulieren sind, erlaubt diese Konzeption die Aufrechterhaltung der klassischen Lexikonkonzeption, wonach konkrete Lexikoneinträge nur das enthalten, was nicht durch Regeln abgeleitet werden kann.

Dazu gehören aufgrund des **Arbitraritätsprinzips** zunächst einmal alle Morpheme einer Sprache. Daneben muß es freilich aufgrund des Phänomens der Lexikalisierung auch komplexe Lexikoneinträge geben. Dazu gehören z. B. Wörter, die bestimmte Affixe aufweisen, aber dennoch offenbar nicht unter die einschlägigen Regeln fallen, im Beispielbereich etwa Wörter mit unikalen Morphemen wie *begehren*, *urbar* sowie lexikalisierte Bildungen wie *befolgen*, *kostbar* etc. Ein besonders instruktives Beispiel ist das Wort *befehlen*. Obgleich es weder wortsyntaktisch noch semantisch irgendwie zu der o. a. Regel paßt, muß es als präfigiertes starkes Verb gekennzeichnet werden, denn das Partizip heißt nicht \**gebefehlt* oder \**gebefohlen*, sondern *befohlen*, eine für Präfixverben typische Bildungsweise (ohne das Partizipialaffix *ge-*). Die lexikalische Repräsentation von *befehlen* muß also einerseits die Struktur "be-fehl-en" spezifizieren, andererseits klar machen, daß *fehl-* hier nicht der gleiche Stamm wie z. B. in *Fehl-er* ist. Es gibt also im Lexikon der Grammatik Einträge für *befehl-*, *befolg-*, *begehr-* etc., da diese Wörter zwar als präfigiert anzusehen sind, aber nicht der o. a. *be-*-Regel folgen. Dagegen gibt es keinen Lexikoneintrag für z. B. *bepflanz-*; für dieses gibt es aber eine lexikalische Repräsentation, da die o. a. Regel auf *pflanz-* anwendbar ist, was in *bepflanzen* resultiert.

## 2.2. Das mentale Lexikon

Die Tatsache, daß die Wortform *sagtest* im Deutschen analysiert werden kann als bestehend aus dem Stamm *sag* und dem Affix *-test*, wobei letzteres die 2. Person Singular des Präteritums bezeichnet, und daß alle regulären deutschen Verben diese Form so bilden, besagt zunächst einmal nichts darüber, ob jemand, der diese Form in einem Satz äußern möchte, in seinem Gedächtnis erst den Stamm *sag* sucht, dann das Suffix für "2. Person Singular des Präteritums regulärer Verben" rauskramt, beide zusammenleimt und schließlich diese Gebilde produziert, oder ob ein Leser sie beim Lesen in ihre Teile zerlegen muß, bevor er sie verstehen kann (vgl. Günther 1988 b), denn "the existence of a logical system in the language does not necessarily imply that the mental structures that develop for dealing with this system will reflect this logic" (Sandra 1994: 245).

Ein zentrales Merkmal linguistischer Theoriebildung ist es ja, daß von der realen Sprachtätigkeit systematisch abgesehen wird, u. a. von Bedingungen der Produktion und Perzeption von Äußerungen sowie von ihrer Modalität. Diese aber spielen eine zentrale Rolle bei psycholinguistischen Fragestellungen. Das Konzept des mentalen Lexikons bezeichnet den bei der menschlichen Sprachverarbeitung beim Sprechen und Hören, Lesen und Schreiben benutzten Speicher an sprachlichen Elementen. Der Ausdruck ist eine Metapher, wonach der Speicher sprachlicher Elemente im Gedächtnis nach Art eines Lexikons organisiert ist; eine gute Zusammenfassung des Forschungsstandes findet sich in Aitchison (1994). Lexikalischer Zugriff (*lexical access*) besteht in der Identifikation eines Reizes mit einer gespeicherten lexikalischen Einheit und ist definiert als derjenige Moment, in dem z. B. beim Lesen die auf dem Papier stehenden Buchstabenfolge *WORT* als das Wort *Wort* im Gedächtnis identifiziert wird. Nach den meisten Konzeptionen stehen in just diesem Moment schlagartig sämtliche Informationen zu dieser Lexikoneinheit zur Verfügung, also seine Bedeutung(en), verwandte Wörter, seine Aussprache, seine grammatischen Eigenschaften (Geschlecht, Flexion, Valenz), etc.

Wie sind morphologisch komplexe Wörter im mentalen Lexikon gespeichert, und wie erfolgt lexikalischer Zugriff: auf morphologisch komplexe Wörter direkt, oder muß das Wort zuerst morphologisch zerlegt werden? Die Konsequenz aus beiden Ansätzen ist deutlich:

Nach der oben schon erwähnten sog. **Dekompositionshypothese** enthält das Lexikon nur Stämme, Affixe und Wortbildungsregeln, und es muß immer erst zerlegt werden (Taft & Forster 1975); nach der Theorie vollständiger Einträge (sog. *full listing hypothesis*, Butterworth 1983) kann direkt auf die vollständig gespeicherten komplexen Wörter zugegriffen werden. Beide Ansätze sind dem symbolverarbeitenden Ansatz verpflichtet. Alternative Vorstellungen, die seit dem Beginn der 80er Jahre im konnektionistischen Umfeld publiziert worden sind, gehen dagegen von einem allgemeinen Erregungsgeschehen aus, in dem der magische Moment des lexikalischen Zugriffs als solcher entfällt; vielmehr ergibt er sich in einem ständig sich wandelnden Aktivationsgeschehen.

## 3. Beobachtungsbereiche und Forschungsmethoden

Der Frage nach der mentalen Repräsentation und kognitiven Verarbeitung morphologisch komplexer sprachlicher Einheiten wird auf verschiedene Weise nachgegangen. Zusammenhängende Theorien lassen sich nur durch die Integration einer Vielzahl von Befunden aus verschiedenen Beobachtungsbereichen formulieren; dies sind u. a. linguistische morphologische Theorien, die Beobachtung und Analyse von Sprechfehlern und Sprachstörungen (Aphasien), die Beobachtung und Analyse des Spracherwerbs und seiner Störungen sowie die experimentelle Untersuchung der Produktion und Perzeption von Wörtern. Eine Darstellung von Theorien und Befunden in den genannten psycholinguistischen Bereichen geben die fünf folgenden Artikel (Art. 163–167). Es soll deshalb hier nur die jeweilige Forschungslogik gekennzeichnet und dazu eine Beschreibung der wichtigsten experimentellen Methoden gegeben werden.

### 3.1. Beobachtungsbereiche

Die Analyse von **Sprechfehlern** gehört seit dem Ende des 19. Jahrhunderts zu den wichtigsten Quellen psycholinguistischer Theoriebildung aufgrund der schon in einer klassischen Studie vorgetragenen Erkenntnis, "daß man sich nicht regellos verspricht ... (die Sprechfehler) müssen durch konstante psychische Kräfte bedingt sein" (Meringer & Mayer 1895: 9). Weil Sprechfehler strukturiert sind und klassifiziert werden können, lassen sie Rückschlüsse auf den normalen

Sprachproduktionsprozeß zu. Sprechfehler, in denen morphologische Elemente falsch angeordnet sind, vgl. *wenn ein begründeter Bedacht versteht, im Wolken nordig oder er hat mich zum Skileiden eingelaufen* (Beispiele aus Levelt 1989 und Leuninger 1993) werden als Beleg dafür angesehen, daß bei der Sprachproduktion komplexe Wörter aus den Morphemen zusammengesetzt werden. Für eine systematische Aufarbeitung der Sprechfehlerliteratur vgl. Levelt (1989); eine sehr knappe Zusammenfassung theoretischer Grundüberlegungen findet sich in Leuninger (1993) (s. auch Art. 164).

Auch die Analyse von **Sprachstörungen** folgt der Logik, durch die Erklärung der Störungsursache den nicht direkt beobachtbaren Sprachmechanismus zu modellieren (s. Art. 167). Dabei werden spezifische sprachliche Ausfälle zum Beleg bestimmter theoretischer Konzepte der psycholinguistischen Morphologieforschung herangezogen. Das Phänomen des **Agrammatismus** hat besondere Aufmerksamkeit gefunden, weil ein Aspekt der Symptomatik, das Weglassen von Flexionsaffixen, diese aphasischen Störung *prima facie* als eine flexionsmorphologische Störung zu kennzeichnen scheint (Goodglass & Berko 1960). Einer sehr vielbeachteten Theorie zufolge handelt es sich aber um eine phonologisch bedingte Störung (Kean 1979). Neuere Arbeiten sehen eher eine syntaktische Störung, d. h. die morphologischen Ausfälle sind Folge und nicht Ursache des Agrammatismus (z. B. Grodzinsky 1990; Art. 167). Ein anderes Syndrom ist die sog. phonologische **Dyslexie** (die Patienten können u. a. Wörter lesen, aber keine Pseudowörter). Caramazza et al. (1988) interpretieren die Ergebnisse einer Untersuchung solcher Patienten als Beleg für ihren Ansatz eines Vollformenlexikons, in dem Wörter morphologisch zerlegt sind (s. 4.1).

In der **Spracherwerbsforschung** gehört die Entwicklung der Morphologie zu den zentralen Themen. In allen Sprachen der Welt beginnen Kinder mit sog. Einwortsätzen zu sprechen, d. h. unflektierten Einzelwörtern. Flexionsformen werden in einer gewissen Reihenfolge erworben, wobei die Komplexität der signalisierten Bedeutung und die des sprachspezifischen Formensystems, die Eindeutigkeit der Laut-Bedeutungs-Zuordnung und andere Faktoren eine Rolle spielen (Slobin 1985, Hrsg.; Art. 165). Die Fähigkeit von Kindern, erworbene Flexionsregularitäten zu (über)generalisieren und auch auf unpas-

sende Stämme sowie auf Nonsensewörter anzuwenden (z. B. *singte, bringte, beste* statt *sang, brachte, fegte* – Berko 1958, ein Klassiker der psycholinguistischen Morphologieliteratur) wird in der Regel als Beleg für eine morphologische Dekomposition im o. g. Sinne interpretiert. Die Fähigkeit zur Wortbildung durch Derivation bildet sich später heraus; auch hier kommt es zu Übergeneralisierungen in Fällen, in denen in der Erwachsenensprache ein bestimmter anderer Ausdruck konventionalisiert ist (wie das o. g. *er beste* statt *fegte*). Unklar ist, welche Relevanz Erwerbsdaten für die Modellierung mentaler Repräsentationen und den Zugriff auf morphologisch komplexe Wörter beim Erwachsenen haben (vgl. Sandra 1994: 260–264).

Die experimentelle Untersuchung der Produktion und Perzeption morphologisch komplexer Wörter hat seit der initialen Studie vor 20 Jahren (Taft & Forster 1975) einen stürmischen Aufschwung genommen. Denn eine psychologische oder psycholinguistische Theorie kann nur in dem Maße als validiert angesehen werden, als sie empirisch – und das heißt experimentell – bestätigt ist; vgl. zu den methodischen Grundlagen Günther (1989). Eine Zusammenfassung der Befunde und Argumentationen in der älteren experimentellen Morphologieforschung bietet Henderson (1985); eine problemorientierte neuere Darstellung findet sich bei Sandra (1994) (s. auch Art. 163). Im folgenden sollen die wichtigsten Untersuchungsmethoden aufgezählt werden.

### 3.2. Methoden

In der psycholinguistischen Morphologieforschung stehen bislang Experimente zur visuellen Worterkennung im Vordergrund. In solchen Experimenten wird der Versuchsperson ein einzelnes Wort visuell dargeboten. Bei den in der älteren Literatur beschriebenen Experimenten bedient man sich eines Tachistoskops. Mithilfe dieses Geräts ist es möglich, visuelle Reize beliebig kurz darzubieten. Die Aufgabe der Versuchsperson besteht darin, den nur als kurzen Lichtblitz gezeigten Reiz (schriftlich oder mündlich) zu benennen. Ausgewertet wird meist entweder die Fehlerzahl oder die Anzahl wiederholter Darbietungen, die nötig ist, um den Reiz korrekt zu identifizieren. Diese Methode wurde z. B. bei der Untersuchung der Verarbeitung englischer flektierter Wortformen angewandt (Gibson & Guinet 1971). Eine andere Aufgabenstellung besteht darin, gleichzeitig zwei

Reize darzubieten und danach zu fragen, ob diese gleich oder verschieden sind. Sie fand z. B. bei der Untersuchung starker und schwächer Verbflection Verwendung (Jarvella & Snodgrass 1974). Ebenfalls häufig ist die Messung der Zeit, die eine Versuchsperson braucht, um ein unter verschiedenen Bedingungen gezeigtes Wort auszusprechen (z. B. Schreuder et al. 1989).

Die überwiegende Mehrzahl der in der neueren wissenschaftlichen Literatur publizierten Experimente zur Verarbeitung morphologisch komplexer Wörter aber bedient sich der Technik der **lexikalischen Entscheidung**. Hier bekommt die Versuchsperson auf einem Bildschirm eine Buchstabenfolge dargeboten und hat die Aufgabe, so schnell wie möglich zu entscheiden, ob diese Folge ein Wort ist oder nicht; gemessen wird die Reaktionszeit. Die Popularität dieser Aufgabenstellung in den letzten 20 Jahren hat mit dem Konstrukt des lexikalischen Zugriffs zu tun. Zur Untersuchung der Vorgänge beim lexikalischen Zugriff können im Grunde nur solche Aufgaben herangezogen werden, bei denen man sicher sein kann, daß auch tatsächlich lexikalischer Zugriff erfolgt (Coltheart 1978). Diese Sicherheit gibt es z. B. nicht, wenn die Versuchsperson die Aufgabe hat, von einem Paar von Wörtern zu sagen, ob diese gleich oder verschieden sind: Das kann sie tun, ohne ihr mentales Lexikon zu konsultieren, indem sie einfach Buchstaben vergleicht. Anders bei der lexikalischen Entscheidung, jedenfalls beim Einhalten bestimmter Randbedingungen: Daß die Buchstabenfolge *FEST* ein deutsches Wort ist, *DEST* dagegen nicht, kann man nur entscheiden, wenn man *FEST* im mentalen Lexikon gefunden hat, da die Buchstabenfolge *DEST* durchaus ein deutsches Wort sein könnte. Es ist typisch für die lexikalische Entscheidungsaufgabe, daß die Reaktionszeiten auf Wörter kürzer sind als die auf sog. Pseudowörter wie *DEST*, die orthographisch und phonologisch möglich, aber in der Sprache nicht realisiert sind.

Die Logik solcher Experimente besteht darin, die Reaktionszeit auf zwei verschiedene Typen von Wörtern miteinander zu vergleichen, von denen die Verarbeitung des einen Typs durch einen morphologischen Zerlegungsprozeß nach der Theorie behindert /gefördert würde, die des anderen aber nicht. Ein Beispiel bieten sog. pseudopräfigierte Wörter: Das bekannteste englische Beispiel ist *sublime*, ein deutsches Beispiel wäre *gering*. Wenn man annimmt, daß automatisch

zerlegt wird, so erhält man mit *sub-lime* (bzw. *ge-ring*) neben dem **Pseudopräfix** einen Bestandteil, der zwar im mentalen Lexikon aufzufinden ist (*lime*, *ring*), aber nicht Stamm des gezeigten Wortes ist; es muß reanalyisiert werden, wodurch es zu Reaktionszeit-Verlangsamungen im Vergleich zu "echt" präfigierten Wörtern kommt (Taft 1981). Eine quasi umgekehrte Logik findet sich in Experimenten, die sich zusätzlich der sog. **Priming-Technik** bedienen. Werden einzelne Reize während des Experiments zweimal gezeigt, so zeigt ein Wort bei der Wiederholung eine signifikant kürzere Reaktionszeit verglichen mit der ersten Darbietung (Scarborough et al. 1977). Zwischen der ersten Darbietung (*prime*) und der zweiten (*target*) können unterschiedlich viele andere Kontrollreize eingefügt werden. Häufig werden die Reize auch ohne Pause hintereinander dargeboten, wobei nur auf den zweiten Reiz reagiert wird. Dieser Ansatz des Bahnens durch Wiederholung (*repetition priming*) hat in der Morphologieforschung der letzten Jahre eine besondere Rolle gespielt. Es konnte nämlich gezeigt werden, daß ein Bahnungseffekt auch dann auftritt, wenn nicht die gleiche Wortform als Bahnungsreiz dargeboten wird, sondern eine morphologisch verwandte Form, z. B. *walked* vor *walk*; dies wurde als Beleg dafür angesehen, daß in der Tat der Stamm wiederholt verarbeitet wird und nicht zwei einzelne Formen (Stanners et al. 1979). Neuerdings spielt auch die Modalität eine Rolle; so wird die Reaktionszeit auf einen visuellen Reiz gemessen, dem ein auditives Prime vorausging (*cross modal priming*, vgl. z. B. Marslen-Wilson et al. 1994; für eine Diskussion der wichtigsten früheren Ergebnisse vgl. Monsell 1985, für neuere Daten Sandra 1994 und Art. 163).

Eine andere Untersuchungstechnik ist die Beobachtung der Augenbewegungen beim Lesen (Inhoff & Rayner 1996). Durch verschiedene Meßmethoden läßt sich zeigen, ob morphologische Eigenschaften der gelesenen Wörter das Augenbewegungsmuster tangieren, z. B. ob bei pseudopräfigierten Wörtern abweichende Bewegungsmuster zu beobachten sind (Lima 1987). Ein Vorteil dieser Methodik besteht darin, daß Sätze oder auch Texte gelesen werden können, nicht nur einzelne Wörter (s. 4.3). – Da die meisten psycholinguistischen Experimente zur Verarbeitung morphologisch komplexer Wörter nach wie vor im Bereich der visuellen Worterkenn-

nung angesiedelt sind, kann für weitere Methoden einschließlich kritischer Diskussion auf Smith (1996) verwiesen werden.

#### 4. Einige Grundprobleme der Forschung

Die psycholinguistische Morphologieforschung ist am Beginn des 21. Jahrhunderts nur 30 Jahre alt. Anders als in anderen Bereichen, in denen während der 80er Jahre ein systematisches Abschütteln der Dominanz linguistischer Theorien zugunsten einer Synthese psychologischer und linguistischer Theoriebildung notiert werden konnte (Tannenhaus 1988), ist in der Morphologieforschung deren Verhältnis nach wie vor unklar (Henderson 1988). Dies hat eine besonders ausgeprägte Divergenz der Ansätze zur Folge. Im folgenden sollen die wichtigsten Fragestellungen und Hauptprobleme der gegenwärtigen Diskussion knapp kritisch gekennzeichnet werden.

**4.1. Dekomposition vs. Vollformenlexikon**  
 Dies ist die klassische Fragestellung: Werden morphologisch komplexe Wörter bei der Sprachwahrnehmung erst zerlegt, bei der Produktion erst zusammengesetzt, ehe sie verstanden/produziert werden können, oder sind sie als "Ganze" für die Wahrnehmung und Produktion verfügbar; enthält das mentale Lexikon ganze Wörter, deren morphologische Struktur die Verarbeitung nicht beeinflusst (Butterworth 1983), oder nur Stämme und Affixe, deren struktureller Zusammenhang durch Dekompositionsregeln bestimmt ist (Taft & Forster 1975)? Praktisch alles, was zwischen 1975 und 1985 in der psycholinguistischen Literatur zu morphologischen Aspekten der Sprachverarbeitung publiziert wurde, ist dieser Frage gewidmet (vgl. Henderson 1985).

Die aus der linguistischen Diskussion (Halle 1973) in die Psychologie importierte Dichotomie von **Dekomposition** vs. **Vollformenlexikon** war ursprünglich durch Ökonomieerwägungen motiviert. Es schien unrealistisch, daß alle Formen im menschlichen Gedächtnis aufgelistet sein sollten; die linguistisch basierten Regeln boten eine Möglichkeit, den Speicherbedarf zu reduzieren. Betrachtet man die Metaphorik wissenschaftlichen Sprachgebrauchs, so zeigt sich eine enge Bindung an den jeweils erreichten Stand der Technik. Das Streben nach Speicherplatzmi-

nimierung der frühen Dekompositionsmodelle entsprach den Kleincomputern der ersten Generation: So verblüffend (damals) ihre Geschwindigkeit schien, so klein war der Speicher – es stand für maximal 6000 Wörter Speicherplatz zur Verfügung, und nur etwa das Fünffache ließ sich auf Diskette speichern. Trotzdem gab es Disketten mit Lexika von knapp 200000 Wörtern, die geschickt organisiert waren und nur durch spezielle Software mit "Auspakeregeln" lesbar waren.

Demgegenüber ist großer Speicher in Form von Festplatten oder CD-ROM heute überhaupt kein Problem mehr. Statt der Reduzierung von Speicherplatz ist heute Geschwindigkeit des Zugriffs auf große Datens Mengen gefragt. In weniger als 300 Millisekunden ruft der durchschnittliche Erwachsene praktisch fehlerfrei aus einem Wortschatz von 50000 bis 400000 Wörtern das richtige Wort ab (Überblick zu solchen Zahlen bei Aitchison 1994: 5–8) und fügt es in syntaktische und semantische Strukturen ein. Auch bislang für unvorstellbar gehaltene Datenmengen scheint das menschliche Gedächtnis aufnehmen zu können, wenn sie strukturiert abgelegt werden (s. 4.2). Sofern also Prozeduren wie die morphologische Zerlegung allein durch Ökonomieüberlegungen motiviert sind, sind sie abzulehnen, da sie Zeit kosten. Befürworter morphologischer Dekomposition müssen deshalb zeigen, daß ihr Modell schnellere Verarbeitung ermöglicht als ein Vollformenlexikon.

Es ist instruktiv, diese Entweder-Oder Diskussion in Parallelle mit der Frage des phonologischen Rekodierens beim Lesen zu stellen: Involviert der Lesevorgang notwendig eine Phase des phonologischen Rekodierens, oder liest zumindest der erfahrene Leser "direkt", d.h. ohne die Notwendigkeit phonologischer "Übersetzung"? Weil es empirische Argumente für beide Auffassungen gab, entwickelte sich die Idee des Pferderennenmodells (*horse race model*), womit eine über 10 Jahre dauernde Diskussion (nicht unbedingt voll befriedigend) abgeschlossen wurde: Danach wird immer beides versucht, nämlich phonologisches Rekodieren und direkter Zugriff; was perzipiert wird, ist (primär) von der Geschwindigkeit der Routen abhängig (Humphreys & Evett 1985; die Schwächen einer solchen Konzeption werden in Henderson (1982) ausführlich diskutiert).

Dem entspricht das Ergebnis der Diskussion über *Decomposition* vs. *Full Listing*: Extremen Vertretern der einen Richtung (Taft &

Forster 1975) und der anderen (Butterworth 1983) entgegnen etwa Laudanna & Burani (1985) oder Frauenfelder & Schreuder (1992), daß auch hier ein Wettbewerb stattfinde: Danach sind im Lexikon sowohl Stämme und Affixe als auch volle Wörter enthalten; zwei parallele Prozesse suchen entweder nach dem einen oder dem anderen, und der schnellere obsiegt.

#### 4.2. Lexikalischer Zugriff und mentale Repräsentation

Betrachtet man die Literatur, so ist bemerkenswert, daß die Frage nach der Rolle morphologischer Strukturen fast ausschließlich als Problem des lexikalischen Zugriffs diskutiert worden ist (Sandra 1994). Dabei wurden paradoxerweise entgegengesetzte Positionen für den gleichen Sachverhalt bemüht. Das Modell von Forster (1976) sieht einen Hauptspeicher aller Wörter vor; nur für den lexikalischen Zugriff muß zerlegt werden. In der Tat ist bei der Weiterentwicklung zum sog. BOSS-Modell (vgl. Taft 1991) nicht mehr die rein morphologische Zerlegung wichtig, sondern nurmehr eine linguistisch völlig unmotivierte Zwitterstruktur aus anlautender Silbe und Morphem. Caramazza et al. (1988) dagegen sehen alle Wörter in der Zugriffsdatei; der Hauptspeicher bietet die Wörter in zerlegter Form. Frauenfelder & Schreuder (1992) schließlich nehmen eben zur Erklärung der Geschwindigkeit doppelte Repräsentation (und doppelten Zugriff) an.

Vom linguistischen Standpunkt aus betrachtet ist diese Vorstellung besonders problematisch: Sie verlegt systematisches morphologisches Wissen in unklare psychologische Prozesse. Eine alternative Konzeption könnte so aussehen, daß zwar das Lexikon volle Einträge enthält, diese aber morphologisch zerlegt und angeordnet sind (Caramazza et al. 1988); gemäß den Anordnungsprinzipien des Lexikons sind damit implizit auch die Morpheme gelistet. Bei der Verarbeitung liegt dann weder prä- noch postlexikalische Dekomposition vor: Es ist Bestandteil des sprachlichen Wissens, daß *Haus-tür*, *freund-lich*, *sag-t* etc. Einheiten sind, die aus zwei Teilen bestehen. Die zerlegte Struktur der Lexikoneinträge impliziert dabei nicht, daß morphologische Strukturen immer in die Syntax "mitgenommen" werden (z. B. bei Argumentvererbung, s. Art. 83), schließt das aber auch nicht aus (vgl. zu dieser Frage Anderson 1992 und die Rezension durch Cardstairs 1993).

Es ist offensichtlich, daß konnektionistische Modelle eine solche Konzeption systematisch inhärieren; aber auch symbolverarbeitende Ansätze könnten sie aufnehmen. In der psychologischen Diskussion war die Möglichkeit der internen Strukturiertheit von vollständig gespeicherten Wörtern lange Zeit nicht gesehen worden. Alle Befunde, die einen Einfluß morphologischer Struktur auf die Verarbeitung belegen (z. B. der Pseudopräfixeffekt, s. 3.2), lassen sich in einem *Full Listing*-Modell mit interner morphologischer Strukturierung gut erklären: Der Verarbeitungsvorteil entsteht hier dadurch, daß die morphologische Information bei der Verarbeitung genutzt wird; dieser automatische Vorgang führt bei pseudopräfigierten Wörtern (*gering*) zu Verlangsamungen.

Dies ist weniger trivial als es zunächst scheint. Beispielsweise ist faktisch unbegrenzter Speicher wenig nützlich, wenn die gespeicherten Daten nicht auch schnell abrufbar sind. Wenn das Suchsystem, wie bei einer Tonkassette, nur linear abfragen kann, weil die Daten eindimensional linear und nicht weiter strukturiert angeordnet sind, ist kaum zu verstehen, wie solche Geschwindigkeiten erreicht werden sollen. Abgesehen von konnektionistischen Modellen sind im Allgemeinen lineare Abfragemechanismen vorgesehen worden (prototypisch: Forster 1976; vgl. kritisch Sandra 1994). Neuerdings ist die Forschung zum mentalen Lexikon stärker mit der Ermittlung seiner Organisationsprinzipien befaßt (Aitchison 1994). Offenbar spielt sowohl die links/rechts Orientierung als auch die von rechts nach links (Reimfähigkeit) eine Rolle. Primärassoziationen, Clusterbildung, semantisches Priming etc. zeigen, daß auch die semantischen Beziehungen zwischen den Wörtern Anordnungsprinzipien bilden – und es gibt noch viele andere (vgl. Aitchison 1994). Auf die alte Metapher des mentalen Lexikons als Bibliothek angewendet bedeutet das, daß z. B. alle Einheiten, die mit *pe* anfangen, in einem Regal stehen; die mit *pel* stehen links von denen mit, z. B., *pet*. Die reimenden Einheiten stehen vor bzw. hintereinander (*Sprung*, *Leitung*, *jung* vor *Geiz*, *Reiz*, *Schweiz*, aber hinter *prima*, *firma*, *Rheuma*) und die semantische Dimension ist in der Vertikalen ausgedrückt: Die Pflanzenwörter (z. B. *Pelargonie*, *Petunie*) stehen unter den Tieren (*Pelikan*, *Perlhuhn*), etc.

Es ist offensichtlich, daß die Metapher zusammenbricht, sobald mehr als drei Dimensionen auftreten; dies aber ist der Fall: mor-

phologische, syntaktische, konnotative Merkmale gehören ebenfalls zur Systematik des mentalen Lexikons. Im Sinne herkömmlicher Datenbanksysteme läßt sich diese vieldimensionale Anordnung der Einheiten als Indizierung begreifen: Jede Einheit ist u. a. sowohl bezüglich ihrer linksläufigen als auch ihrer rechtsläufigen Organisation, ihrer semantischen Kategorie, ihrer syntaktischen Klasse und ihrer alphabetischen Position indiziert. Alle Einheiten könnten intern strukturiert sein. Eine solche Zerlegtheit läßt die Indizierung der einzelnen Lexikoneinheiten nach Morphemen zu. Im konnektionistischen Ansatz würde man davon reden, daß ein Eintrag mit allen gleich beginnenden, allen reimenden, allen semantisch ähnlichen Einträgen etc. Verknüpfungen aufweist, und zwar unterschiedlich starke nach Maßgabe der Ähnlichkeit. Es ist ungeklärt, inwieweit dies lediglich Notationsvarianten ein und derselben Struktur sind oder aber inhaltlich verschiedene Lösungen.

#### 4.3. Flexion vs. Wortbildung, Produktivität, Lexikalisierung

Die Unterscheidung zwischen Flexion und Wortbildung ist in den frühen psycholinguistischen Arbeiten praktisch nicht berücksichtigt worden (Henderson 1985). Obgleich die exakte Grenzziehung im Einzelfall problematisch ist, so daß bisweilen von einem Kontinuum statt einer Dichotomie gesprochen wird (Bybee 1985), ist in der Mehrzahl der Fälle die linguistische Zuordnung problemlos: Flexion ist automatisch, verändert Wortart und lexikalische Bedeutung nicht, Wortbildung dagegen ist idiosynkratisch und verändert häufig Wortart und Bedeutung.

In Analogie zur segmentalen Ebene visueller Wortsprachkennung, für die eine Ebene abstrakter Buchstabenrepräsentationen angenommen wird (Coltheart 1981), wird von vielen Forschern im Falle der Flexion eine abstrakte Repräsentation angenommen (z. B. Marslen-Wilson et al. 1994), aus der dann bei der Äußerung die konkrete Form gebildet wird – eine Form des Paradigmas steht im Lexikon, die anderen werden daraus für die konkrete Verwendung „abgeleitet“. Experimentelle Daten legen nahe, als solche Basisform die lexikalische Zitierform anzusehen (Günther 1988a). Dabei ist nicht Speicherplatzökonomie der Grund für eine solche Annahme (s. 4.1). Es ist vielmehr der Umstand, daß auch Wörter, die der Sprecher nicht kennt oder selbst erfindet, sofort flektiert

werden können (Mugdan 1977), also bereits bevor sie Teil des mentalen Lexikons sind. Teilweise wird auch davon ausgegangen, daß häufige Formen als Ganze gespeichert und abgerufen, seltene jedoch abgeleitet werden (z. B. Frauenfelder & Schreuder 1992). Für die Unterscheidung von Flexion und Derivation spricht der Umstand, daß Bahnungseffekte zwischen Wortformen desselben Lexems immer wieder gefunden wurden (also z. B. zwischen engl. *walking* und *walked*, zuerst von Stanners et al. 1979); solche Effekte gibt es nicht zwischen derivierten Bildungen (wie *friendly* und *friendship*, Marslen-Wilson et al. 1994).

Es ist bekanntlich möglich, aus den vorhandenen Morphemen gänzlich neue Kombinationen zu bilden, die jedoch von den Hörern sofort verstanden werden wie das *betüpfelbar* (s. 2.1), das man sofort zu *unbetüpfelbar* oder gar *Unbetüpfelbarkeitssymptomatik* erweitern kann, die ebenfalls direkt verstehbar sind. Man spricht davon, daß die Affixe wie *be-*, *-bar*, *-keit*, *un-* oder die nominale Komposition im Deutschen produktiv sind. Dieser alte Begriff linguistischer Wortbildungsforschung ist erst in jüngster Zeit Gegenstand psycholinguistischer Studien geworden (Baayen 1992; Frauenfelder & Schreuder 1992). Dabei ist es bemerkenswert, daß sich in der Linguistik Stimmen mehrten, dieses Konzept als grundsätzlich außerhalb des Bereichs linguistischer Morphologieforschung, als reine Angelegenheit der Sprachverwendung anzusehen (Motsch 1988).

Zwar ist das Vorkommen gänzlich „neuer“ Bildungen, d. h. solcher, die man nie gehört oder gesehen hat, erheblich seltener als die Literatur suggeriert, doch scheinen Hörer und vor allem Leser solche Wörter, wenn sie ordnungsgemäß gebildet sind, ohne Probleme zu verstehen. Die Produktivität von Wortbildungsmustern stellt ein zentrales Problem für Listenmodelle dar, die die Prozeßkomponente auf den reinen Zugriff beschränken wollen. Typisch dafür ist der Ansatz von Aitchison (1994), Neubildungen durch einen völlig unzureichend spezifizierten *lexical tool kit* zu erfassen, einer Art Zusatzeinrichtung, die keinen Bezug auf die im Lexikon enthaltene Morphologie nimmt (vgl. kritisch Berg 1989).

Der Ansatz von Baayen (1992) zur Berücksichtigung von Produktivität in einem psycholinguistischen Modell besteht darin, Häufigkeiten von Token in Relation zur Zahl der

vorkommenden Typs zu stellen; die so ermittelten Werte determinieren in seinem Zwei-Wege-Modell, ob ein Wort dekomponiert wird oder lexikalischer Zugriff auf das ganze Wort erfolgt (d. h. also eine Zwei-Wege-Theorie der morphologischen Dekomposition); eine besondere Rolle spielen dabei ad-hoc-Bildungen, die definitionsgemäß nicht in einer Liste aufgeführt sein können. Auch in der kritischen Weiterentwicklung des Modells durch Frauenfelder & Schreuder (1992) spielt Vorkommenshäufigkeit eine zentrale Rolle; allerdings berücksichtigen sie auch die **Durchsichtigkeit** von Wortbildungen. Hinzugefügt muß allerdings werden, daß bisher experimentelle oder andere empirische Daten fehlen, die belegen, daß eine Neubildung wie *betupfelbar* tatsächlich auf die gleiche Weise erkannt bzw. produziert wird wie etwa *bemalbar*, d. h. daß bei der Produktion oder Erkennung von Neubildungen derselbe Mechanismus aktiv wird wie bei "alten" Bildungen (vgl. auch Sandra 1994).

Auch die Lexikalisierung von Wörtern, quasi die umgekehrte Dimension, hat in der psycholinguistischen Morphologieforschung bislang kaum Beachtung gefunden. So werden etwa in Taft (1981) als präfigierte Stimuli ausschließlich englische Wörter herangezogen, die als völlig lexikalisiert gelten müssen (z. B. *approach*, *reproach*). Es scheint jedoch aus linguistischer Sicht äußerst unwahrscheinlich, daß, wie auch immer sie im mentalen Lexikon repräsentiert sein und daraus abgerufen werden mögen, lexikalisierte Bildungen wie *Zeitung*, *Handschuh* oder *anfangen* auf die gleiche Weise verarbeitet werden wie durchsichtige Wörter wie *Bergung*, *Hauschuh* oder *abfangen*. In der Tat weisen Ergebnisse von Marslen-Wilson et al. (1994) in diese Richtung: Zwischen lexikalierten Bildungen und dem Stamm treten keine Bahnungseffekte auf, wohl aber bei durchsichtigen Bildungen: Engl. *government* bahnt *govern*, aber *department* hat keine bahnende Auswirkung auf *depart*.

#### 4.4. Untersuchungstechniken und Modalitäten

Die Mehrzahl der psycholinguistischen Untersuchungen zur Rolle der Morphologie in der Sprachverarbeitung bedient sich der visuellen lexikalischen Entscheidungsaufgabe. Diese Methode ist, wie verschiedene Befunde der letzten 10 Jahre nahelegen, in besonderer

Weise strategieabhängig: Die Versuchsperson versucht, aus den gebotenen Reizen bestimmte Prinzipien für ihre Antwortstrategie abzuleiten. So läßt sich der Pseudopräfixeffekt zum Verschwinden bringen, wenn alle Wörter des Experiments präfigiert sind (Rubin et al. 1979; Günther 1987); Bahnungseffekte werden um so stärker, je mehr "gute" Primes im Material auftauchen, etc. Es ist deshalb denkbar, daß eine Versuchsperson beim lexikalischen Entscheidungsexperiment Dinge tut, die sie bei der "normalen" Produktion und Perzeption von Sprache gerade vermeidet. Dafür spricht u. a. die Reaktionszeit: Geht man davon aus, daß die durchschnittliche Zugriffszeit zu einem Wort 220 msec beträgt (vgl. Aitchison 1994: 8), so klafft zwischen diesem Wert und den ca. 500 msec lexikalischer Entscheidungsaufgaben eine Lücke, die nicht primär der Auslösung und Exekution der motorischen Reaktion zugeschrieben werden kann. Es liegt nahe, an teilautomatisierte metasprachliche Determinanten der Reaktionszeiten zu denken.

Dies wird verstärkt durch die Tatsache, daß es sich hier (wie in vielen anderen Aufgaben) um die Erkennung einzelner Wörter handelt. Dieser Umstand mag Strategien und Verhaltensweisen fördern, die in der "normalen" Sprachproduktion und -perzeption fehlen. In der Tat ist die Erkennung isolierter Wörter eine Aufgabe, die selbst in auditiven Experimenten Schriftlichkeit involviert oder zumindest voraussetzt; es ist eine in genuiner Mündlichkeit nicht vorkommende Tätigkeit, mit einzelnen Wörtern zu hantieren. Es mag also der Fall sein, daß ein Großteil der experimentellen Befunde zur Morphologie eine Ebene schriftbezogener Metasprachlichkeit betreffen, die in der "normalen" Sprachproduktion und -rezeption eine untergeordnete Rolle spielt. Das dürfte insbesondere für Befunde gelten, in denen lexikalierte Bildungen dennoch morphologische Effekte zeigen (z. B. Taft 1981). Sandra (1994: 264 f.) zitiert Überlegungen, wonach morphologische Effekte simple Konsequenzen orthographischer Redundanz sind.

Wichtiger scheint die Beobachtung, daß entwickelte Alphabetschriften grundsätzlich zur Morphologisierung tendieren. Ein wesentliches Ergebnis der neueren Orthographieforschung ist die Beobachtung des Prinzips der Morphemkonstanz in der geschriebe-

nen Sprache, die es im Gesprochenen so nicht gibt (s. Art. 175). In der Tat ist es ja ein wesentliches Geschäft von Phonologie und Morphophonemik, verschiedene Lautformen auf eine und dieselbe lexikalische Einheit zurückzuführen. Im Geschriebenen (Deutsch oder Englisch) ist dagegen eine solche Verschiedenheit die (eigens zu erklärende) Ausnahme. Das Morphemkonstanzprinzip ist im Englischen, der (bislang) hauptsächlich untersuchten Sprache in der psycholinguistischen Morphologieforschung, besonders ausgeprägt; es gibt gute Gründe anzunehmen, daß der Zusammenhang von *sign* und *signal*, von *concern* (Verb) und *concern* (Sustantiv) etc. dem Kind erst nach Erwerb der geschriebenen Sprache zur Verfügung steht. Unter diesen Umständen ist es bemerkenswert, daß morphologische Fragestellungen in erster Linie in schriftlichen Aufgabenstellungen untersucht worden sind. Bis ca. 1985 war die psycholinguistische Morphologieforschung vollständig auf die visuelle Domäne beschränkt; auch der jüngste Überblicksartikel beschränkt sich ausdrücklich auf diesen Bereich (Sandra 1994); Untersuchungen gesprochener Sprache wie Tyler et al. (1988) bleiben die Ausnahme.

#### 4.5. Die Rolle der Diachronie

Die Frage nach der mentalen Repräsentation morphologisch komplexer sprachlicher Einheiten ist im derzeitigen Stadium kaum auch nur ansatzweise zu beantworten. Dies liegt in einer besonderen Diskrepanz der verschiedenen Forschungsansätze. Im Gegensatz etwa zur Erforschung der Satzverarbeitung ist das Verhältnis zwischen linguistischen und psychologischen Ansätzen nach wie vor ungeklärt (Henderson 1988). Auf der einen Seite sind vor allem die frühen Modelle stark von linguistischen Überlegungen zur Struktur des Lexikons der Grammatik beeinflußt, deren Zielsetzung die systematische Ausblendung von Aspekten der Sprachtätigkeit bedingte, dem wesentlichen Forschungsgegenstand psychologischer Forschung. Paradoxe Weise hat die Konzentration (um nicht zu sagen Reduktion) auf das (psychologische) Problem des lexikalischen Zugriffs Modelle hervorgebracht, die (zumindest für dieses Problem) sprachliche Strukturen als möglicherweise irrelevant erscheinen lassen (Sandra 1994). Auf der anderen Seite sind bis in jüngste Zeit klassische Fragestellungen sprachwissen-

schaftlicher Forschung (Produktivität, Lexikalisierung) dort praktisch nicht zur Kenntnis genommen worden. Dabei könnte gerade dieser Bereich zum Modellfall für das sinnvolle Miteinander linguistischer und psychologischer Forschung werden. Das folgende Zitat kennzeichnet den Sachverhalt, wenn Henderson (1982: 65) zu bedenken gibt, "how much clarifications of a psychological problem can be achieved before we have done a single experiment. Indeed, it seems to me folly to take up a psychological position before examining in some detail what the language allows and encourages. The linguistic evidence does not provide the psychological answers, but it goes a long way toward clarifying the questions". Die ältere linguistische Literatur zu Flexion und Wortbildung und ihren historischen Veränderungen wimmelt geradezu von psychologisierenden Erklärungen. Insofern scheint die systematische Auswertung diachroner Morphologie- und Wortbildungsforschung auch und gerade für die psycholinguistische Modellierung der mentalen Repräsentation morphologisch komplexer Wörter von Interesse: Welche Typen von Veränderungen gibt es, und wie lassen sich diese auf Aspekte rezenter Modelle abbilden?

#### 5. Zitierte Literatur

- Aitchison, Jean (1994), *Words in the Mind: An Introduction to the Mental Lexicon*. London: Blackwell
- Anderson, Stephen R. (1992), *A-Morphous Morphology*. Cambridge: Cambridge University Press
- Baayen, Harald (1992), "Quantitative Aspects of Morphological Productivity". In: Booij & van Marle (Hrsg.), 109–149
- Berg, Thomas (1989), "Rezension zu J. Aitchison, *Words in the Mind*". *Lingua* 79, 342–346
- Berko, Joan (1958), "The Child's Learning of English Morphology". *Word* 14, 150–177
- Booij, Geert & Marle, Jaap van (1992, Hrsg.), *Yearbook of Morphology 1991*. Dordrecht: Kluwer
- Butterworth, Brian (1983), "Lexical Representation". In: Butterworth, Brian (Hrsg.), *Language Production, Bd. II: Development, Writing and Other Language Processes*. London: Academic Press, 257–294
- Bybee, Joan L. (1985), *Morphology – A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins

- Caramazza, Alfonso & Laudanna, Alexander & Burani, Cristina (1988), "Lexical Access and Inflectional Morphology". *Cognition* 28, 297–332
- Carstairs, Andrew (1993), "Morphology without Word-internal Constituents: A Review of Stephen R. Anderson's A-Morphous Morphology". In: Booij, Geert & van Marle, Jaap (Hrsg.), *Yearbook of Morphology* 1992. Dordrecht: Kluwer. 209–233
- Carstairs, Andrew & Stemberger, Joseph P. (1988), "A Processing Constraint on Inflectional Homonymy". *Linguistics* 26, 601–618
- Clahsen, Harald & Rothweiler, Monika (1992), "Inflectional Rules in Children's Grammars: Evidence from German Past Participles". In: Booij & van Marle (Hrsg.), 1–34
- Coltheart, Max (1978), Lexical Access in Simple Reading Tasks. In: Underwood, Geoffrey (Hrsg.), *Strategies of Human Information Processing*. London: Academic Press, 151–217
- Coltheart, Max (1981), "Disorders of Reading and their Implications for Models of Normal Reading." *Visible Language* 15, 245–286
- Derwing, Bruce L. & Skousen, Royal (1989), "Morphology in the Mental Lexicon: A New Look at Analogy". In: Booij, Geert & Marle, Jaap van (Hrsg.), *Morphology Yearbook* 2, 55–71
- Ellis, Andrew W. (1985, Hrsg.), *Progress in the Psychology of Language*, Bd. 1–2. London: LEA
- Forster, Kenneth I. (1976), "Accessing the Mental Lexicon". In: Wales, Roger J. & Walker, Edwin (Hrsg.), *New Approaches to Language Mechanisms*. Amsterdam: North-Holland, 139–174
- Frauenfelder, Uli & Schreuder, Rob (1992), "Constraining Psycholinguistic Models of Morphological Processing and Representation: The Role of Productivity". In: Booij & van Marle (Hrsg.), 165–193
- Gibson, Eleanor J. & Guinet, L. (1971), "Perception of Inflections in Brief Visual Presentations of Words". *Journal of Verbal Learning and Verbal Behaviour* 10, 182–189
- Goodglass, H. & Berko, Joan (1960), "Agrammatism and Inflectional Morphology". *Journal of Speech and Hearing Research* 3, 257–267
- Grodzinsky, Yosef (1990), *Theoretical Perspectives on Language Deficits*. Cambridge/MA: MIT Press
- Günther, Hartmut (1987), "Zur Repräsentation morphologisch komplexer Wörter im internen Lexikon". In: Asbach-Schnitker, Brigitte & Roggenhofer, Johannes (Hrsg.), *Studien zur Wortbildung und Historiographie der Linguistik. Festschrift für H. E. Brekle*. Tübingen: Narr, 195–209
- Günther, Hartmut (1988 a), "Oblique Word Forms in Visual Word Recognition". *Linguistics* 28, 583–600
- Günther, Hartmut (1988 b), "Kreativität, Produktivität, Lexikon – Zur Kompatibilität und Inkompatibilität linguistischer und psychologischer Konzepte in der Morphologieforschung". In: Knobloch, Clemens (Hrsg.), *Kommunikation und Kognition*. Münster: Nodus, 157–172
- Günther, Hartmut (1989), "Experimentelle Morphologieforschung". In: Günther (Hrsg.), 9–26
- Günther, Hartmut (1989, Hrsg.), *Experimentelle Untersuchungen zur deutschen Flexionsmorphologie*. Hamburg: Buske
- Günther, Hartmut & Ludwig, Otto (1996, Hrsg.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung*, Bd. II. Berlin et al.: de Gruyter
- Halle, Morris (1973), "Prolegomena to a Theory of Word Formation". *Linguistic Inquiry* 4, 3–16
- Hare, Mary & Elman, Jeff. L. & Daugherty, K. G. (1995), "Default Generalisation in Connectionist Networks". *Language and Cognitive Processes* 10, 601–630
- Henderson, Leslie (1982), *Orthography and Word Recognition in Reading*. London usw.: Academic Press
- Henderson, Leslie (1985), "Towards a Psychology of Morphemes. In: Ellis (Hrsg.), Bd. I, 15–72
- Henderson, Leslie (1988), "Bericht über den Workshop 'Linguistic and Psychological Approaches to Morphology', Cambridge (1987)". *Linguistische Berichte* 116, 357–363
- Humphreys, Glyn W. & Evett, Lindsay J. (1985), "Are there Independent Lexical and Nonlexical Routes in Word Processing? An Evaluation of the Dual Route Theory of Reading." *Behavioural and Brain Sciences* 8, 689–740
- Inhoff, Albrecht W. & Rayner, Keith (1996), "Das Blickverhalten beim Lesen". In: Günther & Ludwig (Hrsg.), 942–957
- Jarvela, Robert J. & Snodgrass, John G. (1974), "Seeing Ring in Rang and Retain in Retention: On Recognizing Stem Morphemes in Printed ". *Journal of Verbal Learning and Verbal Behaviour* 14, 590–598
- Kean, Marie-Louise (1979), "Agrammatism: A Phonological Deficit?". *Cognition* 7, 69–84
- Laudanna, Alessandro & Burani, Cristina (1985), "Address Mechanisms to Decomposed Lexical Entries". *Linguistics* 25, 575–592

- Leuninger, Helen (1993), *Reden ist Schweigen, Silber ist Gold – Gesammelte Versprecher*. Zürich: Ammann
- Levelt, Willem J. M. (1989), *Speaking: From Intention to Articulation*. Cambridge/MA: MIT Press
- Lima, Susan D. (1987), "Morphological Analysis in Sentence Reading". *Journal of Memory and Language* 26, 84–99
- Marslen-Wilson, William & Komisarjevsky Tyler, Lorraine & Waksler, Tachelle & Older, Lianne (1994), "Meaning and the English Mental Lexicon". *Psychological Review* 101, 1–32
- Meringer, Rudolf & Mayer, Karl (1895), *Versprechen und Verlesen: eine psychologisch-linguistische Studie*. Stuttgart: Göschen (Nachdruck Amsterdam 1982)
- Monsell, Stephen (1985), "Repetition and the Lexicon". In: Ellis (Hrsg.), 147–195
- Motsch, Wolfgang (1988), "On Inactivity, Productivity and Analogy in Derivative Processes". In: Motsch, Wolfgang (Hrsg.), *The Contribution of Word-Structure-Theories to the Study of Word Formation* (Linguistische Studien Reihe A 179), 1–30
- Mugdan, Joachim (1977), *Flexionsmorphologie und Psycholinguistik*. Tübingen: Narr
- Olsen, Susan (1986), *Wortbildung im Deutschen: Eine Einführung in die Theorie der Wortstruktur*. Stuttgart: Kröner
- Paul, Hermann (§1920), *Prinzipien der Sprachgeschichte*. Halle: Niemeyer [1880]
- Pylyshyn, Zenon Walter (1984), *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge/MA: MIT
- Rubin, Gary S. & Becker, C. A. & Freeman, R. H. (1979), "Morphological Structure and its Effect on Visual Word Recognition". *Journal of Verbal Learning and Verbal Behaviour* 18, 757–767
- Rumelhart, David E. & McClelland, James L. (1986), "On Learning the Past Tense of English Verbs". In: Rumelhart, David E. & McClelland, James L. (Hrsg.), *Parallel Distributed Processing*. Cambridge/MA: MIT Press
- Sandra, Dominiek (1994), "The Morphology of the Mental Lexicon: Word Structure Viewed from a Psycholinguistic Perspective". *Language and Cognitive Processes* 9, 227–269
- Scarborough, D. L. & Cortese, C. & Scarborough, H. S. (1977), "Frequency and Repetition Effects in Lexical Memory". *Journal of Experimental Psychology: Human Perception and Performance* 3, 1–17
- Scheerer, Eckart (1992), *Mentale Repräsentationen: Tutorial Paper*. Bielefeld: ZiF, Arbeitsgruppe Brain & Mind
- Schreuder, Rob & Jarvella, Robert & Job, Raimo & Sandström, Göran (1989), "Morphologische Strukturen im mentalen Lexikon". In: Günther (Hrsg.), 27–40
- Slobin, Dan (1985, Hrsg.), *The Crosslinguistic Study of Language Acquisition*, Bd. 1–2. Hillsdale: LEA
- Smith, Philipp T. (1996), "Methods in Reading Research". In: Günther & Ludwig (Hrsg.), 932–942
- Smolensky, P. (1988), "On the Proper Treatment of Connectionism". *Behavioural and Brain Sciences* 11, 1–74,
- Stanners, R. F. & Neiser, J. J. & Hernon, W. P. & Hall, R. (1979), "Memory Representation for Morphologically Related Words". *Journal of Verbal Learning and Verbal Behaviour* 18, 399–412
- Taft, Marcus (1981), "Prefix Stripping Revisited". *Journal of Verbal Learning and Verbal Behaviour* 20, 289–297
- Taft, Marcus (1988), "A Morphological Decomposition Model of Lexical Representation". *Linguistics* 26, 657–668
- Taft, Marcus (1991), *Reading and the Mental Lexicon*. London: LEA
- Taft, Marcus & Forster, Kenneth I. (1975), "Lexical Storage and Retrieval of Prefixed Words". *Journal of Verbal Learning and Verbal Behaviour* 14, 638–647
- Tanenhaus, Michael K. (1988), "Psycholinguistics: An Overview". In: Newmeyer, Frederic (Hrsg.), *Linguistics. The Cambridge Survey*, Bd. III. Cambridge: Cambridge University Press, 1–37
- Tyler, Lorraine K. & Marslen-Wilson, William & Rentoul, James & Hanney, Peter (1988), "Continuous and Discontinuous Access in Spoken Word Recognition: The Role of Derivational Prefixes", *Journal of Memory and Language* 27, 368–381

Hartmut Günther, Köln (Deutschland)

## 163. Speech production and perception

1. Introduction
2. Processing models
3. History of the question
4. Important variables
5. Sources of evidence
6. Results of empirical studies
7. Conclusion
8. References

### 1. Introduction

Controversy surrounds the role of morphology in speech perception and production. The crucial issue is the following: do morphemes affect speech processing? That is, are words such as *girls*, *boys*, *girlish*, *boyish*, treated as if they were simple sequences of sound, or does their internal make-up affect the way in which they are produced or perceived?

A common assumption is that items which are derived by a regular rule are (dis)assembled in the course of speech, in line with Chomsky and Halle's claim (1968: 12) that "regular variations are not matters for the lexicon which should contain only idiosyncratic items". The dissimilarity between regular and irregular forms has been confirmed by neural imaging (Jaeger et al. 1996; Marslen-Wilson & Tyler 1998). But a straightforward division between the two types is problematical, for two reasons.

First, it is unclear which forms should be regarded as 'regular' and which 'irregular'. There seems to be a cline, with regular items at one end, and irregular at the other. In the middle, a plethora of semi-regular forms are found (Bybee & Slobin 1982; Bybee & Moder 1983; Bybee 1995 a; Pinker 1999). They are governed by rules, but rules with a number of conditions attached. For example, English has a productive suffix *-ness* which forms abstract nouns from adjectives e.g. *happiness*, *slowness*. Yet if the adjective ends in [l], the preferred ending is *-ity*, not *-ness* e.g. *mobility*, *futility*. Such examples do not easily fit into either the category of 'regular' or 'irregular' behaviour.

Second, common sense cannot decide the issue: there is a trading relationship between processing and storage which has to be discovered. Words ready-assembled in the mental lexicon would presumably allow faster processing, but would take up considerable

memory capacity. (Dis-)assembling words would slow down processing, but require less memory. Our current knowledge of the mind is insufficient to allow a priori judgments.

### 2. Processing models

Three main processing models have been discussed intensively over the last quarter century – though this oversimplifies the numerous proposals found:

- (a) Whole-word look-up, often known as the **full-listing hypothesis** (e.g. Butterworth 1983). According to this model, speakers/hearers look up whole words, which they routinely treat as indivisible sequences. Morphological affixes are firmly attached to their stems in the mental lexicon: *girls*, *girlish*, *boys*, *boyish* are four different lexical entries. Morphological divisions are potentially detectable by speakers/hearers, but this analytic ability is not brought into use during normal speech processing.
- (b) On-the-spot (dis)-assemblage, sometimes referred to as the **decompositional hypothesis** (e.g. Taft et al. 1986). According to this model, speakers/hearers routinely (dis)-assemble complex words. Morphological affixes are detached from their stems, and morphological (de)composition is part of the normal processing procedure. The stems *girl* and *boy* constitute the main lexical entries, and morphological affixes are listed as sub-entries or additional notes.
- (c) **Mixed processing** (e.g. Laudanna & Burani 1985; Burani & Caramazza 1987). According to this model, processing can be carried out either via the full form, or via the component morphemes. Lexical entries have a double format, in which each word occurs both as a whole and in morphemes. A variant of this, which overlaps with (a), is whole-word look-up with automatic 'word-parsing': morphological affixes are attached to their stems, but morphological make-up is automatically accessed whenever the word is considered, as if, metaphorically, the lexical entry was stored in hyphenated form: *girl-s*, *boy-s*, *girl-ish*, *boy-ish*.

These viewpoints all assume that morphological make-up is potentially available during processing. The critical question is the extent to which (de)composition is bound to occur. Schreuder & Baayen (1995) summarise similarities and differences between these and other newer variants, such as AAM (Augmented Addressed Morphology) model (Burani & Laudanna 1992), MRM (Morphological Race Model) (Baayen et al. 1997).

### 3. History of the question

The role of morphology in speech processing has been discussed primarily by psychologists, since theoretical linguists have traditionally been more interested in a speaker's 'competence' (internal system, or 'representation') than in 'performance' (how that system is used, or 'access') (Sandra 1994).

The morphology question surfaced in psychology in the 1970s, largely as a result of a **prefix-stripping** claim made by Taft & Forster (1975). They discovered that readers rejected non-words such as *\*juvenile* relatively slowly. This slowness, they suggested, might be because *-juvenile* is listed as an entry in the mental lexicon, though with an additional note specifying the obligatory attachment of *re-*. *\*Juvenile* was therefore initially assumed by readers to be a word, and was eventually rejected only after the extra notes had been checked. Words, the researchers concluded, are normally stored as stems, with prefixes which have to be stripped off in comprehension. This claim was backed up with further experiments: 'pseudo-prefixes' words – words which look as if they have a prefix, but in fact do not, such as *precipice* – take extra processing time because readers mistakenly strip off the prefix, and look for a non-existent *\*cipice* (Taft 1981). Later, the notion of stripping was extended to include suffixes also. The assumption that the mental lexicon contains primarily stems which can be accessed only after affixes have been stripped off was strongly maintained by some (e.g. Taft et al. 1986) – though Taft has now modified his views (Taft 1994).

**Affix-stripping** claims were however disputed by others even from the start. According to its detractors, the experimental task may have induced special strategies, word-stress was ignored, and affix-stripping effects, if they exist, may involve at most only written words. (Useful summaries of the early years

of this dispute can be found in Butterworth 1983; Cutler 1983; Henderson 1985.) Early work tended to conflate evidence from different processes, different stages of processing, different media, different languages, different types of morphology, and different degrees of productivity.

Later work on the topic showed increasing sophistication, as researchers refined their methodological tools, and showed that it may be unrealistic to rely on one global model for morphological processing.

### 4. Important variables

Over the years, psycholinguists have become increasingly aware of the large number of variables which need to be considered. The following dozen are perhaps the most basic. They can be divided broadly into linguistic factors and processing factors:

---

#### Linguistic factors

---

- (a) inflection vs. derivation vs. compounding
  - (b) prefixes vs. suffixes vs. infixes vs. circumfixes
  - (c) morphemes vs. syllables
  - (d) isolation vs. polysynthesis (morphemes per word)
  - (e) agglutination vs. fusion (segmentability)
  - (f) bound vs. free forms
- 

#### Processing factors

---

- (g) perception vs. production
  - (h) speech vs. reading/writing
  - (i) representation vs. process
  - (j) access vs. selection
  - (k) productive vs. unproductive processes
  - (l) frequent vs. infrequent words
- 

Tab. 163.1: Important variables

#### 4.1. Linguistic factors

- (a) Inflection vs. derivation vs. compounding. In the literature, the main distinction discussed is that between inflection (e.g. *drive-s*) and derivation (e.g. *driv-er*) (cf. Art. 38). Compounding has received less attention. In general, inflectional morphemes provide extra information about existing words (e.g. *The drivers loaded the vans with bananas*) while derivational

- morphemes form new words from existing ones (e.g. *The driver unloaded the perishable goods*). This distinction is not necessarily always clearcut (e.g. English *-ly* as in *quickly* is a well-known borderline case), and in some languages the boundary is more blurred than others. But in languages such as English, differences of treatment are mostly clear: derivational morphemes precede inflectional ones, as in *driv-er-s*, not *\*drive-s-er*; *nois-i-est*, not *\*nois-est-y* (Aronoff 1976; Spencer 1991). In consequence, inflection and derivation must be analysed separately wherever possible.
- (b) Prefixes vs. suffixes vs. infixes vs. circumfixes (cf. Art. 54–55). Affixes are traditionally divided into prefixes, suffixes, infixes, and (sometimes) circumfixes. In practice, almost all psycholinguistic work has related to prefixes (e.g. *un-*, *dis-* as in *uncover*, *discover*) and suffixes (e.g. *-ful*, *-ous*, as in *harmful*, *perilous*). Suffixes are considerably more common than prefixes in the languages of the world (Cutler et al. 1985; Hawkins & Cutler 1988), and beginnings of words are treated differently from endings (e.g. Marslen-Wilson 1989; Segui & Zubizaretta 1985).
  - (c) Morphemes vs. syllables. The syntactic unit ‘morpheme’ interacts with the phonological unit ‘syllable’. It sometimes coincides (e.g. *care-less*) and sometimes clashes (e.g. *lead-er* vs. *lea-der*). A further possibility is that morphemes may not have any status as ‘morphemes’, but may simply represent familiar phonic or orthographic chunks (Seidenberg & McClelland 1989).
  - (d) Isolation vs. polysynthesis. The number of morphemes per word varies from one (isolation, e.g. *sing*, *seven*) to several (polysynthesis, e.g. *polyunsaturated*: *poly-un-satur-ate-d*). This variation occurs partially within a single language, but more importantly, across languages, where languages vary from low to high on an index of synthesis (Comrie 1989).
  - (e) Agglutination vs. fusion. The segmentability of morphemes varies from straightforward divisibility (agglutination, e.g. *unwanted*: *un-want-ed*) to complete interweaving (fusion, e.g. *bit* ‘bite.PAST’). This variation occurs both within and across languages, where languages score differently on an index of fusion (Comrie 1989).
  - (f) Bound vs. free forms. Morphemes may be free (able to occur on their own as words, e.g. *sing*) or bound (obligatorily attached to another, e.g. *-ed* in *wanted*). This leads to degrees of affixation, e.g. *good#ness* has a suffix less tightly attached than *erot+ic*, since *good* is a word but *\*erot* is not. Further complications are noted in Giegerich (1999), Plag (2003).
- #### 4.2. Processing factors
- (g) Perception vs. production. Speech recognition, both auditory and visual, is known to be partly a matter of informed guesswork: hearers/readers reach conclusions based on outline clues. Some morphological information may therefore be redundant, and less carefully attended to. But production requires attention to the full form. It is therefore unreasonable to extrapolate from one process to the other.
  - (h) Speech vs. reading/writing. In speech, stress pattern and syllable structure are available to guide processing, whereas orthographic forms involve purely visual shapes. For example, the orthographic forms *decide*, *decapitate*, *decontaminate*, *decorate* all have a visual similarity for the first three letters, yet the spoken forms provide different, and possibly stronger, clues to their morphological make-up. The reverse happens with some inflections: for example, phonologically conditioned allomorphs of the English plural morpheme are spelled with an *s*, but pronounced either [s] (*cats*), [z] (*dogs*) or [iz] (*horses*). Discussions of morphology therefore need to be modality-specific (Chialant & Caramazza 1995).
  - (i) Mental representation vs. process. Considerable discussion as to the general architecture of the ‘mental lexicon’ has taken place. Some researchers assume the existence of a single ‘master-file’ with word representations based on some abstract form sometimes referred to as a *lemma*, which must be altered in some way when used in speech (Marslen-Wilson 1989; Bock & Levelt 1994; Marslen-Wilson 1993; Marslen-Wilson & Zhou 1999). This *lemma* level (if it exists) is intermediate between lexical-semantic representations and modality-specific lexical representations. The morphemic representations within this hypothesized master-file must be distinguished from the ac-

- tual forms used in speech processing. Others have argued that it is unnecessary to postulate a modality-neutral lemma-level of representation (Caramazza 1997).
- (j) Access vs. selection. Recent discussions mostly assume that perception and production are two-stage processes. Lexical access involves multiple activation of lexical possibilities, followed by a narrowing down and final selection (e.g. Marslen-Wilson 1989; Aitchison 2003). Treatment of morphology may therefore alter, depending on the stage in progress.
  - (k) Productive vs. unproductive processes. Productive processes which currently can be used to form new words (e.g. *non-* as in *non-modular*, *non-slip*) need to be distinguished from historical unproductive ones (e.g. *in-* as in *impermeable*, *indecent*). Productive morphemes are more easily attached/detached, and productivity also (usually) implies transparency and regularity (Bauer 2001).
  - (l) Frequent vs. infrequent words. Morphemes within frequently used words may become welded together (e.g. *would*, *should* where the final *-d* was in origin a past tense marker). This may mean that similarly structured words are not necessarily processed in the same way. Bybee (1985; 1995 a; 1995 b) pointed out that words of high frequency are more likely to be stored whole, while regular words of lower frequency are more likely to be formed by combination. The issue of frequent vs. infrequent words also overlaps with that of familiar vs. unfamiliar words (Chialant & Caramazza 1995).

The above variables indicate that morphological processing is a complex issue: numerous interacting factors cause complications even within a single language, and there is potentially considerable variation across languages.

## 5. Sources of evidence

Broadly, evidence which might decide the issues can be divided into two types: experimental (preferred by psychologists) and naturalistic (favoured by linguists). Ideally, the results of the two types would coincide.

### 5.1. Experimental evidence

The majority of experiments have probed speech perception, with either auditory or (more usually) visual stimuli. Experiments

have typically involved the measurement of response times ('response latencies'), on the assumption that if a morphologically complex word (e.g. *bluish*) takes longer to recognize than a simple word (e.g. *yellow*), then the complex word has probably involved some type of decomposition.

A **lexical decision task** is a widely-used technique ("Tell me as quickly as possible if this is a word or not"). The subject responds either verbally, or by pressing a button. This task can arguably be made subtler by the use of **priming** (prior presentation of another word). In its simplest form, the prime-word is the same as the 'target' (word on which the decision is being made). Such **repetition priming** speeds up response to the target, suggesting that previous access to a word in the mental lexicon facilitates its later retrieval. Any strong facilitatory effect is usually regarded as evidence that the prime-word and target are closely connected, and may even share a lexical entry. It is therefore one proposed way of checking closeness between morphologically related pairs (e.g. *creep* and *crept*, *deduce* and *induce*). A common technique is to compare the time taken to recognize a repeated word (e.g. *keep* and *keep*) with that taken to recognize a word with an extra affix (e.g. *keep* and *kept*).

**Cross-modal priming** is a further (controversial) refinement. It involves presenting the prime-word in a different modality from the target (e.g. an auditorily presented word followed by a visually presented one), on the (controversial) assumption that different modalities obligatorily involve accessing a neutral-between-modalities version of the word in the mental lexicon (Marslen-Wilson 1989). **Masked priming** is another variant, in which the prime-word is preceded by hash marks (#) equal to the number of letters in it, which supposedly makes memorization of the prime-word less likely (Forster & Davis 1984).

But various doubts have been raised about priming, since facilitatory effects may simply be due to one-off strategies devised for the experimental task: priming is unable to distinguish between online access and post-access procedures, and similar or related words can in some cases have inhibitory effects, rather than facilitatory ones (Monsell 1985). Furthermore, it is often impossible to tell whether a delay in recognition is due to independent representation (e.g. *keep* and *kept* might have separate lexical entries) or de-

composition (*kept* might need to be decomposed).

**Gating** is a further variant of a basic lexical decision task. This involves chopping off the word at successive points, until a 'recognition point' is reached. Its value in morphological processing is that it taps online procedures, and can (supposedly) reveal which parts of the word are essential for identification.

Overall, therefore, psycholinguistic experimentation has increased in sophistication over the years, though it is still something of a blunt instrument in that the results can often be interpreted in a variety of ways.

### 5.2. Naturalistic evidence

These data involve deviations or variations mostly in production, which fall into a variety of categories.

(a) **Slips of the tongue.** These are involuntary deviations from a speaker's intended utterance. They are rule-governed, in that such errors show recurring patterns. Morphologically, their value rests on the assumption that morphological errors might represent assemblage which occurs in the course of speech production. e.g. *concision* for 'precision' might be taken as evidence that *con-* was mistakenly added to *-cision* instead of *pre-* – though care needs to be taken to eliminate alternative explanations (e.g. *concision* might be a blend of *conciseness* and *precision*, or might be a late example of the archaic word *concision*).

A further problem is that it is not always possible to distinguish normal production routines from 'back-up' procedures, which come into play when the normal mechanisms break down. An error such as *deduceful* for 'deductive' could suggest that it is normal to assemble a derivative of *deduce* when required, or it might indicate that the speaker could not think of the 'standard' word, and so performed an emergency word-formation operation.

(b) The errors of **aphasics**. These are mistakes made by patients suffering from serious speech disorders. They are often similar in nature to ordinary 'slips of the tongue', though tend to be more numerous and more bizarre. The link between aphasics and normal speakers was pointed out by Freud (1953: 13): "The paraphasia in aphasic patients does not dif-

fer from the incorrect use and the distortion of words which the healthy person can observe in himself in states of fatigue and divided attention". Researchers hope to find patients with selective impairments, with some types of morphological affixes omitted, but others retained. However, care must be taken to consider alternative explanations, especially as people with damaged brains may adopt unpredictable compensatory strategies.

(c) Normal speech variation. Individuals are sometimes inconsistent in their treatment of inflection of the same lexical items. This is partly due to sociolinguistic factors, such as a shift from casual to formal styles, and partly uncertainty as to how to inflect particular words, as when *obscurer* interchanges with *more obscure* and *more obscurer*. Such inconsistency may reflect assemblage procedures.

In conclusion, evidence on morphological processing is rarely clearcut, whether it is experimental or naturalistic. The data can often be interpreted in more than one way. It is therefore necessary to look for convergence: areas in which different types of evidence apparently collude to provide a plausible account. These two sources can sometimes be supplemented by the observations of theoretical linguists. However, care must be taken when using the latter, as the aim of many theoretical linguists is to describe phenomena in a maximally economical way: this may not be the way the mind works.

## 6. Results of empirical studies

The majority of work on morphological processing has been carried out on English, so English will feature prominently in the account below, even though conclusions drawn from English cannot automatically be carried over to other languages.

### 6.1. Inflectional morphology

English inflectional morphology is relatively straightforward, in that a fairly small number of inflectional suffixes are found, which are all easily detachable.

Various clues suggest that these inflections are added on in the course of speech production:

(a) Slips of the tongue involving phrasal verbs (verbs composed of a base form

- and a separable particle). Examples such as *point outed*, *wash upped* instead of *pointed out*, *washed up* suggest that the basic lexical item is *point out*, *wash up*, and that the ending has been added on in the wrong place (Aitchison 1987; <sup>3</sup>2003).
- (b) Correct morphophonological form on slips of the tongue such as: *Take the freezes out of the steaker* (*steaks out of the freezer*). Here the plural morpheme which was originally intended for *steaks* has been adapted to its new stem *freeze*.
  - (c) Ability of normal speakers to immediately use any new lexical item with appropriate inflections, e.g. *they blanded out* from the recently popularized verb *bland out* ‘to become conformist’.
  - (d) Ability of fluent aphasics to inflect nonsense words appropriately, e.g. *She wikses a zen from me, with a pair of loyses or whatchemecalllem* (Butterworth & Howard 1987).
  - (e) Frequent failure of agrammatic aphasics to inflect words, except in the case of words which are more usually found in their inflected form (e.g. *eyes*).
  - (f) Inconsistent inflection of phrasal nouns such as *mother-in-law*, as in *mother-in-laws* (more usually *mothers-in-law*) suggests a lexical item *mother-in-law*, with inflection added on in the course of speech.
  - (g) Sociolinguistic variation between inflected and uninflected forms e.g. *we see*, *we sees*.

Overall, the evidence suggests that regular inflections are added as utterances are produced, though words which are more usually found in their inflected form may be fully listed with the inflection already added. A similar conclusion was reached by Chialant & Caramazza (1995). This is in line with arguments by Cutler et al. (1985) that the greater the syntactic role played by the suffix, the more likely it is that stem and affix will be separately processed.

Relatively little information is available about the comprehension of spoken inflected forms, though decomposition is supported by an aphasic patient who was found to have a selective deficit for inflectional morphology (Tyler & Cobb 1987).

Most of the information about written language relates to comprehension, where the evidence is consistent with decomposition. An early study showed that when subjects

were briefly shown a visual representation of inflected forms, they tended to perceive them as two units (Gibson & Guinet 1971). Priming studies have shown that prior exposure to a regularly inflected form of a word (*sings*) speeds up recognition time of the base *sing* as much as prior presentation of the word itself, though this is not true of irregular inflections (*sang*) (Stanners et al. 1979). This suggests that the subject is indeed accessing the same word, and ignoring the suffix. However, questions have been raised about whether these effects are genuine morphemic ones, rather than a by-product of visual or semantic similarity. Some studies have therefore tried to tease out the effects of formal similarity (e.g. *ribbon – rib*), semantic similarity (e.g. *pain – ache*), and morphemic relatedness (e.g. *sings – sing*, *sang – sing*) (Henderson et al. 1984; Napps & Fowler 1987; Napps 1989). The results suggest that “Priming among morphemic relatives cannot be explained as the combined effects of semantic and formal priming ... morphemic priming is a separate dimension along which two words can be related” (Napps 1989: 736).

Taken as a whole, evidence for (de)composition is fairly strong in the case of English inflectional suffixes, and further evidence is summarized in Cutler (1983). However, these English results are not necessarily generalizable to inflection in other languages – though Dutch and German appear to behave similarly (Jarvella & Meijers 1983). But in fusional languages, such as Serbo-Croat, inflectional forms appear to be fully listed: one inflected form seems to be at the centre of a constellation of related forms. A basic entry, the ‘nucleus’ has a cluster of ‘satellites’ around it (Lukatela et al. 1980; Lukatela et al. 1987; Feldman & Fowler 1987). This suggests that the English results may not be specific to inflectional morphology, but to the transparency and segmentability of inflectional morphemes in English, in particular to the fact that they are added to existing free forms.

Inflectional prefixes (sometimes regarded as part of a circumfix) exist in Dutch and German, and the evidence here is for (de)-composition (Jarvella & Meijers 1983). The prefix *ge-*, for example, is transparent and easily detachable.

## 6.2. Derivational morphology: prefixes

Evidence for prefix-stripping in English derivational morphology is slim and unreliable. One early study of production found a

number of apparently switched prefixes (e.g. *advice* for *device*) suggesting that prefixes are added on in the course of speech (Fay 1977). However, this study did not look at prefix maintenance, nor did it control for stress, or tease out the different types of prefixes. Another study which analysed speech errors involving words containing common initial sequences (e.g. *con-*, *de-*, *in-*, *dis-*) found that initial sequences were mostly maintained, and that almost half of the examples of apparent prefix switching were possibly blends e.g. *displayed* for *delayed* in: *The trains were disrupted and displayed* (Aitchison 1983–1984). A small residue was more plausibly interpreted as unstressed syllables which were ‘faint’ in memory, rather than detached prefixes (e.g. *deferred* for *referred*). In short, for all examples of prefix change “it is possible to suggest alternative, and arguably more plausible explanations than one which involves the notion of a mental lexicon with detached prefixes” (Aitchison 1983–1984:70).

However, this study did not cover productive prefixes before word boundaries (e.g. *un-* as in *unhappy*) – though the rarity of reported errors such as \**dishappy*, \**untasteful* perhaps indicates that prefixes are firmly attached to words, even when they are combined with free lexemes, since the productive possibilities are wide-ranging. Attested examples of wrongly attached negatives such as: *I disregard this as precise* (*I regard this as imprecise*) do not result in non-words, so cannot be taken as firm evidence for on-the-spot attachment. In short, there is no evidence of prefix attachment for existing words – but plenty of evidence that it is possible to coin new words with the use of productive prefixes (e.g. recently coined *unbundle* ‘to sell subsidiary companies’).

In spoken speech perception, Taft et al. (1986) claim to have found evidence of prefix stripping in a task which asked subjects to make lexical decisions in the case of various types of auditorily presented non-words: \**dejoyce* (real prefix, real stem) vs. \**tejoyce* (non-prefix, real stem) vs. \**dejouse* (real prefix, non-stem) vs. \**tejouse* (non-prefix, non-stem). They found that response latencies were slow in the case of real prefixes, and interpreted this as evidence that the prefix had been stripped, and the stem examined. Responses were slowest of all if the non-word contained a real prefix and a real stem, suggesting that the stem had been found in the mental lexicon, before realization dawned

that the prefix-stem combination was non-existent. But this ingenious task indicates only that speakers try to make sense of a word in any way they can before discarding it: it possibly probes strategies for coping with oddities, rather than normal processing.

A gating procedure (chopping off the word at various points until a ‘recognition point’ is reached) was used by Tyler et al. (1988) in a lexical decision task. They concluded that the presence of a prefix, whether stressed or unstressed, did not slow down processing: there was no evidence that lexical access was delayed until the stem could be identified, noting that words are simply recognized too early for the system to be waiting that long.

The general conclusion therefore is that prefix (de)composition occurs in speech production and recognition primarily as a back-up strategy, when the ‘normal’ processes have failed.

Extensive literature exists on the topic of visual recognition of prefixes, spawned by the original Taft & Forster (1975) prefix-stripping paper. As already noted, this has been heavily criticized by various researchers (e.g. Henderson et al. 1984). In general, Taft & Forster’s claim that certain pseudo-prefixes (e.g. *pre-* from *precipice*) are stripped off is thought by many to be a task-induced strategy, or a post-access phenomenon. At the most, prefix stripping is a preliminary attempt to divide a visual word into chunks, rather than a standard decoding procedure.

Relatively little information exists on the written production of prefixed words, though one puzzling study reported that a dysgraphic patient made more errors in writing non-affixed words than in productively-affixed ones (Badecker et al. 1990).

Overall, prefix stripping in English appears to be unlikely, though may be used as an occasional reading strategy. Of course, English does not necessarily behave in the same way as other languages. Reports of reading experiments suggest that Italian, like English, has its prefixes attached, but that Dutch derivational prefixes may be detachable (Jarvella & Meijers 1983; Jarvella et al. 1987).

### 6.3. Derivational morphology: suffixes

The unpredictability of English suffixes indicates that the majority are likely to be on fully listed words. Even productive and supposedly regular suffixes may show a high degree of unpredictability, as Henderson (1985)

points out: he draws attention to agentive *-er*, and notes irregularities even in a small and circumscribed area such as manufacturers/vendors (Henderson 1985: 39):

(1)	<i>bake</i> (V)	<i>baker</i>	<i>bakery</i>
	<i>butcher</i> (V)	<i>butcher</i>	( <i>butchery</i> ?)
	<i>fruit</i> (N)	<i>fruiterer</i>	?
	<i>hat</i> (N)	<i>hatter</i>	?
	<i>boot</i> (N)	?	?

In addition, there are various ‘baseless derivatives’. For example, *confabulation* is apparently ‘baseless’, in that there is no stem *confablate* or *-fication*. Such examples count against the proposal that composition is a normal procedure.

In speech production, slips of the tongue show relatively few ‘pure’ suffix errors. In one study, overall maintenance of the suffix was 81 Percent in a set of malapropisms (similar sound errors) which contained one of six common suffixes as target, error or both (e.g. *cyclonic* for ‘syphonic’, *actual* for ‘active’) (Aitchison 1987). Where suffix errors occurred, the malapropism usually involved other alterations as well (e.g. *malicious* for ‘malignant’, *relative* for ‘relevant’). The few examples of ‘pure’ suffix change could better be explained by procedures such as blending (e.g. *audial* for ‘auditory’, probably from ‘audio-visual’ or ‘auditory + visual’, *contential* for ‘contentious’, when the subject reported hovering between *contentious* and *controversial*) (vgl. Art. 91). An alternative explanation for some errors is a back-up procedure when the target is not wholly available: this is also a proposed explanation for a number of suffix errors such as *orangine* for ‘orangier’ made by a French aphasic (Pillon et al. 1991). According to the authors, the patient could only partially retrieve the desired phonological form, and produced errors which sometimes looked like derivational suffix errors, but were in fact provisional attempts to fill in the gaps in his partially retrieved form.

Experimental evidence is somewhat inconclusive (Cutler 1983), though tends to support the suggestion that initial access does not necessarily involve decomposition: it takes no longer to recognize a complex word such as *dusty* than a monomorphemic one such as *fancy* (Manelis & Tharp 1977). But when subjects were presented with two words, of which one was suffixed and the other not (e.g. *printer*, *slander*), the task took longer than when the words both had suffixes, or both lacked them. This possibly indi-

cates that suffix-stripping is an optional strategy which can be utilized if necessary. A variety of studies have shown close relationships between the base word and derived forms, such as *select* and *selection*, *destroy* and *destruction*, (e.g. Stanners et al. 1979), but none have produced definitive evidence for (de)-composition at an early stage. An authoritative survey (Henderson 1989) comments on priming studies that “findings suggest that priming by morphological relatives occurs at a stage beyond that of the word detectors” (Henderson 1989: 377).

In all, evidence suggests that initial access involves whole-word look-up, but that word-formation and word-parsing routines are available if required.

But full listing of words with derivational suffixes is not generalizable to all languages. In Turkish and other languages with a massive amount of productive agglutination, there must at least be some online parsing, though it may be misleading to refer to this as ‘suffix-stripping’ (Hankamer 1989).

## 7. Conclusion

Three models were outlined at the beginning of this article:

- (a) whole-word look-up: the full-listing hypothesis;
- (b) on-the-spot (dis)assemblage: the decompositional hypothesis;
- (c) mixed processing.

Empirical and experimental studies suggest that English inflectional suffixes are (de)composed in the course of speech, but that established complex words have their prefixes and suffixes firmly attached, though word parsing is an available back-up option which enables new words to be formed. The mixed processing model possibly represents a mixture of stages: whole-word look-up (initial access) may precede and overlap parsing and word formation. These findings are specific to English, and are not necessarily generalizable to other languages.

However, two points at least need stressing about these conclusions.

First, when morphological processing was first discussed, studies of the mental lexicon were in their infancy. Since then, it has become clear that the mind is capable of massive parallel processing (see summary in Aitchison <sup>3</sup>2003). It may be that considerably

more covert parsing activity is going on in parallel than we realise (cf. Bergman et al. 1988).

Second, there is clearly an interrelationship between transparency and segmentability of morphemes, and productivity. It may be more revealing to look at the interaction between these than to look at a straightforward split between inflection and derivation (Hay 2002). Also an aphasic patient showed an inability to handle morphologically complex words involving suffixes that were phonologically and semantically transparent, whether inflected or derivational (Tyler et al. 1990).

As researchers continue to work on an increasingly wide range of languages, and use increasingly sophisticated research tools, there is hope that many of the outstanding puzzles will soon be solved.

## 8. References

- Aitchison, Jean (1983–1984), “The Mental Representation of Prefixes”. *Osmania Papers in Linguistics* 9–10, 61–72 (Nirmala Memorial Volume)
- Aitchison, Jean (1987), “Reproductive Furniture and Extinguished Professors”. In: Steele, Ross & Treadgold, Terry (eds.), *Language Topics: Essays in Honour of Michael Halliday*. Amsterdam: Benjamins, 3–14
- Aitchison, Jean (³2003), *Words in the Mind: An Introduction to the Mental Lexicon*. Oxford: Blackwell
- Aronoff, Mark (1976), *Word Formation in Generative Grammar*. Cambridge/MA: MIT Press (Linguistic Inquiry Monograph 1)
- Baayen, R. Harald & Dijkstra, Ton & Schreuder, Robert (1997), “Singulars and Plurals in Dutch: Evidence for a Parallel Dual-Route Model”. *Journal of Memory and Language* 37, 94–117
- Badecker, William & Hillis, Argye & Caramazza, Alfonso (1990), “Lexical Morphology and Its Role in the Writing Process: Evidence from a Case of Acquired Dysgraphia”. *Cognition* 35.3, 205–243
- Bauer, Laurie (2001), *Morphological Productivity*. Cambridge: Cambridge University Press
- Bergman, M. W. & Hudson, P. T. W. & Eling, P. (1988), “How Simple Complex Words Can Be: Morphological Processing and Word Representations”. *The Quarterly Journal of Experimental Psychology* 40.A, 41–72
- Bock, Kay & Levelt, Willem (1994), “Language Production: Grammatical Encoding”. In: Gerns-
- bacher, Morton A. (ed.), *Handbook of Psycholinguistics*. San Diego/CA: Academic Press, 945–984
- Burani, Cristina & Caramazza, Alfonso (1987), “Representation and Processing of Derived Words”. *Language and Cognitive Processes* 2, 217–227
- Burani, Cristina & Laudanna, Alessandro (1992), “Units of Representation of Derived Words in the Lexicon”. In: Frost, Ram & Katz, Leonard (eds.), *Orthography, Phonology, Morphology, and Meaning*. Amsterdam: Elsevier, 361–376
- Butterworth, Brian (1983), “Lexical Representation”. In: Butterworth, Brian (ed.), *Language Production, Vol. II*. London: Academic Press, 257–294
- Butterworth, Brian & Howard, David (1987), “Paragrammatisms”. *Cognition* 26, 1–37
- Bybee, Joan L. (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins
- Bybee, Joan L. (1995 a), “Regular Morphology and the Lexicon”. *Language and Cognitive Processes* 10, 425–455
- Bybee, Joan L. (1995 b), “Diachronic and Typological Properties of Morphology and Their Implications for Representation”. In: Feldman (ed.), 225–246
- Bybee, Joan L. & Moder, Carol Lynn (1983), “Morphological Classes as Natural Categories”. *Language* 59, 251–270
- Bybee, Joan L. & Slobin, Dan I. (1982), “Rules and Schemas in the Development and Use of the English Past Tense”. *Language* 58, 265–289
- Caramazza, Alfonso (1997), “How Many Levels of Processing Are There in Lexical Access?”. *Cognitive Neuropsychology* 14.10, 177–208
- Chialant, Dorian & Caramazza, Alfonso (1995), “Where is Morphology and How Is It Processed? The Case of Written Word Recognition”. In: Feldman (ed.), 55–76
- Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York: Harper and Row
- Comrie, Bernard (²1989), *Language Universals and Linguistic Typology*. Oxford: Blackwell
- Cutler, Anne (1983), “Lexical Complexity and Sentence Processing”. In: Flores d'Arcais & Jarvella (eds.), 43–80
- Cutler, Anne & Hawkins, John A. & Gilligan, Gary (1985), “The Suffixing Preference: a Processing Explanation”. *Linguistics* 23, 723–758
- Ellis, Andrew W. (1985, ed.), *Progress in the Psychology of Language, Vol. 1–2*. London: Erlbaum

- Fay, D. (1977), "Prefix Errors". Paper presented at the 4th Salzburg International Linguistics Meeting, August 1977
- Feldman, Laurie B. (1995, ed.), *Morphological Aspects of Language Processing*. Hove: Erlbaum
- Feldman, Laurie B. & Fowler, Carol A. (1987), "The Inflected Noun System in Serbo-Croatian: Lexical Representation of Morphological Structure". *Memory and Cognition* 15, 1–12
- Flores d'Arcais, Giovanni B. & Jarvella, Robert J. (1983, eds.), *The Process of Language Understanding*. Chichester: Wiley
- Forster, Kenneth I. & Davis, Carol (1984), "Repetition Priming and Frequency Attenuation in Lexical Access". *Journal of Experimental Psychology: Language, Memory and Cognition* 10, 680–698
- Freud, Sigmund (1953), *On Aphasia*, translated by Erwin Stengel. New York: International University Press [1891]
- Gibson, Eleanor J. & Guinet, Lynn (1971), "Perception of Inflections in Brief Visual Presentations of Words". *Journal of Verbal Learning and Verbal Behavior* 11, 182–189
- Giegerich, Heinz J. (1999), *Lexical Strata in English: Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press
- Hankamer, Jorge (1989), "Morphological Parsing and the Lexicon". In: Marslen-Wilson (ed.), 392–408
- Hawkins, John A. & Cutler, Anne (1988), "Psycholinguistic Factors in Morphological Asymmetry". In: Hawkins, John A. (ed.), *Explaining Language Universals*. Oxford: Blackwell, 280–317
- Hay, Jennifer (2002), "From Speech Perception to Morphology: Affix Ordering Revisited". *Language* 78, 527–555
- Henderson, Leslie (1985), "Towards a Psychology of Morphemes". In: Ellis (ed.), Vol. I, 15–72
- Henderson, Leslie (1989), "On Mental Representation of Morphology and Its Diagnosis by Measures of Visual Access Speed". In: Marslen-Wilson (ed.), 357–391
- Henderson, Leslie & Wallis, Julie & Knight, Denise (1984), "Morphemic Structure and Lexical Access". In: Bouma, Herman & Bouwhuis, Don G. (eds.), *Attention and Performance, Vol. X*. London, Hillsdale/NJ: Erlbaum, 211–226
- Jaeger, Jeri J. & Lockwood, Alan H. & Kemmerer, David L. & Valin Jr., Robert D. van & Murphy, Brian W. & Khalak, Hanif G. (1996), "A Positron Emission Tomographic Study of Regular and Irregular Verb Morphology in English". *Language* 72, 451–497
- Jarvella, R. & Job, R. & Sandström, G. & Schreuder, R. (1987), "Morphological Constraints on Word Recognition". In: Allport, Alan & Mackay, Donald G. & Prinz, Wolfgang & Scheerer, E. (eds.), *Language, Perception and Production: Relationships between Listening, Speaking, Reading and Writing*. London: Academic Press, 245–262
- Jarvella, Robert J. & Meijers, Guust (1983), "Recognizing Morphemes in Spoken Words: Some Evidence for a Stem-Organized Mental Lexicon". In: Flores d'Arcais & Jarvella (eds.), 81–112
- Laudanna, Alessandro & Burani, Cristina (1985), "Address Mechanisms to Decomposed Lexical Entries". *Linguistics* 23, 775–792
- Lukatela, Georgije & Gligorijevic, B. & Kostic, Aleksandar & Turvey, Michael T. (1980), "Representation of Inflected Nouns in the Internal Lexicon". *Memory and Cognition* 8, 415–423
- Lukatela, Georgije & Carello, Claudia & Turvey, Michael T. (1987), "Lexical Representation of Regular and Irregular Inflected Nouns". *Language and Cognitive Processes* 2.1, 1–17
- Manelis, Leon & Tharp, David A. (1977), "The Processing of Affixed Words". *Memory and Cognition* 5, 190–195
- Marslen-Wilson, William (1989, ed.), *Lexical Representation and Process*. Cambridge/MA: MIT Press
- Marslen-Wilson, William (1989), "Access and Integration: Projecting Sound onto Meaning". In: Marslen-Wilson (ed.), 3–24
- Marslen-Wilson, William (1993), "Issues of Process and Representation in Lexical Access". In: Altmann, Gerry & Shillcock, Richard (eds.), *Cognitive Models of Speech Processing: The Second Sperlonga Meeting*. Hove: Erlbaum, 187–210
- Marslen-Wilson, William & Tyler, Loraine K. (1998), "Rules, Representations, and the English Past Tense". *Cognitive Sciences* 2.11, 428–435
- Marslen-Wilson, William & Zhou, Xiaolin (1999), "Abstractness, Allomorphy, and Lexical Architecture". *Language and Cognitive Processes* 14.4, 321–352
- Monsell, Stephen (1985), "Repetition and the Lexicon". In: Ellis (ed.), Vol. II, 147–195
- Napps, Shirley E. (1989), "Morphemic Relationships in the Lexicon: Are They Distinct from Semantic and Formal Relationships?". *Memory and Cognition* 17.6, 729–739
- Napps, Shirley E. & Fowler, Carol A. (1987), "Formal Relationships among Words and the Organization of the Mental Lexicon". *Journal of Psycholinguistic Research* 16, 257–272

- Pillon, Agnese & Partz, Marie-Pierre de & Raison, Anne-Marie & Seron, Xavier (1991), "L'orange, c'est le fruitier de l'orangine: A Case of Morphological Impairment?". *Language and Cognitive Processes* 6, 137–167
- Pinker, Steven (1999), *Words and Rules: The Ingredients of Language*. London: Weidenfeld & Nicolson
- Plag, Ingo (2003), *Word-Formation in English*. Cambridge: Cambridge University Press
- Sandra, Dominiek (1994), "The Morphology of the Mental Lexicon: Internal Word Structure Viewed from a Psycholinguistic Perspective". *Language and Cognitive Processing* 9.3, 269–327
- Schreuder, Robert & Baayen, R. Harald (1995), "Modelling Morphological Processing". In: Feldman (ed.), 131–154
- Segui, Juan & Zubizarreta, Maria-Luisa (1985), "Mental Representation of Morphologically Complex Words and Lexical Access". *Linguistics* 23, 759–774
- Seidenberg, Mark S. & McClelland, James L. (1989), "A Distributed, Developmental Model of Word Recognition and Naming". *Psychological Review* 96, 523–568
- Spencer, Andrew (1991), *Morphological Theory*. Oxford: Blackwell
- Stanners, Robert F. & Neiser, James J. & Painton, Scot (1979), "Memory Representation for Prefixed Words". *Journal of Verbal Learning and Verbal Behavior* 18, 733–743
- Taft, Marcus (1981), "Prefix Stripping Revisited". *Journal of Verbal Learning and Verbal Behavior* 20, 289–297
- Taft, Marcus (1994), "Interactive-Activation as a Framework for Understanding Morphological Processing". In: Sandra, Dominiek & Taft, Marcus (eds.), *Morphological Structure, Lexical Representation and Lexical Access*. Hove: Erlbaum, 271–294
- Taft, Marcus & Forster, Kenneth I. (1975), "Lexical Storage and Retrieval of Prefixed Words". *Journal of Verbal Learning and Verbal Behavior* 15, 607–620
- Taft, Marcus & Hamblly, Gail & Kinoshita, Sachiko (1986), "Visual and Auditory Recognition of Prefixed Words". *Quarterly Journal of Experimental Psychology* 38A, 351–366
- Tyler, Lorraine K. & Behrens, Susan & Cobb, Howard & Marslen-Wilson, William (1990), "Processing Distinctions between Stems and Affixes: Evidence from a Non-Fluent Aphasic". *Cognition* 36, 129–153
- Tyler, Lorraine K. & Cobb, Howard (1987), "Processing Bound Grammatical Morphemes in Context: The Case of an Aphasic Patient". *Language and Cognitive Processes* 2, 245–262
- Tyler, Lorraine K. & Marslen-Wilson, William D. & Rentoul, James & Hanney, Peter (1988), "Continuous and Discontinuous Access in Spoken Word-Recognition: The Role of Derivational Prefixes". *Journal of Memory and Language* 27, 368–381

*Jean Aitchison, Oxford (Great Britain)*

## 164. Speech errors

1. Introduction
2. Speech errors and linguistic theory
3. The structure of the lexicon
4. Rules and accommodations
5. Conclusion
6. References

### 1. Introduction

A basic assumption of all linguistic **speech errors** research is that an analysis of these data can reveal something about both the **mental representation** of language knowledge – the grammar – and the ways in which we access

this grammar in speech production and comprehension (cf. Bierwisch 1982; Cutler 1988; Fromkin 1971; 1973, ed.; 1980, ed.; 1988; Garrett 1988; Stemberger 1982). Speech errors have been defined as involuntary deviations from intended utterances, as exemplified in the following (T = target; E = error):

- (1) (a) Feature  
T: *Cedars of Lebanon*  
E: *Cedars of Lemadon*
- (b) Segment  
T: *speech production*  
E: *preach seduction*

- (c) Word  
T: *tend to turn out*  
E: *turn to tend out*
- (d) Stem  
T: *The cost of flying*  
E: *The fly of costing* (stem)
- (e) Inflectional suffix morpheme  
T: *all the phones rang*  
E: *all the phone rangs*
- (f) Derivational prefix  
T: *imprecise*  
E: *disprecise*
- (g) Syntactic constituents  
T: *my sister went to the Grand Canyon*  
E: *the Grand Canyon went to my sister*

More recently, the deviant language errors produced by post-brain damaged aphasic patients have also been used as additional evidence for grammatical and production models (cf. Buckingham 1980; Dressler & Stark 1988, eds.).

There are two major kinds of **errors**. The first category involves linguistic units, e.g. disordering, deletion, or addition of intended features, segments, morphemes, words, syntactic phrases, which show what units must be represented or derived in the grammar, as illustrated in examples (1 a–g) above.

There are also errors which involve grammatical rules, for example, the application of a rule which should not apply as in (2):

- (2) T: *I know where they are*  
E: *I know where they're*  
(wrong contraction)

or failure to apply a rule which should be applied as in (3):

- (3) T: *But when will you leave?*  
E: *But when you will leave?*  
(Failure to invert subject and auxiliary)

or a possible wrong application of a derivational or inflectional morphological rule resulting in non-occurring lexical items as in (4):

- (4) (a) T: *he's a New Yorker*  
E: *he's a New Yorkan*  
(b) T: *he swam in the pool*  
E: *he swimmed in the pool*

Speech errors suggest that rules which have been posited as part of the grammar (i.e. as part of linguistic competence (Chomsky

1965)) may also be applied in linguistic performance. As will be discussed further below, some rules seem to be applied in order to prevent a proliferation of errors, referred to as accommodation (Fromkin 1971; Garrett 1980 b) illustrated in (5).

- (5) T: *cow tracks (/s/ plural)*  
E: *track cows (/z/ plural)*

(Note that /kaws/ is phonologically well-formed as shown by the word *mouse* /maws/ but if the intended /s/ was not “changed” to /z/ the result would have been morphologically ill-formed).

The majority of the examples are taken from the English UCLA corpus, but such errors occur in all languages obeying the same constraints.

## 2. Speech errors and linguistic theory

The attempt to understand (and explain) the kinds of normal and **aphasic speech** errors that occur would not be possible without general linguistic concepts. All theories of grammar (however they may differ in detail, concepts or goals) posit a number of basic components: phonology, morphology, the lexicon, syntax, and semantics. Questions of interest concern the autonomy of these components, the ways in which they interact, their units and forms of their representation, and the rules or principles which constrain the wellformedness of structures. It is interesting that spontaneously produced speech errors reveal deviations in the units and rules of all these components, occurring at different stages of the production of an utterance (which parallel but may not be identical to the stages in a derivation of a sentence).

Speech error data have proved insightful regarding various linguistic hypotheses concerning all aspects of the grammar; morphology – the structure of words and the organization of the lexicon is no exception.

Morphology occupied a major role in pre-generative linguistics (cf. e.g. Bloomfield 1933; 1939; Marchand 1969; Nida 1949). In the earliest versions of generative grammar (Chomsky 1957), however, the grammar contained no morphological component and no lexicon, and all word-formation, including inflectional and derivational morphology and compounding, resulted from syntactic transformations. With the publication of *Aspects of the Theory of Syntax* (Chomsky 1965) a

lexicon was included as a separate component of the grammar but word formation (compounding and affixation) was still considered a syntactic process. With the extension of Chomsky's (1970) lexicalist hypothesis by Jackendoff (1972), Halle's (1973) seminal *Prolegomena to a Theory of Word Formation*, and Aronoff's (1976) detailed model of word formation, morphology became an important area of study and theoretical development within the generative framework (Mohanan 1986; Kiparsky 1982; Siegel 1979; Pesetsky 1985; Anderson 1982; 1988).

Psycholinguistic models have also been concerned with the processing and accessing of lexical information which in many cases require assumptions on morphological representation and rules. Some of the attempts to address these issues have included analyses of speech errors; what follows is a discussion of some of these studies.

### 3. The structure of the lexicon

Traditionally, morphology concerned the structure of words and the rules of word formation. The nature of the lexicon itself was not part of morphology but rather the concern of lexicographers. With the inclusion of a lexicon as a component of the grammar, it was no longer possible to discuss morphological processes separate from the structure of the lexicon and the nature of lexical representation. Thus for example, one now needs to know whether lexical morphemes and/or words as well as derivational and inflectional morphemes are included in the lexicon, or in separate lexicons, whether complex words are derived by rule or listed in lexical paradigms, and how one should represent morphologically complex lexical entries. Analyses of speech errors have been used to address such questions.

#### 3.1. Sub-lexicons

Most psycholinguistic models of the lexicon partition entries into sub-components or **modules** (cf. Allport & Funnell 1981; Forster 1976; Fromkin 1985; Morton 1969). Thus, each lexical entry will have its phonological, orthographic, syntactic, and semantic representations in separate but intermeshed sub-lexicons. These particular sub-components have not been discussed in theories of morphology, although their inclusion does not seem to present any problems for any of the

current theories. Word substitution errors and blends provide some evidence for such partitioning, particularly for processing models. Although an isomorphism between a grammatical model and a psycholinguistic model is an open question, when such a parallel can be maintained the elegant solution is to do so.

Examples (6 a) through (6 c) illustrate that in **word substitution** errors, the substituting words often fall into the same semantic class as the substituted words in that they share semantic features (including the polar feature values of antonyms). This may be due to their having semantic 'addresses' close to each other in the semantic sub-lexicon, or due to their being activated because of their shared features.

- (6) (a) T: *don't burn your fingers*  
E: *don't burn your toes*
- (b) T: *my boss's wife*  
E: *my boss's husband*
- (c) T: *a horse of another color*  
E: *a horse of another race*

Examples (7 a–b) show that substituted words may instead be phonologically similar to the substituting words, again suggesting either listings according to phonology in a phonological sub-lexicon or activation of words with shared phonology.

- (7) (a) T: *white Anglo-Saxon protestant*  
E: *white Anglo-Saxon prostitute*
- (b) T: *Universal City*  
E: *University City*

It is interesting that word substitution errors made by aphasics and acquired dyslexics also show the same kind of semantic and/or phonological similarity (Fromkin 1987).

The possible causal effect of phonological similarity is discussed by Butterworth (1981) in an analysis of **word blends**. 51 of 65 blends that he analyzed involved pairs of presumed words which are phonologically similar in shared phonemes, syllable structure, initial segment, or stress pattern as in the following:

- (8) (a) *trying/striving* > *strying*
- (b) *draft/breeze* > *dreeze*
- (c) *terrible/horrible* > *herrible*
- (d) *grizzly/ghastly* > *grastly*
- (e) *slick/slippery* > *slickery*

Fay & Cutler (1977) in their study of **mala-propisms** – word substitution errors in which the substituted word is phonologically but

not semantically similar to the intended word – conclude that:

“(a) There is a single dictionary used for production and comprehension. (b) In this dictionary, words are arranged by phonemic structure, in a left-to-right manner, and based on a distinctive feature system. (c) The major partitioning of the dictionary, however, seems to be by number of syllables, with stress pattern as a second categorization within syllable categories. (d) Words may also be arranged by syntactic category.” (Fay & Cutler 1977: 515f.)

Hurford (1981), after analyzing the same set of malapropisms, concluded instead that:

“a. Malapropisms tend significantly to resemble their targets at both extremities of the lexical item ... b. Items listed in the mental lexicon are whole words, with suffixes already attached to stems.” (Hurford 1981: 423)

Hurford argues for this position on the basis of the similarity between target and error to the right of the first wrong segment which he concludes “is due ... almost entirely to a coincidence of such derivational suffixes” (Hurford 1981: 420).

There is, however, other evidence to suggest that in addition to whole words being listed in the lexicon (Halle 1973; Aronoff 1976), stems and/or morphemes must also be listed together with productive rules of word formation which may be incorrectly applied as illustrated by examples (4a) and (4b) above and the derivational errors shown in (9a–b) and inflectional errors in (9c–d). If the lexicon contained only actual words there would be no way to account for the possible but non-occurring words which are produced in speech errors.

- (9) (a) T: *we often go to the movies*  
E: *we oftently go to the movies*
- (b) T: *motivate*  
E: *motify*
- (c) T: *I already took a bath*  
E: *I already tooken a bath*
- (d) T: *I meant to say*  
E: *I meaned to say*

As Fay & Cutler (1977) and Hurford (1981) point out, there is strong evidence that words are listed in the lexicon according to various phonological criteria. This however does not contradict the need for also including a semantic sub-lexicon in which words are listed by semantic classes and/or features. There is also evidence for an orthographic sub-lexicon which is beyond the scope of this article.

### 3.2. Two lexicons

While the arguments for sub-lexicons revolve solely on processing considerations, there have been proposals for the bifurcation of words and/or morphemes into two separate lexicons based on grammatical and theoretical concerns (Stockwell et al. 1972). Golston (1991) presents evidence from speech errors in support of his claim that the grammar includes two lexicons, one consisting of content words and derivational affixes, and the second (which he calls a phrasicon) consisting of function words and inflectional affixes. In his model which he posits as being both a model of the grammar and a speech production processing model, he suggests that different kinds of speech errors occur at specific stages in production and further that other kinds of errors cannot occur. His model is similar to but not identical with Garrett’s (1980a; 1980b) model which also makes claims on the basis of the kinds of errors which occur at different stages or levels.

There is indeed much evidence that the linguistic processing mechanisms differentiate between affixes and lexical content words and/or stems. Berg (1987) summarizes the distinction (cf. also Fromkin 1973, ed.; Garrett 1975; MacKay 1979; Nooteboom 1969):

“Stems and affixes do not interact in tongue slips ... Prefixes and suffixes do not communicate with each other ... Inflectional and derivational suffixes, in turn, stay within their respective subcategorizations ... Lexical morpheme errors respect word class membership to a lesser extent than do independent word errors.” (Berg 1987: 13f.)

Golston points out that errors which consist solely of inappropriately selected function words can be accounted for by a two lexicon model in which the selection of such words occurs after the selection of content words as in the following errors from the UCLA corpus (Fromkin 1988):

- (10) (a) T: *John's going, isn't he?*  
E: *John's going, isn't it?*
- (b) T: *That's what we're going to be doing today.*  
E: *That's what we're going to be do today.*
- (c) T: *I wouldn't be surprised, would you?*
- (d) E: *I wouldn't be surprised, do you?*
- (e) E: *when everyone had left*  
T: *when everyone was left*

### 3.2.1. Word and stem errors

Garrett (1980 b) points out that affixes and lexical stems exhibit different patterns in speech errors further supporting the notion of separate lexicons for each and specific stages at which errors involving each occur. Word stems ('open class' lexical content morphemes) participate in exchanges (spoonerisms or reversals) often leaving their inflectional morphemes behind (referred to as stranding) as shown in the following:

- (11) (a) T: *the cost of the purchase*  
E: *the purchase of the cost*
- (b) T: *your keys are in my purse*  
E: *my keys are in your purse*
- (c) T: *does Jack smoke?*  
E: *does smoke Jack?*
- (d) T: *sweeping streets*  
E: *streeting sweeps*
- (e) T: *wine racks*  
E: *rine (Rhine?) wacks*

(Note that word exchanges do not always respect syntactic categories)

If pronouns are included in the first lexicon (that containing lexical formatives, and derivational morphemes, but not grammatical words or inflectional morphemes) it is not surprising to find that pronouns are involved in exchanges such as in (11 b) above.

### 3.2.2. Derivational exchanges

There are also errors involving exchanges of derivational affixes:

- (12) (a) T: *thinking and cognition*  
E: *thintktion and cogniting*
- (b) T: *colder and windy*  
E: *coldy and winder*
- (c) T: *bloody students*  
E: *bloodent studies (/studiz/)*

There are however no attested errors involving the exchange of inflectional affixes (when they function as bound morphemes, not as phonological units such as syllables). That is, errors such as (13) do not seem to occur:

- (13) T: *the girls are playing*  
E: \**the girling are plays*

### 3.2.3. Stranding errors

Errors such as (12 b) and (12 c) above are called stranding errors in that inflectional affixes are left behind or 'stranded' when the lexical stems exchange. Other examples are given in (14 a) to (14 c).

- (14) (a) T: *rules of word-formation*  
E: *words of rule formation*
- (b) T: *my check cashed*  
E: *my cash checked*
- (c) T: *a sweeter flute*  
E: *a fluter sweet*

Derivational morphemes rarely are stranded but such stranding errors do also occur as shown in (15).

- (15) T: *return your call*  
E: *recall your turn*

Garrett's model accounts for such stranding errors by suggesting that lexical stems are selected and inserted into syntactic planning structures or frames which contain these syntactic elements.

### 3.2.4. Shifts

The shift or movement of affixes from the intended place in the utterance with attachment to another word is also observed, as in the following:

- (16) (a) T: *print outs*  
E: *prints out*
- (b) T: *Larry Hyman's paper*  
E: *Larry's Hyman paper*
- (c) T: *when someone comes up to me*  
E: *when someones come up to me*

Again it should be noted that while derivational affixes seldom shift, grammatical words and minor grammatical categories (e.g. adverbs, intensifiers, determiners), like inflectional morphemes often do:

- (17) (a) T: *I can hear her*  
E: *I can her hear*
- (b) T: *he's been around a long time*  
E: *he's been a long around time*
- (c) T: *Why do we have to?*  
E: *Why do have we to?*
- (d) T: *I really hate to correct exams*
- (e) E: *I hate to really correct exams*

## 4. Rules and accommodations

It has often been pointed out that children's errors during the acquisition stage – *mouses* for *mice* or *bringed* for *brought* for example – reveal more about their grammar construction and the rules they are forming than do their well formed utterances (cf. Art. 165). This is also true of adult speech as shown by the errors they produce, as in (18 a–e):

- (18) (a) T: *the last I knew about it*  
E: *the last I knowned about it*
- (b) T: *he had to have it*  
E: *he haved to have it*
- (c) T: *I meant to say*  
E: *I meanded to say*
- (d) T: *Rosa always dated shrinks*  
E: *Rose always date shrank*
- (e) T: *they'd be breastfed*  
E: *they'd be breastfeeded*

Such possible forms (non-existent in the normal lexicon) suggest that a speaker may utilize regular inflectional rules applying them incorrectly at times to exceptional forms (although it is still possible that there may also be, in the normal situation, the selection of the correctly specified inflected form).

Such rules also surface in the process known as **accommodation** described and illustrated in (5) above and in the following:

- (19) (a) T: *I don't know that I'd know one if I heard it*  
E: *I don't know that I'd hear one if I knew it*
- (b) T: *an eating marathon*  
E: *a meating arathon*
- (c) T: *the museum that we saw*  
E: *the museum that saw us – we saw*
- (d) T: *I thought I was finishing your beer*  
E: *I thought you were finishing my beer*

In the Swedish example (20) the plural suffix *-ar* is changed to *-er* to accommodate to the morpholexical rules of the language.

- (20) Swedish (Linell 1982)  
T: *två dagar i rad*  
E: *två rader i dag*

## 5. Conclusion

The discussion above aimed at showing how questions concerned with the nature and structure of the lexicon and morphological representation and rules have been investigated through an analysis of spontaneously produced errors of speech. While many questions remain to be answered and some only through normal linguistic analysis, performance data such as speech errors can contribute to our understanding of some of the issues. Just from those examined here we can draw some tentative conclusions.

- (a) The lexicon appears to be subdivided into sub-lexicons, each containing a different set of representations. Thus a lexical entry like "coat" might be represented as in Tab. 164.1:

Sub-Lexicon	Representation
phonological	/kot/
orthographic	coat/COAT
syntactic	[ <sub>N</sub> coat] + sub-categorization, etc.
semantic	'coat'

Tab. 164.1: Sub-lexicons

- (b) In the phonological sub-lexicon, words and/or morphemes appear to be listed according to initial phonemes, number of syllables, and stress patterns.
- (c) Lexical and grammatical morphemes, and derivational and inflectional affixes function differentially and at different stages in production. This argues for, at the very least, distinct specifications and possibly two separate lexicons which are accessed at different stages in production.
- (d) Both derivational and inflectional word formation rules exist which when misapplied result in possible but non-occurring lexical items. Speech error data cannot determine whether in addition to the application of such rules, derived and inflected words also occur in the one or two lexicons and their sub-components, although it appears that they do.

Further cross-linguistic analyses of speech errors should make further contributions to questions concerning the nature and access of the lexicon and morphological representations and rules.

## 6. References

- Allport, D. A. & Funnell, E. (1981), "Components of the Mental Lexicon". *Philosophical Transactions of the Royal Society of London B* 295, 183–196
- Anderson, Stephen R. (1982), "Where's Morphology". *Linguistic Inquiry* 13, 571–612
- Anderson, Stephen R. (1988), "Morphological Theory". In: Newmeyer, Frederick (ed.), *Linguistics: The Cambridge Survey I*. Cambridge: Cambridge University Press, 146–191
- Aronoff, Mark (1976), *Word Formation in Generative Grammar*. Cambridge/MA: MIT Press

- Berg, Thomas (1987), *A Cross-linguistic Comparison of Slips of the Tongue*. Bloomington, Indiana: Indiana University Linguistics Club
- Bierwisch, Manfred (1982), "Linguistics and Language Error". In: Cutler, Anne (ed.), *Slips of the Tongue and Language Production*. Berlin: Mouton de Gruyter, 29–72
- Bloomfield, Leonard (1933), *Language*. New York: Holt
- Bloomfield, Leonard (1939), "Menomini Morphophonemics". *Travaux du Cercle Linguistique de Prague*, 105–115
- Buckingham, Henry W. Jr. (1980), "On Correlating Aphasic Errors with Slips of the Tongue". *Applied Psycholinguistics* 1, 199–220
- Butterworth, Brian (1981), "Speech Errors: Old Data in Search of New Theories". *Linguistics* 19, 627–662
- Chomsky, Noam (1957), *Syntactic Structures*. The Hague: Mouton
- Chomsky, Noam (1965), *Aspects of the Theory of Syntax*. Cambridge/MA: MIT Press
- Chomsky, Noam (1970), "Remarks on Nominalization". In: Jacobs, Roderick A. & Rosenbaum, Peter S. (eds.), *Readings in English Transformational Grammar*. Waltham/MA: Blaisdell, 184–221
- Cutler, Anne (1988), "The Perfect Speech Error". In: Hyman, Larry M. & Li, Charles N. (ed.), *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin*. London, New York: Routledge, 209–223
- Dressler, Wolfgang U. & Stark, John A. (1988, eds.), *Linguistic Analyses of Aphasic Language*. New York: Springer-Verlag
- Fay, Donald & Cutler, Anne (1977), "Mala-prisms and the Structure of the Mental Lexicon". *Linguistic Inquiry* 8, 505–520
- Forster, Keith I. (1976), "Accessing the Mental Lexicon". In: Wales, Roger J. & Walker, Edward (eds.), *New Approaches to Language Mechanisms*. Amsterdam, New York: North Holland Publishing Co., 257–288
- Fromkin, Victoria A. (1971), "The Non-Anomalous Nature of Anomalous Utterances". *Language* 47, 27–52
- Fromkin, Victoria A. (1973, ed.), *Speech Errors as Linguistic Evidence*. The Hague: Mouton
- Fromkin, Victoria A. (1980, ed.), *Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen, and Hand*. New York: Academic Press
- Fromkin, Victoria A. (1985), "Evidence in Linguistics". In: Robins, Robert H. & Fromkin, Victoria A. (eds.), *Linguistics and Linguistic Evidence*. Newcastle upon Tyne: Grevatt & Grevatt, 18–38
- Fromkin, Victoria A. (1987), "The Lexicon: Evidence from Acquired Dyslexia". *Language* 63, 1–22
- Fromkin, Victoria A. (1988), "Grammatical Aspects of Speech Errors". In: Newmeyer, Frederick J. (ed.), *Linguistics: The Cambridge Survey II*. Cambridge: Cambridge University Press, 117–138
- Garrett, Merill F. (1975), "The Analysis of Sentence Production". In: Bower, Gordon H. (ed.), *The Psychology of Learning and Motivation, Vol. IX*. New York: Academic Press, 133–177
- Garrett, Merill F. (1980 a), "The Limits of Accommodation". In: Fromkin, Victoria A. (ed.), *Errors in Linguistic Performance*. New York: Academic Press, 263–271
- Garrett, Merill F. (1980 b), "Levels of Processing in Sentence Production". In: Butterworth, Brian (ed.), *Language Production, Vol. I*. New York: Academic Press, 177–220
- Garrett, Merill F. (1988), "Processes in Language Production". In: Newmeyer, Frederick J. (ed.), *Linguistics: The Cambridge Survey III*. Cambridge: Cambridge University Press, 69–96
- Golston, Chris (1991), *Both Lexicons*. Ph.D. dissertation, UCLA
- Halle, Morris (1973), "Prolegomena to a Theory of Word-Formation". *Linguistic Inquiry* 4, 3–16
- Hurford, Jim R. (1981), "Malapropisms, Left-to-Right Listing, and Lexicalisms". *Linguistic Inquiry* 12, 419–423
- Jackendoff, Ray (1972), *Semantic Interpretation in Generative Grammar*. Cambridge/MA: MIT Press
- Kiparsky, Paul (1982), *Explanation in Phonology*. Dordrecht: Foris
- Linell, Per (1982), "Speech Errors and the Grammatical Planning of Utterances". In: Koch, Wolfgang & Platzack, Christer & Tottie, Gunnar (eds.), *Textstrategier i tal och skrift*. Stockholm: Almqvist & Wiksell International, 134–151
- MacKay, Donald G. (1979), "Lexical Insertion, Inflection, and Derivation: Creative Processes in Word Production". *Journal of Psycholinguistic Research* 8, 477–498
- Marchand, Hans (1969), *The Categories and Types of Present-Day English Word-Formation*. Munich: Beck
- Mohanan, Karuvannur (1986), *The Theory of Lexical Phonology*. Dordrecht: Reidel
- Morton, John (1969), "The Interaction of Information in Word Recognition". *Psychological Review* 76, 165–178

Nida, Eugene (1949), *Morphology: The Descriptive Analysis of Words*. Ann Arbor: University of Michigan Press

Nooteboom, Sieb G. (1969), "The Tongue Slips into Patterns". In: Sciarone, A. G. & Essen, A. J. van & Raad, A. A. van (eds.), *Nomen Society: Leyden Studies in Linguistics and Phonetics*. The Hague: Mouton, 114–132

Pesetsky, David (1985), "Morphology and Logical Form". *Linguistic Inquiry* 16, 193–246

Siegel, Dorothy (1979), *Topics in English Morphology*. New York: Garland

Stemberger, Joseph Paul (1982), "The Nature of Segments in the Lexicon: Evidence from Speech Errors". *Lingua* 56, 235–259

Stockwell, Robert P. & Schachter, Paul & Partee, Barbara H. (1972), *The Major Syntactic Structures of English*. New York: Holt, Rinehart & Winston, Inc.

*Victoria A. Fromkin †, Los Angeles (U.S.A.)*

## 165. First language acquisition

1. Acquiring morphology
2. Inflection
3. Derivation
4. Summary
5. References

### 1. Acquiring morphology

Morphology has been studied for several different reasons by researchers interested in first language acquisition. First, the acquisition of inflections has been used to assess children's stage of language development. It is well established that, across languages, children at first make no productive use of closed class items like clitic pronouns, particles, and adpositions or of inflections marking grammatical categories such as tense, aspect, case, number, or gender. They then begin to add inflections and other grammatical markers, and their words become more complex (measured in morphemes), so that their utterances are longer. Computing the average length of children's utterances in morphemes yields a measure of early grammatical development that is more sensitive than length-in-words (e.g., Brown 1973; Dromi & Berman 1982).

Morphological development has also provided an important domain for testing hypotheses about how children learn. For example, do they learn inflections by rote, one word at a time, or do they abstract over words to come up with generalisations in the form of schemas or rules to apply to unfamiliar instances? If children can add past tense inflections to unfamiliar verbs, they must be relying on some more general principle than

rote learning. This finding is even more robust if children can be shown to inflect non-sense words, items that are non-existent in the established lexicon of their language. Such generalisations can be expressed as rules or schemas; they emerge early in the process of acquisition and show up in both spontaneous over-regularisations and elicited productions (e.g., Berko 1958; Bybee & Slobin 1982).

The acquisition of morphology also offers a testing ground for crosslinguistic comparisons where some of the contributions of typology can be assessed. Do children learn morphological rules or schemas in the same way, with the same ease, in synthetic, agglutinating, and isolating languages? What aspects of typology make for error-free versus error-full paths during acquisition? And what effect does typology have on the acquisition of specific grammatical systems such as tense-aspect or passivisation on the verb, or accusative versus ergative case-marking on the noun? (See, e.g., Berman 1990; Pye 1990; Slobin 1985, ed.; 1992, ed.; 1997, ed.)

Finally, the internal structure of words can be connected to conceptual and semantic development on the one hand, and to vocabulary acquisition on the other. When do children begin to analyse words into component parts? When do they start to make active use of derivational morphology in coining words? Are children sensitive to the productivity of the options available when they construct new words? These are just a few of the questions researchers have explored for derivational morphology and word formation more generally.

When children begin to talk, they typically produce one word at a time. These single words may consist of bare stems only, as in English, where the stem is normally a word too; or they may be “frozen” inflected forms. For example, a child may always produce a particular noun with the same case ending on it. Or they may always produce a verb in one particular form with tense, person, number, and gender marking (e.g., Berman & Armon-Lotem 1996). These initially invariant forms of words and their subsequent treatment by children discovering the relations of stems to inflectional affixes show which elements children focus on in the speech they hear. They consistently attend to stressed syllables and final syllables, for example, and tend to produce those first. Only later do they add other elements to their early words. It takes time for children to identify words or stems and their modulations – the different additions licensed by each word or word-type in a paradigm (Anderson 1985 a; 1985 b). Identification and analysis represent the first step (Peters & Menn 1993; Veneziano & Sinclair 2000). Only once that has been taken can children be credited with knowledge about the forms and meanings of inflectional paradigms or derivational patterns. And only then is it possible for researchers to generalise about children’s acquisition of such inflectional systems as case, number, gender, or tense.

A major source of information about children’s morphological development consists of the **errors** they make en route to the adult system (see also Art. 164). The errors may be errors of omission, where children fail to add the expected inflectional or derivational morpheme to a word-stem. These errors offer indirect information about what children at each stage find difficult and therefore omit from their utterances. Other errors are errors of commission. These provide a more direct source of evidence about how children are organising their morphological paradigms. When children regularise specific word-forms, they reveal both how they have analysed and how they have organised the relevant inflections or options for word formation so far.

In what follows, we look first at some of the findings for inflectional morphology (cf. 2) and then at derivational morphology (cf. 3) in acquisition.

## 2. Inflection

Children begin to add their **first inflections** and to use them in contrasting contexts sometime between 1;6 (one year, six months) and 2;0 (two years). By 2;6 to 3;0, they usually control some case endings, number marking on nouns and verbs, some person marking on verbs as well as some marking of aspect and tense, as well as gender and classifier systems in languages that mark nouns (and their associated modifiers) for these categories (see Art. 97–98). Just which inflections children have acquired by age three varies to a large extent with the typology of the language and with the complexity of the paradigms to be learnt. Cognitive complexity is one major pace-setter in acquisition (children must have established a conceptual category before they look for a means for talking about it), but formal complexity also plays an important role in determining when children master the forms available (Slobin 1973). If a language marks three genders as well as number and case on nouns, with different patterns of case endings for each gender and number, the inflectional system will take longer to learn than a language that marks neither gender nor case. Inflections that combine several elements of meaning (e.g., the Hebrew suffix *-ot* for plural number and feminine gender together) are acquired later than those that bear a clear one-to-one correspondence to a specific grammatical category (e.g., the Hebrew suffix *-a* for feminine gender in the default instance of singular number).

Our discussion focusses on questions of complexity in the acquisition of inflections (2.1), rote learning versus rule application (2.2), agreement marking (2.3), and the linear ordering of morphemes (2.4).

### 2.1. Added complexity

When children begin to add inflections, they modulate the meanings of the bases or stems the inflections are added to. These modulations mark additions to the basic or core lexical meanings of the pertinent terms. In general, “complexity in thought is reflected in complexity of expression” (Clark & Clark 1977: 523; Greenberg 1966). Each inflectional modulation adds **complexity**, whether in the forms of noun affixes to mark number and case, say, or in the forms of verb affixes to mark aspect, tense, person, and number (Bybee 1985). The added morphemes *will* and *-ed* explicitly mark added complexity of meaning

in the contrasts between *sleep* and *will sleep*, or *jump* and *jumped*, just as the plural morpheme *-s* marks the added complexity of ‘more than one’ over ‘one’ in *cats* versus *cat* or the ordinal *-th* marks ordinal rank versus cardinality in *seventh* versus *seven*. But added complexity of meaning is not always visible morphologically, so the various possibilities are: full form/meaning transparency where a given modulation is marked by a single additional morpheme (e.g., the English plural *-s* on nouns or the regular past tense on verbs); added complexity of meaning with no obvious morphological addition (e.g., the English plural noun *sheep* or the past tense form *cut*); or more than one modulation of meaning packed into a single morpheme (e.g., the Hebrew feminine plural on nouns, or Italian first person imperfect past on verbs). The general prediction for acquisition, then, is that the more transparent the modulation, the earlier it should be acquired by children.

Relative semantic complexity appears to account quite closely for order of acquisition in those domains where it is possible to make the relevant comparisons. It accounts for the order in which children acquire verb inflections in English (Brown 1973; see also Berman 1985 and Levy 1983 on Hebrew). But it is difficult to apply such a measure since few inflectional domains have been carefully analysed with respect to meaning, and even fewer have been compared across languages. Complexity of meaning is not the only kind of complexity children encounter during the acquisition of inflections. They must also master complexities of form – the number and kinds of changes required to mark a particular meaning in a language.

Complexity of form can be analysed from different points of view. One depends on which categories are morphologically marked in a language – case, gender, tense, number, aspect, negation, and so on (Anderson 1985a). A second factor that may increase formal complexity is the way the available structures map the relevant meanings – the number of different elements (compare the noun case systems of Hungarian and German, say); the number of case forms that carry more than one meaning within the singular or plural paradigm; or the morphophonemic rules and the extent to which the constituent morphemes in a word form remain readily analysable in the surface of each word. In Finnish, for instance, morpheme boundaries within words are often obscured,

whereas in English they often remain visible. A third factor is the consistency of the form-to-meaning mapping across semantic domains (Slobin 1985, ed.; 1992, ed.; 1997, ed.). Languages may differ greatly in these respects so that parts of the inflectional morphology within a language may be hard for children to acquire (e.g., noun plurals in Arabic, locative marking in Serbo-Croatian) while other inflectional paradigms may be much easier (e.g., plurals in English, locative marking in Hungarian) (Mikes 1967; Omar 1973; Pléh 1998; Slobin 1973). Complexity of meaning and of form must be taken into account when children are acquiring morphology.

## 2.2. Rote or rule?

Do children learn inflections for each word one by one? Or do they make generalisations based on some number of examples that they then apply to unfamiliar forms? That is, do they learn the form each member of every paradigm can have, and add each one to the paradigm as it is acquired? Such a task would demand **rote-learning** on a massive scale for a language like Turkish, for example, where the possible forms for a stem may be counted in the thousands. The alternative to such rote learning – where children could only learn the specific forms they happened to be exposed to – is that they analyse some sets of forms and extract from them systematic **schemas** that they can apply to any candidate stem to construct the desired form.

One of the first studies to tackle this question was carried out by Berko (1958). She presented five- and seven-year-old children with nonsense words used in specific syntactic contexts, followed by an elicitation frame, e.g., “This man likes to biff every day. Yesterday, he \_\_”, or “This is a wug. Now there are two of them. There are two \_\_”, accompanied by line drawings of the scenes or objects supposedly denoted by the nonsense words. Children were asked in this way to supply several inflections including the plural and possessive on nouns, and *-ing* and *-ed* on verbs. The results showed that children could supply appropriate inflections for ‘nouns’ and ‘verbs’ they had never heard before. They could not, therefore, be relying on any rote learning of the forms for such words. Rather, they must have been using some type of rule or schema to supply the requisite ending to express a specific meaning (Bybee & Slobin 1982). This finding is also consistent

with children's spontaneous errors, where production of such regularised forms as English *runned* (for *ran*) or *foots* (for *feet*) also give strong evidence for rule application and against simple rote learning.

At the same time, children are also influenced by the familiarity of the words they are asked to inflect in elicitation tasks. They more readily add inflections to familiar than to unfamiliar words. In Berkó's study, for instance, children succeeded in adding the appropriate allomorph of the plural /-iz/ to the noun *glass*, for *glasses*, but failed to do so with nonsense words ending in sibilants. In addition, children do better in elicitation tasks with nonsense words, when the situation itself calls for an inflected form (because it is pragmatically appropriate) than when the setting appears more test-like. So a child may fail to supply a plural inflection when asked about *wugs* ("Now there are two. There are two \_\_") but succeed when asked which pictures should be put away next ("Which pictures go next?"). Pragmatically appropriate tasks tend to offer a fuller picture of children's abilities than test settings can (Hecht 1983; see also Menn & Bernstein Ratner 2000).

As children learn the regular and irregular patterns of inflection in their language, they are tacitly building paradigms – sets of modulations that apply to particular classes of nouns or verbs. When they begin to identify such patterns, they must keep track of which affix-type goes on which stem-type. For example, in the Romance languages, children must classify verbs by conjugation in order to know which inflections apply. One might imagine that the most frequently heard forms would offer the first model for children, so if certain irregular verbs were the most frequent ones they heard, children should begin by generalising irregular forms. However, children attend not only to token frequencies but also to type frequencies, and their earliest generalisations appear to be based on the most frequent types (Guillaume 1927). In Romance, children consistently begin by choosing the first conjugation as their model when they add inflections to a verb, and so make errors of over-regularisation on verbs from other conjugations, as well on irregular verbs (Clark 1985). And in Semitic, children have to attend to the consonantal structure of the root in order to know which allomorph of an inflection to choose (e.g., Berman 1985).

Children must also store in memory what they hear from other speakers. This enables them to recognise words and phrases when they hear them on subsequent occasions and from different speakers. As a result, they store many conventional forms for comprehension that they themselves may not yet produce. The fact that they have some stored representation of such forms means that these are available when, for instance, children are asked questions about irregular verb forms. In fact, children aged three and older are willing to make judgements about who might say a particular form, adult or child, when presented with *brought* versus *bringed*. They could have both forms stored since they hear language spoken by their peers (users of *bringed*) as well as by adults (*brought*). But they also hear conventional *brought* more often and from many more speakers. Child judgements about which form is "okay" at a certain point begin to favour adult over child usage (Kuczaj 1978; Platt & MacWhinney 1983). A major issue here, as in other areas of language learning, is how children come to give up their own over-regularisations in favour of the conventional irregular forms (Bowerman 1988; Clark 1993; Clark & Chouinard 2000).

Frequency of exposure is one factor here. Since children hear conventional forms more frequently than any non-conventional ones, from the largest number of speakers, those should be the forms that they judge as more adult-like and that they themselves will eventually favour in production over earlier erroneous forms. Indeed, as children get older, it has been argued, their regularisation errors may largely reflect problems they are having in retrieving the appropriate (conventional) form from memory: A child who produces *buyed* in lieu of *bought* has failed to retrieve *bought* and so constructs a past tense form using the regular morpheme *-ed* (see Marcus et al. 1992). Speakers could store irregular forms in memory – these must be learnt by rote – but construct all regular forms when they are needed on the basis of an abstract rule or schema that captures the generalisation, (for English, here) that the past tense is marked by the addition of the morpheme *-ed*. However, other processing considerations have led researchers to argue for storage of some (not necessarily all) regular forms as well as irregular ones in addition to being able to construct regular forms by rule, rather than proposing two quite distinct

types of processing with separate domains (cf. Frauenfelder & Schreuder 1991; Maratsos 2000).

### 2.3. Agreement

One major function of inflectional morphology is to mark **agreement** in a language, to indicate which elements (or groups of elements) belong together and which play a specific grammatical role. For example, gender marking on nouns, articles, and adjectives can indicate which of these elements belong together within an utterance. Case marking on nouns typically distinguishes the grammatical functions of noun phrases as subject, direct object, or indirect (dative) object of the verb. And number agreement on nouns and verbs, along with person agreement, can also indicate which noun phrase is the subject. Languages differ considerably in the extent to which they rely on inflections for these purposes, and hence in the patterns of agreement they exhibit. Acquisition of the ways inflections function syntactically in marking **grammatical concord** has been rather less studied than the acquisition of isolated forms. Available data suggest that children begin working on these systems (e.g., article-noun agreement in French or German, subject-verb and noun-adjective agreement in Hebrew) before age two though children may make many errors both of omission and commission on the way to adult mastery.

In case-marked languages, children's earliest words are case-marked, but they do not vary in form from one context to the next. In other words, the case inflection is treated initially as if it were part of the word. Very soon, however, children begin to contrast two (or more) case-markings, and to use them consistently. In Russian for example, children begin to use accusative marking on nouns following transitive verbs before age two (Gvozdev 1961). Nominative and accusative marking are usually the first two cases to be distinguished (MacWhinney 1978; Smoczynska 1985). Early case marking often interacts with word order, another device for marking grammatical relations in addition to pragmatic discourse structure. But although it is possible that children might rely on word order to mark grammatical relations prior to acquisition of case-marking, the data so far suggest that case marking emerges at least as early as consistent word order (Slobin & Bever 1982).

Case on its own, though, does not provide evidence for agreement. What is needed is subject-case combined with an appropriate form of the verb, agreeing in person and number. This form of agreement, like others, begins to appear between age two and three (e.g., Clahsen 1986; Meisel 1986; Tracy 1984; Vihman 1999, ed.). In Hebrew, for example, children make a number of errors in noun-adjective agreement in the first few months after age two, and often produce masculine singular adjectives with feminine singular nouns. These agreement errors then decrease in number and have typically vanished by age three. What appears critical in the mastery of agreement here is children's prior acquisition of plural inflections for nouns. Once those are in place, Hebrew-speaking children can use the same plural ending on noun and adjective to mark agreement, e.g., *ha-xaruz-im ha-gadol-im nofl-im* 'the-bead-PL the-big-PL fall-PL (the big beads are falling)' with a masculine plural head noun, or *ha-kubiy-ot ha-gadol-ot nofl-ot* 'the-cube-PL the-big-PL fall-PL (the big cubes are falling)', with a feminine plural head noun. That is, children rely on the phonological similarity of the inflections across word classes in marking agreement (Berman 1985; Levy 1983).

Children learning other languages also seem to rely on phonological cues to gender and number marking where possible. In Bantu languages like Sesotho, children must learn that every noun belongs in one of 14 classes, each marked by a prefix. Class prefixes (some of them identical in form to the noun prefixes) must also be added to verbs for both subject and object agreement, to adjectives, and to demonstrative, personal, possessive, and relative pronouns. Noun class markers emerge rather slowly from age two on, with both omissions and some over-use of class 9/10 prefixes (probably because many nouns used at this age come from this class). Adjectives and demonstrative, possessive, and personal pronouns are in use by age two, but often appear with the wrong class marker. By age three, though, children mark agreement with considerable consistency and accuracy. They are aided in this, it has been suggested, by the phonological similarity of class markers to the markers for each noun class (Demuth 1988).

In Romance languages like French, children again make some errors in article-noun agreement once they begin to produce articles, but by age three such errors no longer

occur. It takes French-speaking children slightly longer to master agreement for gender on adjectives and on pronouns. But by about age six, they are consistent in making articles and nouns agree, regardless of the natural gender of the referents. They appear to rely primarily on phonological correlations of word form (in the noun) and article in gender agreement (Clark 1985; Karmiloff-Smith 1979). Reliance on phonological rather than semantic cues also seems to be the norm for Slavic (Gvozdev 1961; Smoczynska 1985) and for German (Mills 1985).

Use of inflections to mark agreement emerges as soon as children begin to produce the pertinent morphology and the relevant syntactic combinations (i.e., subject-verb, article-noun, noun-adjective, etc.). This suggests that they have already represented those parts of the system in memory, and have done so on the basis of the language spoken to them in the community. It also suggests that children do not first use word order, for example, to mark grammatical relations and then switch to inflections. Rather, they will use the agreement patterns conventional in the language they hear, from the start. They may mark them imperfectly, but as they produce more of the relevant inflections, these appear to emerge without significant error. In fact, reliance on case versus word order as clues to grammatical relations varies with the typology of the target language (see Bates et al. 1984; Slobin 1985, ed.; 1992, ed.; 1997, ed.; Slobin & Bever 1982).

#### 2.4. Linear ordering of morphemes

Children appear to make no errors in the **ordering** of bound morphemes within words. When they add inflections to a base or stem, they do so in the right order. At first they add only single inflections, so the only error possible would be to add suffixal inflections as prefixes or prefixal ones as suffixes. But children apparently never make this type of error. Suffixes are added to the end, prefixes to the beginning, in the appropriate order.

The only morpheme order error attested for English occurs in a setting that adult speakers tend to avoid: the formation of the comparative of adverbial forms like *prickly* or *dirtily*. Instead of the adult solution of using *more*, as in *more prickly*, young English-speaking children sometimes try out forms like *prickerly* where they add the comparative degree marker before the adverbial *-ly*, and thus appear to order an inflection be-

fore a derivational morpheme. (However, there is some dispute about whether degree *-er* is an inflection at all.)

Many languages license strings of inflectional morphemes, on the end or the beginning of the stem form. Here children have many more opportunities for errors in order. Yet again, studies of agglutinating and polysynthetic languages such as Turkish, Mohawk, K'iche Maya, Greenlandic Eskimo, and Walpiri report no instances of order errors (cf. Aksu-Koç & Slobin 1985, ed.; Miethun 1988; Slobin 1985, ed.; 1992, ed.; 1997, ed.). This striking absence of error within words is consistent with the operating principle: "Keep the order of morphemes within a word constant across the various environments in which that word can occur" (Slobin 1985: 1231). Children's adherence to morpheme order within words contrasts to some extent with the freedom they use in ordering their words. Use of word order to mark grammatical relations, for example, emerges only slowly. Initially, one- and even two-year-olds typically use word order as a pragmatic device to mark information as given versus new, and only later learn that word order may, in some languages, mark such relations as 'subject-of' (e.g., Bates 1976). Although children rarely make errors in their ordering of constituents after age three to four, they may still make word-order errors within constituents, as when they combine a quantifier like English *some* with a demonstrative and a noun (e.g., *these beads*) to produce *\*these some beads* instead of *some of these beads*.

Children acquire languages of quite different morphological types in similar fashion. In English, they learn a relatively lean system of inflections combined with considerable irregularity in the established vocabulary, but in Turkish they master a rich bound morphology that is highly regular in form and transparent in meaning. And in Hebrew, they learn a synthetic system with rich inflectional marking plus an elaborate system of agreement. In all three languages, children's acquisition of inflections plays a central part in early grammatical development and simple clause structure. What they must add subsequently is mastery of subregularities or minor lexical schemas that are limited to specific sets of items, conjugations, or declensions, as well as lexical exceptions – the unpredictable, idiosyncratic irregularities of the lexicon.

### 3. Derivation

The acquisition of **derivational morphology** has, by comparison, been relatively neglected even though it is relevant both to grammatical form and to lexical structure. Nonetheless, there are extensive corpora of novel word formations available for English, French, German, Hebrew, Hungarian, and Polish (Clark 1993). Innovative words that rely on derivational morphemes are an important source of evidence that children have carried out some morphological analysis and are not simply using ready-made forms from the conventional lexicon. The data available consist of spontaneous innovations in the lexicon, elicited innovations, and elicited glosses of the possible meanings of novel words. Several domains have been studied in detail for the association of a derived morpheme with a specific meaning, e.g., agent and instrument nouns, compound nouns and other nominals, and denominal verbs.

#### 3.1. Conventionality

A major task in language acquisition is to master the conventional lexicon. But children often wish to talk about things for which they have not yet acquired the conventional terms. When this occurs, they may coin a word on the spot to serve their purpose (cf. Art. 80). The conventional lexicon plays a critical role in this because it displays the range of possible word-structures that children can make use of in building word forms to carry new meanings.

Children fill the gaps in their lexicon as needed. For adults, such gaps may be momentary (when speakers fail to retrieve a word on the tip of the tongue and so fill in with a coinage instead) or permanent (when speakers come across a true gap in the lexicon where there is no conventional word for the pertinent meaning). The same distinction exists for children, but their still-limited knowledge of the lexicon may lead them to coin many words where conventional terms already exist. These coinages are effectively ones that, for adult speakers, would be pre-empted by the conventional terms.

**Conventional terms** take priority over **innovative** ones. Speakers expect that if there is a conventional term for the expression of a particular meaning, it must be used for that meaning. If it is not, then the person talking must have some other meaning in mind. As a result, innovative words cannot simply be

produced in place of conventional ones. To be used at all, they must contrast with them in meaning (Clark 1990; 1993; Clark & Clark 1979).

For children, there are many more gaps to fill because their conventional vocabulary is still so small. The result is a large number of illegitimate **coinages** – innovations that are pre-empted (or blocked) by the existence of conventional words with just those meanings – as well as many legitimate ones. Both types of innovation allow us to look in some detail at what children at different stages know about word formation in their first language.

#### 3.2. Structural options for word formation

What do children know, at each stage, about possible form-meaning combinations in their language? **Word formation** is a critical domain for answering this question. The major options available in a language typically include derivation (with or without affixes) and compounding (Anderson 1985 b). When children innovate, they show us what they know about the forms of possible words in the language. This holds true both at the phonotactic level (where children never attempt to impose impossible consonant clusters or syllable structures) and at the level of morpheme combination. That is, children take account of typological features of their first language from a very early point in acquisition, and are sensitive to the kinds of word formation options favoured in the language they are acquiring.

English-speaking children, for example, rely heavily on **zero-derivation** for forming new verbs from as young as age two (e.g., *[to] scale* for ‘weigh’, *[to] oar* for ‘row’; cf. Clark 1982), but Hebrew-speaking children do not. Children acquiring Germanic languages rely extensively on **compounding** to form new nouns (e.g., English *car-smoke* ‘exhaust’, *cup-egg* ‘boiled egg’), again from age two on, but children acquiring Romance languages do not. **Derivational affixes** (e.g., *caker* ‘someone who makes cakes’, *sweeper* ‘broom’; Clark 1993) tend to emerge slightly later, so in languages that favour derivation over compounding, children may not produce many coinages until age three or later. In general, children acquire suffixes before prefixes (Kuczaj 1979), for both inflection and derivation. Other options such as blending or clipping emerge later still, and are usually favoured more by adults than children (Ravid 1990; 1995). Compounding combined

with derivation emerges later than compounding alone. Just as with inflections, children begin with the stem and then add the modulation of meaning, whether through affixation, compounding, or some other means.

Children's earliest choices of word-forms for new meanings appear to be governed by two main factors in addition to typology: (a) the **simplicity** of the form used, with a premium on the fewest changes possible in the base form, and (b) the **transparency** of the meaning expressed, in relation to the meaning of the base. The youngest children favour coinages with zero-derivation (as in the English innovative verbs *[to] sand* 'grind into small pieces' or *[to] rug* 'vacuum a rug', both from two-year-olds) or simple noun + noun compounding as in *plate-egg* 'fried egg' or *baby-towel* 'face-cloth' (also from two-year-olds). The compounds are also transparent to the degree that the head noun denotes the kind of category being talked about (here an egg and a towel). As children analyse the meanings of affixes, their coinages become increasingly transparent: they add the appropriate affix to a base, and use English *-er*, for example, to mark the agent or actor (as in *rainer* '(imaginary) person who makes rain go away' or *crayoner* 'someone who makes pictures with crayons') (Clark & Hecht 1982; Clark & Berman 1984).

In short, children's lexical innovations show that they know a lot about the regularities of word formation in their language, even though they may not yet know the conventional terms for the referents they talk about.

### 3.3. Productivity

Once children have acquired several options related in meaning for forming new words, they have to choose which to use on particular occasions. Their choices appear to be guided by which option is the most **productive** for adult speakers in the same community (cf. Art. 33). Children rely more frequently on the device adults favour than on other options also available in the language. For example, in English, four- and five-year-olds favour agentive *-er* over *-ist* and *-ian*, and, in Hebrew, children favour suffixal *-an* to mark agents over the other options available (Clark & Berman 1984; Clark & Cohen 1984).

Previous research also shows that when adult preferences for one option over others change, children's choices of forms follow the

adults'. Children must therefore be tracking the frequencies of different form types in the input they hear. Shifts in what is productive for adult speakers are paralleled by shifts in children's preferences in the forms chosen for lexical innovations (Chmura-Klekotowa 1970).

Productivity in word formation presents a rather different face to productivity in inflection. In word-formation, some slots in a paradigm may be taken by existing words with the appropriate meanings, so use of the otherwise regular form is pre-empted by the existence of another form with the pertinent meaning. The degree of regularity overall is therefore somewhat lower than in inflectional paradigms. At the same time, the forms of innovative lexical items, coined to contrast with existing form-meaning combinations, offer strong evidence for the relative productivity of specific affixes or compounding patterns. In word formation, then, new formations, whether nonce uses or words that eventually enter the conventional lexicon, provide critical evidence for productivity (Berman 1987; 2000).

### 3.4. Order in word formation

When children coin words, they often make errors, but they rarely, if ever, make mistakes in how they order the morphemes within a new word. That is, derivational morphemes are consistently added to the base or stem before children add any inflections. And where they need to add more than one derivational or inflectional affix to a single stem, these too are placed in the correct order, as in *sweeper-s* (verb + *er* + *s*) for the machines used to sweep streets. This attention to morpheme-order within the word is all the more striking in the face of order errors elsewhere.

Children do make frequent order errors in compounds in some languages. These errors appear to be largely predictable on typological grounds and occur where there is a mismatch of head-modifier ordering in different constructions. For example, in English, the head of the verb phrase is the phrase-initial verb (e.g., *throw* in *throw the ball*). But in compound nouns, the head is in final position (e.g., *climber* in *cliff-climber*). Where the compound contains a verb stem (*climb*), children go through a stage of mis-ordering the modifier and head to produce compounds like a *climb-cliff* or a *climber-cliff* for 'someone who climbs cliffs'. But in languages with the same head-modifier order across con-

structions, children never make such errors (Clark 1993; Clark & Berman 1987).

Children are consistent in adhering to the ordering of morphemes within words. They begin with the base or stem form, add the relevant derivational affixes, and then add any inflections needed. Within words, children rarely if ever make errors in the linear ordering of affixal morphemes.

#### 4. Summary

After initial acquisition of single words, children aged two to three gain command of the syntax of simple clauses. They typically master case-marking with prepositions first, and then inflectional marking of agreement for number, gender, and person. They begin to add the modulations of case, tense, and other inflections before their second birthday, but make systematic errors along the way. Derived forms may also begin to appear between age two and three as children turn to coinage to fill gaps in their vocabulary. But the extent of such early reliance on derivation is limited. This is particularly clear in a language like Hebrew, where nearly all the word-formation processes require use of affixal morphology (Berman 2000). In languages like English, where zero derivation (without affixes) is also an option for word formation, children take advantage of it as young as age two to coin verbs from nouns. They may also begin to use derivational affixes that do not demand any change in the shape of the word stem (e.g., English *-er* in agent nouns like *jumper*) soon afterwards. However, derivational affixes appear to enter in large numbers only after age three. In both inflection and derivation, children regularise. They treat irregular forms as if they were regular, adopting those paradigms with the largest number of members (types) as their templates. The learning of lexical exceptions, whether inflectional or derivational, occurs only after mastery of regular forms.

The next year, from age three to four, is critical in word formation. It is then that children analyse the internal structure of words and become capable of coining new words with derivational affixes. They do this with a range of form-types consistent with the options in the target language (Clark 1993). Indeed, the ease with which children master morphological structure depends heavily on the typology of the language, for example

whether it is synthetic, polysynthetic, isolating, or agglutinating. It also depends to a large extent on the nature of the meaning-to-form mappings for inflectional and derivational paradigms. Finally, some lexical options in word formation are more productive than others, and the more productive they are, the earlier they tend to be acquired. In English and German, for instance, compounding appears much earlier than in Hebrew or French. Other options may be acquired late because they are formally very complex, as is the case for vowel-laxing and stress-shifts with derivation in the Latinate vocabulary of English, or for passive forms and certain derived nominals in Hebrew. Finally, some word-formation types may be used only in formal or written registers, and there, the patterns of usage account for late acquisitions in word-formation (see Jisa et al. 2002).

#### 5. References

- Aksu-Koç, Ayhan & Slobin, Dan I. (1985), "The Acquisition of Turkish". In: Slobin (ed.), *Vol. I*, 839–878
- Anderson, Stephen R. (1985 a), "Inflectional Morphology". In: Shopen (ed.), 150–201
- Anderson, Stephen R. (1985 b), "Typological Distinctions in Word Formation". In: Shopen (ed.), 3–56
- Bates, Elizabeth (1976), *Language and Context*. New York: Academic Press
- Bates, Elizabeth & MacWhinney, Brian & Caselli, Cristina & Devescovi, Antonella & Natale, Francesco & Venza, Valeria (1984), "Crosslinguistic Study of the Development of Sentence Interpretation Strategies". *Child Development* 55, 341–354
- Berko, Jean (1958), "The Child's Learning of English Morphology". *Word* 14, 150–177
- Berman, Ruth A. (1985), "Acquisition of Hebrew". In: Slobin (ed.), *Vol. I*, 255–371
- Berman, Ruth A. (1987), "Productivity in the Lexicon: New-word Formation in Modern Hebrew". *Folia Linguistica* 21, 425–461
- Berman, Ruth A. (1990), "Acquiring an (S)VO Language: Subjectless Sentences in Children's Hebrew". *Linguistics* 28, 1135–1166
- Berman, Ruth A. (2000), "Children's Innovative Verbs versus Nouns: Structured Elicitations and Spontaneous Coinages". In: Menn & Bernstein Ratner (eds.), 69–93

- Berman, Ruth A. & Armon-Lotem, Sharon (1996), "How Grammatical are Early Verbs?". In: Martinet, Claire (ed.), *Actes du Colloque Internationale sur l'Acquisition de la Syntaxe en Langue Maternelle et en Langue Etrangère. Annales Littéraires de l'Université de Franche-Comté*: No. 631, 17–60
- Bowerman, Melissa (1988), "The 'No Negative Evidence' Problem: How do Children Avoid Constructing an Overly General Grammar?" In: Hawkins, John A. (ed.), *Explaining Language Universals*. Oxford (UK): Blackwell, 73–101
- Brown, Roger (1973), *A First Language: The Early Stages*. Cambridge/MA: Harvard University Press
- Bybee, Joan (1985), *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins
- Bybee, Joan & Slobin, Dan I. (1982), "Rules and Schemas in the Development and Use of the English Past Tense". *Language* 58, 265–289
- Chmura-Klekotowa, Maria (1970), "Odbicie tendencji slowotwórczych języka polskiego w neologizmach dzieci". *Prace Filologiczne* 20, 153–159
- Clahsen, Harald (1986), "Verb Inflections in German Child Language: Acquisition of Agreement Markings and the Functions they Encode". *Linguistics* 24, 79–121
- Clark, Eve V. (1982), "The Young Word Maker: A Case Study of Innovation in the Child's Lexicon" In: Wanner, Eric & Gleitman, Lila R. (eds.), *Language Acquisition: The State of the Art*. Cambridge (UK): Cambridge University Press, 390–425
- Clark, Eve V. (1985), "The Acquisition of Romance, with Special Reference to French". In: Slobin (ed.), *Vol. I*, 687–782
- Clark, Eve V. (1990), "On the Pragmatics of Contrast". *Journal of Child Language* 17, 417–431
- Clark, Eve V. (1993), *The Lexicon in Acquisition*. Cambridge (UK): Cambridge University Press
- Clark, Eve V. & Berman, Ruth A. (1984), "Structure and Use in the Acquisition of Word Formation". *Language* 60, 542–590
- Clark, Eve V. & Berman, Ruth A. (1987), "Types of Linguistic Knowledge: Interpreting and Producing Compound Nouns". *Journal of Child Language* 14, 547–567
- Clark, Eve V. & Chouinard, Michelle M. (2000), "Enoncés enfantins et reformulations adultes dans l'acquisition du langage". *Langages* 140, 9–23
- Clark, Eve V. & Clark, Herbert H. (1979), "When Nouns Surface as Verbs". *Language* 55, 767–811
- Clark, Eve V. & Cohen, Sophia R. (1984), "Productivity and Memory for Newly-Formed Words". *Journal of Child Language* 11, 611–626
- Clark, Eve V. & Hecht, Barbara F. (1982), "Learning to Coin Agent and Instrument Nouns". *Cognition* 12, 1–24
- Clark, Herbert H. & Clark, Eve V. (1977), *Psychology and Language*. New York: Harcourt Brace Jovanovich
- Demuth, Katherine A. (1988), "Noun Classes and Agreement in Sesotho Acquisition". In: Barlow, Michael & Ferguson, Charles A. (eds.), *Agreement in Natural Languages: Approaches, Theories, Descriptions*. Stanford, California: Center for the Study of Language and Information (CSLI), 305–341
- Dromi, Esther & Berman, Ruth A. (1982), "A Morphemic Measure of Early Language Development: Data from Modern Hebrew". *Journal of Child Language* 9, 403–424
- Frauenfelder, Uli H. & Schreuder, Robert (1991), "Constraining Psycholinguistic Models of Morphological Processing and Representation: The Role of Productivity". In: Booij, Geert & van Marle, Jaap (eds.), *Yearbook of Morphology 1991*. Dordrecht: Kluwer, 165–183
- Greenberg, Joseph H. (1966), *Language Universals*. The Hague: Mouton
- Guillaume, Paul (1927), "Le développement des éléments formels dans le langage de l'enfant". *Journal de Psychologie* 24, 203–229
- Gvozdov, Aleksandr N. (1961), *Voprosy Izucheniya Detskoj Rechi*. Moskva: Idz. Akademii Pedagogicheskikh Nauk RSFSR
- Hecht, Barbara F. (1983), *Situations and Language: Children's Use of Plural Allomorphs in Familiar and Unfamiliar Settings*. Ph.D. dissertation, Stanford University
- Jisa, Harriet & Baruch, Elisheva & Berman, Ruth & Reilly, Judy & Rosado, Elisa & Verhoeven, Ludo (2002), "Passive Voice Constructions in Written Texts: A Crosslinguistic Developmental Study". In: Berman, Ruth A. & Verhoeven, Ludo (eds.), *Cross-Linguistic Perspectives on the Development of Text-Production Abilities in Speech and Writing. Part 2* (Special issue of *Written Language and Literacy* 5:2), 163–181
- Karmiloff-Smith, Annette (1979), *A Functional Approach to Child Language: A Study of Determiners and Reference*. Cambridge (UK): Cambridge University Press
- Kuczaj, Stan A. (1978), "Children's Judgements of Grammatical and Ungrammatical Irregular Past Tense Verbs". *Child Development* 49, 319–326
- Kuczaj, Stan A. (1979), "Evidence for a Language Learning Strategy: On the Relative Ease of

- Acquisition of Prefixes and Suffixes". *Child Development* 50, 1–13
- Levy, Yonata (1983), "The Acquisition of Hebrew Plurals: The Case of the Missing Gender Category". *Journal of Child Language* 10, 107–121
- MacWhinney, Brian (1978), *The Acquisition of Morphophonology*. Chicago/IL: University of Chicago Press (Monographs of the Society for Research in Child Development 43)
- Maratsos, Michael P. (2000), "More Overregularizations after All: new Data and Discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu". *Journal of Child Language* 27, 183–212
- Marcus, Gary & Ullman, Michael & Pinker, Steven & Hollander, Michelle & Rosen, T. John & Xu, Fei (1992), *Overregularization*. Chicago/IL: University of Chicago Press (Monographs of the Society for Research in Child Development 57)
- Meisel, Jürgen (1986), "Word Order and Case Marking in Early Child Language: Evidence from Simultaneous Acquisition of Two First Languages: French and German". *Linguistics* 24, 123–183
- Menn, Lise & Bernstein Ratner, Nan (2000, eds.), *Methods for Studying Language Production*. Mahwah/NJ: Erlbaum
- Mikes, Melanie (1967), "Acquisition des catégories grammaticales dans le langage de l'enfant". *Enfance* 20, 289–298
- Mills, Anne (1985), "The Acquisition of German". In: Slobin (ed.), *Vol. I*, 141–254
- Mithun, Marianne (1988), "The Acquisition of Polysynthesis". *Journal of Child Language* 16, 285–312
- Omar, Margaret (1973), *The Acquisition of Egyptian Arabic as a Native Language*. The Hague: Mouton
- Peters, Ann M. & Menn, Lise (1993) "False Starts and Filler Syllables: Ways to Learn Grammatical Morphemes". *Language* 69, 742–777
- Platt, Carol B. & MacWhinney, Brian (1983), "Error Assimilation as a Mechanism in Language Learning". *Journal of Child Language* 10, 401–414
- Pléh, Csaba (1998), "Early Spatial Case Markers in Hungarian Children". In: Clark, Eve V. (ed.), *The Proceedings of the 29th Child Language Research Forum [1997]*. Stanford/CA: Center for the Study of Language and Information (CSLI), 211–219
- Pye, Clifton (1990), "The Acquisition of Ergative Languages". *Linguistics* 28, 1291–1330
- Ravid, Dorit (1990), "Internal Structure Constraints on New Word Formation Devices in Modern Hebrew". *Folia Linguistica* 34, 289–347
- Ravid, Dorit (1995), *Language Change in Child and Adult Hebrew*. Oxford (UK): Oxford University Press
- Shopen, Timothy (1985, ed.), *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*. Cambridge (UK): Cambridge University Press
- Slobin, Dan I. (1973), "Cognitive Prerequisites for the Acquisition of Grammar". In: Ferguson, Charles A. & Slobin, Dan I. (eds.), *Studies of Child Language Development*. New York: Holt, Rinehart & Winston, 173–208
- Slobin, Dan I. (1985, ed.), *The Crosslinguistic Study of Language Acquisition, Vol. I–II*. Hillsdale/NJ: Lawrence Erlbaum Associates
- Slobin, Dan I. (1992, ed.), *The Crosslinguistic Study of Language Acquisition. Vol. III*. Hillsdale/NJ: Lawrence Erlbaum Associates
- Slobin, Dan I. (1997, ed.), *The Crosslinguistic Study of Language Acquisition, Vol. IV–V*. Hillsdale/NJ: Lawrence Erlbaum Associates
- Slobin, Dan I. & Bever, Thomas G. (1982), "Children Use Canonical Sentence Schemas: A Crosslinguistic Study of Word Order and Inflections". *Cognition* 12, 229–265
- Smoczyńska, Małgorzata (1985), "The Acquisition of Polish". In: Slobin (ed.), *Vol. I*, 595–686
- Tracy, Rosemarie (1984), "Fallstudien: Überlegungen zum Erwerb von Kasuskategorie und Kasusmarkierung". In: Czepluch, Hartmut & Janssen, Hero (eds.), *Syntaktische Struktur und Kasusrelation*. Tübingen: Niemeyer, 271–313
- Veneziano, Edy & Sinclair, Hermine (2000), "The Changing Status of 'Filler Syllables' on the Way to Grammatical Morphemes". *Journal of Child Language* 27, 461–500
- Vihman, Marilyn M. (1999, ed.), *First Steps in Morphological and Syntactic Development: Cross-Linguistic Evidence*. Special issue, *International Journal of Bilingualism* 3

Eve V. Clark, Stanford (U.S.A.)  
Ruth A. Berman, Tel Aviv (Israel)

## 166. Second language acquisition

1. Morpheme-acquisition order
2. Inherent inflectional morphology
3. Contextual inflection
4. Case marking
5. Derivation
6. Compounding
7. References

### 1. Morpheme-acquisition order

The first developmental studies in second-language (L2) acquisition dealt with the acquisition of a given set of morphemes in L2 English. In particular, it was studied in which order these morphemes were acquired. The reason for carrying out these studies was a discussion on the nature of L2 acquisition in the mid-seventies. The central issue of this discussion was the role of the mother tongue (L1). Dulay & Burt (1974; 1975) argued that L2 acquisition, like L1 acquisition, was essentially a process of “creative construction”. Given their assumption that L2 learners relying on their L1 could only do so if they were applying old habits, they compared L1 and L2 acquisition in order to demonstrate that L1 was not a relevant factor.

In order to test their claim on the irrelevance of L1, Dulay & Burt (1974; 1975) compared morpheme acquisition order in the L2 English of children with differing language backgrounds. It was found that the order of acquisition was independent of the L1. Therefore, since old habits did not play a role in L2 acquisition, it was concluded that L2 acquisition, just like L1 acquisition, was a process of creative construction. A similar conclusion was reached by Bailey et al. (1974) leading to what Krashen (1977) called the ‘Natural Order’ for L2 English represented in Fig. 166.1.

The studies by Dulay & Burt and Krashen were the first of many similar studies on the acquisition of morphemes in L2 English by native speakers of differing L1s (e.g. Larsen-Freeman 1975 and Andersen 1978). In addition, processes of morpheme acquisition in an L2 were compared with morpheme acquisition in other acquisitional and developmental contexts. For example, L2 acquisition was compared with processes of pidginization and creolization as well as diachronic change in language. The results of all the morpheme studies point to a common or nat-

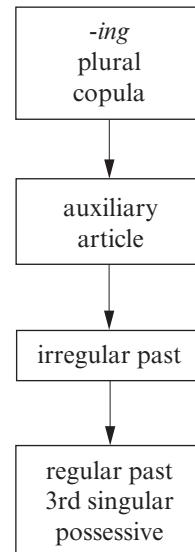


Fig. 166.1: Krashen’s (1977) ‘Natural Order’ for L2 English

ural order of acquisition which is seen as evidence for L2 learning as a process of creative construction. Most of these studies, however, have not paid a great deal of attention to what mental mechanisms could explain the particular orders found.

### 2. Inherent inflectional morphology

As proposed in Booij (1994), two types of inflection should be distinguished, **inherent** and **contextual inflection**. Inherent inflectional morphology as opposed to contextual inflectional morphology serves to express a particular semantic content. Examples of inherent inflection are tense or aspect with verbs and number with nouns.

#### 2.1. Past tense

The acquisition of past tense morphology in English has become an area of heated debate between adherents of connectionist and symbolic models of language acquisition. The question is whether or not there is an essential difference between regular and irregular forms with respect to acquisition, storage, and production. Within a **symbolic framework**, linguistic knowledge of the regularly formed past tense form in English is ade-

quately represented by a morphological rule, whereas irregular forms are stored as associative knowledge (Pinker & Prince 1988). Within the **connectionist view**, on the other hand, there is no principled difference in linguistic status between regular and irregular forms. All types of past morphology are based on associative relations (Rumelhart & McClelland 1986). Notions such as rules and **overgeneralizations** are only to be seen as epiphenomena (Bates et al. 1983).

L2 acquisition of regular and irregular verbs was studied by Lalleman et al. (1997). They wanted to find out whether or not it is possible for advanced learners of L2 Dutch to implicitly acquire the same kind of unconscious linguistic knowledge as L1 speakers. Based on the idea that production by rule is not affected by **frequency** of occurrence, whereas **lexical retrieval** is, they hypothesized that production time for irregular verbs is faster with familiar forms than with unfamiliar forms.

In order to investigate this hypothesis they carried out a production time experiment for past tense formation. The results from this experiment were similar to those from the experiments on L1 English. First, differences in production time were found between regular and irregular past tense forms. Furthermore, high-frequency irregular past tense forms were produced faster than their low-frequency counterparts, while with regular verbs there is no effect of past tense frequency on the production time, neither for L1 nor for L2 speakers. Obviously, native speakers were faster than L2 learners and differences in production time between regular and irregular forms were larger with L2 learners.

Lalleman et al. (1997) also compared production time of prototypical and atypical regular verbs. Prototypical regular verbs are those with a short vowel (except *i*, *e*) plus a geminate consonant in the infinitive form. They are prototypically regular because, as far as their morphology is concerned, there are no irregular counterparts. Examples are Dutch *knallen* 'to bang', *bukken* 'to stoop', *klappen* 'to clap' (an exception is *vallen* 'to fall'). Prototypical regular verbs are produced more quickly and with fewer errors than other regular verbs. Comparing the percentages of irregularization errors in 11 prototypical regular verbs such as *knallen* 'to bang', *bukken* 'to stoop', *klappen* 'to clap', and in 11 atypical regular verbs such as

*kleven* 'to stick' and *huilen* 'to cry', they found that L2 learners irregularize past tense forms in prototypical regular verbs less often (0,3 percent) than in atypical regular verbs (12,0 percent). In this respect L2 learners behave differently from native speakers. This L2 learner behaviour is probably due to the way in which morphological rule formation takes place. Rule formation in morphology depends on type frequency, i.e. the frequency with which different lexical items share the same morphological properties. Since there are classes of high-frequency irregular verbs which share morphological properties with regular verbs, L2 learners are likely to infer that these verbs are to be related on the basis of a particular morphological rule. As long as learners do not know that this apparent regularity is just an epiphenomenon of the lexical properties of an arbitrary list of verbs, L2 learners may treat regular verbs with similar morphology the same way.

## 2.2. Aspect

In child language acquisition, studies by Bronckart & Sinclair (1973), Antinucci & Miller (1976), and Bloom et al. (1980) started a debate on whether past morphology is initially used for coding past tense or perfective aspect. In L1 acquisition this issue has been addressed in connection with what is called the Defective Tense Hypothesis (Weist 1986). In studies on second language acquisition the same question of 'aspect before tense' or 'tense before aspect' became relevant with respect to the development of interlanguage morphology.

With respect to the acquisition of L2 Spanish by native speakers of English, Andersen (1991) found a developmental sequence for encoding tense and aspect with past morphology. Given a semantic classification of verbs into the four verb categories 'states', 'activities', 'telic events', and 'punctual events', Andersen discovered that L2 learners approach the L2 system by linking aspectual marking and lexical meaning: "[t]he Spanish Preterit and Imperfect are initially interpreted as [redundant] markers of inherent aspect" (Andersen 1991: 319). This results in a developmental process in which the Imperfect is first used for states (*tenía* 'had') and then also for activities (*jugaba* 'played'), whereas the Preterit is first used for punctual events (*se partió* 'broke in two') and later for telic events (*enseñó* 'taught x to y') as well (see Fig. 166.2).

imperfect →			← preterit
states	activities	telic events	punctual events
<i>tenía</i> ‘had’	<i>jugaba</i> ‘played’	<i>enseñó</i> ‘taught x to y’	<i>se partió</i> ‘broke in two’

Fig. 166.2: Aspectual/lexical meaning in L2 Spanish (Andersen 1991: 314)

states	activities	telic events	punctual events
<i>tenía</i> ‘had’	<i>jugaba</i> ‘played’	<i>enseñó</i> ‘taught x to y’	<i>se partió</i> ‘broke in two’
<i>tuvo</i> ‘had’	<i>jugó</i> ‘played’	<i>enseñaba</i> ‘taught x to y’	<i>se partía</i> ‘broke in two’

Fig. 166.3: Grammatical aspect in Spanish (Andersen 1991: 314)

Subsequently, learners gradually also acquire the use of the Imperfect with telic and punctual events and the use of the Preterit with activities and states (see Fig. 166.3). This means that in the end they will have acquired the variable use of past tense morphology as a device for marking grammatical aspect according to the perspective the speaker may take with respect to the temporal course of the situation.

Giacalone Ramat (1995) observed a systematic opposition between simple present and past participle morphology in the initial stages of L2 Italian of native speakers of Chinese, Eritrean, Iranian, English, German, and French. An example from Giacalone Ramat (1995) is given in (1).

- (1) *Poi aperta la bottiglia... poi*  
 then open:PAST.PART the bottle then  
*bevi vino.*  
 drink:3.SG wine  
 ‘After he opened the bottle, he drinks wine.’

Giacalone Ramat argued that this opposition has to be interpreted in terms of aspect: the simple present form or the infinitive encodes imperfective aspect, whereas the bare past participle indicates perfect aspect.

Another observation by Giacalone Ramat indicates that L2 learners of Italian initially encode the aspectual opposition imperfective aspect vs. perfect aspect lexically. This explains why in many cases punctual and telic verbs are marked with a past participle form (such as *arrivato* ‘arrived’, *lasciato* ‘left’, *trovato* ‘found’), whereas durative and stative

verbs (such as *stare* ‘stay’) receive imperfect marking. An example is given in (2).

- (2) *Pensi così + + quanto*  
 think:1.SG so if  
*trovato molto bene + stare*  
 find:PAST.PART very well stay:INF  
*qui + va bene.*  
 here go:3.SG well  
 ‘I think so: If I feel O.K. in Italy, I’ll stay here.’

With respect to the acquisition of L2 German, Dutch, and English, it has been noted that, independently of the L1, tense is acquired before aspect. This results from research by Dietrich et al. (1995) on L2 German, by Housen (1993) on L2 Dutch and by Vogel (1987) on L2 English. These authors draw the conclusion that their L2 learners first distinguish between past and non-past. Andersen (1991), however, concludes from a number of studies on the acquisition of L2 English, that here past inflection is first used only with punctual verbs (*left*), then also with telic verbs (*taught*), subsequently for activities too (*ran*), and finally also with verbs referring to states (*had*).

### 2.3. Plural

For L2 learners, plural marking in German is difficult to acquire due to the fact that the German plural system consists of nine plural markers. Studies dealing with L2 learners' plural marking in German are Köpcke (1987) and Wegener (1994; 1995). In Köpcke (1987), plural marking on pseudowords is experimentally investigated in 40 adult native

speakers. The data are compared with L2 learners' data from 80 English speaking American students of German who were studied by Phillips & Bouma (1980). Wegener collected data from two groups of children between 7 and 9 years of age: 4 Turkish children participating in bilingual classes and 4 *Aussiedler* children (3 Polish and 1 Russian) going to transition or regular classes. Both studies focused on processes of rule formation and thus on processes underlying overgeneralization.

First, Wegener found that regularizations are formed with *-Ø*, *-e*, *-(e)n*, *-s*, and "practically never with *-er* and the Umlaut" (Wegener 1994: 277). Furthermore, "the *-(e)n* plural is the most frequently overgeneralized plural" (Wegener 1994: 279). With respect to the order in which these overgeneralizations were acquired, it was found that first *-(e)*, then *-(e)n*, and finally *-s* was learnt.

Köpcke's analysis of the L2 learners' pseudoword production leads to similar findings. The plural morphemes most frequently used were *-(e)n* and *-e*. In one-syllable words with no particular morphological properties, *-(e)n* and *-e* are used most frequently with no obvious gender bias.

To explain the fact that *-en* and *-e* are overgeneralized, Wegener assumes that in dealing with the input data learners may choose between two strategies: a holistic-formulaic or a cognitive-analytic strategy (cf. Art. 163). If learners first acquire plural forms as unstructured wholes and analyze them later, the holistic-formulaic strategy is applied. If they are geared towards the identification of plural markers and rules for plural formation, learners apply a cognitive-analytical strategy. In her data Wegener has found evidence that rule formation is aided by high type frequency. High type frequency means that many different nouns have the same plural marker. High type frequency furthers rule formation because the more heterogeneous the nouns with a particular plural marker, the more productive or applicable this plural marker is. The fact that it is first the plural *-e* that is regularized most frequently is explained by its higher token frequency in the basic lexicon. Later in the acquisition process it is substituted by *-(e)n*, which has the highest type frequency. Furthermore, whereas in the target grammar, use of a particular plural marker depends on markedness and gender of the noun, in the L2 learners' grammar it correlates with word endings: nouns ending

with *-e* have plural *-n* (as in *Jungen* 'boys'), nouns ending with a consonant have plural *-e* (as in *Türe* 'doors', *Elefante* 'elephants'), and nouns with pseudosuffix *-el* have *-Ø* or *-n* (as in *Kugel* or *Kugeln* 'bullets').

### 3. Contextual inflection

Inherent inflectional morphology serves to express a particular semantic content independently of its syntactic context. Contextual inflectional morphology, however, depends on particular properties of syntactic context cf. Booij (1994). In the following sections, two examples of the L2 acquisition of **contextual inflection** will be discussed: the acquisition of finiteness in L2 German and L2 Dutch and adjectival morphology in L2 Dutch.

#### 3.1. Finiteness

In L2 German and Dutch the acquisition of finiteness constitutes an important target. Verb inflection in German and Dutch is context-dependent in the sense that morphology is determined by the person and number of the subject. It is obvious, however, as illustrated in (3), (4), and (5), that use of inflectional morphology also correlates with positional properties.

- (3) (a) *Karl gibt Hans ein Buch.*  
*Karel geeft Hans een boek.*  
 Charles give:3.SG John a book  
 'Charles gives John a book.'
- (b) *Gibt Karl Hans ein Buch?*  
*Geeft Karel Hans een boek?*  
 give:3.SG Charles John a book  
 'Does Charles give John a book?'
- (4) *Ich sehe, daß Karl Hans ein Buch gibt.*  
*Ik zie dat Karel Hans een boek geeft.*  
 I see:1.SG that Charles John a book give:3.SG  
 'I see that Charles gives John a book.'
- (5) *Karl möchte Hans ein Buch geben.*  
*Karel wil Hans een boek geven.*  
 Charles want:3.SG John a book give:INF  
 'Charles wants to give John a book.'

Verb forms showing agreement (i.e. finite) morphology, occur in second (3 a) or first po-

sition (3 b) in main clauses, whereas in subordinate clauses they occur in sentence-final position (4). Verb forms with no inflectional morphology always appear in sentence final position (5).

In the initial stages of the acquisition of L2 German by L1 Italian speakers, learners apparently assume that verbs should occur in first or second position. This is what they know from their L1. Then, (as evidenced by Klein & Carroll 1992: 157 ff.) at some point they learn that lexical verbs can occur both in non-final and final position, without noticing that these positional options correlate with finite vs. non-finite morphology. What they do seem to acquire, however, is that this distributional difference correlates with the absence or presence of an auxiliary. Illustrations of this particular stage of acquisition are utterances from Gina, which were collected within the framework of the European Science Foundation project as reported in Klein & Carroll (1992).

- (6) *gucken in eine bäckerei*  
look:INF in a bakery  
'to look into a bakery'
- (7) *und dann habe eine brot*  
and then have:1.SG a bread  
*gebringen*  
bring:PAST.PART  
'and then I have brought a loaf of bread'

As Gina is able to learn the positional properties of lexical verbs independently of the relevant functional property of finiteness, it is not at all obvious how she will eventually be able to do so. What seems to be crucial to the acquisition of finiteness by learners such as Gina, is the acquisition of particle verbs with a separable particle. An example is *zurückkommen* (come back) in *er kommt morgen zurück* 'he come:3.SG tomorrow back (he comes back tomorrow)'. The acquisition of separable particle verbs focusses learners on the variability of the position and the form of the lexical verb with respect to the particle. See e.g. the position of *kommt* and *gekommen* in (8).

- (8) *kommt zurück vs. ist zurückgekommen*

The acquisition of finiteness in L2 Dutch by L1 speakers of Turkish differs from the acquisition of L2 German by Romance speakers. While the particular properties of the target languages are the same, the developmental processes differ.

It is typical of Turkish learners of Dutch (see Coenen & Klein 1992; Jordens 1988) that in the initial stages they use most of the lexical verbs with non-finite morphology in final position, as in (9). Only a small group of verbs occurs with finite morphology in initial, i.e. first, second or third position, as in (10). The examples in (9) and (10) are taken from Mahmut.

- (9) *En dan die politie gezien.*  
and then that police seen:PAST.PART  
'And then he saw the police.'
- (10) *Ja, dan politie komt naast.*  
yes then police come:3.SG next  
'Yes, then the police comes next.'

Ergün is a good example of a Turkish learner who has reached a more advanced developmental stage. When he begins to use infinitives in final position with modal verbs as in (11) and past participles in final position with or without auxiliaries as in (12), he also uses non-finite verbs in non-final position most of the time as in (13).

- (11) *Hij zegt, "ik wil niet gaan buiten".*  
he say:3.SG I want:1.SG not go:INF outside  
'He says, 'I do not want to go outside'.'
- (12) *Ja dan is terug naar achter politie/politie gaan.*  
yes then be:3.SG back to after police/police go:PAST.PART  
'Yes, then he went after the police again.'
- (13) *Hij gaan werk of zo.*  
he go:PAST.PART work or so  
'He goes to work probably.'

This opposition between (11) and (12), on the one hand, and (13), on the other, indicates that in final position lexical verbs are used to describe the 'non here-and-now', whereas in non-final position lexical verbs are used to refer to the immediate 'here-and-now' situation.

At the point at which Turkish learners discover that for a description of here-and-now situations the lexical verb must be placed in non-final position, they will notice that the target model does not have *-en* morphology. This means that now, faced with positive evidence, they are able to learn the target-like, correct forms.

### 3.2. Adjectival morphology

Adjectival morphology in languages such as French, Dutch, and German depends on relations of **agreement**. In French, for example, depending on grammatical gender and number of the noun, adjectives vary with respect to their morphological endings: *la grande hotel* ‘the big hotel’, *le grand café* ‘the big café’, *les grandes hotels* ‘the big hotels’. In Dutch, adjectival morphology also depends on whether the determiner is definite or indefinite: *het dikke boek* ‘the thick book’, *een dik boek* ‘a thick book’. Furthermore, in German adjectives vary according to grammatical case as well. For example, in the case of a singular noun with masculine gender, adjectives have different morphology according to whether the noun is nominative or accusative: *der dicke Baum* (nominative) ‘the big tree’, *den dicken Baum* (accusative) ‘the big tree’.

An interesting case for the L2 acquisition of adjectival morphology is Dutch. Here, the target system is relatively simple. Only with nouns that have neuter gender (evident from the use of the article *het*) can adjectives have or not have *e*-morphology depending on whether the noun is definite, as in *het dikke boek* ‘the thick book’, or indefinite, as in *een dik boek* ‘a thick book’. Masculine/feminine nouns (with the article *de*), on the other hand, always have *e*-morphology regardless of whether they are definite, as in *de dikke boom* ‘the big tree’, or indefinite, as in *een dikke boom* ‘a big tree’.

With regard to language use there are two ways in which this type of knowledge may be represented. One possibility, as illustrated in (14), is that the rules of adjectival morphology are stored as a property of the noun.

- (14)  $N_{de}$  with definite article: adj. *-e*  
with indefinite article: adj. *-e*
- $N_{het}$  with definite article: adj. *-e*  
with indefinite article: adj. *-Ø*

Another possibility is that the use of *-e* or *-Ø* is stored not as a property of the noun but as a property of the determiner *de* or *het*, such that *de* can be substituted by the options given in (15 a) and *het* by those in (15 b):

- (15) (a) *de de* [ADJ]*e een* [ADJ]*e*  
(b) *het het* [ADJ]*e een* [ADJ]

If the use of adjectival morphology is based on a knowledge system as represented in (15), language learners can be expected to produce

errors such as: \**de (grote) bos*, \**een grote bos*, \**het (dikke) boom*, \**een dik boom*. However, one would not expect to find errors such as \**del\*het dik boom* or \**del\*het groot bos*. Analysis of a set of production errors (Jordens 1992) leads to the conclusion that this is indeed the case. L2 learners of Dutch make many errors as in (16). Here, the errors are due to incorrect gender.

- (16) *Gelukkig mijn vader woont in een rustig wijk.* (correct: *rustige*)  
‘Fortunately, my father lives in a quiet neighbourhood.’
- Hij is een betrouwbare mens.* (correct: *betrouwbaar*)  
‘He is a reliable person.’

Errors based on the use of an indefinite adjectival form in a definite context (\**del\*het dik boom*, \**del\*het groot bos*) do not occur. Moreover, there is evidence that learners using *e*-morphology with the adjective in an indefinite noun phrase use *de* as the definite article, while learners using *Ø*-morphology with the adjective use *het*. Hence, relations of adjectival agreement are learnt as syntactic patterns which constitute the morphological properties of either *de-* or *het*-nouns.

### 4. Case marking

Structural case is usually seen as an example of contextual inflectional morphology, because it is the verb that determines both the number and the kind of case positions that are required. German, however, also allows a choice between dative and accusative morphology, i.e. between *ihm* and *ihn* in sentences as in (17).

- (17) *Der Hund beißt ihm/ihn in das Bein.*  
the dog bite:3.SG 3.SG.M.DAT/ACC in the leg  
‘The dog bites him in the leg.’

Using experimental and corpus data, Zubin (1975; 1977; 1979) has argued that the use of dative vs. accusative morphology is determined by what he calls ‘degree of contribution’. The notion ‘degree of contribution’ is defined by Zubin as “the degree to which a participant is ‘agent-like’ in an event” (Zubin 1975: 174f.).

With respect to subject selection in particular, Ertel (1974; 1977) has shown in several experiments that it is the egocentric bias

which is the determining variable. This means that “symptoms of relative closeness between the subject element and the speaker’s ego” (Ertel 1977: 147) decide which participant is chosen as the subject of the sentence.

In order to explain the use of **case morphology** in the L2 German of Dutch learners, I assumed that here too principles of ego-nearness and degree of contribution played a role (Jordens 1986; 1992). Use of these principles by L2 learners of German explains why in (18) the indirect object is incorrectly marked with nominative case, and why in (19) the grammatical subject is incorrectly marked with accusative case.

- (18) \**Er wurde von uns geholfen.*

he was by us help:PAST.PART  
(correct: *Ihm*)  
‘We helped him.’

- (19) \**Den Sieg ist mir*

the victory:ACC is me  
*entgangen.*  
escape:PAST.PART  
(correct: *Der Sieg*)  
‘Victory deluded me.’

Second, according to the principles underlying subject selection, sentences with a rhematic subject are processed as if they are subjectless. Hence, it can be understood why it is that with this type of subject L2 learners of German typically use incorrect accusative case. An example is given in (20).

- (20) \**es wurde einen*

there become:PAST.3.SG INDEF:ACC.SG.M  
*Fall erwähnt*  
case mentioned  
(correct: *ein Fall*)  
‘a case was mentioned’

Third, object topicalization is a way of expressing ego-nearness of the object. Hence, Dutch learners of German often incorrectly use nominative morphology here. An example is given in (21).

- (21) \**Jeder Republikaner betrachtete*

every Republican consider:PAST.3.SG  
*er als seinen persönlichen Feind.*  
he as his personal enemy  
(correct: *Jeden Republikaner*)  
‘He considered every Republican a personal enemy.’

## 5. Derivation

In French there are two morphological devices to derive nouns from verbs. From transitive verb stems, nouns can be formed with

the affix *-age*. An example is *raffiner* ‘to refine’ – *raffinage*. From intransitive verb stems, nouns can be formed with the affix *-ment*. An example is *embarquer* ‘to embark’ – *embarquement*. In an experiment on the role of error correction, Carroll et al. (1992) investigated whether native speakers of English are able to learn to use these noun formation rules productively in L2 French. Knowledge of noun formation in French was tested by having subjects give the corresponding noun to a given verb, as in the following test items: *Dans cette usine, on RAFFINE le sucre* ‘In this factory, they refine sugar’. *Donc, c'est une usine de ...* ‘So, it's a refinery’. *Le champagne doit PÉTILLER* ‘Champagne should bubble’ ... *est bon signe* ‘This bubbling is a good sign’. “When the subjects in the experimental group made an error, they were given the correct response” (Carroll et al. 1992: 180). The aim was to find out if this procedure would enable adult learners to construct morphological generalizations. However, the results of the experiment were negative. Carroll et al. had to conclude that their experimental subjects were not able to induce “the subcategorization properties of the suffixes *-age* and *-ment*” (Carroll et al. 1992: 186).

Whereas in this particular experiment Carroll et al. (1992) found no evidence that L2 learners of French learn the use of *-age* and *-ment*, there is evidence showing that in other circumstances L2 learners do learn to use L2 derivational properties. This evidence comes from research on the use of compensatory strategies. Poulsis (1995: 147), for example, noted that Dutch learners of English who did not know the lexical verb *to iron* produced the verb *ironize* instead, thus showing that they learnt to use the suffix *-ize* productively as a device to derive verbs from adjectives and nouns, as in *normalize*, *foreignize*, *synchronize*, *womanize*.

It is a matter of investigation why it is that with respect to *-age* and *-ment* L2 learners were not able to discover the underlying generalization, whereas in the case of the suffix *-ize* they were. It seems reasonable to hypothesize that transfer plays a role here. While the difference between *-age* and *-ment* is a property of French which does not exist in English, there is a Dutch equivalent to the English suffix *-ize*. Therefore, as soon as Dutch learners of English discover that the L1 derivational morpheme *-iseren* can have the same function as *-ize* in English (as, for example,

in *normaliseren/normalize, verbaliseren/verbalize*), it may be used as productively as the L1 equivalent.

## 6. Compounding

In compounding, typological properties with respect to head-modifier relations become manifest (see Art. 165). In Dutch, for example, compounds have modifier-head structure. However, in comparing compounds such as *afdelingshoofd* ‘department-head’ with their [noun phrase + prepositional phrase] equivalent *hoofd van de afdeling* ‘head of the department’, it is obvious that, with regard to the structure of complex noun phrases, input data for L2 learners must be confusingly ambiguous: the same meaning can be expressed either by a modifier-head or a head-modifier structure. Interestingly, the same holds for nominal possessive constructions. A nominal possessive with a preposed genitive as in *mijn vaders broer* ‘my father’s brother’ can also be expressed as a [noun phrase + prepositional phrase] construction such as *de broer van mijn vader* ‘the brother of my father’.

Particularly interesting with respect to the acquisition of compounding is a comparison of the acquisition of L2 Dutch by native speakers of Turkish and Moroccan. This is because typologically Turkish is structured in terms of modifier-head relations, whereas Moroccan has underlying head-modifier structure. Since both Turkish and Moroccan learners of Dutch encounter structures which are typologically similar to the way in which the head-modifier relations of their L1 are instantiated, there is reason for each of these learner groups to assume that Dutch is structured like their native language. Evidence from these learners’ data on the acquisition of **compounds** and nominal possessives shows that this is precisely what is found (Broeder 1991: 171 f.). As illustrated by the examples (22) and (23), Turkish informants prefer to use compounds or complex noun phrases with head-modifier structure, while the Moroccan subjects choose the [noun phrase + prepositional phrase] equivalent with modifier-head structure.

- (22) L1 Turkish:  
*sigaretten-winkel*  
 cigarette-shop  
 ‘cigarette shop’

L1 Moroccan:  
*winkel van sigaret*  
 shop of cigarette  
 ‘cigarette shop’

- (23) L1 Turkish  
*vader zus*  
 father sister  
 ‘father’s sister (aunt)’  
 L1 Moroccan:  
*die dochter van ander vrouw*  
 the daughter of other woman  
 ‘the daughter of the other women’

An interesting example discussed recently by Clahsen (1991) is the acquisition of compounding in L2 German. In German, there is a relation between plural formation and compounding: plurals with -Ø (*Lehrer* ‘teachers’), -e (*Schweine* ‘pigs’), -e and Umlaut (*Mäuse* ‘mice’), -er (*Kinder* ‘children’), -er and Umlaut (*Kräuter* ‘herbs’), and -(e)n (*Schrauben* ‘screws’) are accessible to compounding, while plurals with the inflectional -s marker (*Autos* ‘cars’), are not. Examples are: *Lehrermangel* ‘teacher shortage’, *Schweinefleisch* ‘pork’, *Mäusemännchen* ‘little mouseman’, *Kindergarten* ‘kindergarten’, *Kräuterkuchen* ‘spicecake’, *Schraubenzieher* ‘screwdriver’, *Autfriedhof* ‘car cemetery’.

Clahsen observed that native speakers of Spanish, Portuguese, and Italian initially used -e, -er, and -en as plural markers for compounding in L2 German (24 a). Furthermore, in some cases -en was left out in the early stages (24 b), while -s was never used (24 c).

- (24) (a) *schweinefleisch* ‘pork’  
*kindergarten* ‘kindergarten’  
*schraubenzieher* ‘screwdriver’  
 (b) \**krankkasse* (correct: *Krankenkasse*)  
 ‘health insurance’  
 \**tapetenfabrik* (correct: *Tapetenfabrik*)  
 ‘wallpaper factory’  
 (c) *kontonummer* ‘bank account number’

Clahsen assumes innate level ordering to explain why the -s plural – which he regards as the default plural form – does not occur in L2 German compounding and why, at an intermediate stage of the acquisition of German at which -en is seen as the regular plural ending, -en does not occur in compounds either.

Lardiere (1995) investigated the acquisition of compounding in the L2 English of native speakers of Spanish and Chinese. Sub-

jects were asked to produce novel compounds in response to questions such as: "What could you call a person who cleans shoes/protects children/wears pants, etc.". The aim of the experiment was to see whether subjects would produce plurals inside deverbal (synthetic) compounds. In the test questions, subjects were confronted with three types of plurals: regular plurals (*babies, cats*), irregular plurals (*children, mice*), and pluralia tantum (*pants, people*). Most of the 15 Spanish subjects used both regular and irregular plurals inside these compounds, there were only 2 subjects who correctly produced irregular plurals in compounds while omitting regular -s plurals. The Chinese speakers, on the other hand, while tending to omit the regular plural -s in compounds, also often used singular forms for irregular plurals and even for pluralia tantum. From this Lardiere concludes that "there is no evidence of the engaging of a universal constraint prohibiting regular plurals in compounds" (Lardiere 1995: 42). Furthermore she notes that given the fact that Spanish has plural objects in compounds, as in *un lavaplatos* 'dishwasher' and *un abrelatas* 'can-opener', whereas Chinese has no plural inflection at all, "the subjects' L1 clearly seems to be having some kind of influence on their responses" (Lardiere 1995: 42).

One may wonder how it is possible that Clahsen and Lardiere come to such contradictory conclusions. In order to understand why this is the case, one has to acknowledge that English and German are linguistically different as far as compounding is concerned.

First, in English, the -s plural is the default while in German, contrary to Clahsen's beliefs, it is not. As has been pointed out by Wegener (1994: 270), the -s plural in German has marked status, whereas due to high type frequency -e and -en are the default plural markers. Second, in German the -s occurring in compounds such as *Arbeitsamt* 'employment service' and *Brechungsindex* 'refractive index' is a linking morpheme and not a plural marker. Third, contrary to English, German never has the plural -s in compounds, whether in deverbal, as in *Autoverkauf* 'car sale' or in root compounds (see German *Trambahn* 'tramway' vs. English *weapon(s)* analysis).

As far as second language acquisition is concerned, in L2 English learners are confronted with the regular plural -s in root

compounds. Generalization from plurals in root compounding explains why L2 learners may incorrectly use -s in deverbal compounds as well.

In L2 German, the plural -s is marked. Given the status of the plural -s in German, the differences between Spanish and Italian learners of German can be explained (see Clahsen 1991: 59). Due to their L1 only Spanish and Portuguese learners use -s initially incorrectly as in \*grupps 'groups' (correct: *Gruppen*), \*saches 'things' (correct: *Sachen*). However, as soon as they are confronted with more and richer input, these s-plurals die out quickly and are never used in compounds. Clahsen (1991: 62) also observed that at a certain point in L2 development the -n is sometimes left out in compounds like \*krankkasse 'health insurance', \*tapetefabrik 'wallpaper factory'. Taking into account that Clahsen only found the examples given, these errors seem to be generalization errors in compounding due to the productive use of the singular form, and the plural -e and -en morphology. Overgeneralization of -e in compounds may be due to the high type frequency of -e plurals in isolation compared to their occurrence in compounds as well as, in the case of native speakers of Italian, to L1 influence.

## 7. References

- Andersen, Roger W. (1978), "An Implicational Model for Second Language Research". *Language Learning* 28, 221–282
- Andersen, Roger W. (1991), "Developmental Sequences: The Emergence of Aspect Marking in Second Language Acquisition". In: Huebner, Thom & Ferguson, Charles A. (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*. Amsterdam: John Benjamins, 305–324
- Antinucci, Francesco & Miller, Ruth (1976), "How Children Talk about What Happened". *Journal of Child Language* 3, 169–189
- Bailey, Nathalie & Madden, Carolyn & Krashen, Stephen D. (1974), "Is There a Natural Sequence in Adult Second Language Learning?". *Language Learning* 24, 235–243
- Bates, Elizabeth & MacWhinney, Brian & Smith, Stan (1983), "Pragmatics and Syntax in Psycholinguistic Research". In: Felix, Sascha W. & Wode, Henning (eds.), *Language Development at the Crossroads*. Tübingen: Narr, 11–30
- Bloom, Lois & Lifter, Karin & Hafitz, Jeremie (1980), "Semantics of Verbs and the Development

- of Verb Inflection in Child Language". *Language* 56, 386–412
- Booij, Geert E. (1994), "Against Split Morphology". In: Booij, Geert E. & van Marle, Jaap (eds.), *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 27–50
- Broeder, Peter J. F. J. (1991), *Talking about People: A Multiple Case Study on Adult Language Acquisition*. Amsterdam, Lisse: Swets & Zeitlinger
- Bronckart, Jean-Paul & Sinclair, Hermine (1973), "Time, Tense, and Aspect". *Cognition* 2, 107–130
- Carroll, Susanne & Swain, Merrill & Roberge, Yves (1992), "The Role of Feedback in Adult Second Language Acquisition: Error Correction and Morphological Generalizations". *Applied Psycholinguistics* 13, 173–198
- Clahsen, Harald (1991), "German Plurals in Adult Second Language Development". In: *Theorie des Lexikons: Arbeiten des Sonderforschungsbereichs 282, Nr. 4* (Universal Grammar in Second Language Acquisition), 48–66
- Coenen, Josée & Klein, Wolfgang (1992), "The Acquisition of Dutch". In: Klein, Wolfgang & Perdue, Clive (eds.), *Utterance Structure: Developing Grammars Again*. Amsterdam: John Benjamins, 189–224
- Dietrich, Rainer & Klein, Wolfgang & Noyau, Colette (1995), *The Acquisition of Temporality in a Second Language Acquisition*. Amsterdam: John Benjamins
- Dulay, Heidi C. & Burt, Marina K. (1974), "Natural Sequences in Child Second Language Acquisition". *Language Learning* 24, 37–53
- Dulay, Heidi C. & Burt, Marina K. (1975), "Creative Construction in Second Language Learning and Teaching". In: Burt, Marina K. & Dulay, Heidi C. (eds.), *New Directions in Second Language Learning, Teaching, and Bilingual Education*. Washington, D.C.: TESOL, 21–32
- Ertel, Suitbert (1974), "Satzsubjekt und Ich-Perspektive". In: Eckensberger, Lutz H. & Eckensberger, Ute H. (eds.), *Bericht über den 28. Kongreß der deutschen Gesellschaft für Psychologie in Saarbrücken. Vol. I: Wissenschaftstheorie und Psycholinguistik*. Göttingen: Verlag für Psychologie, Hogrefe, 129–139
- Ertel, Suitbert (1977), "Where Do the Subjects of Sentences Come From?". In: Rosenberg, Sheldon (ed.), *Sentence Production: Developments in Research and Theory*. Hillsdale/NJ: Erlbaum, 141–167
- Giacalone Ramat, Anna (1995), "Tense and Aspect in Learner Italian". In: Bertinetto, Pier M. & Bianchi, Valentina & Dahl, Östen & Squartini, Mario (eds.), *Temporal Reference, Aspect and Actionality. Vol. II: Typological Perspectives*. Torino: Rosenberg & Sellier, 289–309
- Housen, Alex (1993), "L2 Acquisition of Verb Morphology: A Case Study". In: Kettemann, Bernhard & Wieden, Wilfried (eds.), *Current Research in European Second Language Acquisition Research*. Tübingen: Gunter Narr, 195–212
- Jordens, Peter (1986), "Production Rules in Interlanguage: Evidence from Case Errors in L2 German". In: Kellerman, Eric & Sharwood Smith, Michael (eds.), *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon, 91–109
- Jordens, Peter (1988), "The Acquisition of Word Order in L2 Dutch and German". In: Jordens, Peter & Lalleman, Josine (eds.), *Language Development*. Dordrecht: Foris, 149–180
- Jordens, Peter (1992), "The Cognitive Function of Case Marking". In: Gass, Susan M. & Selinker, Larry (eds.), *Language Transfer in Language Learning*. Amsterdam: John Benjamins, 138–175
- Klein, Wolfgang & Carroll, Mary (1992), "The Acquisition of German". In: Klein, Wolfgang & Perdue, Clive (eds.), *Utterance Structure: Developing Grammars Again*. Amsterdam: John Benjamins, 123–188
- Köpcke, Klaus-Michael (1987), "Die Beherrschung der deutschen Pluralmorphologie durch muttersprachliche Sprecher und L2 Lerner mit englischer Muttersprache". *Linguistische Berichte* 107, 23–43
- Krashen, Stephen D. (1977), "Some Issues Relating to the Monitor Model". In: Brown, H. Douglas & Yorio, Carlos A. & Crymes, Ruth H. (eds.), *Teaching and Learning English as a Second Language: Trends in Research and Practice. On TESOL '77*. Washington D.C.: TESOL, 144–158
- Lalleman, Josine A. & van Santen, Ariane J. & van Heuven, Vincent J. (1997), "L2 Processing of Dutch Regular and Irregular Verbs". *ITL, Review of Applied Linguistics* 115–116, 1–26
- Lardiere, Donna (1995), "L2 Acquisition of English Synthetic Compounding is Not Constrained by Level-Ordering (and Neither, Probably, is L1)". *Second Language Research* 11, 20–56
- Larsen-Freeman, Diane (1975), "The Acquisition of Grammatical Morphemes by Adult ESL Students". *TESOL Quarterly* 9, 409–430
- Phillips, Betty S. & Bouma, Lowell (1980), "The Acquisition of German Plurals in Native Children and Non-native Adults". *International Review of Applied Linguistics* 18, 21–29

- Pinker, Steven & Prince, Alan (1988), "On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition". *Cognition* 28, 73–193
- Poulisse, Nanda (1995). "Strategies". In: Jordens, Peter & Lalleman, Josine (eds.), *Investigating Second Language Acquisition*. Berlin: Mouton de Gruyter, 135–163
- Rumelhart, David E. & McClelland, James L. (1986), "On Learning the Past Tenses in English Verbs". In: McClelland, James L. & Rumelhart, David E. & the PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. II: Psychological and Biological Models*. Cambridge Mass.: Bradford Books, MIT Press, 216–271
- Vogel, Thomas (1987), *Die Entwicklung von Tempus und Aspekt im natürlichen Zweitspracherwerb*. Universität Kiel: Englisches Seminar (Arbeitspapiere zum Spracherwerb 33)
- Wegener, Heide (1994), "Variation in the Acquisition of German Plural Morphology by Second Language Learners". In: Tracy, Rosemarie & Lattey, Elsa (eds.), *How Tolerant is Universal Grammar? Essays on Language Learnability and Language Variation*. Tübingen: Max Niemeyer, 267–294
- Wegener, Heide (1995), "The German Plural and Its Acquisition in the Light of Markedness Theory". In: Pishwa, Hanna & Maroldt, Karl (eds.), *The Development of Morphological Systematicity: A Cross-linguistic Perspective*. Tübingen: Gunter Narr, 245–261
- Weist, Richard M. (1986), "Tense and Aspect". In: Fletcher, Paul & Garman, Michael (eds.), *Language Acquisition: Studies in First Language Development*. Cambridge Mass.: Cambridge University Press, 356–373
- Zubin, David A. (1975), "On the Distributional Properties of Surface Morphology and their Consequences for Semantic Analysis". *Columbia University Working Papers on Linguistics* 2, 189–218
- Zubin, David A. (1977), "The Semantic Basis of Case Alternation in German". In: Fasold, Ralph W. & Shuy, Roger W. (eds.), *Studies in Language Variation*. Washington D.C.: Georgetown University Press, 88–99
- Zubin, David A. (1979), "Discourse Function of Morphology: The Focus System in German". In: Givón, Talmy (ed.), *Syntax and Semantics. Vol. XII: Discourse and Syntax*. New York: Academic Press, 469–504

*Peter Jordens, Amsterdam (The Netherlands)*

## 167. Sprachstörungen

1. Einleitung
2. Aphasie
3. Sprachentwicklungsauffälligkeiten
4. Zitierte Literatur

### 1. Einleitung

Schon Francis Bacon hat erkannt, daß die Untersuchung des Pathologischen das Studium des Normalen fördert. In diesem Sinne können **Sprachstörungen** von Erwachsenen sowie **Sprachentwicklungsauffälligkeiten** bei Kindern psycho- und neurolinguistische Einblicke in das Wesen der Morphologie bieten, insbesondere in Prinzipien ihrer Speicherung und/oder Verarbeitung.

Doch können Daten von Sprachstörungen keine direkte Evidenz für oder gegen Hypothesen eines linguistischen Modells liefern. Dazu müssen solche Daten erst im Lichte verschiedener, aber miteinander zu verbin-

dender theoretischer Modellbereiche interpretiert werden:

- (a) eines neuropsychologischen bzw. neuro-psycholinguistischen Modells der entsprechenden sprachlichen Störung bzw. Entwicklungsauffälligkeit,
- (b) eines psycholinguistischen Modells normaler Sprachverarbeitung und -speicherung und unauffälligen Spracherwerbs,
- (c) eines linguistischen (insbesondere: morphologischen) Modells,
- (d) einer Brückentheorie, welche die Bereiche a) bis c) kombiniert bzw. interdisziplinär integriert (cf. allgemein Botha 1979).

Einen (inzwischen überholten) klassischen Ansatz dieser Art stellen die Arbeiten von Luria dar (Luria 1970 und später). Da man aber in sonstigen Arbeiten derartig vollständig integrierte Ansätze selten findet, bleiben Schlußfolgerungen aus Sprachstörungen für

morphologietheoretische Fragestellungen oft hypothetisch.

Theoretische Fragen, auf welche sich ein Morphologiemodell Antworten erwarten kann, betreffen den differentiellen, nicht-chaotischen Charakter spezifischer Sprachstörungen und Sprachentwicklungsauffälligkeiten. Diese können oder könnten erwartungsgemäß einzelne Bereiche ausschließlich oder vorwiegend (oder zumindest stärker) betreffen, z. B. a) die Grammatik, aber nicht Pragmatik oder Diskurs bzw. umgekehrt (s. 2.1), b) die Syntax oder die Morphologie oder das Lexikon (s. 2.2, 2.4, 3.1, 3.2) bzw. Morpheme gegenüber Wörtern, c) innerhalb der Morphologie die Flexion oder die Wortbildung (s. 2.4, 3.1), d) auch abstraktere Entitäten eines Morphologiemodells (s. 2.2, 3.1.2).

In diesem Beitrag beschränken wir uns auf die wesentlichen, in der Literatur am besten dokumentierten Sprachstörungen und Sprachentwicklungsauffälligkeiten mit Schwerpunkten in der Morphologie. Erwähnenswert wären überdies u. a. Tiefendyslexie (cf. Bayer & De Bleser 1989; Coltheart et al. 1980; Lecours et al. 1989); Demenz (cf. Blanken 1986; Hamanaka et al. 1985; Schwartz 1990, Hrsg.); Schizophrenie (cf. Kraus & Mundt 1991, Hrsg.; Lecours 1993). Eine allgemeine Bibliographie geben Dittmann & Tesak (1993), einen allgemeinen Überblick Blanken et al. (1993, Hrsg.).

## 2. Aphasie

### 2.1. Einleitung

**Aphasie** ist eine erworbene, differentielle Sprachstörung bedingt durch eine zentrale Störung des sog. Sprachzentrums, d. h. der sprachrelevanten Verarbeitungs- und Speicherungsareale im Gehirn. Diese liegen in der linken Hirnhälfte – außer bei manchen Linkshändern, bei denen die rechte Hirnhälfte die sprachdominante ist. Läsionen dieser Areale der sprachdominanten Hirnhälfte bewirken eine Aphasie, entsprechende Läsionen der nicht-sprachdominanten Hirnhälfte rufen nach gängiger Meinung Störungen der Pragmatik hervor. Die Fragestellung, wie die Grammatik und Pragmatik verbindende Morphopragmatik, etwa der Diminutive (cf. Dressler & Merlini Barbaresi 1994), bei Hirnläsionen gestört ist, hat in diesem Zusammenhang einen eigenen Stellenwert. Morphologie als wesentlicher Teil der Grammatik

stand stets im Zentrum der Interessen aphasologischer Forschung (cf. die Überblicke in Cholewa 1993 und Badecker & Caramazza 1993).

Bei der immer wieder umstrittenen Klassifizierung aphasischer Syndrome sind zumindest folgende Unterscheidungen wichtig:

- (a) Die Ätiologie der Aphasie, etwa Traumata (z. B. Schußwunden – in ihrer Wirkung oft eng eingegrenzt – spielten besonders nach dem Ersten Weltkrieg eine große Rolle in der klassischen deutschen Aphasielehre, die sich bei Jakobson (1941) niederschlug, oder nach dem Zweiten Weltkrieg bei Luria (1970); Schlaganfälle (vaskuläre Ätiologie, mit meist weniger umgrenzten Ausfällen); Tumore.
- (b) Die neurologische Lokalisierung der Läsion (z. B. im Frontallappen bei **Broca-Aphasie**, im Temporallappen bei **Wernicke-Aphasie**) spielt heute wegen der – in einer Brückentheorie unterzubringenden – schwierigen und mehrdeutigen Verbindung von Lokalisation und anderen Datenbereichen eine geringere Rolle als
- (c) die Symptomatik. So unterscheidet man aufgrund der Flüssigkeit der Sprachproduktion allgemein zwischen flüssiger Aphasie (*fluent aphasia*) und nicht-flüssiger Aphasie (*non-fluent aphasia*). Unter den flüssigen Aphasien ist die Wernicke- oder sensorische Aphasie mit ihrem **Paragrammatismus** in unserem Zusammenhang am wichtigsten (cf. 2.3). Kennzeichnend für die nicht-flüssige Broca- oder motorische Aphasie ist der **Agrammatismus**, benannt nach den dominanten Beeinträchtigungen in der grammatischen Produktion (cf. 2.2).
- (d) Einen weiteren Klassifizierungsfaktor bildet der Schweregrad der Störung. Zur Abgrenzung „pathologischer“ von „normaler“ Sprache sind neben Vergleichen von Aphasikern mit korrespondierenden (Alter, Bildung, Geschlecht, etc.) ungestörten Kontrollpersonen auch Vergleiche von aphasischen Fehlern mit alltäglichen sprachlichen Fehlleistungen (*speech errors*; s. Art. 164) wichtig, um festzustellen, ob sie sich nur quantitativ oder auch qualitativ unterscheiden (cf. Dressler 1987; Stemberger 1984).

### 2.2. Agrammatismus

Nach dem ursprünglichen Konzept der motorischen Aphasie handelte es sich dabei um ein reines Produktionsdefizit ohne Störung des

Sprachverständnisses (cf. z. B. Miceli et al. 1983; Nespolous et al. 1988; Tesak 1990). Die Sprachproduktion ist durch einen sog. Telegrämmstil oder Agrammatismus gekennzeichnet. Zu den klinischen Charakteristika des Agrammatismus gehören: Die langsame, stockende Produktion von Kurzsätzen, das Fehlen satzgrammatischer Funktionswörter (Artikel, Präpositionen, Auxiliarverben, nicht aber satzverbindender Konjunktionen) und das Fehlen sowie der Ersatz von Flexionsmarkern (cf. Menn & Obler 1990, Hrsg.), wie im folgenden Beispiel:

“ja! ich äh immer äh äh denken äh ihr? ihr Buch? aha! richtig! dann immer denken! und äh äh sprechen ... und äh und jemand äh zu Hause mein Mann und ich Einladung äh eingeladen.” (aus Peuser 1978: 127)

Inzwischen ist allgemein sichergestellt, daß die Mehrheit – aber eben nicht alle – der Patienten mit Agrammatismus parallele Grammatikdefizite auch beim Verständnis aufweisen (Ausnahmefälle ohne Verständnisprobleme in Kolk et al. 1985; Miceli et al. 1983; Nespolous et al. 1988).

Traditionell wurde der Agrammatismus als eine zentrale Störung der Syntax interpretiert. Inzwischen variieren die Meinungen über die Natur der zugrundeliegenden Störung sehr. Einerseits bestehen Divergenzen über den primär betroffenen sprachlichen Bereich: (a) die Störung betrifft sowohl Syntax als auch Morphologie (z. B. Miceli et al. 1983); (b) die Störung betrifft nur die Morphologie, indem der Zugriff zu den morphologischen Funktionswörtern und Affixen geschädigt ist (z. B. Bradley et al. 1980), was von Kean (1979) auf (c) phonologische Gründe zurückgeführt wird (cf. Nespolous & Dordain 1990); (d) die primäre Störung liegt am Schnittpunkt von Morphologie und Syntax, nämlich bei der Kongruenz, besonders wenn sie über die Nominalphrase hinausgeht (z. B. Bayer et al. 1987); (e) die primäre Störung betrifft die Syntax (nur sekundär die Morphologie) und besteht in einem Defizit der funktionalen Kategorien: Es fehlen die Kategorien selbst oder ihre syntaktisch-semantischen Merkmale, etwa Finitheit, Tempus, etc. (z. B. Caplan 1985; Grodzinsky 1984; Ouhalla 1993; s. auch Hackl 1996).

Andererseits siedeln immer mehr Ansätze die zugrundeliegende Störung nicht im sprachlichen Wissenssystem sondern in der Verarbeitung an. Dafür spricht die Tatsache, daß Agrammatiker bei Grammatikalitätsur-

teilen – auch in der Morphologie – wesentlich besser abschneiden als in Produktion und Verständnis (cf. Caramazza & Hillis 1989; Linebarger et al. 1983; Nespolous et al. 1988). Wichtige Informationen müssen demnach erhalten geblieben sein. Linebarger et al. (1983; s. auch Schwartz et al. 1987) schlagen eine Störung in der Übertragung syntaktischer Funktionen auf thematische Rollen vor (“*mapping*-Hypothese). Für Friederici (z. B. 1988) ist Agrammatismus die Folge eines pathologischen Verfalls der syntaktischen Information: Zwar ist diese vorhanden aber bereits nicht mehr parat im Zeitrahmen, in dem die morphologische Verarbeitung folgen sollte. Auch Kolk und Kollegen (z. B. Haarmann & Kolk 1994; Hofstede & Kolk 1994; Kolk et al. 1985; Kolk & Heeschen 1990) gehen in ihrer “Adaptionstheorie” und “*strategic variation*-Hypothese (cf. Isserlin 1922) von Beschränkungen in der Verarbeitung aus, an die sich Agrammatiker aufgabenspezifisch adaptieren – durch Vereinfachungen und Verwendung unterschiedlicher Register (Bastiaanse 1995).

### 2.3. Paragrammatismus

**Paragrammatismus** charakterisiert die Produktion bei der sensorischen oder Wernicke-Aphasie. Sie ist im Gegensatz zum Agrammatismus flüssig und in ihrer syntaktischen Komplexität vergleichbar mit normaler Sprache. Allerdings kommt es gehäuft zu Verschränkungen im Satz, die aus einem Zusammenbruch des Monitoring und aus anderen Defiziten in der Sprachverarbeitung resultieren dürften. Eine paradigmatische Selektionsstörung ist als Ursache für die typisch auftretenden Vertauschungen vorgeschlagen worden (Jakobson 1964). Ähnlich wie in Syntax und Lexikon werden auch Flexionsformen – ohne Rücksicht auf die syntaktische Struktur – miteinander vertauscht, und zwar sowohl in der Spontansprache als auch in Tests aller Art. So wie Wortvertauschungen zumeist im selben Wortfeld verbleiben, so gehören die miteinander vertauschten Affixe gewöhnlich derselben Klasse an, z. B. Suffixe werden nicht mit Präfixen vertauscht, verbale nicht mit nominalen Flexionsaffixen, und selten Wortbildungs- mit Flexionsaffixen, wie z. B. in *be-schlagen* für *ge-schlagen* (cf. Bayer et al. 1987; Butterworth & Howard 1987; Dressler & Denes 1987). Hingegen sind Substitutionen zwischen Suffixen derselben Klasse gut bezeugt, z. B. zwischen Flexionssuffixen in *ge-fang-t* statt dem Partizipium

des starken Verbs *ge-fang-en*. Wernicke-Aphasiker produzieren auch oft **morphologische Neologismen**, d. h. nichtexistierende, aber meist potentiell mögliche (legale) Wörter, mit Hilfe von Wortbildungsregeln oder Analogie (cf. Buckingham 1993). Auch **Kontaminationen** kommen vor, wie z. B. *plötzartig* ← *plötzlich* & *blitzartig*. Obwohl solche **Paraphasien** auch bei Broca-Aphasia auftreten (z. B. legal *patriarch-isch* statt *patriarch-alisch*, illegal *Hochschul-schaft* für *Hochschüller-schaft*), sind sie doch typischer für Wernicke-Aphasiker.

Eine Sonderform der sensorischen Aphasia ist die **Jargonaphasie**, bei der – jargonartig – viele Pseudowörter erzeugt werden, die in ihrer phonologischen Gestalt aktuellen Wörtern ähneln, aber zu diesen weder in morphologischen Beziehungen stehen, noch durch phonologische Substitutionen aus ihnen abgeleitet werden können. Diese Jargonwörter können allerdings korrekt flektiert werden (cf. Buckingham 1993; Badecker & Caramazza 1993: 184), z. B. Engl. Plural *chpicters*. Auch produktive Anwendung von Wortbildungsregeln ist bezeugt, z. B. in ital. *fratellismo* ‘Bruderschaft’ als Abstrakta-Bildung von *fratello* ‘Bruder’ (cf. Panzeri et al. 1990).

#### 2.4. Differentielle Störungen der Morphologie

Bei morphologischen Paraphasien als typischen Substitutionsfehlern bei Aphasia stellt sich die Frage, ob es sich um Substitutionen ganzer Wörter bzw. Wortformen oder einzelner Morpheme handelt (cf. Badecker & Caramazza 1993). Hier bietet die unterschiedliche Fehleranfälligkeit verschiedener morphologischer Konstruktionen gute Anhaltspunkte. Zu den Faktoren, die bei der variablen Fehleranfälligkeit eine Rolle spielen, gehört einerseits die Vorkommensfrequenz (*token frequency*), aber eben nicht nur diejenige von Wörtern bzw. Wortformen, sondern auch die von Morphemen (cf. Badecker & Caramazza 1993: 182). Andererseits spielt auch die morphosemantische und morphosyntaktische Transparenz komplexer Wörter eine Rolle (cf. Dressler & Denes 1987). Z. B. waren italienische Broca- und Wernicke-Aphasiker deutlich besser beim Verstehen und der korrekten Identifizierung von morphosemantisch transparenten Komposita des Typs *porta-lettere* ‘Briefträger’ als von opaken des Typs *mangia-preti* ‘scharfer Antiklerikaler’ (wörtlich ‘Priesterfresser’). Ebenso war die Produktion von morphotaktisch transparen-

ten Derivationen wie *vend-i-tore* ‘Verkäufer’ (zu *vendere* ‘verkaufen’) deutlich besser als von opaken wie *scrittore* ‘Schriftsteller’ (zu *scrivere* ‘schreiben’). Diese Auswirkungen sowohl von Frequenz als auch von Transparenz setzen Morphemzerlegung, d. h. die Verarbeitung der einzelnen Morpheme, voraus (cf. Cholewa 1993: 106–110).

Erwähnenswert sind auch differentielle Störungen von Wurzeln und Affixen (cf. Micieli 1995; Tyler et al. 1990) sowie von Flexion und Derivation (cf. Badecker & Caramazza 1993: 183 f.; Cholewa 1993: 114–118). So ist bei Broca-Aphasia die Flexion, nicht aber die Derivation beeinträchtigt (cf. Cholewa & De Bleser 1995; Laine et al. 1994; Menn & Obler 1990, Hrsg.; Tyler & Cobb 1987). Hingegen kommen bei der Wernicke-Aphasia und der **amnestischen Aphasie** (*anomia*, charakterisiert durch Wortfindungsstörungen) viel mehr Fehler in der Wortbildung vor (cf. Kudo 1992). Die isolierte Bewahrung der Morphologie einschließlich der Morphosyntax ohne Zugang zur lexikalischen Semantik bei sensorisch-transkortikaler Aphasia ist in Dogil et al. (1995) dokumentiert.

#### 2.5. Verschiedene Sprachen – verschiedene Störungen?

Die Tatsache, daß Agrammatismus signifikante Störungen der **Flexionsmorphologie** hervorruft, hat zur Folge, daß Sprachen mit reicher und komplexer Flexionsmorphologie, wie die flektierenden (z. B. Polnisch) oder die introflektierenden (z. B. Hebräisch) bei Agrammatismus ganz offensichtlich stärker betroffen sind als solche mit geringer Flexionsmorphologie, wie z. B. Chinesisch, aber auch Englisch. Sprachvergleichende Daten und deren Analyse zeigen solche Unterschiede (cf. Friederici et al. 1991 (Vergleich Deutsch, Französisch, Niederländisch); Ha-verkort 1993 (Japanisch, Englisch); Menn & Obler 1990, Hrsg. (Daten aus 14 Sprachen); Sonderheft 2 des *Nordic Journal of Linguistics* 16 (1993) zu skandinavischen Sprachen). Aufschlußreich sind auch Untersuchungen bilingualer Patienten (z. B. Laine et al. 1994).

Typologische Unterschiede in der einzelsprachlichen Flexionsmorphologie erklären auch Unterschiede in der relativen Häufigkeit von Auslassung versus Substitution von Flexionsendungen, da eine generelle Tendenz darin besteht, markierte durch unmarkierte Flexionsformen zu ersetzen. Je nachdem welche Form der unmarkierten Flexion vorliegt (Nullform/Grundformflexion versus eine spe-

zifische flektierte Form) kommt es entweder zu Auslassungen bzw. zu Substitutionen (z. B. die Auslassung der Verbendung *-s* bei der '3.SG.PRÄS' im Englischen versus der Ersatz der markierten ("starken") Partizipform *ge-fang-en* durch die weniger markierte ("schwache") Form *ge-fang-t*, nicht aber durch die endungslose Form *ge-fang*; cf. Stark & Dressler 1990: 367; Mimouni & Jarema 1995). Oberflächlich vergleichbare morphologische und morphonologische Phänomene, welche in typologisch unterschiedlichen Sprachen eine verschiedene Rolle spielen, sind erwartungsgemäß verschieden gestört: Z. B. ist der deutsche Umlaut in Flexion und Wortbildung bei Agrammatismus stark gestört, die ungarische Vokalharmonie, weil regelmäßig, so gut wie gar nicht (cf. Dressler et al. 1996).

### 3. Sprachentwicklungsauffälligkeiten

#### 3.1. Entwicklungs dysphasie

##### 3.1.1. Deskriptive Charakterisierung

Die Bezeichnung **Entwicklungs dysphasie** (*Spezifische Sprachentwicklungsstörung – SSES*; *developmental dysphasia, specific language impairment – SLI, developmental language disorder*) wird in der Literatur auf eine heterogene Gruppe von Kindern angewandt, bei denen isoliert die Sprachentwicklung verspätet einsetzt und mit einer Beeinträchtigung in der Entwicklung der Sprachproduktion und/oder des Sprachverständnisses einhergeht. Eine mentale Retardierung, eine Hörbehinderung, psychische Störungen wie z. B. Autismus, psychosoziale Beeinträchtigungen oder eine nachweisbare neurologische Störung liegen dabei nicht vor. Nach ihrem Erscheinungsbild lassen sich homogeneren Untergruppen unterscheiden: Einerseits eine vorwiegend expressive vs. eine expressiv-rezeptive Form; andererseits (a) eine rein morphologische/syntaktische Beeinträchtigung (Bishop 1994; Leonard 1989; van der Lely 1993; 1996) vs. (b) eine Form mit zusätzlichen artikulatorischen und/oder phonologischen Defiziten (Gopnik & Crago 1991) und (c) eine Form mit zusätzlichen semantisch-pragmatischen Auffälligkeiten (cf. Bishop & Adams 1989; Mills et al. 1992). Bis heute ist ungeklärt, ob es sich bei diesen Erscheinungsformen um unterschiedliche Beeinträchtigungen mit unterschiedlichen Ätiologien oder bloß um eine einzige Beeinträchtigung in ihren verschiedenen Gradausprägungen bzw. zu unterschiedlichen Entwicklungszeitpunkten

handelt (cf. Bishop et al. 1995). Auch wenn die Ätiologie/n für die Entwicklungs dysphasie bis jetzt unbekannt ist/sind, konnte zumindest in bestimmten Fällen eine erbliche Komponente nachgewiesen werden (cf. Tomblin 1997; in Zwillingsuntersuchungen, cf. Bishop et al. 1995; in Familien mit mehreren dysphasischen Generationen, cf. Gopnik & Crago 1991; Plante 1991; Tallal et al. 1989; 1991; Tomblin 1989; zu dysphasischen Familien siehe speziell Matthews 1994, Hrsg., Paradis 1997, Hrsg.; zur Dysphasie im Allgemeinen siehe auch Leonard 1998, Kamhi 1993, Menyuk 1993).

Als Hauptmerkmal aller Formen der Entwicklungs dysphasie (und deshalb auch zentraler Forschungsgegenstand) werden in der Literatur vorwiegend speziell ausgeprägte Probleme mit Flexionsmorphologie, grammatischen Funktionswörtern und Morphosyntax genannt (cf. Watkins & Rice 1994, Hrsg. für eine Übersicht der diesbezüglichen Literatur). Diese äußern sich in einem nicht-altersadäquaten, eingeschränkten Gebrauch von Flexionsmarkern und Funktionswörtern, die häufig nicht vollständig erworben werden. "Fehlertypen" und Erwerbsabfolgen sind in der Regel mit denen von jüngeren unauffälligen Kindern vergleichbar. Konkret werden in der Literatur folgende Charakteristika der Sprache von dysphasischen Kindern allgemein hervorgehoben:

- Auslassungen und die Verwendung von unmarkierten Formen sind typisch;
- Übergeneralisierungsfehler (z. B. durch Anwendung einer regelmäßigen Regel) treten laut den meisten Autoren nie auf (u. a. Gopnik 1994), laut anderen gibt es sie wohl, jedoch kommt es nie zu kategorieübergreifenden Fehlern (z. B. Leonard & Bortolini et al. 1992; Rice et al. 1995; Rothweiler & Clahsen 1994; Smith-Lock 1993);
- frequente sowie unregelmäßige Formen werden relativ besser beherrscht;
- neben extrem betroffenen Bereichen – wie z. B. Subjekt-Verb-Kongruenz, Auxiliare, Kopula, Artikel sowie Kasus- und Kongruenzmarkern in der Nominalphrase – zeigen sich andere morphologische Aspekte vergleichsweise weniger betroffen und dem jeweiligen Sprachentwicklungsstand des Kindes zumindest angemessen – wie etwa die Pluralbildung von Nomina, die Vergangenheitsformen von unregelmäßigen Verben, die Partizip-

bildung, der durative Aspekt auf *-ing* bei englischen Verben (cf. u. a. Bartke 1998; Clahsen et al. 1992; Goad 1998; Gopnik 1990; Gopnik & Crago 1991; Leonard 1989; Oetting & Rice 1993; Rice 1994; Rothweiler & Clahsen 1994; Ullman & Gopnik 1994).

Wie in 2.5 bei den Aphasien bereits angesprochenen, gibt es auch beim Erscheinungsbild der Dysphasie eine Abhängigkeit von typologischen Unterschieden in der einzelsprachlichen Flexionsmorphologie. So konnten z. B. die fürs Englische allgemein betonten, massiven Schwierigkeiten mit der Subjekt-Verb-Kongruenz bzw. der Verbflexion (d. h. anstelle von flektierten, finiten Verbformen werden vorwiegend Infinitivformen oder eventuell Verbstämme verwendet) zwar für viele andere – aber nicht für alle – untersuchten Sprachen bestätigt werden (cf. Miller & Leonard 1998): z. B. für Deutsch (u. a. Clahsen 1988; Rice et al. 1997), für Englisch (u. a. Rice et al. 1995), für Französisch (Gopnik 1990), für Inuktitut (Crago & Allen 1994), für Japanisch (Fukuda & Fukuda 1994), für Niederländisch (Bol & Kuiken 1990), für Schwedisch (Hansson 1992); nicht aber für Italienisch (Leonard & Bortolini et al. 1992; Bortolini & Leonard 1996) und Hebräisch (Dromi et al. 1993; Leonard & Dromi 1994). Diese sprachspezifischen Unterschiede decken sich mit entsprechenden sprachvergleichenden Daten von unauffälligen Kindern (cf. dazu auch Wexler 1994). Leonard (1998) widmet dem Thema "SLI across languages" sogar ein eigenes Kapitel.

Spezifische Auffälligkeiten in der Wortbildung wurden in der Literatur zur Dysphasie bislang negiert bzw. nicht erwähnt. Einzelne Untersuchungen berichten hier jedoch über ähnliche Probleme wie in der Flexionsmorphologie und vermuten demnach bei der Dysphasie ein allgemeines Unvermögen "to abstract bound morphemes (be they affixes or roots) and represent them as such" (Dalachakis & Gopnik 1995: 202).

### 3.1.2. Hypothesen über das der Dysphasie zugrundeliegende Defizit

Hypothesen über das der Dysphasie ursächlich zugrundeliegende Defizit gibt es in der Literatur viele und divergierende, zum Teil allerdings durchaus komplementäre (einen detaillierteren Überblick und Gegenüberstellung bieten u. a. Bishop 1992; 1997; Crago & Allen 1994; Gopnik 1994; Leonard 1998): Ei-

nerseits werden unterschiedliche, spezifische Probleme beim Erwerb des grammatischen Systems, andererseits mehr allgemeine Lern- und/oder Verarbeitungsdefizite postuliert.

Grammatikspezifische Ansätze erklären die morphologischen Auffälligkeiten der Dysphasie entweder als ein Defizit in der grammatischen Kongruenz ("grammatical agreement deficit"-Hypothesen) oder als ein morphologisches Musterbildungsdefizit. Das Kongruenzdefizit wird unterschiedlich bewertet: Es ist ein generelles (Clahsen 1988; Rothweiler & Clahsen 1994; Clahsen & Hansen 1997) oder besteht darin, daß syntaktisch-semantische Merkmale (wie Finitheit, Tempus, Aspekt etc.) fehlen und daher nicht markiert werden können ("feature deficit"-Hypothese, Gopnik 1990; Gopnik & Crago 1991). Merkmale mögen zwar vorhanden sein, aber ihre Abhängigkeiten – entweder insgesamt oder partiell – werden unzureichend verwaltet ("feature checking deficit"-Hypothese, Ullman & Gopnik 1994; ausschließlich "Spec-Head"-Kongruenz: Rice 1994; van der Lely 1994; 1996), oder Merkmale werden fälschlich als optional behandelt ("extended optional infinitive"-Hypothese, Rice et al. 1995; 1997). Dem morphologischen Musterbildungsdefizit wird entweder ein "paradigm formation"-Defizit zugrunde gelegt, oder es wird dafür eine "missing-rule"-Hypothese bzw. eine "impaired morphological rule-construction"-Hypothese vorgeschlagen (Gopnik 1992; 1994; Gopnik & Crago 1991; Ullman & Gopnik 1994; Gopnik et al. 1997). Demnach werden morphologisch komplexe Wörter unanalysiert gespeichert, keine morphologischen Regeln entwickelt und morphologische Defizite durch konzeptuell basierte Rekategorisierungen kompensiert.

Hypothesen, die die dysphasischen Probleme auf allgemeine Lern- und/oder Verarbeitungsdefizite zurückführen, nehmen einerseits ein nicht-sprachspezifisches, allgemeines Muster- und Regelbildungsdefizit an, woraus ein "unique language learning" (Connell 1987; Swisher et al. 1995) bzw. ein abweichendes sprachliches System resultiert (Restrepo et al. 1992). Andererseits werden spezifische Defizite in der auditiv-perzeptuellen Verarbeitung, Speicherung und Wiederabrufbarkeit vermutet: Zu geringe phonetische Substanz der morphologischen Marker (in der sog. "surface"-Hypothese, u. a. Leonard 1989; 1992; Leonard et al. 1987; Leonard & McGregor & Allen 1992; Leonard & Bortolini et al. 1992; cf. auch Fellbaum et al. 1995);

zu beschränktes phonologisches Gedächtnis (Gathercole & Baddeley 1990) oder Schwierigkeiten in der prosodischen Verarbeitung (Bortolini & Leonard 1996) werden als mögliche Ursachen für die Defizite genannt. In ihrer ursprünglichen Formulierung hat sich Leonards „surface“-Hypothese inzwischen – speziell auch im Sprachvergleich – als unhaltbar erwiesen (s. Leonard 1998: 246 ff. für eine kritische Auseinandersetzung damit). Die im Sprachvergleich beobachteten zwischensprachlichen Unterschiede bei *specific language impairment* werden neuerdings im Rahmen von sog. „morphological richness“ Ansätzen dahingehend erklärt, daß “[...] children with SLI devote their limited processing resources to a dominant property of the language and by necessity, allow other aspects of the language to falter“ (Leonard 1998: 257).

### 3.2. Mentale Retardierung

Zahlreiche Studien sprachunauffälliger Kinder (u. a. Schaner-Wolles & Haider 1987) und mental retardierter Personen liefern starke empirische Evidenz für die Modularitätsthese (cf. Fodor 1983; Chomsky 1982), d. h. für eine Unabhängigkeit bzw. Dissoziation von grammatischen und allgemein-kognitiven Fähigkeiten.

#### 3.2.1. Sprachliche Sonderbegabung bei mentaler Retardierung

Trotz mentaler Retardierung können ausgesprochen fortgeschrittene grammatische Fähigkeiten – auch morphologische – erworben werden, wie eine Reihe von Fallstudien sog. „hyperverbaler“ Personen (engl. *hyper-linguistic*) in der Literatur belegen: z. B. Laura (= Marta), Rick und Antony (cf. u. a. Curtiss 1988; Yamada 1990); Christopher (Smith & Tsimpli 1995); das Mädchen D. H. (Cromer 1994); Personen mit Williams-Betten-Syndrom (cf. u. a. Bellugi & Marks et al. 1988; Clahsen & Almazan 1998; Levy & Hermon 2000; und den Literaturüberblick in Schaner-Wolles 2000a); Françoise mit Down-Syndrom (Rondal 1994; 1995). Besondere Beachtung in morphologischer Hinsicht verdient Christopher, ein sog. „linguistic savant“, der im Alter von 33 Jahren trotz mentaler Behinderung neben Englisch als seiner Erstsprache zumindest 18 weitere Sprachen (genetisch und typologisch gestreut) erworben hatte (darunter eine zu Testzwecken eigens entwickelte Kunstsprache *Epun* mit SVO Struktur, aber einem äußerst komplexen Kongruenzsystem). Diese besondere Begabung

charakterisieren Smith und Tsimpli wie folgt: “One aspect of Christopher’s exceptionality resides precisely in his sensitivity to and learning of complex morphology” (Smith & Tsimpli 1995: 124).

Mehr oder weniger ausgeprägte Diskrepanzen zwischen unterschiedlichen grammatischen Bereichen dürften – wie auch bei der Entwicklungs dysphasie – bei mentalen Retardierungen typisch sein. Anders als bei Christopher ist bei einer solchen disharmonischen Entwicklung die Morphologie häufig selektiv beeinträchtigt: z. B. fortgeschrittene Syntax versus flexionsmorphologische Probleme bei Rick (cf. auch Cromer 1988); gute flexionsmorphologische Leistungen in der Produktion versus schwache bei Verständnistests bei Laura (Yamada 1990: 35–39); ausgeprägte morphologische Schwächen auch beim Down-Syndrom (cf. 3.2.2).

#### 3.2.2. Down-Syndrom

Abgesehen von wenigen Ausnahmefällen wie Françoise, einer damals 34-jährigen Frau mit Down-Syndrom (DS), deren grammatische (auch flexionsmorphologische) Fähigkeiten Rondal (1994; 1995) als normal oder quasi-normal bezeichnet, ist die Sprachentwicklung bei Down-Syndrom verzögert und erreicht auch im Erwachsenenalter höchstens den Entwicklungsstand eines 6–7-jährigen unauffälligen Kindes. Es zeigen sich Disharmonien zwischen den einzelnen sprachlichen Bereichen, wobei morphologische und speziell morphosyntaktische Fähigkeiten am stärksten zurückbleiben (cf. z. B. Fowler 1988; 1990; Fowler et al. 1994; Rondal 1994; 1995; Schaner-Wolles 1992; 1994; 2000 a; 2000 b). In der Entwicklung der einzelnen sprachlichen Fähigkeiten unterscheiden sich Personen mit Down-Syndrom jedoch qualitativ nicht von unauffälligen jüngeren Kindern. Dies gilt auch für die Morphologie (cf. Ergebnisse aus Flexionsmorphologie, Morphosyntax und Derivationsmorphologie von 82 deutschsprachigen Kindern, Jugendlichen und Erwachsenen mit Down-Syndrom in Schaner-Wolles 1989; 1992; 1994; Schaner-Wolles & Dressler 1986).

### 3.3. Hörbehinderung

Beim Erwerb oraler Sprache kommt es aufgrund des auditiven Wahrnehmungsdefizits generell zu einer zeitlichen Verzögerung. Speziell der Erwerb von unbetonten Flexionsendungen ist davon betroffen, da sie auditiv besonders schlecht oder überhaupt nicht wahr-

genommen werden können (cf. Cantwell & Baker 1987; de Villiers et al. 1994; Mogford-Bevan 1993; Wisch 1990.) Aber sogar gehörlosen Kindern ist es möglich, die wichtigsten morphologischen Strukturen ihrer Muttersprache vollständig zu erwerben, wobei weniger der Grad des Hörverlustes als vielmehr die Frühförderung den Erfolg bestimmt (cf. Peltzer-Karpf 1994).

Beim erstsprachlichen Erwerb einer Gebärdensprache (s. Art. 143) verläuft die morphologische Entwicklung parallel zu der von oralen Sprachen bei nicht-hörbehinderten, unauffälligen Kindern (cf. Newport & Meier 1985; Bellugi & van Hoek et al. 1988).

#### 4. Zitierte Literatur

- Badecker, William & Caramazza, Alfonso (1993), "Disorders of Lexical Morphology in Aphasia". In: Blanken et al. (Hrsg.), 181–186
- Bartke, Susanne (1998), *Experimentelle Studien zur Flexion und Wortbildung: Pluralmorphologie und lexikalische Komposition im unauffälligen Spracherwerb und im Dysgrammatismus*. Tübingen: Niemeyer
- Bastiaanse, Roelien (1995), "Broca's Aphasia: A Syntactic and/or a Morphological Disorder? A Case Study". *Brain and Language* 48, 1–32
- Bayer, Josef (1987, Hrsg.), *Grammatik und Kognition: Psycholinguistische Untersuchungen*. Opladen: Westdeutscher Verlag [Linguistische Berichte, Sonderheft 1]
- Bayer, Josef & Bleser, Ria De (1989), "Lexikalische Morphologie und Tiefendyslexie: Eine Fallbesprechung". In: Günther (Hrsg.), 123–154
- Bayer, Josef & Bleser, Ria De & Dronsek, Claudia (1987), "Form und Funktion von Kasus bei Agrammatismus". In: Bayer (Hrsg.), 118–163
- Bellugi, Ursula & Hoek, Karen van & Lillo-Martin, Diane & O'Grady, Lucinda (1988), "The Acquisition of Syntax and Space in Young Deaf Signers". In: Bishop & Mogford (Hrsg.), 132–149
- Bellugi, Ursula & Marks, Shelly & Bahrle, Amy & Sabo, Helene (1988), "Dissociation between Language and Cognitive Functions in Williams Syndrome". In: Bishop & Mogford (Hrsg.), 177–189
- Bishop, Dorothy V. M. (1992), "The Underlying Nature of Specific Language Impairment". *Journal of Child Psychology and Psychiatry* 33, 3–66
- Bishop, Dorothy V. M. (1994), "Grammatical Errors in Specific Language Impairment: Competence or Performance Limitations?". *Applied Psycholinguistics* 15, 507–550
- Bishop, Dorothy V. M. (1997), *Uncommon Understanding: Development and Disorders of Language Comprehension in Children*. Hove: Psychology Press
- Bishop, Dorothy V. M. & Adams, Catherine (1989), "Conversational Characteristics of Children with Semantic-Pragmatic Disorder. 2: What Features Lead to a Judgement of Inappropriacy?". *British Journal of Disorders of Communication* 24, 241–264
- Bishop, Dorothy & Mogford, Kay (1988, Hrsg.), *Language Development in Exceptional Circumstances*. Edinburgh: Churchill Livingstone [= 1993 Hove, Hillsdale: Lawrence Erlbaum]
- Bishop, Dorothy V. M. & North, T. & Donlan, C. (1995), "Genetic Basis of Specific Language Impairment: Evidence from a Twin Study". *Developmental Medicine and Child Neurology* 37, 56–71
- Blanken, Gerhard (1986), *Gestörtes Sprachverhalten bei seniler Demenz und Aphasie*. Dissertation Universität Freiburg
- Blanken, Gerhard & Dittmann, Jürgen & Grimm, Hannelore & Marshall, John C. & Wallesch, Claus-W. (1993, Hrsg.), *Linguistic Disorders and Pathologies*. Berlin: de Gruyter
- Bol, Gerard & Kuiken, Folkert (1990), "Grammatical Analysis of Developmental Language Disorders: a Study of the Morphosyntax of Children with Specific Language Disorders, with Hearing Impairment and Down's Syndrome". *Clinical Linguistics & Phonetics* 4, 77–86
- Bortolini, Umberta & Leonard, Laurence B. (1996), "Phonology and Grammatical Morphology in Specific Language Impairment: Accounting for Individual Variation in English and Italian". *Applied Psycholinguistics* 17, 85–104
- Botha, Rudolf P. (1979), "External Evidence in the Validation of Mentalistic Theories: a Chomskyan Paradox". *Stellenbosch Papers in Linguistics* 2, 1–38
- Bradley, Dianne C. & Garrett, Merrill F. & Zurif, Edgar B. (1980), "Syntactic Deficits in Broca's Aphasia". In: Caplan, David (Hrsg.), *Biological Studies of Mental Processes*. Cambridge/MA: MIT Press, 269–286
- Buckingham, Hugh W. (1993), "Disorders of Word-Form Processing in Aphasia". In: Blanken et al. (Hrsg.), 187–197
- Butterworth, Brian & Howard, David (1987), "Paragrammatism". *Cognition* 26, 1–37
- Cantwell, Dennis P. & Baker, Lorian (1987), *Developmental Speech and Language Disorders*. New York, London: Guilford Press

- Caplan, David (1985), "Syntactic and Semantic Structures in Agrammatism". In: Kean (Hrsg.), 125–152
- Caramazza, Alfonso & Hillis, Argye E. (1989), "The Disruption of Sentence Production: Some Dissociations". *Brain and Language* 36, 625–650
- Cholewa, Jürgen (1993), "Störung der lexikalisch-morphologischen Wortverarbeitung bei Aphasie: Ein Literaturüberblick". *Neurolinguistik* 7, 105–134
- Cholewa, Jürgen & Bleser, Ria De (1995), "Neurolinguistische Evidenz für die Unterscheidung morphologischer Wortbildungsprozesse: Dissoziationen zwischen Flexion, Derivation und Komposition". *Linguistische Berichte* 158, 259–297
- Chomsky, Noam (1982), *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge/MA: MIT Press
- Clahsen, Harald (1988), *Normale und gestörte Kindersprache: Linguistische Untersuchungen zu Syntax und Morphologie*. Amsterdam: Benjamins
- Clahsen, Harald & Almazan, M. (1998), "Syntax and Morphology in Williams Syndrome". *Cognition* 68, 167–198
- Clahsen, Harald & Hansen, Detlef (1997), "The Grammatical Agreement Deficit in Specific Language Impairment: Evidence from Therapy Experiments". In: Gopnik (Hrsg.), 141–160
- Clahsen, Harald & Rothweiler, Monika & Woest, Andreas & Marcus, Gary F. (1992), "Regular versus Irregular Inflection in the Acquisition of German Noun Plurals". *Cognition* 45, 225–255
- Coltheart, Max & Patterson, Karalyn E. & Marshall, John C. (1980), *Deep Dyslexia*. London: Routledge
- Connell, Phil J. (1987), "An Effect of Modeling and Imitation Teaching Procedures on Children with and without Specific Language Impairment". *Journal of Speech and Hearing Research* 30, 105–113
- Crago, Martha B. & Allen, Shanley E. M. (1994), "Morphemes Gone Askew: Linguistic Impairment in Inuktitut". In: Matthews (Hrsg.), 206–215
- Cromer, Richard (1988), "Differentiating Language and Cognition". In: Schiefelbusch, Richard & Lloyd, Lyle (Hrsg.), *Language Perspectives: Acquisition, Retardation and Intervention*. Austin/TX: Pro-Ed, 128–156
- Cromer, Richard (1994), "A Case Study of Dissociations between Language and Cognition". In: Tager-Flusberg (Hrsg.), 141–153
- Curtiss, Susan (1988), "The Special Talent of Grammar Acquisition". In: Obler, Loraine & Fein, Deborah (Hrsg.), *The Exceptional Brain*. New York: Guilford Press, 364–386
- Dalalakis, Jenny & Gopnik, Myrna (1995), "Lexical Representation of Greek Compounds: Evidence from Greek Developmentally Language Impaired Individuals". In: MacLaughlin & McEwen (Hrsg.), 192–203
- Dittmann, Jürgen & Tesak, Jürgen (1993), *Neurolinguistik*. Heidelberg: Groos (Studienbibliographien Sprachwissenschaft 8)
- Dogil, Grzegorz & Haider, Hubert & Schaner-Wolles, Chris & Husmann, Renate (1995), "Radical Autonomy of Syntax: Evidence from Sensory Transcortical Aphasia". *Aphasiology* 6, 577–602
- Dressler, Wolfgang U. (1987), "A Linguistic Classification of Phonological Paraphasias". In: Dressler & Stark (Hrsg.), 1–22
- Dressler, Wolfgang U. & Denes, Gianfranco (1987), "Word Formation in Italian-Speaking Wernicke's and Broca's Aphasics". In: Dressler & Stark (Hrsg.), 69–81
- Dressler, Wolfgang U. & Merlini Barbaresi, Lavinia (1994), *Morphopragmatics*. Berlin: Mouton de Gruyter
- Dressler, Wolfgang U. & Stark, Jacqueline A. (1987, Hrsg.), *Linguistic Analyses of Aphasic Language*. New York: Springer
- Dressler, Wolfgang U. & Stark, Jacqueline A. & Pons, Christiane & Kiefer, Ferenc & Kiss, Kata琳 & Mészáros, Éva (1996), "Cross Language Analysis of German- and Hungarian-Speaking Broca's Aphasics' Processing of Selected Morphonological and Morphological Features: A Pilot Study". *Working Papers, Hungarian Institute for Linguistics, Hungarian Academy of Sciences* 3
- Dromi, Esther & Leonard, Laurence B. & Shterman, Michal (1993), "The Grammatical Morphology of Hebrew-Speaking Children with Specific Language Impairment: Some Competing Hypotheses". *Journal of Speech and Hearing Research* 36, 760–771
- Fellbaum, Christiane & Miller, Steven & Curtiss, Susan & Tallal, Paula (1995), "An Auditory Processing Deficit as a Possible Source of SLI". In: MacLaughlin & McEwen (Hrsg.), 204–215
- Fodor, Jerry A. (1983), *The Modularity of Mind*. Cambridge/MA: MIT Press
- Fowler, Anne E. (1988), "Determinants of Rate of Language Growth in Children with Down Syndrome". In: Nadel, Lynn (Hrsg.), *The Psychobiology of Down Syndrome*. Cambridge/MA: MIT Press, 217–245

- Fowler, Anne E. (1990), "Language Abilities in Children with Down Syndrome: Evidence for a Specific Syntactic Delay". In: Cicchetti, Dante & Beeghly, Marjorie (Hrsg.), *Children with Down Syndrome: A Developmental Perspective*. Cambridge: Cambridge University Press, 302–328
- Fowler, Anne E. & Gelman, Rochel & Gleitman, Lila R. (1994), "The Course of Language Learning in Children with Down Syndrome: Longitudinal and Language Level Comparisons with Young Normally Developing Children". In: Tager-Flusberg (Hrsg.), 91–140
- Friederici, Angela D. (1988), "Agrammatic Comprehension: Picture of a Computational Mismatch". *Aphasiology* 2, 279–284
- Friederici, Angela D. & Weissenborn, Jürgen & Kail, Michèle (1991), "Pronoun Comprehension in Aphasia: a Comparison of Three Languages". *Brain and Language* 41, 289–310
- Fukuda, Shinji & Fukuda, Suzy (1994), "Familial Language Impairment in Japanese: A Linguistic Investigation". In: Matthews (Hrsg.), 150–177
- Gathercole, Susan E. & Baddeley, Alan D. (1990), "Phonological Memory Deficits in Language Disordered Children: Is there a Causal Connection?". *Journal of Memory and Language* 29, 336–360
- Goad, Heather (1998), "Plurals in SLI: Prosodic Deficit or Morphological Deficit?". *Language Acquisition* 7, 247–284
- Gopnik, Myrna (1990), "Feature Blindness: A Case Study". *Language Acquisition* 1, 139–164
- Gopnik, Myrna (1992), "Theoretical Implications of Inherited Dysphasia". In: Levy, Yonata (Hrsg.), *Other Children, other Languages*. Hillsdale/NJ: Erlbaum, 331–381
- Gopnik, Myrna (1994), "Impairments of Syntactic Tense in a Familial Language Disorder". In: Matthews (Hrsg.), 67–80
- Gopnik, Myrna (1997, Hrsg.), *The Inheritance and Innateness of Grammars*. New York: Oxford University Press
- Gopnik, Myrna & Crago, Martha B. (1991), "Familial Aggregation of Developmental Language Disorder". *Cognition* 39, 1–50
- Gopnik, Myrna & Dalalakis, Jenny & Fukuda, Suzy E. & Fukuda, Shinji (1997), "Familial Language Impairment". In: Gopnik (Hrsg.), 111–140
- Grodzinsky, Yosef (1984), "The Syntactic Characterization of Agrammatism". *Cognition* 16, 99–120
- Günther, Harmut (1989, Hrsg.), *Experimentelle Studien zur deutschen Flexionsmorphologie*. Hamburg: Buske
- Haarmann, Henk J. & Kolk, Herman H. J. (1994), "On-line Sensitivity to Subject-Verb Agreement Violations in Broca's Aphasics: The Role of Syntactic Complexity and Time". *Brain and Language* 46, 493–516
- Hackl, Martin (1996), "Verb Second and Agrammatism: Evidence from Two Case Studies". *Wiener Linguistische Gazette* 53–54, 23–62
- Hamanaka, Toshihiko & Kanemoto, Kosuke & Ohigashi, Yoshitaka & Hadano, Kazuo (1985), "Paraphasia and Related Disorders in Primary Degenerative Dementia". *Studia Phonologica* 19, 11–17
- Hansson, Kristina (1992), "Swedish Verb Morphology and Problems with its Acquisition in Language Impaired Children". *Scandinavian Journal of Logopedics & Phoniatrics* 17 (1), 23–29
- Haverkort, Marco (1993), "Universalgrammatik, Parameter und Agrammatismus". *Linguistische Berichte* 143, 60–80
- Hofstede, Ben T. M. & Kolk, Herman H. J. (1994), "The Effects of Task Variation on the Production of Grammatical Morphology in Broca's Aphasia: a Multiple Case Study". *Brain and Language* 46, 493–516
- Howell, S. Catherine & Fish, Sarah A. & Keith-Lucas, Thea (2000, Hrsg.), *Proceedings of the 24<sup>th</sup> Annual Boston University Conference on Language Development, Bd. II*. Somerville/MA: Cascadilla Press
- Isserlin, M. (1922), "Über Agrammatismus". *Zeitschrift für die Gesamte Neurologie und Psychiatrie* 75, 322–410
- Jakobson, Roman (1941), *Kindersprache, Aphasie und allgemeine Lautgesetze*. Uppsala: Almqvist
- Jakobson, Roman (1964), "Towards a Linguistic Typology of Aphasic Impairments". In: de Reuck, Anthony V. S. & O'Connor, Maere (Hrsg.), *Disorders of Language*. London: Churchill, 21–42
- Kamhi, Alan G. (1993), "Children with Specific Language Impairment (Developmental Dysphasia): Perceptual and Cognitive Aspects". In: Blanck et al. (Hrsg.), 625–640
- Kean, Mary-Louise (1979), "Agrammatism: a Phonological Deficit?". *Cognition* 7, 69–83
- Kean, Mary-Louise (1985, Hrsg.), *Agrammatism*. New York: Academic Press
- Kolk, Herman H. J. & Grunsven, Marianne M. F. van & Keyser, Antoine (1985), "On Parallelism between Production and Comprehension in Agrammatism". In: Kean (Hrsg.), 165–208
- Kolk, Herman H. J. & Heeschen, Claus (1990), "Adaption Symptoms and Impairment Symptoms in Broca's Aphasia". *Aphasiology* 4, 221–232

- Kraus, Alfred & Mundt, Christoph (1991, Hrsg.), *Schizophrenie und Sprache*. Stuttgart: Thieme
- Kudo, Takayaki (1992), "Word Formation in Aphasia: Evidence from Japanese *kanji* Words". *Journal of Neurolinguistics* 7, 197–216
- Laine, Matti & Niemi, Jussi & Koivuseikä-Sallinen, P. & Ahlsén, Elisabeth & Hyönä, J. (1994), "A Neurolinguistic Analysis of Morphological Deficits in a Finnish-Swedish Bilingual Aphasics". *Clinical Linguistics & Phonetics* 8, 177–200
- Lecours, André Roch (1993), "Glossomania and Glossolalia in Schizophasia and their Linguistic Kinships to the Jargonaphasias". In: Blaken et al. (Hrsg.), 543–549
- Lecours, André Roch & Lupien, Sonia & Bub, Daniel (1989), "Semic Extraction Behavior in Deep Dyslexia: Morphological Errors". *Tapuscrits* 22 (Centre du Recherche du Centre Hospitalier Côte-de-Neiges, Montréal), 42–51
- Lely, Heather K. J. van der (1993), "Clinical and Research Congruence in Identifying Children with Specific Language Impairment". *Journal of Speech and Hearing Research* 36, 580–591
- Lely, Heather K. J. van der (1994), "Canonical Linking Rules: Forward versus Reverse Linking in Normally Developing and Specifically Language-Impaired Children". *Cognition* 51, 29–72
- Lely, Heather K. J. van der (1996), "Specifically Language Impaired and Normally Developing Children: Verbal Passive vs. Adjectival Passive Sentence Interpretation". *Lingua* 98, 243–272
- Leonard, Laurence B. (1989), "Language Learnability and Specific Language Impairment in Children". *Applied Psycholinguistics* 10, 179–202
- Leonard, Laurence B. (1992), "Specific Language Impairment in Three Languages: Some Cross-Linguistic Evidence". In: Fletcher, Paul & Hall, David (Hrsg.), *Specific Speech & Language Disorders in Children*. London: Whurr, 118–126
- Leonard, Laurence B. (1998), *Children with Specific Language Impairment*. Cambridge/MA: MIT Press
- Leonard, Laurence B. & Bortolini, Umberta & Casselli, M. Cristina & McGregor, Karla K. & Sabbadini, Letizia (1992), "Morphological Deficits in Children with Specific Language Impairment: The Status of Features in the Underlying Grammar". *Language Acquisition* 2, 151–179
- Leonard, Laurence B. & Dromi, Esther (1994), "The Use of Hebrew Verb Morphology by Children with Specific Language Impairment and Children Developing Language Normally". *First Language* 14, 283–304
- Leonard, Laurence B. & McGregor, Karla K. & Allen, Georg D. (1992), "Grammatical Morphology and Speech Perception in Children with Specific Language Acquisition". *Journal of Speech and Hearing Research* 35, 1076–1085
- Leonard, Laurence B. & Sabbadini, Letizia & Leonard, Jeanette S. & Volterra, Virginia (1987), "Specific Language Impairment in Children: A Cross-Linguistic Study". *Brain and Language* 32, 233–252
- Levy, Yonata & Hermon, Shula (2000), "Morphology in Children with Williams Syndrome: Evidence from Hebrew". In: Howell et al. (Hrsg.), 498–509
- Linebarger, Marcia C. & Schwartz, Myrna F. & Saffran, Eleonor M. (1983), "Sensitivity to Grammatical Structure in so-called Agrammatic Aphasics". *Cognition* 13, 361–392
- Luria, Aleksander R. (1970), *Traumatic Aphasia*. The Hague: Mouton
- MacLaughlin, Dawn & McEwen, Susan (1995, Hrsg.), *Proceedings of the 19th Annual Boston University Conference on Language Development*, Bd. I. Somerville/MA: Cascadilla Press
- Matthews, John (1994, Hrsg.), *Linguistic Aspects of Familial Language Impairment*. Montreal: McGill University [Special Issue of the *McGill Working Papers in Linguistics* 10]
- Menn, Lise & Obler, Loraine (1990, Hrsg.), *Agrammatic Aphasia*. Amsterdam: Benjamins
- Menyuk, Paula (1993), "Children with Specific Language Impairment (Developmental Dysphasia): Linguistic Aspects". In: Blaken et al. (Hrsg.), 606–625
- Miceli, Gabriele (1995), "The Summation Hypothesis and Its Implications for Output Morphological Processes". In: Dressler, Wolfgang U. & Burani, Cristina (Hrsg.), *Crossdisciplinary Approaches to Morphology*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften, 71–89
- Miceli, Gabriele & Mazzucchi, Anna & Menn, Lise & Goodglass, Harold (1983), "Contrasting Cases of Italian Agrammatic Aphasia without Comprehension Disorder". *Brain and Language* 19, 65–97
- Miller, Carol & Leonard, Laurence (1998), "Deficits in Finite Verb Morphology: Some Assumptions in Recent Accounts of Specific Language Impairment". *Journal of Speech, Language and Hearing Research* 41, 701–707
- Mills, Anne E. & Pulles, Astrid & Witten, Francine (1992), "Semantic and Pragmatic Problems in Specifically Language Impaired Children: One or Two Problems?". *Scandinavian Journal of Logopedics & Phoniatrics* 17, 51–57

- Mimouni, Zohra & Jarema, Gonia (1995), "Agrammatic Aphasia in Arabic". *Clasnet Working Papers 3*
- Mogford-Bevan, Kay (1993), "Language Acquisition and Development with Sensory Impairment: Hearing-Impaired Children". In: Blanken et al. (Hrsg.), 660–679
- Nespoulous, Jean-Luc & Dordain, Monique (1990), "Agrammatism: a Disruption of the Phonological Processing of Grammatical Morphemes?". In: Nespoulous & Villiard (Hrsg.), 270–277
- Nespoulous, Jean-Luc & Dordain, Monique & Perron, Cécile & Ska, Bernadette & Bub, Daniel & Caplan, David & Mehler, Jacques & Lecours, André Roch (1988), "Agrammatism in Sentence Production without Comprehension Deficits: Reduced Availability of Syntactic Structures and/or of Grammatical Morphemes? A Case Study". *Brain and Language* 33, 273–295
- Nespoulous, Jean-Luc & Villiard, Pierre (1990, Hrsg.), *Morphology, Phonology, and Aphasia*. New York: Springer
- Newport, Elissa L. & Meier, Richard P. (1985), "The Acquisition of American Sign Language". In: Slobin, Dan I. (Hrsg.), *The Crosslinguistic Study of Language Acquisition. Bd. I*, Hillsdale/NJ: Erlbaum, 881–938
- Oetting, Janna B. & Rice, Mabel L. (1993), "Plural Acquisition in Children with Specific Language Impairment". *Journal of Speech and Hearing Research* 36, 1236–1248
- Ouhalla, Jamal (1993), "Functional Categories, Agrammatism, and Language Acquisition". *Linguistische Berichte* 143, 3–36
- Panzeri, Maria & Semenza, Carlo & Ferreri, Tiziana & Butterworth, Brian (1990), "Free Use of Derivational Morphemes in an Italian Jargonaphasic". In: Nespoulous & Villiard (Hrsg.), 72–94
- Paradis, Michael (1997, Hrsg.), *Genetic Dysphasia*. Elsevier [Special Issue *Journal of Linguistics* 10 (2–3)]
- Peltzer-Karpf, Annemarie (1994), *Spracherwerb bei hörenden, sehenden, hörgeschädigten, gehörlosen und blinden Kindern*. Tübingen: Narr
- Peuser, Günter (1978), *Aphasia*. München: Fink
- Plante, Elena (1991), "MRI Findings in the Parents and Siblings of Specifically Language-Impaired Boys". *Brain and Language* 41, 67–80
- Restrepo, Maria Adelaide & Swisher, Linda & Plante, Elena & Vance, R. (1992), "Relations among Verbal and Nonverbal Cognitive Skills in Normal Language and Specifically Language-Impaired Children". *Journal of Communication Disorders* 25, 205–219
- Rice, Mabel L. (1994), "Grammatical Categories in Children with Specific Language Impairments." In: Watkins & Rice (Hrsg.), 69–89
- Rice, Mabel L. & Noll, Karen Ruff & Grimm, Hannelore (1997), "An Extended Optional Infinitive Stage in German-Speaking Children with Specific Language Impairment". *Language Acquisition* 6, 255–295
- Rice, Mabel L. & Wexler, Ken & Cleave, Patricia L. (1995), "Specific Language Impairment as a Period of Extended Optional Infinitive". *Journal of Speech and Hearing Research* 38, 850–863
- Rondal, Jean A. (1994), "Exceptional Cases of Language Development in Mental Retardation: The Relative Autonomy of Language as a Cognitive System". In: Tager-Flusberg (Hrsg.), 155–174
- Rondal, Jean A. (1995), *Exceptional Language Development in Down Syndrome*. Cambridge: Cambridge University Press
- Rothweiler, Monika & Clahsen, Harald (1994), "Dissociations in SLI Children's Inflectional Systems: A Study of Participle Inflection and Subject-Verb-Agreement". *Scandinavian Journal of Logopedics & Phoniatrics* 18, 169–179
- Schaner-Wolles, Chris (1989), "Plural vs. Komparatiververb im Deutschen". In: Günther (Hrsg.), 155–186
- Schaner-Wolles, Chris (1992), *Spracherwerb trotz Down-Syndrom: Eine Vergleichende Studie über die modulare Organisation der Grammatik*. Habilitationsschrift, Universität Wien
- Schaner-Wolles, Chris (1994), "Intermodular Synchronization: On the Role of Morphology in the Normal and Impaired Acquisition of a Verb-Second Language". In: Tracy, Rosemarie & Lattey, Elsa (Hrsg.), *How Tolerant is Universal Grammar?* Tübingen: Niemeyer, 205–224
- Schaner-Wolles, Chris (2000 a), "Sprachentwicklung bei geistiger Retardierung: Williams-Beuren-Syndrom und Down-Syndrom". In: Grimm, Hannelore (Hrsg.), *Sprachentwicklung*. [Enzyklopädie der Psychologie]. Göttingen: Hogrefe, 663–685
- Schaner-Wolles, Chris (2000 b), "Within-Language Dissociations in Mental Retardation: Williams-Beuren and Down Syndrome". In: Howell (Hrsg.), 633–644
- Schaner-Wolles, Chris & Dressler, Wolfgang U. (1986), "On the Acquisition of Agent/Instrument Nouns and Comparatives by Normal Children and Children with Down's Syndrome". *Acta Linguistica Academiae Scientiarum Hungaricae* 35, 133–149

- Schaner-Wolles, Chris & Haider, Hubert (1987), "Spracherwerb und Kognition – Eine Studie über interpretative Relationen". In: Bayer (Hrsg.), 41–80
- Schwartz, Myrna F. (1990, Hrsg.), *Modular Deficits in Alzheimer-Type Dementia*. Cambridge/MA: MIT Press
- Schwartz, Myrna F. & Linebarger, Marcia C. & Saffran, Eleonor M. & Pate, D. S. (1987), "Syntactic Transparency and Sentence Interpretation in Aphasia". *Language and Cognitive Processes* 2, 85–113
- Smith, Neil & Tsimpli, Ianthi-Maria (1995), *The Mind of a Savant. Language Learning and Modularity*. Oxford: Blackwell
- Smith-Lock, Karen M. (1993), "Morphological Analysis and the Acquisition of Morphology and Syntax in Specifically-Language-Impaired Children". *Haskins Laboratories Status Report on Speech Research* 114, 113–138
- Stark, Jacqueline A. & Dressler, Wolfgang U. (1990), "Agrammatism in German: Two Case Studies". In: Menn & Obler (Hrsg.), 281–441 & 1458–1547
- Stemberger, Joseph P. (1984), "Structural Errors in Normal and Agrammatic Speech". *Cognitive Neuropsychology* 1, 281–313
- Swisher, Linda & Restrepo, Maria Adelaida & Plante, Elena & Lowell, Soren (1995), "Effects of Implicit and Explicit 'Rule' Presentation on Bound-Morpheme Generalization in Specific Language Impairment". *Journal of Speech and Hearing Research* 38, 168–173
- Tager-Flusberg, Helen (1994, Hrsg.), *Constraints on Language Acquisition: Studies of Atypical Children*. Hillsdale/NJ: Lawrence Erlbaum
- Tallal, Paula & Ross, Randal & Curtiss, Susan (1989), "Familial Aggregation in Specific Language Impairment". *Journal of Speech and Hearing Disorders* 54, 167–173
- Tallal, Paula & Townsend, Jeanne & Curtiss, Susan & Wulfeck, Beverly (1991), "Phenotypic Profiles of Language-Impaired Children Based on Genetic/Family History". *Brain and Language* 41, 81–95
- Tesak, Jürgen (1990), "Agrammatismus: Ergebnisse und Probleme der Forschung". *Neurolinguistik* 4, 1–41
- Tomblin, J. Bruce (1989), "Familial Concentration of Developmental Language Impairment". *Journal of Speech and Hearing Disorders* 54, 287–295
- Tomblin, J. Bruce (1997), "Epidemiology of Specific Language Impairment". In: Gopnik (Hrsg.), 91–110
- Tyler, Lorraine K. & Behrens, Susan & Cobb, Howard & Marslen-Wilson, William (1990), "Processing Distinctions between Stems and Affixes: Evidence from a Non-Fluent Aphasic Patient". *Cognition* 36, 129–153
- Tyler, Lorraine K. & Cobb, Howard (1987), "Processing Bound Morphemes in Context: the Case of an Aphasic Patient". *Language and Cognitive Processes* 2, 245–262
- Ullman, Michael & Gopnik, Myrna (1994), "Past Tense Production: Regular, Irregular and Nonsense Verbs". In: Matthews (Hrsg.), 81–118
- Villiers, Jill de & Villiers, Peter de & Hoban, Esme (1994), "The Central Problem of Functional Categories in the English Syntax of Oral Deaf Children". In: Tager-Flusberg (Hrsg.), 49–74
- Watkins, Ruth V. & Rice Mabel L. (1994, Hrsg.), *Specific Language Impairment in Children*. Baltimore: Brookes
- Wexler, Ken (1994), "Optional Infinitives, Head Movement and the Economy of Derivations". In: Lightfoot, David & Hornstein, Norbert (Hrsg.), *Verb Movement*. Cambridge: Cambridge University Press, 305–350
- Wisch, Fritz-Helmut (1990), *Lautsprache und Gebärdensprache*. Hamburg: Signum Verlag
- Yamada, Jeni E. (1990), *Laura: A Case for the Modularity of Language*. Cambridge/MA: MIT Press
- Chris Schaner-Wolles, Wien (Österreich)*  
*Wolfgang U. Dressler, Wien (Österreich)*

## XX. Morphologie in der Praxis Morphology in practice

### 168. Field work

1. Introduction
2. Getting to the language
3. Collecting the data
4. Identifying the meaning
5. Resistant problems
6. Technical helps
7. References

#### 1. Introduction

In this article methods of gathering and evaluating data will be presented. Such data serve as basis for the morphological analyses which are discussed in the other articles in this volume. Morphological analysis can be done from many points of view, and the results vary greatly according to what kinds of questions are asked of the data and in what framework the analysis is taking place. But any and all analysis has to be based on solid linguistic data, that which is phonologically correct, natural in language use, and for which the meaning of words, parts of words or other morphemes is clear.

Some methods of how to arrive at such a basis in a hitherto little or unknown language will be presented. Their main merit may well be that they have been tried and found useful.

Much of what has been written on the subject is found in scattered articles and in handbooks on language learning (for a partial bibliography see Goyvaerts 1986), or in introductory textbooks on linguistic analysis or translation. Not all will be mentioned here, but enough to contain the basic information in detail.

#### 2. Getting to the language

##### 2.1. The physical set-up

Many (but by no means all) “insufficiently studied languages” (Kibrik 1977: 1) are found in what is commonly called the Third World.

Thanks to air and road service many of them are fairly easily accessible, though travel to some of them is still an adventure. This is true not only for means of transport, but also for living conditions in general and, maybe most importantly, hygienic eating and drinking facilities. The investigator intending to spend only a few weeks in the area in which the language is spoken should therefore first investigate the physical aspect of his stay (Healey 1975, ed.) and prepare accordingly. He is wise if he strictly observes recommendations given him by authorities and other foreigners living in the area.

##### 2.2. The language assistant(s)

If possible a selection of potential language assistants (also known as “informants”, Samarin 1967) should be available. It is always good to work with more than one. The qualification of such assistants is of utmost importance. Even if the ideal persons may not readily be found, here is a partial list of their qualifications, some of them contradictory.

They should be native speakers of the language, preferably not from a “mixed marriage” where each parent speaks a different language, nor having lived out of the language area for too long as this may impair a speaker’s fine sense of feeling for his language. The best assistant would be a monolingual person who never set foot outside his area.

The assistants should be neither too young nor too old. Before puberty they may not yet have full command of the language. One too old might make a good story teller but be no longer capable of giving any other type of linguistic data.

They should not be ashamed of their language. An educated assistant might have gone through a school system in which use of the mother tongue was punished. Perhaps the official language was learned by means

of punishment. Use of the language may be forbidden and involve risks. Cases are not unknown of people “hiding” their language from foreigners by purposely giving out false information.

They should be reasonably bilingual in the official language of the country. If this language is foreign to both the investigator and the assistants, sources of misunderstanding abound, both because of linguistic and cultural differences.

There is also the matter of remuneration: the salary of a primary school teacher or of a secretary is a good basis on which to calculate it. The amount may have to depend on the educational background of the assistant. It is good to be informed about existing regulations and practices by other foreigners living in the area.

### 2.3. The investigator

The attitudes the investigator carries with him have a lot to do with his effectiveness in his work. In many – but by no means all – cultures the foreigner is received with a fair amount of friendship, coloured by previous experience with other foreigners. It is very important that he respects local traditional and civil authorities (as well as higher placed people in the cities). If he cannot develop a certain affinity for the people whose language he has come to study he might better change the language.

His education and position will usually place him in a higher social category than many of the people speaking the language he has come to study. Paternalism or arrogance is quickly felt by them.

His assistants and their families are real people with real needs – and he might be in a better position to do something about such needs than anyone else. On the other hand friendship also involves receiving. Respect and honest appreciation of the other person are often recognized keenly.

## 3. Collecting the data

### 3.1. The intermediate language

As already mentioned, data is best collected monolingually. This would mean that the investigator learns at least the rudiments of the language first, something that can be achieved superficially in a matter of weeks (thorough learning will take at least a year or more). Eliciting data monolingually enhances

the quality, for it keeps the assistant in the right context: he does not have to translate but can react from within the language in which the data is to be gathered.

The use of an intermediate language could also be avoided by collecting through an interpreter who himself fully understands what data is needed, so that he can trigger it in the language under investigation.

### 3.2. The questionnaire

Good preparation of the kind of data to be collected often includes a questionnaire in the intermediate language, to be translated into the language under investigation (Weiers 1980). Our knowledge of general traits and peculiarities of a language family allows us to prepare such elicitation tools quite well.

If a questionnaire is used to elicit straightforward translation, there are liable to be many pitfalls. Any data elicited without a natural context is easily misunderstood even in the intermediate language. And if misunderstanding is not the problem, a translation of what sounds natural in the official language might yield grammatical but quite unnatural data in the language under investigation – thus the data is falsified. The best data is not elicited via translation, but by a stimulus in the language under investigation.

### 3.3. The transcription

For a good transcription it is best not to use the orthography (provided there is one) unless it has been developed on the basis of sound phonological principles. A poor orthography might easily underdifferentiate the sounds, yielding homographs which are not homophones, causing different morphemes to appear to be identical. Thus, the investigator might look for similarities in meaning for words that have nothing to do with each other.

The investigator should have some practical knowledge of phonetic transcription including ear training so that he can accurately represent on paper what he hears. This training should include tone differentiation, should he work in a tone language. He should also have some basic training in phonological analysis. For a purely phonetic transcription might lead him astray in a different direction: he might not recognize that two utterances are the same though different phonetically and end up with more morphemes than warranted.

If the investigation is done through a translator or by a literate assistant himself, there might be no alternative to using the orthography he is familiar with. Being aware of the pitfalls might be enough to circumvent them.

#### 4. Identifying the meaning

No analysis is possible without being able to identify the meaning of the morphemes under investigation. One of the most important principles to remember is that the meaning changes according to context (Pike & Pike 1977: 173 f.).

##### 4.1. Lexical meaning

The lexical meaning is relatively easy to discover as it involves the first three of the basic semantic categories (Nida 1964: 62): the meaning of the nouns (for things), verbs (for events), and descriptives (for attributes) in a language. It is good to be aware of the fact that the name of any of the above is not necessarily a simple noun, it may be a derivation, a compound or a phrase, and yield all sorts of material for morphological analysis.

One word of caution: even though it may be obvious that a word is morphologically complex, a compound word or even a phrase, it is by no means obvious that the assistant can give a morpheme-by-morpheme or a word-by-word translation. To him more often than not the whole expression has one meaning; the break-down has to be supplied by the investigator.

Eliciting vocabulary with lexical meaning is best done by pointing and enacting on the one hand and by analogy on the other (Beekman 1968: 361 f.), like for example: Many *bees* are a *swarm*, what are many *dogs*? The opposite of *good* is *bad*; what else can it be? What is the opposite of *new*?

##### 4.2. Grammatical meaning

The last of the four basic categories, the relationships, are far more difficult to deal with. For they encode the grammar of the language (apart from word order). This category includes all sorts of affixes as well as the particles, many of them “pesky” ones (see 5.2; Grimes 1975: 93), because they are so hard to define. No direct questioning will yield them in the language, and only the rarest of assistants is able to give a translation, once they are encountered. Even then a translation is

only a starting point to find the “real” meaning or function. Who, for example, could easily fully define the use of *still* or *yet* in English?

The best chances for an equivalent in the intermediate language are found in such meaning bearers as pronouns, conjunctions, adpositions (prepositions or postpositions), singular/plural differences, negatives, question words, and maybe verb tenses. But translations are at best approximations. They do yield a beginning point for the analysis which will eventually bring to light the full meaning and use of the morpheme.

##### 4.3. A simple methodology

A quite simple method for identifying the morphemic forms and a first indication as to their meaning is that of substitution of recurring particles (Elson & Pickett 1983).

It consists of collecting a number of utterances (complex words, maybe) which contain the morpheme in question. Hopefully a meaning can be supplied for the complete form, maybe other forms without the morpheme can be found and their meaning ascertained. A comparison of the paradigm will yield the various forms of the morpheme(s) under investigation as in (1) from Tzeltal (Mexico):

(1)	<i>k'ab</i> 'hand'	<i>hk'ab</i> 'my hand'
	<i>akan</i> 'leg'	<i>kakan</i> 'my leg'
	<i>lumal</i> 'land'	<i>alumal</i> 'your land'
	<i>inam</i> 'wife'	<i>awinam</i> 'your wife'
	<i>k'op</i> 'language'	<i>sk'op</i> 'his language'
	<i>ar'el</i> 'work'	<i>yat'el</i> 'his work'

(Nida 1949: 16)

This tried and true method in the tradition of the guess-and-check procedures (Longacre 1964: 11, 101 f.) still serves to recognize the majority of the morphemes in any language, including their various forms which yield the morpho-phonemic rules. It is not quite so simple to identify their full meaning. For this it is necessary to find all the possible occurrences of the form under investigation and study their meaning nuances which might be caused by the presence or absence of that form.

#### 5. Resistant problems

##### 5.1. Homophones

A certain number of words and particles in any language sound exactly the same. Are they the same, or different morphemes? For

the lexicon this question is not quite so crucial; it is not that hard to distinguish *a fish* from *to fish*, one a noun the other a verb. If one of the pair is a grammatically functioning particle or an affix (and if a lexical meaning cannot be ascertained it always is), this question becomes relevant for the analysis and gains quite a different importance.

For example, in Kaingang (Brazil), are *mu-* 'go:PL' and *mu-* 'action in progress' the same morpheme or two different ones? Or, asked differently: in the utterance under examination, is the occurrence of *mu-* a verb or a particle or an affix? This difference might be all important in deciding on the "real" meaning of the utterance (as over against the general and vague translation provided by the assistant or guessed at by the investigator) and on the analysis of the structure of that utterance – be it part of the morphology or the syntax.

If the language is full of such homophones, this is one of the thornier questions to resolve. Certain verbs have a tendency to become particles (or, in an earlier stage, auxiliaries). If both members of the homophonous pair are particles or affixes, the solution is none the easier.

### 5.2. Pesky particles

Another difficult problem is that of particles for which the meaning is difficult to ascertain. The first question to decide is whether such particles belong into the syntax or the morphology. They often have some function within discourse rather than within sentence or phrase and are therefore difficult to pin down. Nevertheless their identification is important to the understanding of the grammar. There is only one way to study them: within natural texts.

### 5.3. Text as context

All morphemes can best be studied within the framework of text analysis, for this is where they occur most naturally. In order to do this different types of texts (Longacre 1976: 197 f.) need to be collected, recorded on tape, and transcribed if possible. Such texts are best charted according to different points of view (Wiesemann et al. 1992; Seminar für Sprachmethodik 1999): certainly a basic chart needs to be made, in which the clauses are grouped around the verbs in the text. Then an information chart lists the different types of information transmitted by the verbs in several columns, including highlighted information.

Texts charted this way are the best sources of contexts for affixes or particles for which the precise function is difficult to determine.

## 6. Technical helps

### 6.1. Recorders

The recording technique is one of the most useful helps in field work (Bouquiaux & Thomas 1976, eds.: 27 f.). Today analog (cassette recorders) and digital (MiniDisc, DAT, etc.) recorders are of such high quality and at the same time so small that they can be carried almost everywhere and used even in adverse circumstances, with or without electricity. A separate microphone improves the quality and is more versatile than the built-in one, particularly if suprasegmental sound features are to be observed. A moving coil microphone can stand difficult climates better than a condenser microphone (Wiesemann 1991: 195 f.). In order to transcribe tape recorded data it is necessary to listen to small stretches and return to the same utterance or part of the utterance several times (Gudsinsky 1967: 9). It should be possible to rewind a bit of the tape while the play button is held down.

### 6.2. Data books, lists

It is useful to gather the data in notebooks. To keep track of the chronological development of the data-gathering each page is numbered, each session dated, and the name of each assistant written down. It is not advisable to erase any of the attempts at writing down the data; corrections are rather written in under the original entry.

At first the phonetic notations will be full of errors; it takes a while for the ear to adjust to the sounds in the language. Even after a while doubts about individual sounds are bound to persist. These can be double checked with the help of lists for contrastive listening, a type of differential drill (Brewster & Brewster 1976: 309 f.). Words with what are either identical or very similar sounds are gathered in a list. The sounds to be compared should preferably appear in the same position in each word. As the assistant pronounces each word, the investigator listens to just that troubling sound. The lists may need to be rearranged several times, the assistant may be consulted as to which words, in his opinion, have the same sound and which sounds are different.

The morphological analysis may be done in the notebook or on separate sheets. Individual sheets are not helpful for data gathering but can easily be rearranged for analysis purposes.

### 6.3. Morphology files

In order to keep data handy and the analysis under control, a shoe box or something similar with subdivisions and a whole set of  $3 \times 5"$  very thin cardboard slips is very useful. On each slip a word is written (normally in the left hand corner), a grammatical designation (in the middle), and a gloss (in the right hand corner). In this way the lexicon can be gathered and kept in alphabetical order, the parts of words, their forms and meanings are easily retrievable. Filing with the help of punch cards (Day 1976) is even more versatile.

### 6.4. Computer programs

PCs are small and so powerful today, that they constitute another very useful technical tool for morphological analysis in almost any field situation. Even a small data bank can take the place of the morphology file. Specially developed programs (Weber et al. 1988; Wimbish 1990; Antworth 1990) assist in morphological analysis, glossing text and pulling out the contexts of whatever morpheme is to be studied (cf. also Art. 172):

- Shoebox and LinguaLinks are data base systems for field linguists and anthropologists. Among several facilities for the field worker both programs provide also lexical databases and tools for morphological analysis and interlinear transcription of words.
- PC-KIMMO is designed for generation and morphological parsing of words. The program is an implementation of the model of two-level morphology (see Koskenniemi 1983) and uses a lexicon, a module for the declaration of morphonological rules, and a module for the declaration of morphotactic rules.

## 7. References

- Antworth, Evan L. (1990), *PC-KIMMO: A Two-Level Processor for Morphological Analysis*. Dallas/TX: Summer Inst. of Linguistics
- Beekman, John (1968), "Eliciting Vocabulary, Meaning and Collocations". *Notes on Translation*.
- (Dallas/TX: Summer Inst. of Linguistics) 1.29: 1–11 [reprinted in: Healey (1975, ed.)]
- Bouquiaux, Luc & Thomas, Jacqueline M. C. (1976, eds.), *Enquête et description des langues à tradition orale, vol. 1–3*. Paris: SELAF
- Brewster, E. Thomas & Brewster, Elizabeth S. (1976), *Language Acquisition Made Practical: Field Methods for Language Learners*. Colorado Springs: Lingua House
- Day, A. Colin (1976), *Getting the Edge on Card Files*. High Wycombe, Bucks.: Summer Inst. of Linguistics
- Elson, Benjamin F. & Pickett, Velma B. (1983), *Beginning Morphology and Syntax*. Dallas/TX: Summer Inst. of Linguistics
- Goyvaerts, Didier L. (1986), *Language and History in Central Africa*. Antwerpen: Univ. Instelling Antwerpen (Antwerp Papers in Linguistics 44)
- Grimes, Joseph E. (1975), *The Thread of Discourse*. The Hague: Mouton (Janua Linguarum Series Minor 207)
- Gudschinsky, Sarah C. (1967), *How to Learn an Unwritten Language*. New York etc.: Holt, Rinehart and Winston
- Healey, Alan (1975, ed.), *Language Learner's Field Guide*. Ukarumpa, Papua New Guinea: Summer Inst. of Linguistics
- Kibrik, Alexandre E. (1977), *The Methodology of Field Investigations in Linguistics*. The Hague: Mouton (Janua Linguarum Series Minor 142)
- Koskenniemi, Kimmo (1983), *Two-level Morphology: A General Computational Model for Word-Form Recognition and Production*. Ph.D. dissertation, University of Helsinki
- Longacre, Robert E. (1964), *Grammar Discovery Procedures: A Field Manual*. The Hague: Mouton (Janua Linguarum Series Minor 33)
- Longacre, Robert E. (1976), *An Anatomy of Speech Notions*. Lisse: de Ridder
- Nida, Eugene A. (2<sup>nd</sup>1949), *Morphology: The Descriptive Analysis of Words*. Ann Arbor: Univ. of Michigan Press [1<sup>st</sup>1946]
- Nida, Eugene A. (1964), *Toward a Science of Translating*. Leiden: Brill
- Pike, Kenneth L. & Pike, Evelyn G. (1977), *Grammatical Analysis*. Dallas/TX: Summer Inst. of Linguistics (SIL Publications in Linguistics 53)
- Samarin, William J. (1967), *Field Linguistics: A Guide to Linguistic Field Work*. New York etc.: Holt, Rinehart and Winston
- Seminar für Sprachmethodik (1999), *Textanalyse und Grammatik, ein sprachwissenschaftliches Lehrbuch*. Bonn: Verlag für Kultur und Wissenschaft

Weber, David J. & Black, H. Andrew & McConnel, Stephen R. (1988), *AMPLE: A Tool for Exploring Morphology (Version 1.0)*. Dallas/TX: Summer Inst. of Linguistics

Weiers, Michael (1980), *Linguistische Feldforschung: Ein Leitfaden*. Wiesbaden: Harrassowitz

Wiesemann, Ursula (1991), *Verstehen und verstanden werden: Praktisches Handbuch zum Zweitspracherwerb*. Bad Liebenzell: Verlag der Liebenzeller Mission

Wiesemann, Ursula & Nsémé, Clédor & Vallette, René (?1992), *Manuel d'analyse du discours*. Yaoundé: Collection PROPELCA (26) [?1984 with different pagination]

Wimbish, John S. (1990), *SHOEBOX: A Data Management Program for the Field Linguist (Version 1.2)*. Waxhaw: JAARS Computer Services (Summer Inst. of Linguistics)

*Ursula Wiesemann, Burbach-Holzhausen  
(Germany)*

## 169. Interlinear morphemic glossing

1. Basic concepts
2. Prerequisites of morphological analysis
3. Principles of interlinear glossing
4. Boundary symbols
5. Typographic conventions
6. Summary
7. References

### 1. Basic concepts

#### 1.1. Purpose

Given an object language L1 and a meta-language L2, then an **interlinear morphemic gloss** (IMG) is a representation of a text in L1 by a string of elements taken from L2, where, ideally, each morph of the L1 text is rendered by a morpheme of L2 or a configuration of symbols representing its meaning, and where the sequence of the units of the gloss corresponds to the sequence of the morphs which they render. Its primary aim is to make the reader understand the grammatical structure of the L1 text by identifying aspects of the free translation with meaningful elements of the L1 text. The ultimate purpose may be to aid the reader in grasping the spirit of the language, to control the linguistic argument the author is making by means of the L1 example or to scan a corpus for a certain gloss in order to find relevant examples.

#### (1) Latin

<i>exeg-i</i>	<i>monumentum</i>
<i>implement\PRF-1.SG</i>	<i>monument.N:ACC.SG</i>
<i>aer-e</i>	<i>perennius</i>
<i>ore.F-ABL.SG</i>	<i>lasting:CMPR:ACC.SG.N</i>

'I have executed a monument more durable than ore' (Hor. Od. 3, 30, 1)

(1) illustrates the typical use of an IMG. The first line of (1) contains the L1 text line; the second line contains the IMG, and the third line contains an idiomatic translation into L2.

Interlinear morphemic glossing is at the intersection of different communicative purposes. On the one hand, it is a kind of translation that accompanies the original. In this sense, it is comparable to the arrangement that one finds in synoptic editions of original and translation. On the other, it is a kind of linguistic analysis. In this sense, it competes with a fragment of a grammar. Its hybrid character leads to a number of problems and to different styles in interlinear morphemic glossing.

The aim of the following treatment is a standardization of an aspect of linguistic methodology on the basis of widespread usage as developed in the 20<sup>th</sup> century. To the extent that linguistics is a science, its methods are susceptible and in need of standardization. Interlinear morphemic glossing has to do with the representation of linguistic data, comparable in this with a phonetic transcription. Just as the latter has been successfully standardized by the IPA, so interlinear morphemic glossing should be standardized.

This will be done in the present article in the form of a set of rules, which are listed in section 6.1. Such a standardization only concerns linguistic science. Linguistic data are often presented to a lay public, with the purpose of education, entertainment or divulgence of the achievements of our science. Here some kind of interlinear glossing may be necessary, too. However, scientific formal-

ism tends to damage rather than serve the good cause. An example how interlinear glossing has been handled in a book directed to a non-specialist public is quoted in the next section (Finck 1909). The present article is biased in favor of a more formalized treatment, on the assumption that it will be easier to derive a less formal representation from the proposals made here than the other way round. The treatment is, however, not fully formal, since it focuses on interlinear glossing in printed texts. In the annotation of texts by markup languages for automatic retrieval, the same conceptual problems, but very different technical problems arise which will not be dealt with here.

Data are commonly quoted from sources in which they are already provided by an analysis. In linguistic publications, it has been wide-spread usage to quote data together with their IMG and their translation, even if their form or language is different from the one used in the quoting context. That is, such composite data representations have been treated as indecomposable blocks. Such scruples do not seem to be warranted. Primary data may be quoted and provided with the quoting author's analysis and translation (cf. Bickel et al. 2004: 1).

## 1.2. Precursors

Interlinear glossing has precursors in the descriptive tradition which link it up not with some kind of morphological representation, but with efforts to bring out the spirit of the language. The point there is not to provide a formal representation of a piece of linguistic data, but to render the language-specific construal of the world intelligible. To this end, literal translations were provided. For instance, G. Gabelentz (1901: 460), in a passage arguing that the personal verb suffixes in Semitic languages are possessive pronouns, gives the following Arabic example: “*ya-kft-ka-hùm* er genügt dir gegen sie (eig. er-genügt-dein-ihr)”.

The IMG is a late-comer in linguistics. Early grammars were intended as primers, the user was expected to work through them and learn the morphemes; so no glossing was necessary. Many scientific grammars, e.g. of Latin, Greek, Arabic etc., were meant for the initiated who needed no glossing either (not seldom even the free translations were spared). Even comparative studies, historical or typological, left the analysis of the examples of diverse languages to the reader. H. C.

Gabelentz, in the middle of a discussion of Lule, Osage and other languages, presents the following passage:

“Im Dakota (meine Grammatik der Dakotasprache § 34) dient die 3 Pers. Plur. Act. dazu, das Passivum auszudrücken, sogar wenn ein Actor im Singularis hinzuzudenken ist, z. B. *Jesus Jan eñ hi q ix Jordan watpa ohna baptizapi*, Jesus kam zu Johannes und sie tauften ihn (st. er wurde getauft) im Jordanfluss.” (H. C. Gabelentz 1861: 465)

Here the reader who does not have the grammar mentioned on his desk is given no chance.

Pace Gabelentz, IMGs are needed when two conditions coincide: the level of analysis is above morphology, and the reader is not expected to be familiar with the languages under discussion (which is generally the case in typology, but not in descriptive or historical-comparative linguistics). W. v. Humboldt (1836 [1963]: 534) invented his own device to help the reader identify L2 meaningful elements with L1 morphemes. He gives the following example from Classical Nahuatl:

1	2	3	4	5	6	7	8	9
ni-	c-	chihui	-lia	in	no-	piltzin	ce	calli
1	3	2	4	5	6	7	8	9
ich	mache	es	für	der	mein	Sohn	ein	Haus

While dispensing with the IMG proper, this method fails for L1 elements which cannot be rendered by L2 words.

Beside the literal translation illustrated above, G. Gabelentz (1901) uses a variety of techniques. He also has interlinear glosses, as when he says: ‘Der Satz “Ich bin Dein Sohn” heißt im Maya:

*a – meyen – en.*  
Dein Sohn ich,’  
(Gabelentz 1901: 383)

and occasionally (e.g. Gabelentz 1901: 400) he uses Latin as L2 in IMGs.

Finck (1909) is one of the first linguistic publications that illustrate the working of a language with a sizable text provided with a free translation and an IMG. The following sentence from his Turkish text (Finck 1909: 83) illustrates his glossing style (s. Fig. 169.1).

As may be seen, these forerunners have no grammatical category labels yet. Finck glosses Turkish *-m* ‘GEN’ by Germ. *der* because this word displays a morphological trace of the genitive. Similarly, Turkish *-up* ‘GER’ is glossed by *-enderweise*, maybe the closest to a gerund that German can muster. This pro-

<i>xodža - da</i>			
<i>Meister - auch</i>	<i>Der</i>	<i>Meister</i>	
<i>esbāb - un</i>	<i>dzümle - si - ni</i>	<i>ates - e</i>	<i>warf nun sämt-</i>
<i>Kleider - (der)</i>	<i>Gesamtheit - ihre - die</i>	<i>Feuer - zu</i>	<i>liche Kleider ins</i>
<i>vur - up</i>	<i>yak - ar</i>		<i>Feuer und ver-</i>
<i>werf = enderweise</i>	<i>verbrenn = end</i>		<i>brannte sie.</i>

Fig. 169.1

cedure is a tribute to the non-specialist readership that the booklet aims at, but necessarily falsifies the working of the language by attributing lexical meanings to its grammatical morphemes.

It took a long time until interlinear morphemic glossing became firmly established. In Bloomfield's *Language*, of 1933, examples abound, but they are presented like this: "Some languages have here one word, regardless of gender, as Tagalog [kapa'tid]; our *brother* corresponds to a Tagalog phrase [kapa'tid na la'la:kil], where the last word means 'male', and our *sister* to [kapa'tid na ba'ba:jil], with the attribute 'female'" (Bloomfield 1933: 278).

IMGs that fulfill most of the requirements set out below appear first in the sixties of the 20<sup>th</sup> century. From the eighties on, they become standard in publications dealing with languages whose knowledge is not presupposed. Editors and publishers increasingly require them even for languages like Latin, French and German that used to be well-known to linguists. The development is towards (not only translating, but even) glossing every language except English. This is apparently a symptom of a global development in which every language except English becomes exotic.

Good IMGs are relatively costly, both for the scientist and for the typesetter. Authors and publishers are therefore not too eager to produce them (well). There is at least one software on the market that aids the linguist in generating systematic IMGs for his texts, the interlinearizer that comes with the program Shoebox, from the Summer Institute of Linguistics (cf. Simons & Versaw 1988; Art. 168).

Since IMGs are fairly recent in linguistics, they have seldom been treated by linguistic methodology. The first treatise of the present subject is Lehmann (1982). Subsequent work

includes Simons & Versaw (1988), Lehmann et al. (1994), Lieb & Drude (2000), Bickel et al. (2004). They have been freely made use of in the present treatment.

### 1.3. Levels of representation

Interlinear morphemic glossing must be seen in the larger context of representation of linguistic data and, even more comprehensive, of the documentation of a language (cf. Lieb & Drude 2000). On such a background, an isolated example given in a descriptive context is a particularly constrained case of the edition and annotation (also called 'markup' for technical purposes) of a piece of primary linguistic data for posterity. In other words, a general-purpose edition of a linguistic corpus is a kind of maximum model, subject to the full set of rules for explicitness, detail and elaboration, from which the quotation of an isolated example in the context of some grammatical discussion represents a subset delimited by considerations of feasibility, usefulness and the like.

Every linguistic representation of some piece of raw data, even if it limits itself to a phonetic transcription, involves some linguistic analysis (Lehmann 2004). Insofar, no sharp boundary is to be drawn between the sheer representation of data and their analysis. Bearing this in mind, we can speak of various levels at which linguistic data may be represented. Presupposing spoken language data, at least the following are relevant:

- (a) raw data recording (video or audio tape),
- (b) phonetic transcription of the utterance,
- (c) orthographic representation of the utterance,
- (d) morph(ophon)emic representation of the utterance,
- (e) IMG of the utterance,
- (f) free translation of the utterance into the background language,

- (g) descriptive and explanatory comment on pragmatic or cultural aspects of the utterance.

This set may be supplemented by even more representations (cf. Lieb & Drude 2000). There may be a phonological representation distinct from both levels (b) and (d). There may be a syntactic representation, e.g. in the form of a labeled bracketing. And there may be a semantic representation instead of, or in addition to, representation (f). In such representations, the portion of linguistic analysis is probably even stronger than in the seven levels enumerated.

The raw data have a temporal structure which is projected onto a spatial line in written representations. These representations are synchronized more or less closely. For instance, representation (f) generally matches L1 sentences, units of level (g) may be associated with L1 units of any size, and representation (e) may match representation (d) morpheme by morpheme. This has different consequences for the typographic layout. For instance, units of level (g) may be associated with the running text by making full use of a multidimensional display, while representation (f) may be in a lateral column at the same height as its original, as is usual in synoptic editions and also practiced in the example from Finck (1909) given in section 1.2. Other representations should be arranged in lines one of which is beneath the other and runs in parallel with it.

For the purposes of descriptive and typological grammatical analysis and exemplification, the seven-level set is generally reduced to only three. What may be called the ‘canonical trilinear representation’ of linguistic examples involves:

- a representation of L1 at one of the levels (b), (c) or (d),
- an IMG in L2 (level e),
- an idiomatic translation into L2 (level f).

An IMG will seldom be paired with a phonetic representation, because this serves phonetics, while an IMG serves grammar. They therefore form an unequal pair. If both are required, they will generally be mediated by another representation, a morphophonemic or orthographic one.

It makes a difference for the glossing whether L1 is rendered in a morphophonemic representation or in conventional orthography. In the former case, the rules of

orthography do not apply, and the linguist may dress up the representation in such a way that a biunique mapping onto the IMG is facilitated. In the latter case, morpheme boundaries may be obscured by the orthography, and there will be delimiters such as blanks, hyphens and punctuation marks which do not necessarily represent grammatical boundaries and may interfere with the glossing. However, the choice between an orthographic and a scientific representation of a text is generally a higher-order choice which cannot depend on glossing requirements. In particular, an example may be quoted unchanged from a primary source (think of Sanskrit examples). It may then not be possible to insert boundary symbols and the like in the L1 text. Glossing conventions therefore have to be adjusted to use with orthographic representations.

If the first line representing the L1 text differs too much from a morphophonemic representation, then it is advisable to expand the canonical trilinear representation by an additional morphophonemic representation. It will then be this line that the IMG refers to.

The two languages involved will be called L1 and L2 throughout. However, it should be clear that the relationship between them is asymmetric: L1 is the object language, L2 is the metalanguage. The symbols occurring in an IMG have a different status from the elements of the text line that they gloss: For present purposes, the L1 text line consists of morphs, while the IMG consists of names of L2 morphemes and of grammatical categories (cf. section 3.2). There can, thus, be no question of “mirroring” the structure of the L1 expression by the sequence of the L2 elements. Instead, an element in an IMG serves as a kind of mnemonic hint to the meaning or function of its corresponding L1 element.

#### 1.4. Delimitation

The complete set of representations rendering an L1 text may be sufficient to derive a grammatical description from it (as postulated in Lieb & Drude 2000, § 1.1). However, given its inherent restrictions, an IMG cannot by itself compensate for a grammar (or just a morphology). Apart from the form of presentation, the most important substantive difference between a grammatical description and an IMG lies in the fact that the grammar treats of categories in the sense of classes, while the IMG identifies individual morphemes. For instance, a grammar treats of

the verbal category of aspect. An IMG contains a gloss for an individual aspect morpheme, e.g. *PERF*, neglecting the question of whether this is actually an aspect morpheme or rather a tense morpheme, and also leaving unanswered questions concerning other members of the paradigm, let alone the construction and use of the *PERF* morpheme. Some of these kinds of information may be given in other representations, e.g. a syntactic representation.

By the same token, the IMG does not indicate the syntactic category of a word form. For instance, the IMG of Germ. *laufend* is ‘run:PART.PRS’, showing that the form contains a morpheme whose function it is to mark a present participle. The gloss is not ‘run (part.prs)’ or anything of the sort, meaning that *laufend* is a present participle. While the latter is true, it is not the task of an IMG to give this information.

Moreover, the type of morphological unit is not an object of an IMG. Thus, concepts like ‘stem’, ‘root’, ‘prefix’ do not appear in IMGs. Such information may, to a large extent, be inferred from a proper IMG, since the gloss of a root differs typographically from the gloss of a grammatical formative.

Similarly, an IMG cannot replace a lexicon. Here again, elements appearing in an IMG are but names of elements appearing in the L1 line. They are not meant to exhaust the meaning of such an element.

Finally, an IMG is not meant to replace an idiomatic translation. Thus, it cannot and should not render closely the sense of an L1 item in the given context. An IMG is regularly accompanied by a free translation which fulfills precisely this purpose.

## 2. Prerequisites of morphological analysis

Interlinear glossing might appear to be just an elementary form of representing data. As a matter of fact, it presupposes a morphological analysis. The following analytic problems are directly reflected by the glosses.

### 2.1. Unmarkedness and zero morphemes

Where the L1 text contains a morph, the IMG contains an element rendering it. Where the L1 text contains nothing, the issue of rendering it is complicated by markedness theory. Germ. *Herr* may be glossed by ‘master’ or by ‘master(NOM.SG)’. Latin *mone-*

may be glossed by ‘warn-3.SG’ or by ‘warn (IND.ACT)-3.SG’ (according to R16). Moreover, one may believe that such forms contain zero morphemes and put thus: *Herr-Ø* ‘master-NOM.SG’, *mone-Ø-Ø-t* ‘praise-IND-ACT-3.SG’. All of these IMGs are formally correct. The choice among them is not a matter of appropriate glossing, but of morphological theory. For interlinear glossing, only the general rule R1 is relevant.

### 2.2. Allomorphy

If the L1 representation to be glossed corresponds to standard orthography, the analyst has no decisions to make in its regard. Otherwise, a good option for the representation (as well as for any writing system) is a morphophonemic representation which steers a middle course as far as allomorphy is concerned: Phonologically conditioned allomorphy is resolved (ignored), morphologically conditioned allomorphy is not resolved (is rendered).

The IMG, on the other hand, shows morphemes, not allomorphs. In order to understand what this implies, consider three examples. Modern Yucatec Maya expresses complete and incomplete aspect by suffixes on transitive and (one conjugation class of) intransitive verbs as follows:

	aspect valence	completive	incomplete
transitive		-ah	-ik
intransitive		-Ø	-Vl

Tab. 169.1: Aspectual suffixes in Yucatec Maya

For instance, *t-u hats'-ah* ‘PAST-SBJ.3 beat-CMPL (he beat it)’. One might think that the table contains four morphemes. Actually, however, transitivity is inherent in the verb stem and conditions allomorphy in the aspect suffix. The conditioning factor should not make part of the gloss. That is, the correct gloss for *-ah* is not ‘TR.CMPL’, but simply ‘CMPL’. See also 4.5.

Yucatec Maya also has personal clitics that precede nouns as possessive cross-reference markers and verbs as subject cross-reference markers. If the noun or verb starts with a vowel, a glide is inserted in its front. The choice between the two glides *w* and *y* is morphologically conditioned: If the pronoun is of first person singular or of second person, it

is *w*; if the pronoun is of third person, the glide is *y*. For instance, *in watan* ‘POSS.1.SG Ø:wife (my wife)’, *u yatan* ‘POSS.3.SG Ø:wife (his wife)’. It is also possible to regard the noun forms modified by the initial glide as stem allomorphs, in which case the glide would not even receive the gloss by ‘Ø’. However, in the third person, a pronominal clitic followed by the glide can be omitted. Thus, *yatan* by itself means ‘his wife’. (Historically, the glide is indeed a reflex of an older cross-reference marker). We therefore have *u y-atan* ‘POSS.3 Ø-wife’ ~ *y-atan* ‘POSS.3-wife’, and we face the problem that the same element is not even a morph in one context, but a full-fledged morpheme in another. Whatever the correct morphological analysis may be, the IMG presupposes it and brings it out.

Last, consider gender marking in a language such as Latin (cf. Art. 48). *Puellae bonae* means ‘good girls’, *pueri boni* ‘good boys’. Apart from motion, gender is inherent in a noun stem. It is, however, recognizable by the declension suffixes. Nevertheless, the gloss of the morph in question does not contain the conditioning category. The noun forms will be glossed ‘girl.F:NOM.PL’, ‘boy.M:NOM.PL’, implying that gender is a category of the stem, not of the suffix. What about the adjectives? Gender is not inherent in an adjective stem. We may therefore gloss them by ‘good:NOM.PL.F’ and ‘good:NOM.PL.M’. Then one and the same element would be a morpheme on adjectives, but a conditioned allomorph on nouns, and therefore it would get two different glosses. Since two different glosses for the same element are not admissible in interlinear glossing (R4), this would entail that there are two homonymous declension suffixes *-ae* in Latin, which is obviously undesirable. We may stop this consideration here, since the problem is obviously not one of glossing, but one of morphological analysis. R2 codifies the convention that IMG expressions represent morphemes, not allomorphs.

### 3. Principles of interlinear glossing

#### 3.1. General

In the canonical trilinear representation, one L1 text line is matched by two L2 lines, the IMG and the free translation. This entails a division of labor between the two L2 representations. The free translation is the idio-

matic semantic equivalent of the L1 line, the IMG is a representation of its morphological structure. There is consequently no need for the translation to be particularly literal, just as there is no need for the IMG to repeat the morphemes that appear in the translation. For instance, a polysemous L1 item will be rendered by its contextual sense in the free translation, but by its generic meaning in the IMG (R8). Unnecessary parallelism between the two L2 lines is redundant; the trilinear canonical representation offers an occasion to provide additional information.

In principle, the degree of detail displayed in an IMG depends on the purpose the example with its gloss is meant to serve. However, the author cannot foresee the purposes to which others will want to use his examples. A morphological detail that is not at stake in the current discussion may be essential for the argument another linguist may wish to base on the example. For this reason, the principle is to allow for as much precision and detail as seems tolerable (R3). The following rules specify the properties of a complete IMG. They do not exclude less detailed IMGs where they suffice. Cf. R13 and R23 for possibilities of underspecifying morphological structure.

The IMG of a morpheme is some sort of name for it, a name that alludes to its meaning or function and is insofar mnemonic or, at least, more helpful to the non-specialist than the L1 morph itself. It must therefore have a certain recognition value. R4, which actually is a tightening of R1, therefore requires that given a particular L1 morpheme, its gloss will be the same in all contexts; and apart from full synonymy, no two morphemes of L1 will have the same gloss. These points will be elaborated in the following subsections.

#### 3.2. Glossing vocabulary

Glosses are taken from a language L2 that serves as a metalanguage of L1. L2 is based on a natural language – in this article, English –, but with far-reaching deviations from natural language use. The glossing vocabulary consists of the following kinds of symbols:

- vocables:
  - L2 morphemes and stems
  - grammatical category labels
  - boundary symbols

The difference between the two kinds of vocables is the following: Morphemes and stems

are taken from natural L2 vocabulary and are meant to be translation equivalents (in a sense to be made precise below) of L1 items. For instance, the notation “Germ. *Schreibtisch* ‘write-table (desk)’” is to be interpreted thus: The German word form *Schreibtisch* ‘desk’ consists of two morphs, of which *schreib-* means ‘write’ and *tisch* means ‘table’. Grammatical category labels, on the other hand, are taken from scientific terminology and are meant to categorize the function of L1 items. For instance, “Germ. *schreib-en* ‘write-INF (write (inf.))’” is to be interpreted thus: The German word form *schreiben* ‘write (inf.)’ consists of two morphs, of which *schreib-* means ‘write’, while *-en* is an infinitive marker (that is, *-en* does not mean ‘infinitive’; it is the German word *Infinitiv* which means ‘infinitive’). To bring out this essential difference between the two kinds of IMG vocables, L2 morphemes and stems are written in straight orthography, while grammatical category labels are written in (small) capitals (R29).

A grammatical category label represents (i.e. is the name of) the value of a grammatical category (the latter being taken, technically, as a parameter or attribute). For instance, the label ‘ACC’ is the name of the value ‘accusative’ of the morphological category ‘case’. Just like a grammatical category label is a name of a value of a grammatical category, what is called ‘L2 morphemes and stems’ are actually names of L2 morphemes and stems. In the following, we will abide by the simpler way of speaking. The choice and use of vocables are treated in the following subsections; boundary symbols are treated in section 4.

### 3.3. Lexemes

An L1 lexeme is, in principle, glossed by an L2 lexeme (R5(a)). Sometimes more than one L2 word is necessary, for instance in Germ. *fabulieren* ‘invent.stories’. However, profusion is to be avoided. Adjectives that do not require a copula in predicative function are often glossed by adding a copula, e.g. West Greenlandic *anurli* ‘windy’ is glossed ‘be.windy’ in Fortescue (1984:65). This is only correct if a word of this class requires an attributor in attributive function. Otherwise it wrongly implies that there is no difference between adjectives and verbs, and it tends to obscure the fact that the language does not use a copula with adjectival predicates.

L1 cardinal numerals are glossed by Arabic numbers. An issue arises for proper names, which are often not glossed at all. However, there is no room here for an exception to the general rule: a proper name is rendered by its counterpart in L2. Some proper names have conventional counterparts that are specific to L2; Engl. *John* corresponds to Germ. *Hans*, and Engl. *Munich* corresponds to Germ. *München*. These then appear in the IMG. Whenever there is no such language-specific convention, the counterpart of an L1 name is usually the same word in L2.

If L2 is English, no problem arises for the form in which L2 lexemes are quoted in the IMG. In other languages, lexemes have a citation form in conformity with L2 conventions. If this is an inflected form, like the nominative for nouns or the infinitive for verbs, then it is excluded from an IMG by R5(b), and instead the bare stem must be used. The reason is that such a gloss would seem to imply that there is a nominative, or an infinitive, in the L1 line where actually just a stem is being glossed.

### 3.4. Grammatical formatives

L1 morphs are, in principle, glossed by citation forms of L2 morphemes. However, interlinear morphemic glossing crucially revolves around grammatical properties of L1 items. These will differ between L1 and L2. Even if, in a number of cases, the L2 stem appearing in a gloss has the same grammatical properties as the L1 morph that it represents, this cannot be expected and therefore not be relied upon. For instance, Latin *eum* could be glossed by Engl. *him*, and at the typological level, they do share a number of features. However, *eum* is accusative and can thus not be indirect object, while *him* is the form for direct and indirect object. Therefore, grammatical items of L1 are generally not glossed by grammatical items of L2, but by a configuration of symbols taken from the scientific metalanguage and representing their grammatical features, i.e. by grammatical category labels (R6). Thus, Latin *eum* may be glossed by ‘ANA:ACC.SG.M’.

No bound grammatical or derivational morphemes should appear in IMGs. Free grammatical morphemes may be used to render free grammatical morphemes. However, use of those in the second column of Tab. 169.2 is discouraged unless L1 happens to exhibit the same ambiguity as English:

word class	instead of	use
copulas, auxiliaries	<i>be</i> <i>have</i> (except to mean ‘possess, own’)	COP, PASS, PROG ... PF, OBLG ...
prepositions	<i>by</i> <i>with</i> <i>for</i> <i>as</i> <i>from</i> <i>to</i> <i>of</i>	AG, ERG ... INST, COM, ASSOC ... BEN, DEST ... EQT, ESS ... ABL, DEL ... DAT, ALL, DEST, TERM, INF ... GEN, ASSOC ...
subordinators	<i>that</i> <i>if</i>	COMP, SR (, D3) INT, COND.SR
relativizers	<i>that</i> <i>who</i> <i>which</i>	REL REL.HUM.NOM ... REL.NHUM.NOM ...

Tab. 169.2: Free grammatical morphemes

Some morphemes are extremely deeply entrenched in the semantic or pragmatic system of the language and simply have no translation equivalent in L2. Two common ways out are a) to repeat the significans of the item in the gloss, and b) to indicate the class of the item instead of its meaning. Thus, we find the German modal particle *eben* glossed either as ‘EBEN’ or as ‘PTL’. Both glosses are inadequate. If there is no translation equivalent in natural L2, then the linguist has a specialized metalanguage to describe such functions. For the sake of an IMG that is not devoted to modal particles in particular, a gloss like ‘REAFF’ (reaffirmed) will be fully sufficient and more helpful than either of the aforementioned.

A gloss is a proper name of an L1 morpheme. It does not give information on the grammatical class of the morpheme in question other than what is implied by the name itself. If a gloss is ‘ACC’, one assumes that the morpheme belongs to the grammatical class of the case morphemes. It is the task of the grammar to clarify whether or not this implication is correct in a particular case. The gloss will not be ‘CASE.ACC’ or anything of this sort. For the same reason, the gloss of the perfective aspect is simply ‘PRFV’ and not ‘PRFV.ASP’, and so on.

From this it follows that the gloss will not be ‘ASP’ either. In the literature, one frequently encounters glosses such as ‘PTCL’ (particle), ‘AGR’ (agreement), ‘ART’ (article). If

L1 possesses only one particle, agreement morpheme (hardly imaginable) or article (this is possible), then these glosses are sufficient. In all other cases, this kind of gloss is not helpful because it does not give the information on the meaning or function of the morpheme that a gloss is supposed to give. Moreover, the whole glossing becomes inconsistent, as some glosses name particular morphemes, while others name the class a morpheme belongs to. More on this in section 3.9.1.

### 3.5. Ambiguity

Each morpheme of L1 should be recognizable by its gloss. The reader is supported in this task if glosses are consistent within one publication. It will rather confuse him if Yucatec Maya *k'in* is once glossed ‘sun’ and the next time ‘day’. **Polysemy** is resolved in the idiomatic translation. The gloss renders neither the contextual sense nor the full meaning range of an item. Naturally, this does not apply to **homonymy**. Homonymous L1 morphs represent different morphemes and therefore receive different glosses. This is stipulated by R7, which follows from R4.

If the senses of an item are reducible to a **Gesamtbedeutung**, then this should be used in the gloss (R8). For instance, the Turkish dative/allative suffix *-a* is glossed by ‘DAT’. The **Gesamtbedeutung** rather than the **Grundbedeutung** should appear in the gloss, because it has better chances to fit all the diverse

contexts in which the item occurs. Sometimes, there is either no Gesamtbedeutung, or if there is, L2 does not have a term for it. In cases like Yucatec Maya *k'iin* 'sun, day', there are various alternatives. First, the Grundbedeutung may be used as the gloss; thus Yucatec Maya *k'iin* 'sun'. However, if all the occurrences of a polysemous morpheme in a particular publication reflect the same (derived) reading, then generally no useful purpose is served if it is consistently glossed by its basic meaning. For instance, all the occurrences of Yucatec Maya *k'iin* in a particular text might mean 'day'. Then this would be the appropriate gloss. Finally, any kind of reduction may seem misleading. Then two or even more senses may be indicated in the gloss, separated by a slash, e.g. Yucatec Maya *k'iin* 'sun/day'. (2) illustrates the same convention.

### (2) Korean

<i>Toli-nun kae-hako cal</i>
Toli-TOP dog-ADD often/well
<i>non-ta.</i>
play:PRS-DECL

'Toli likes to play with the dog.'

**Syncretism** often involves extensive polysemy and/or homonymy. If it should be made explicit in an IMG, then e.g. the gloss for Lat. *ancillae* would have to be 'maid.F:GEN.SG/DAT.SG/NOM.PL'. This may be appropriate if the discussion in the context deals with syncretism. Otherwise, only the category actually required by the context may be shown, e.g.:

### (3) Latin

<i>ancillae</i>	<i>orant</i>
maid.F:NOM.PL	pray:3.PL
'the maids	'pray'

In other words, in cases of syncretism the last two options of R8 must be resorted to.

A whole paradigm of markers may be used in two clearly distinct functions. For instance, a set of cross-reference markers may combine with a verb to reference its subject, and with a noun to reference its possessor. Here again, the two alternatives mentioned are open: either gloss the verb markers by 'SBJ' and the noun markers by 'POSS', or gloss them by 'SBJ/POSS' in both positions (which is, actually, never done). A third alternative – one that is actually resorted to in Mayan linguistics; cf. Art. 170, section 6.1.2 – is to coin

a concept and a term for a paradigm that is used in these two functions and use this in the IMG.

### 3.6. Features and functions

As remarked in section 1.4, an IMG cannot fill the place of a grammar. In particular, the **grammatical category label** that represents a morpheme in the gloss cannot possibly represent the full functionality of that morpheme. It can only serve as a mnemonic identifier for the reader. We just saw that the full polysemy of an item cannot be accounted for in a gloss. The same goes for functional information associated with a morphological position. If the slot filler is a verb agreement affix or cross-reference marker, then its meaning is in the sphere of person, number and gender. Consider conjugation endings as in Germ. *lieb-e* 'love-SBJ.1.SG', *lieb-st* 'love-SBJ.2.SG', *lieb-t* 'love-SBJ.3.SG'. The information that these suffixes cross-reference the subject is functional information associated with the morphological slot. It must be given in the grammar; the IMG may simply read *lieb-e* 'love-1.SG' etc.

The same would apply, in principle, if the verb cross-references more than one of its dependents. Here, however, it has become customary to distinguish the references of the cross-reference markers by indicating their syntactic function, as in (4).

### (4) Swahili

<i>ni-li-mw-on-a</i>	<i>m-toto</i>
SBJ.1.SG-PST-OBJ.CL.1-see	CL.1-child
'I saw the/a child'	

The information that the initial prefix references the subject, while the one following the tense prefix references the direct object must be contained in the grammar. The task of the gloss is to identify the particular element, not to specify the rules of its use. Insofar, adding functional information concerning the morphological slot itself – 'SBJ' and 'OBJ' in (4) – is a service to the reader that may be useful, but that also clutters up the gloss (cf. R3).

The distinction between morphological categories and syntactic or semantic functions is also relevant in the domain of case and valence. The frequent confusion among syntactic/semantic functions, cases and valence-derivational functions also manifests itself in glossing habits. One frequently encounters glosses such as Turkish *ates-in* 'fire-

POSS' instead of 'fire-GEN', *ates̩-e* 'fire-io' instead of 'fire-DAT' or '... -send-DAT ...' instead of (5). The quality of the glossing reflects the quality of the morphological analysis.

(5) Swahili

<i>Musa a-li-ni-andik-ia</i>
Musa SBJ.CL.1-PST-OBJ.1.SG-send-APPL
<i>barua</i>
letter

'Musa sent me a letter'

### 3.7. Derived stems

The morpho-semantic structure of a derived stem may be completely regular and transparent, as in Germ. *wolk-ig* 'cloud-ADJVR (cloudy)', or it may be opaque, as in Germ. *heil-ig* 'salvation-ADJVR (holy)'. If the discussion focuses on word-formation, then both of these words will be glossed as indicated. If the internal structure of stems is of no relevance, then it will not be shown in the L1 text line, and consequently the glosses can reduce to 'cloudy' and 'holy', respectively.

For opaque complex stems, morphological segmentation plus corresponding gloss often amounts more to etymology than to morphological analysis. It also unnecessarily obscures the correspondence of the gloss to the idiomatic translation. This should be borne in mind before one carries it through as a general principle in text editions.

In an ideal methodological situation, an IMG is taken from a lexicon, where the gloss constitutes one of the fields in the microstructure of each lexical entry. The German lexicon may contain, e.g., the three entries *Huf* 'hoof', *Eisen* 'iron' and *Hufeisen* 'horse-shoe'. If the latter occurs in an L1 text, then it may either be analyzed or not. In the former case *Huf* and *Eisen* will be looked up in the lexicon and will be matched by their glosses, while in the latter case *Hufeisen* will be looked up and be glossed accordingly.

### 3.8. Submorphemic units

There are two kinds of submorphemic units: parts of morphemes with a sound-symbolic value and strings of phonemes inserted between morphemes for euphonic or similar reasons. The former kind is not generally subjected to morphemic analysis and may therefore be left out of consideration here. The latter kind may be illustrated by the second element in forms such as French *a-t-il* 'has he' and Germ. *Weihnacht-s-gans* 'Christmas goose'. If the submorphemic unit is not

at stake in the context, then the first choice is to abstain from an analysis by regarding the submorphemic unit as part of a stem alternant: *Weihnachts-gans* 'Christmas-goose'. The second choice is to render the submorphemic unit by Ø, e.g. *a-t-il* 'has-Ø-he'. A euphonic submorphemic unit may be glossed by 'EU' instead of 'Ø'.

### 3.9. Grammatical category labels

#### 3.9.1. General

As was said in 3.4, the gloss for a grammatical item is generally not a grammatical item of L2, but a grammatical category label (R6). For instance Yucatec Maya *yàan* is not rendered by 'be', but by 'EXIST', one of the reasons being that L2 'be' is a copula, while Yucatec Maya *yàan* is not. While this poses few problems for such categories for which the European grammaticographic tradition possesses terms, it does pose a problem for certain classes of semi-grammaticalized items such as function verbs and coverbs. Coverbs are words which are grammaticalized from verbs to minor parts of speech, mostly adpositions. If they function as the latter, they may express a semantic role. In Mandarin, for instance, *yòng* has the lexical meaning 'use' and the grammatical meaning 'INSTR', as in (6).

(6) Chinese

<i>Tā yòng shǒu zǒu lù.</i>
he use/INSTR hand walk road
'He walks on his hands.'

This kind of problem is not solved by putting the lexical meaning in upper case ('USE'), since 'use' is neither a grammatical concept in L2 nor a term of the grammatical metalinguage. Applying R8 in such cases would imply opting in favor of the Gesamtbedeutung of the item, which in such cases is the grammatical meaning. The gloss would then be 'INSTR' (or some more language-specific grammatical category which may suit better this particular function). The problem remains, however, that the same word can occur as the sole predicate of a clause, in the meaning 'use' (e.g. *tā yòng shǒu* 'he uses his hand'). An IMG 'INSTR' would be hardly intelligible there. The alternative of only using the Grundbedeutung – 'use' in (6) and throughout – would be in conflict with the principle that morphological analysis must be kept distinct from etymology. Here the third alternative offered by rule R8 may be re-

label	intended meaning	comment
A	transitive subject	in morphemic glosses, the abbreviation is ERG
ADV	adverb	specify meaning
AGR	agreement	specify agreement categories
AGT	agent	this is not a value of a morphological category
ART	article	only if it has no determinative properties
ASP	aspect	specify particular aspect
AUX	auxiliary	only if there is only one auxiliary morpheme in the language
CARD	cardinal	only if it is a morpheme or grammatical feature
CLF	classifier	this is a word class
CLT	clitic	this is neither a morphological category nor a value of one
EP	epenthetic	has no morphological status, should not be separated in the first place
EVID	evidential	specify particular evidential
PAT	patient	this is not a value of a morphological category
PREP	preposition	this is a word class
PTL	particle	this is (at best) a word class
TNS	tense	specify particular tense

Tab. 169.3: Labels to be avoided

sorted to, viz. providing both meanings in the gloss of each occurrence of the item, thus: *yòng* ‘use/INSTR’.

An IMG identifies an L1 morpheme. It names a value, not a parameter. Mentioning the name of the generic category in the gloss instead of the specific value is nevertheless widespread usage. One finds both Japanese *yom-i* and *yon-de* glossed by ‘read-CONV’ (converb), which hinders the reader in his attempt to keep the converb forms apart. One finds Onondaga *wa?ha-ye?kwa-hni:-nu?* ‘he bought tobacco’ glossed as ‘TNS:he/it-tobacco:buy-ASP’ (Woodbury 1975: 10), which is of no use for somebody studying the interdependence of incorporation with tense and aspect.

IMGs not seldom contain labels that do not correspond to the principles introduced so far. Sometimes, elements without morphological status are separated and glossed. Sometimes, the parameter instead of the particular value of a grammatical category is identified. Sometimes, syntactic or semantic instead of morphological information is given. Here is an incomplete list of labels that have repeatedly been found in glosses but which should be avoided.

### 3.9.2. List of grammatical categories and their glossing labels

No list of grammatical category labels can be complete. The list following in Tab. 169.4

(which incorporates the list in Lehmann et al. 2019) only contains the most widespread categories. When more than one abbreviation is mentioned, they are given in the order of preference. To the extent that these abbreviations are or become wide-spread, they get the status of linguistic abbreviations like ‘NP’, which need not be defined when used. If a publication uses labels not contained in the following list, it must explain them in an individual list of abbreviations.

Grammatical category labels are subject to two conflicting requirements: they must be both distinct and short. The former requirement takes precedence. It is, for instance, not possible to use ‘COMP’ in one and the same publication to mean both ‘completive’ and ‘complementizer’. The list in Tab. 169.4 avoids such clashes. However, in an individual publication that has nothing to do with complementation, the aspect may, of course, be abbreviated by ‘COMP’ (instead of ‘CMP(L)’, as in the list). Parenthesized parts of an abbreviation are only necessary if a distinctness conflict arises.

Tab. 169.4 contains only such terms which may appear in an IMG. In other publications, similar lists of terms for syntactic categories and functions and for semantic and pragmatic functions may be found.

‘Cross-reference position’ means a morphological slot, usually on a verb, occupied by pronominal elements that agree with or

value	abbrev.	category	comment
1 <sup>st</sup> person	1	person	
2 <sup>nd</sup> person	2	person	
3 <sup>rd</sup> person	3	person	
abessive	(PRV) (AVERS)		use ‘privative’ and ‘aversive’
ablative	ABL	local case	‘from’ (= separative)
absolute	ABSL	nominal	free non-incorporated form of noun
absolutive	ABS	grammatical case or cross-reference position	in ergative system
abstract	ABSTR	nominal	
accusative	ACC	grammatical case	
action nominalizer	ACNNR	deverbal nominal derivation	
active	ACT	1. voice; 2. case or cross-reference position	2. in active system
actor	ACR	grammatical case or cross-reference position	
actor topic	A	voice	
additive	ADD	case	
addressee-honorific	2HON	honorification	
addressee-humble	2HML	honorification	
adelative	ADEL	local case	
adessive	ADESS	local case	
adhortative	(HORT)		use ‘hortative’
aditive	(ALL)		use ‘allative’
adjectiv(al)izer	ADJR	derivational or syntactic mood	
admonitive	ADM		
adverbializer	ADVR	derivational or syntactic	
adversative	ADRVS	interpropositional relation	‘whereas’
affirmative	AFFMT	opposite to negative	normally unmarked
agent nominalizer	AGNR	deverbal nominal derivation	
agentive	AG		
alienable	AL	possessive attribution morpheme	
allative	ALL	local case	‘to’
allocutive	ALLOC	honorification	kind of addressee-honorific
anaphoric	ANA	pronominal	
andative	AND	deictic	
animate	AN		
anterior	ANT	tense	relative tense
anticausative	ACAUS	deverbal verb derivation	= deagentic, blocking of actor argument
antipassive	APASS	voice	
aorist	AOR	tense-aspect	perfective past (as opposed to imperfect)
applicative	APPL	deverbal verb derivation	subtypes may be distinguished by APPL.REC, APPL.INST etc.
apprehensional	APPR	interpropositional relation	‘lest’
assertive	ASRT	modality	subtype of declarative: high degree of commitment
associative	ASS(OC)	adnominal case	‘with, à’
assumed	ASSUM	evidential	
attenuative	ATTEN	deverbal verb derivation	
attributor	AT	nominal	links an attribute to the head
auditory	AUD	evidential	
augmentative	AUG	denominal nominal derivation	
auxiliary	AUX		if it is the only auxiliary root
benefactive	BEN	case	‘for’
cardinal	CARD	numeral	if marked grammatically use ‘privative’
caritive	(PRV)		

Tab. 169.4: *continued*

value	abbrev.	category	comment
causative	CAUS	deverbal verb derivation	
circumstantial	CIRC	interpropositional relation	'in, by'
clamative	(EXCL)		use 'exclamative'
classifier	CLF	nominal	followed by class identifier, e.g. HUM
cohortative	(HORT)		use 'hortative'
collective	COLL		
comitative	COMIT	case	'with, in the company of'
common	COMM	gender	either masc. or fem.; cf. 'human' and 'animate'
comparative	CMPR	degree of comparison	
complementizer	COMP	subordinator	= SR
completive	CMP(L)	aspect	normally = perfective
conative	CNTV	mood	
concessive	CONC	interpropositional relation	'although'
conditional	COND	interpropositional relation; mood	'if'; 'would'
conjectural	CONJC	evidential	
conjunctive	CONJ	interpropositional relation	of non-finite predicate
connector, -ive	CONN		if there is only one
consecutive	CONSEC	interpropositional relation	'so that'
construct	CONST	nominal	construct state
converb	(GER)		use 'gerund'
continuous	CONT	aspect/aktionsart	
copula	COP		
crastinal	CRAS	tense	if there is only one
dative	DAT	grammatical case	tomorrow
deagentive	(ACAUS)		use 'anticausative'
deitive	(OBLG)		use 'obligative'
declarative	DECL	sentence-type	normally unmarked
deferential	DEFR	honorification	~ speaker-humble
definite	DEF	determination	
deictic of 12 person	D12	determination	
deictic of 1 <sup>st</sup> person	D1	determination	
deictic of 2 <sup>nd</sup> person	D2	determination	
deictic of 3 <sup>rd</sup> person	D3	determination	
delative	DEL	local case	'down from'
demonstrative	DEM	determination	
dependent verb form	(SUBJ)		use 'subjunctive'
desiderative	DES	deverbal verb derivation	
destinative	DEST	local case; also on non-finite verb forms (= supine)	'to'; if typically for human destinations, use 'benefactive'
determiner	DET	pronominal	will normally be DEF, INDEF, GNR, SPEC, NSPEC
detransitivizer	DETR	deverbal verb derivation	see also 'anticausative' and 'introversive'
different subject	DS		
diminutive	DIM	denominal noun derivation	
direct	DR	voice	vs. inverse
direct evidential	DIREV	evidential	
direct object	DO	cross-reference position	
directional	DIR	case or verb derivation	'towards'; use AND and VEN for deictic directionals
distal	DIST	determination	remote from deictic center
distributive	DISTR	nominal or verbal	
donative	DON		auxiliary of benefactive construction

Tab. 169.4: *continued*

value	abbrev.	category	comment
dual	DU, DL	number	
dual exclusive	DE	number	
dual inclusive	DI	number	
dubitative	DUB	mood	
durative	DUR	aktionsart	
dynamic	DYN	aktionsart	vs. stative
egressive	EGR	aktionsart	
elative	ELAT	local case	'out of'
emphasizer/emphatic	EMPH	funct. sentence perspective	e.g., class of pronoun
equative	EQT	1. case; 2. predicative	'as'; feature/marker of adjective in no- nominal clause
ergative	ERG	grammatical case or cross-ref- erence position	in ergative system
essive	ESS	case	'as'; see also 'transformative'
evidential	EVID	verbal	if there is only one
exclamative	EXCL	mood	
exclusive	(DE, PE)		use 'dual exclusive', 'plural exclusive'
exist(ential)	EXIST	grammatical verb	
experiential	EXPER	aspect	
extrafocal	EXFOC	verbal	status of subordinate clause of cleft-sentence
extraversive	EXTRV	deverbal verb derivation	transitivization by addition of un- dergoer
factitive	FACT	denominal/deadjectival verb derivation	A-FACT NP 'make NP A'
familiar	FAM	pronominal	
feminine	F	gender	
fientive	(INCH, PROC)	denominal verb derivation	use 'inchoative, processive'
finite	FIN	verbal	
first person dual inclusive	12		if treated as a quasi-singular; otherwise 'dual inclusive'
focus	FOC	funct. sentence perspective	
formal	FRM	mood	
frequentative	FREQ	aktionsart	multiple times on several occa- sions
future	FUT	tense	
generic	GNR	determination	
genitive	GEN	grammatical case	
gerund	GER	verbal	verbal adverb or converb
gerundive	(OBLG)	use 'obligative'	~ customary
habitual	HABIT	aktionsart	use 'habitual', 'generic' use 'habitual', 'past'
habitual-generic			
habitual-past			
hesitative	HESIT	funct. sentence perspective	yesterday's past
hesternal	HEST	tense	today's future
hodiernal future	HODFUT	tense	today's past
hodiernal past	HODPST	tense	
honorific	HON	honorification	1 <sup>st</sup> person imperative
hortative	HORT	mood	
human	HUM		comprises 'speaker-humble, ad- dresser-humble, referent-humble'
humble	HML	honorification	
hypocoristic	HCR	affect	
hypothetical	HYP	mood	'into'
illative	ILL	local case	
immediate	IMM	tense	specifier of other tenses

Tab. 169.4: *continued*

value	abbrev.	category	comment
immediate/imminent future	IMMFUT	tense	
immediate past	(RECPST)		use 'recent past'
imperative	IMP	mood	
imperfect	IMPF	tense-aspect	imperfective past; vs. aorist
imperfective	IPFV	aspect	
impersonal	IMPR		only if formally distinct from the specific persons
impersonal passive	IPS	voice	passive without promotion to subject in active system
inactive	INACT	grammatical case or cross-reference position	
inalienable	INAL	nominal	possessive attribution morpheme or feature
inanimate	INAN		use 'ingressive'
inceptive	(INGR)		N/A-INCH 'become N/A'
inchoative	INCH		use 'dual inclusive', 'plural inclusive'
inclusive	(DI, PI)		normally = imperfective
incomplete, non-completive	INCMP(L)	aspect	
inconsequential	INCONS	interpropositional relation	
indefinite	INDEF	determination	
independent	INDEP	mood	only if distinct from indicative
indicative	IND	mood	
indirect object	IO	cross-reference position	
inessive	INESS	local case	'inside'
inferential	INFR	mood or evidential	
infinitive	INF	verbal	
ingressive	INGR	aktionsart	
injunctive	INJ	mood	
instructive	(MAN)		use 'manner'
instrument nominalizer	INSTNR	deverbal nominal derivation	
instrumental	INST(R)	case	
intensive	INTS	verbal	often aktionsart
interrogative	INT	sentence type	particle or morphological category
intransitive	INTR	verbal	morpheme or grammatical category
intransitive subject	S	cross-reference position	only if opposed to both A and P; use SBJ otherwise
introversive	INTRV	deverbal verb derivation	blocking of undergoer argument vs. direct
inverse	INV	usually verbal	
invisible	INVS	determination	
irrealis	IRR	mood	
iterative	ITER	aktionsart	several times on one occasion
jussive	JUSS	mood	3 <sup>rd</sup> ps. imperative or dependent mood
lative	LAT	local case	'to ~ from ~ via'
ligature	LIG	nominal	
linker	LNK	nominal	links subconstituents of a phrase, typically an NP; properly includes 'attributor'
locative	LOC	local case	
locative topic	LT	voice	
manner nominalizer	MANNR	deverbal nominal derivation	
masculine	M	gender	
masculine personal	MHUM	gender	

Tab. 169.4: *continued*

value	abbrev.	category	comment
logophoric	LOG	pronominal or verbal	
malefactive	MAL	deverbal verb derivation	
manner	MAN	case	
medial	MED	determination	also on non-finite verbs medial distance from deictic center
medial	MEDV	verbal	verb form in a chain
mediative	MEDT	case	'between, among; by means of'
mediopassive	MEDP	voice	
middle	MID	voice	excludes passive
motivative	MTV	case	'by'; sometimes called 'causal'
narrative	NARR	tense	
near future	NRFUT	tense	after 'immediate future'
negative	NEG		
neuter	N	gender	
nominalizer	NR	deverbal nominal derivation or syntactic subordination	see also the more specific ones
nominative	NOM	grammatical case	
non-	N		e.g. NPST
non-finite	NFIN	verbal	
non-future	NFUT	tense	
non-human	NHUM	gender	
non-masculine personal	NM	gender	
non-past	NPST	tense	< 3
non-plural	NPL	number	> 1; only if there is a plural for > 2
non-singular	NSG	number	
non-specific	NSPEC	determination	
non-visual	NVIS	evidential	
non-volitional	NVOL	verbal	non-eye-witness
noun class n	CLn		where n is a number or a feature
object	OBJ	cross-reference position	
obligative	OBLG	mood	
oblique	OBL	case	
obviative	OBV	person	vs. proximate
optative	OPT	mood	
ordinal	ORD	numeral	
participle (marker)	PTCP	verbal	
partitive	PRTV	case	
passive	PASS	voice	
past	PST	tense	
patient nominalizer	PATNR	deverbal nominal derivation	
patient topic	PT	voice	
paucal	PAU	number	
pejorative	PEJ	affect	
perfect	P(R)F	tense-aspect	
perfective	PFV	aspect	
pergressive	(PERL)		use 'periative'
periative	PERL	local case	'through'
place nominalizer	LOCNR	deverbal nominal derivation	
pluperfect	PLUP	tense	past or perfect of a past
plural	PL	number	
plural exclusive	PE	number	
plural inclusive	PI	number	
pluritive	(PL)		plural of a singulative; use 'plural'
polite	(FRM)		use 'formal'
positional	POSIT	verbal	
positive	(AFFM)		use 'affirmative'
possessive	POSS	possessive adjective, pronoun and cross-reference position	not for an adnominal case relation; that is GEN or AT

Tab. 169.4: *continued*

value	abbrev.	category	comment
postcrastinal	POCRAS	tense	
postelative	POSTEL	local case	'from behind'
posterior	POST	relative tense	
postessive	POSTESS	local case	'behind'
post-hodiernal	POHOD	tense	future after today
potential	POT	mood	
precative	PREC	mood	for requesting
predicative	PRED	nominal	predicative form
present	PRS	tense	
preterite	(PST)		use 'past'
pre-hesternal	PRHEST	tense	past before yesterday
primary object	PO	cross-reference position	
privative	PR(I)V	case	'without'
processive, -ual	PROC	denominal verb derivation	
progressive	PROG	aspect	
prohibitive	PROH	mood	negative imperative
prolative	PROLAT	local case	'along, by (way of)'
propriete	PROPR	case or denominal derivation	'having, provided with'
prospective	PROSP	tense-aspect	'going to'; opposite of perfect
proximal	PROX	determination	near the deictic center
proximate	PRX	person	vs. obviative
punctual	PNCT	aspect or aktionsart	
purposive	(DEST)		use 'destinative'
quality nominalizer	QUALNR	deverbal nominal derivation	
quotative	QUOT		marking indirect speech
realis	RLS	mood	vs. unrealis
recent past	RECPST	tense	= immediate past
reciprocal	REC(P)	voice or pronominal	
reduplicative			gloss by function
referent-honorific	3HON	honorification	
referent-humble	3HML	honorification	
referentive	RFR	case	'about'
reflexive	R(E)FL	voice or pronominal	
reinforcement	(INTNS)		use 'intensive'
relational(izer)	RELL	nominal	
relative	REL	subordinative and/or pronominal	in relative clause
relative	(RFR)		use 'referentive'
remote	(DIST)		use 'distal'
remote past	REMPST	tense	
repetitive	REP	aktionsart	only if distinct from iterative
reportative	RPRT	evidential	
resultative	RES	aspect or aktionsart	
reversive	RVRS	aktionsart	
same subject	SS		
secondary object	SO	cross-reference position	
semelfactive	SMLF	aktionsart	
sensory	SENS	evidential	
separative	(ABL)		use 'ablative'
sequential	SEQ	interpropositional relation	vs. simultaneous
simultaneous	SIM	interpropositional relation	vs. sequential
singular	SG	number	restricted
singulative	SGT	nominal	vs. collective
sociative	SOC	verbal	'together'
speaker-honorific	1HON	honorification	
speaker-humble	1HML	honorification	
specific	SPEC	determination	
speculative	SPECL	evidential	
stative	STAT	aktionsart	

Tab. 169.4: *continued*

value	abbrev.	category	comment
subelative	SUBEL	local case	'from under'
subessive	SUBESS	local case	'under'
subject	SBJ	cross-reference position	
subjunctive	SUBJ	mood	
sublative	SUBL	local case	'to under'
subordinator	SR	interpropositional relation	only for the single universal subordinator ('that') use super-lative
superdirective	(SUPRL)		'from above'
superrelative	SUPEL	local case	'above'
superessive	SUPESS	local case	
superlative	SUP	degree of comparison	
super-lative	SUPL	local case	'to above'
terminative	TERM	local case or aktionsart	'up to'
topic	TOP	funct. sentence perspective	
transformative	TRNSF	case	'becoming'; dynamic counterpart of essive
transitive	TR	verbal	morpheme or grammatical category
transitive patient	P	cross-reference position	only if opposed to both s and A; use OBJ otherwise
transitive subject	A	cross-reference position	only if opposed to both s and P; use ERG otherwise
transitivizer	TRR	deverbal verb derivation	
translative	TRNSL	local case	'across'
trial	TRL	number	only if distinct from paucal
undergoer	UGR	cross-reference position	
unrestricted	(PL)		use 'plural'
unspecified	UNSPEC	person	unspecified argument of relational base use 'assertive', 'declarative'
validator			
venitive	VEN	deictic	
verbalizer	VR, VBZ	verb derivation	
visible	VS	determination	
visual	VIS	evidential	eyewitness
vocative	VOC	case	
volitional, volitive	VOL	verbal	
zero	Ø		making no contribution to sentence meaning

Tab. 169.4: Grammatical category labels

refer to a dependent in a specific syntactic function. 'Case' means a case relator that may take the form of a case affix or an adposition. Verb derivational morphemes get these glosses only if they are homonymous with nominal case relators.

#### 4. Boundary symbols

##### 4.1. Basic rules

Rules R1 and R4 guarantee correspondence between units in the L1 text and in the IMG.

They do not, however, insure that the vertical alignment works in a mechanical way. This is desirable in certain contexts such as automatic parsing. It can be guaranteed in a fully formalized representation, which would then take the form of a table (s. Lieb & Drude 2000). In less formal situations, it cannot be fully guaranteed because there may be good reasons not to insert morpheme boundaries in the L1 text while still representing each morph by a separate gloss (cf. R13). Correspondence of boundary symbols in the L1

and the IMG lines is therefore not generally an equivalence, but only an implication: boundary symbols in the L1 line are matched by corresponding boundary symbols in the IMG (R9). We will review the kinds of boundaries and their delimiters in turn.

The **word boundary** is shown by a blank in L1. This is repeated in the IMG, and conversely there is a blank in an IMG only if there is a corresponding blank in the L1 line. This particular rule (R10) is therefore stricter than R9. R10 prohibits two situations: a word being rendered by a sequence of two words; and a sequence of two words being rendered by one word. The first situation will be discussed in section 4.5. Sometimes a sequence of two L1 units (words or morphemes) corresponds to one L2 unit. In principle, this situation should not arise in the IMG because each of the L1 units should have its own gloss. However, it is possible that either the L1 units have no meaning in isolation or else mean something totally different than their combination, the latter being idiomaticized. In such cases, glossing them separately might give a misleading impression of the workings of the grammar. When the bisected L1 unit forms an orthographic unit (e.g. a compound), one may simply dispense with the analysis (cf. section 3.7). For instance, instead of Germ. *be-komm-en* ‘APPL-come-INF’, one can write *bekomm-en* ‘get-INF’. If the orthography requires a boundary, as in Yucatec Maya *le kah* ‘when’, the first choice is to gloss the items separately (in this case, ‘DEF SR’) and to leave the semantic interpretation to the idiomatic translation. The second choice is to indicate the semantic unity of the two L1 items typographically by replacing the blank by a boundary symbol that does not interfere with the orthography, e.g. by an underscore: *le\_kah* ‘when’ (R11). If L1 orthography links the two items by another symbol that is also an IMG boundary symbol, as in Engl. *vis-à-vis* ‘facing’, no satisfactory solution is known.

Apart from special cases to be noted, the **morpheme boundary** is shown by a hyphen in L1 (R12). This is repeated in the IMG; and here again the converse applies, too. Apart from the *vis-à-vis* type exception, this does not pose any problems. It does, however, happen that the L1 text contains a combination of two morphemes, but no boundary is shown between them. Various motivations for this are conceivable, be it that two morphemes are fused in a portmanteau morph,

be it that the position of the boundary is not clear or irrelevant, be it that the analyst does not want to disfigure L1 orthography with boundary symbols. In such cases, a colon in the IMG is a hint at a morpheme boundary existing, but not shown in the L1 line (R13). The purpose of R13 is to allow the analyst to forgo a segmentation while still saving R1 and insuring uniqueness of the other boundary symbols. Several examples may be seen in (1). The colon is also used to render a portmanteau morph, e.g. French *au* ‘DAT: DEF’. More on this in section 4.5.

Special symbols may be introduced to distinguish kinds of morpheme boundaries. For instance, the use of the plus sign to signal a boundary in **compounding**, as in German *Weihnachts+gans* ‘Christmas+goose’ is rather widespread; and occasionally it is also found in derivation, as in German *wolk+ig* ‘cloud+ADJVR (cloudy)’ (R14).

No orthography distinguishes **clitic boundaries** from word and morpheme boundaries. If L1 is represented in conventional orthography, then the simplest solution for an IMG is not to distinguish them either. Thus French *je le sais* ‘I know it’ will be glossed as ‘SBJ.1.SG DO.3.SG.M know.SG’, while Latin *itaque* ‘and so’ will be glossed by ‘so:and’. If clisis is important or the L1 representation is non-orthographic, then the clitic boundary will be shown by an equal sign both in the L1 text and in the IMG, thus: *ita=que* ‘so=and’ (R15).

If a **zero morph** or **morpheme** is represented in L1 by Ø (cf. section 2.1), no special measures need be taken. If it is not there represented, then its gloss is enclosed in parentheses (R16), like this: Lat. *timor* ‘fear.M (NOM.SG)’. In this example, a stem is accompanied by two (complexes of) grammatical category labels, ‘M’ and ‘NOM.SG’. The first is separated by a period because it corresponds to an inherent feature of the stem. The second is enclosed in parentheses because it corresponds to a separate morpheme.

#### 4.2. Discontinuity

Discontinuous units – words or morphemes – are like bisected units in that one semantic unit is represented by two expression units. However, they present the added difficulty that their parts are not adjacent, so the IMG has to make it explicit what belongs together. For a **discontinuous stem** or **affix**, diverse so-

lutions have been proposed in the literature. Among them is the proposal (Bickel et al. 2004) to repeat the same gloss under each part of the discontinuous item. However, this seems misleading, as the syntagmatic co-occurrence of synonymous L1 items is not at all rare – e.g. in hypercharacterization – and must be distinguished from discontinuity. An unambiguous solution for a **circumfix** is to set it off by angled brackets, like this: Germ. *ge>lauf<en* ‘<PART.PRF>run’ (run (part. prf.))’ (R17). Discontinuous words are rare. The first choice is to try and gloss each part independently, as done for the German circumposition *um ... willen* ‘for’ in (7).

## (7) German

<i>um unser-es</i>	<i>Heil-es</i>	<i>willen</i>
for our-GEN.SG	salvation-GEN.SG	sake
‘for (the sake of) our salvation’		

The second choice is to treat them by the same formalism as for circumfixes. Consider the case of preverbs. In several Indo-European languages, they may be distantiated from their host verb to yield a discontinuous verb stem. There are two options for glossing such discontinuous compounds: If the compounding is relatively transparent, one may prefer to provide the preverb and the base each with its gloss. If the compound is completely lexicalized, this might be misleading, and so it may be preferable to treat it as a discontinuous morpheme in the gloss, as in (8).

## (8) German

<i>es hör&gt;t</i>	<i>jetzt &lt;auf</i>
it <stop>3.SG	now
‘it stops now’	

**Infxes**, too, require a special boundary symbol in order to insure that the root bisected by them is perceived as a unit. This is achieved enclosing them in angled brackets as shown in (9)–(10) (R18).

## (9) Latin

<i>vi&lt;n&gt;c-o</i>
conquer<PRS>-1.SG
‘I conquer’

## (10) Indonesian

<i>t&lt;el&gt;unjuk</i>
<AGNR>point
‘forefinger’

The gloss of a left-peripheral infix precedes the gloss of its host, the gloss of a right-peripheral infix follows it (Bickel et al. 2004).

## 4.3. Reduplication

Reduplicative segments may have the same kinds of grammatical functions as affixes, and sometimes they are formally not easily distinguished from affixes. Therefore they must be glossed just like affixes, but at the same time they must be formally distinguished from affixes. This is achieved by providing the same kind of gloss for them as for grammatical formatives, but separating them by a tilde (R19; Bickel et al. 2004), as in (11)–(12).

## (11) Ancient Greek

<i>gé~graph-a</i>
PRF~write-1.SG
‘I have written’

## (12) Yucatec Maya

<i>k'úa~k'as</i>
INTNS~bad
‘wicked’

## 4.4. Other morphological processes

Morphological processes not covered by the above conventions comprise transfixation, internal modification, metathesis, subtraction and suprasegmental processes (cf. ch. VIII). These are like infixation in not being peripheral to the base, but they differ from it in that the grammatical meaning in question is not associated with a single string of segments which, if subtracted, leaves the base. The notation recommended here distinguishes them from the other morphological processes, but not from each other. Such a morpheme can hardly be signaled in the L1 representation. In the IMG, its gloss follows the gloss of the base, separated by a backslash (R20). An example of transfixation is the Arabic broken plural, as in *bujūt* ‘house\PL (houses)’. Apophony, metapophony, e.g. German *säng-e* ‘sing\IRR-1/3.SG (I/he would sing)’, and tone shift, as in Yucatec Maya. *háats* ‘beat\INTROV (beat (unspec. object))’ are treated in the same way.

## 4.5. Semantic and grammatical features

The gloss of a grammatical morph often consists of a set of symbols. They are separated by a period, as in Germ. *Tisch-es* ‘table-GEN.SG’ (R21). The same rule applies in the

situation mentioned in section 3.3, where an L1 lexeme is glossed by more than one L2 words. These, too, are separated by a period, as in Germ. *fabulier-en* ‘invent.stories-INF’.

Lexical stems fall into grammatical classes. Noun stems, for instance, have gender; verb stems have valence. If such grammatical categories are covert, this information is not deducible from (the gloss of) the lexical meaning. It therefore makes sense to represent it in the gloss of the stem. The Latin example *puellae* ‘girl.F:NOM.PL’ of section 2.1 shows how this may be done for gender. The same would be possible with transitivity. Instead of Yucatec Maya *hats'-ah* ‘beat-CMPL’ as shown in section 2.2, we might put ‘beat.TR-CMPL’. It does not seem necessary to have a rule here beyond R3 and R21.

The period between values of different morphological categories cumulated in one morpheme is dispensable between person, gender and number, provided the resulting letter sequence is unambiguous. Thus, Latin *lauda-mus* may be glossed as ‘praise (PRS.IND)-1.PL’ or ‘praise (PRS.IND)-1PL’.

Sometimes the period is used as a general-purpose symbol to hide the lack of an analysis, including the function of the colon as regulated by R13. This is not recommendable if – as is usually the case – the period is also used in the function regulated by R21. Given R21, the notation Lat. *orant* ‘pray.3.PL’ would imply that *orant* consists of a single morph. An IMG should at least make the distinction between a morph and a grammatical feature of a morph. In other words, if the author knows the number and order of morphs in an L1 form, then he should indicate them. If the author does not even know so much, he should probably not use the example. Still, in emergency situations, R23 may be viable, which allows for linking IMG elements by an underscore without any implications for L1 morphological structure. This would allow for putting *orant* ‘they\_pray’.

#### 4.6. Composite categories

Two cross-reference categories may share a morphological slot, as in (13).

#### (13) Mayali

*Kamak kan-bolk-bukka-n*  
good SBJ.2&OBJ.1-country-show-NPST  
*ke.*  
your  
'It is good that you will show me your country.' (Evans 1997: 400)

In principle, the case is analogous to one declension suffix showing both number and case. However, when actor and undergoer cross-reference is cumulated in one morpheme, sticking to R21 would lead to obscurity. Instead, information on the two dependents should be separated by ‘&’ or by ‘>’ (R22). The ‘greater than’ sign has two advantages here: it is iconic, and it dispenses with the use of function labels such as ‘SBJ, OBJ, ACR, UGR’ (simply ‘2>1’ in (13)). It has the disadvantage that the same symbol is used for discontinuous and infix material, which may lead to conflicts.

This case must be kept distinct from a portmanteau morph, viz. when two cross-reference categories that generally each have their own morphological slot fuse in one morph occasionally. There R13 applies.

#### 4.7. Constituency

The IMG abides at the level of morphology. The text may be represented at other levels in addition, if this is desired. Still, IMGs are used most frequently in publications on syntax, where not only morphological, but also syntactic properties of the examples are at stake. Very often it suffices to identify one constituent in the example, for instance the prepositional phrase or the relative clause that is the subject of analysis. Then no harm is done, but on the contrary the reader is helped in scanning the example, if constituency is shown by brackets. Thus in (14), the relative clause is identified by the bracketing.

#### (14) Yucatec Maya

*le máak chowak u ho'l-e'*  
DEF person [long poss.3 head]-D3  
'the person who has long hair'

In principle, this may be done either in the L1 line or in the IMG (it need not be repeated in both). However, since the IMG line is the one that contains the grammatical analysis, the bracketing seems more natural there (R24). In principle, an IMG may even be combined with a labeled bracketing; but above some rudimentary level, this will soon lead to illegibility.

#### 5. Typographic conventions

IMGs obey a number of typographic conventions all of which aim at facilitating the reader's task. First, if there are more lines of linguistic representation (cf. section 1.3), for instance one of syntactic constituency or lines

that show syntactic, semantic or pragmatic functions of the construction, then these follow the IMG, as stipulated in R25. Second, words (neither larger nor smaller units) of L1 are left-aligned with their glosses (R27). Further, since IMGs are generally longer than the L1 text they render, they are printed in a smaller type-face (R28), and grammatical category labels are abbreviated (R29).

Here is an example of a publication which does not observe these rules (Monod-Becquelin 1976: 138 on Trumai):

šyšyk letsi k'ate šy hai-ts šyšy-ka-ke  
“avec du piment, je rends le poisson piquant  
(regarde)”  
// piment / avec / poisson / *actualis.* / 1ère pers.  
erg. / piquant-causatif-marque d'adjectivisa-  
tion //

Furthermore, since IMG lines are not sentences, the relevant orthographic rules of punctuation, initial capitalization and syllabification do not apply (R30–R32).

## 6. Summary

Instead of a prose summary, a list of the rules and symbols proposed follows:

### 6.1. Rules

#### 6.1.1. Glossing principles

- R1. With the exceptions specified below, there is a symbol or a configuration of symbols in the IMG if and only if there is a morph in the L1 text that it corresponds to.
- R2. The IMG represents morphemes, not allomorphs. Therefore, the gloss of a grammatically conditioned allomorph does not contain the grammatical category conditioning it.
- R3. An IMG should be as precise and detailed as tolerable. The limits of precision and detail are defined by practical considerations of complexity and intelligibility.
- R4. There is a biunique mapping of individual L1 morphemes onto glosses.
- R5. (a) An L1 lexeme is glossed by L2 lexemes.  
(b) L1 stems are glossed by L2 stems.
- R6. The gloss of a grammatical morph is a configuration of grammatical category labels each of which represents the

value of a grammatical category. A grammatical morph should not be glossed by an L2 bound morpheme. It may be glossed by an L2 word if that has the same function as the L1 morph.

- R7. Homonymy is resolved in the IMG, polysemy is preferably not.
- R8. The gloss of a polysemous L1 item should represent, in the order of decreasing preference,
  - its Gesamtbedeutung,
  - its Grundbedeutung,
  - the set of its senses,
  - its contextual sense.

#### 6.1.2. Boundary symbols

- R9. Apart from R30, there is a boundary symbol of a certain type in the IMG if there is a corresponding boundary symbol in the L1 text. More strictly, there is a blank, hyphen, plus, equal sign, angled bracket and tilde in an IMG if and only if there is an identical symbol in the L1 text corresponding to it.
- R10. A word boundary is shown binniquely by a blank ( ).
- R11. Two successive orthographic L1 words which must be glossed by one L2 word are linked by an underscore ( \_ ).
- R12. A morpheme boundary is generally shown by a hyphen (-).
- R13. A morpheme boundary not shown in the L1 text is indicated by a colon (:) in the IMG. This applies also to portmanteau morphs.
- R14. A boundary in a compound stem, and possibly also in a derived stem, may be shown by a plus sign (+).
- R15. A clitic boundary may be shown by an equal sign (=).
- R16. A gloss of a zero morpheme or allo-morph is enclosed in round parentheses ( () ).
- R17. The string enclosed in a discontinuous L1 item P1 ... P2 is enclosed in inverted angled brackets (P1> ... <P2). In the IMG, P1 receives a gloss enclosed in angled brackets; P2 is not glossed.
- R18. An infix is enclosed in angled brackets both in the L1 text and in the IMG. The gloss of a left-peripheral infix precedes the gloss of its host, the gloss of a right-peripheral infix follows it.
- R19. A reduplicative segment is glossed like an affix (i.e. by a configuration of

- grammatical category labels) and separated from its source by a tilde (~).
- R20. A grammatical meaning expressed by a non-segmentable morphological process (transfixation, internal modification, metathesis, subtraction, suprasegmental process) is not set off in the L1 representation. Its gloss follows the gloss of the base, separated by a back-slash (\).
- R21. Elements of an IMG that represent components of one L1 morph are separated by a period (.)
- R22. As a special case of R21, components of one L1 cross-reference morph that have distinct reference are separated by the ampersand ('&') or, where no conflict with R17 and R18 arises, by the greater-than sign ('>') for actor and undergoer cross-reference.
- R23. An L1 word form whose morphological structure is not represented in the IMG may be represented by a set of symbols whose status as representing morphs or features is ignored and whose sequence has no implications as to L1. Such symbols that jointly correspond to an L1 word form are joined by an underscore (\_).
- R24. If constituent structure is to be displayed, square brackets ([ ]) can be inserted in the IMG.

## 6.2. Symbols

L1	IMG	meaning
x y	x y	word boundary between x and y
x_y	z	x and y are two orthographic words, but one lexical word
z	x_y	x and y jointly render z without morphological analysis
x-y	x-y	morpheme boundary between x and y
x+y	x+y	x and y form a compound or a derivative stem
x=y	x=y	x and y are joined by clisis
z	x/y	x and y are alternative meanings of ambiguous z
xy	x:y (x)	morpheme boundary between x and y not shown in the L1 text x does not have a significans in the L1 text
a<x>b	ab<x>	x is an infix in ab
x>a<y	<xy>a	xy is a circumfix around a
z	x\y	y is a non-segmentable morphological process on lexeme x
z	x.y	x and y are semantic or grammatical components of z
z	x&y (x>y)	x and y are grammatical components of z cross-referencing two different dependents
x	[x]	x is a syntactic constituent
x	[x]Y	x is a syntactic constituent of category Y

### 6.1.3. Typographic conventions

- R25. The IMG is in the line immediately below the corresponding L1 text line.
- R26. The distance between an L1 text line and the line immediately preceding it is greater than that between it and the IMG line belonging to it.
- R27. Each L1 word form is left-flush with the L2 word or complex of symbols rendering it. If such an arrangement is impossible, the following is a minimum requirement: If there is, in an IMG, an equivalent to an element of an L1 text line, it is contained in the line immediately below that line.
- R28. The IMG is printed in a smaller typeface than the L1 text. If this is impossible, then at least grammatical category labels are in small capitals.
- R29. Grammatical terms appearing in IMGs are abbreviated, without a period at the end, and set in (small) capitals.
- R30. There is no punctuation in an IMG. Parentheses including optional material in the L1 line are not repeated in the IMG, either (cf. R16).
- R31. There is no sentence-initial uppercase in an IMG.
- R32. There is no syllabication either in the L1 line or in the IMG.

## 7. References

### 7.1. Specialized literature

- Bickel, Balthasar & Comrie, Bernard & Haspelmath, Martin (2004), *The Leipzig Glossing Rules. Conventions for Interlinear Morpheme by Morpheme Glosses*. Leipzig: Max-Planck-Institut für Evolutionäre Anthropologie
- Lehmann, Christian (1982), “Directions for Interlinear Morphemic Translations”. *Folia Linguistica* 16, 199–224
- Lehmann, Christian (2004), “Data in Linguistics”. *Linguistic Review* 21.3/4, 175–210
- Lehmann, Christian & Bakker, Dik & Dahl, Östen & Siewierska, Anna (2<sup>nd</sup>1994), *EUROTyp Guidelines*. Strasbourg: Fondation Européenne de la Science (EUROTyp Working Papers)
- Lieb, Hans-Heinrich & Drude, Sebastian (2000), *Advanced Glossing: A Language Documentation Format*. Berlin: Technische Universität (Working Paper)
- Simons, Gary F. & Versaw, Larry (1988), *How to use IT. A Guide to Interlinear Text Processing*. Dallas, Tx.: Summer Institute of Linguistics (Revised edition, Version 1.1)

### 7.2. Sources of examples

- Bloomfield, Leonard (1933), *Language*. New York etc.: Holt, Rinehart & Winston
- Evans, Nicholas (1997), “Role or Cast? Noun Incorporation and Complex Predicates in Mayali”. In: Alsina, Alex & Bresnan, Joan & Sells, Peter (eds.), *Complex Predicates*. Stanford: CSLI Publications, 379–430
- Finck, Franz Nikolaus (1909), *Die Haupttypen des Sprachbaus*. Leipzig: B. G. Teubner [Nachdr. d. 3., unveränd. Aufl.: Darmstadt: Wiss. Buchgesellschaft, 1965]
- Fortescue, Michael (1984), *West Greenlandic*. London etc.: Croom Helm (Croom Helm Descriptive Grammars)
- Gabelentz, Hans Conon von der (1861), “Über das Passivum. Eine sprachvergleichende Abhandlung”. *Abhandlungen der philologisch-historischen Classe der Königlich-Sächsischen Gesellschaft der Wissenschaften* 8, 449–546
- Monod-Becquelin, Aurore (1976), “Classes verbales et construction ergative en trumai”. *Amérindia* 1, 117–143
- Woodbury, Hanni (1975), “Onondaga Noun Incorporation: Some Notes on the Interdependence of Syntax and Semantics”. *International Journal of American Linguistics* 41, 10–20

*Christian Lehmann, Erfurt (Germany)*

## 170. Grammaticography

1. General problems
2. Grammaticographic problems in morphology
3. Structure of a grammar
4. Descriptive devices
5. References

### 1. General problems

#### 1.1. Basic concepts

Like many scientific and, especially, linguistic terms, the term ‘grammar’ is ambiguous in designating both a sector of the object area of linguistics and also a scientific account of this sector. Grammar<sub>2</sub> is, thus, at a meta-level with respect to grammar<sub>1</sub>, which is its object. The grammar<sub>1</sub> of a language is the systematic part of the way the language maps meanings

onto expressions. The grammar<sub>2</sub> of the linguist is a scientific representation of this part of a language – in the form of a book, a computer program or some other suitable medium –, whose locus within a comprehensive presentation of a language will be characterized more precisely in 3.

**Grammaticography** is an activity related to grammar<sub>2</sub>. This term, too, is ambiguous in the same way. At the object level, grammaticography is (the practice – experience or art – of) grammar writing. At the meta-level, it is the investigation of methodological principles that reconcile this practice with linguistic theorizing. Because of the analogous ambiguity, the term ‘lexicography’ has been coupled with the term ‘meta-lexicography’; and the same could be done for ‘grammaticogra-

phy'. **Metagrammaticography** starts from two ends: On the one hand, it takes stock of successful grammars and distills from them the grammaticographic principles that they follow. On the other hand, it takes successful models of language description as a theoretical basis and deduces from them requirements for an adequate grammar<sub>2</sub>. Grammaticography is related to investigation of grammar just as lexicography is related to lexicology (investigation of the lexicon).

Metalexicography has had a firm position in (applied) linguistics for several centuries. While grammaticography in the sense of 'production of grammars' goes back to antiquity, metagrammaticography is a fairly recent discipline. The earliest treatment known to us is Gabelentz (1891/1901, Zweites Buch, esp. VI. Capitel): "Die Darstellung der Einzelsprache". The term 'grammaticography' does not appear much earlier than Cherubim (1973). This disproportion between lexicography and grammaticography is not easy to account for. In systematic terms, neglect of grammaticography is simply unjustified and detrimental both to the linguistic discipline and to the quality of actual grammars. In terms of the history of linguistics, there has apparently been a disequilibrium between lexicon and grammar: Research into the former started out as a practical concern, i.e. as lexicography; and only towards the end of the 19<sup>th</sup> century did lexicology in the modern sense begin. Research into grammar started out as **theory of grammar** with the modists of the 13<sup>th</sup> century. The beginnings of modern grammaticography with Gabelentz (1891) were taken up by O. Jespersen, but remained largely inconsequential otherwise. Theory of grammar<sub>1</sub> got an even stronger position in the second half of the 20<sup>th</sup> century. The complete neglect of linguistic methodology during the period of the dominance of generative grammar included the complete neglect of grammar<sub>2</sub>; the concept or problem of grammaticography did not surface in linguistics. As a consequence, most grammars<sub>2</sub> actually published, while incorporating conceptions of theories of grammar<sub>1</sub> en vogue at their time, do not in their general organization reflect any theoretical foundation at all.

In assessing the relationship between the practice of grammar writing and linguistic theorizing, we may recognize that the last third of the 20<sup>th</sup> century has brought a significant progress in grammar writing. Apart from many noteworthy monographs, the

series *Lingua Descriptive Studies* alias *Croom Helm Descriptive Grammars*, *Cambridge Studies in Linguistics*, *Mouton Grammar Library* and *London Oriental and African Language Library* may be singled out, as they are devoted to comprehensive accounts of hitherto underdescribed languages. This progress in the quality of descriptive grammars can be attributed not so much to some particular model or theory of grammar, but, rather, to a general agreement to keep grammars comprehensible for those outside a particular framework and/or jargon. Roughly since the forties of the 20<sup>th</sup> century, a certain amount of grammatical descriptions of minority languages had appeared that applied some model of structural or – less often – transformational linguistics. It was soon recognized that these grammars were not even used by linguists, let alone by laymen. Since the seventies of the 20<sup>th</sup> century, most comprehensive grammars stick to what their authors call "traditional grammar". It was recognized that "traditional" descriptive grammars remain up-to-date and readable much longer than any model-oriented grammars. Such grammars also do not strive for the precision sometimes attributed to formal descriptions. Neglect of precision in itself is not a virtue, but historically it must be seen as a reaction to a kind of hollow precision that served to conceal lack of insight.

In theory, grammaticography and theory of grammar inform each other. In practice, mutual awareness leaves much to be desired. Still, good grammar writing does apply the achievements of linguistic theorizing. Some new concepts and approaches (or more elaborated old ones), presentational techniques, empirical domains to be looked at and included in a grammar, etc. have established themselves as part of recent grammaticographic tradition. To the extent that these innovations are picked up from various models of grammar, this approach is eclectic. Grammars have to be both consistent and comprehensive, and these goals are sometimes in conflict. The **eclecticism** often practiced in contemporary grammaticography may appear as a symptom of inconsistency, but is more properly seen as a consequence of the fact that none of the available theories of grammar suits all languages and all empirical domains equally well. Natural languages comprise heterogeneous phenomena which favor different models and methods. If a grammarian is able to choose the most pow-

erful approach for each descriptive domain and to combine them in the compilation of a comprehensive grammar, this is a virtue rather than a vice. Moreover, this approach to writing grammars is actually the only way to determine the borders of applicability of theoretical concepts and models and their relations to one another, and thus to achieve theoretical progress. In this sense, metagrammaticography provides an interface between the practice of grammar writing and the theory of grammar. On the other hand, metagrammaticography can be viewed as a sort of self-reflection inside the field of descriptive linguistics, which can help to accumulate positive experiences and most suitable (technical) solutions to be applied in future grammatical descriptions.

Like most disciplines concerned with the practical application of scientific insights, and like metalexicography in particular, metagrammaticography is at least as much a prescriptive discipline as it is a descriptive one. It is descriptive insofar as it is faced with actually published grammars and with the esteem these are held in, a datum which it has to take into account lest it become escapist. It is, however, prescriptive in that it converts both the principles distilled from grammaticographic practice and the theorems of grammatical theory into instructions of grammar writing (cf., e.g., Mosel 2003, Noonan to app.). Metagrammaticography is part of the methodology of linguistics. Consequently, the present treatment will account for observable practice, but will also derive recommendations from it.

## 1.2. Universalism and particularism

The description of a language has two opposite tasks: to bring out the uniqueness of this language and to render it comparable with other languages. To achieve the former task, the description must not identify the categories and operations of the language with those of other languages, but rather describe the language in its own terms. To achieve the latter, the description has to characterize the language in terms of a linguistic type, presupposing and referring to universal parameters of human language as a background against which the peculiarities of the language stand out. To do this, it must describe the language in the same terms as other languages. The two tasks seem, thus, irreconcilable. Consequently, many existent grammatical descriptions sin by an overemphasis on either **particularism**,

**universalism**, which renders the language different from anything that may be familiar to the reader and, thus, unintelligible, or on **universalism**, which converts the language into just another instance of something well-known to the reader and, thus, uninteresting. Both extremes miss their target (cf. also 1.6.1). Many of the grammaticographic principles discussed below revolve around this problem.

The purpose of a language description is not only to tell the reader the facts to any desired degree of detail, but, at a more general level, to convey to him an impression of what this language is like. This evaluation proceeds against the general background of ‘le langage’, shared by author and reader, and possibly against the comparative background of other languages, too. In the former respect, the grammar will take for granted many aspects of the language that it shares with all other languages (e.g. the sheer existence of word formation, and its most general cognitive and communicative functions). In the latter respect, many grammars, especially textbooks, even presuppose as deserving no comment those features of the language that it shares with the metalanguage of the description. Thus, few grammars of French will bother to mention that it lacks infixation or that clause structure follows the accusative model of fundamental relations.

## 1.3. Purpose and kinds of grammar

Grammars may be classified according to a variety of parameters. The first distinction is between a comparative grammar and a grammar limited to one language. There are various kinds of comparative grammars according to the purpose and method of comparison: a **general comparative grammar** is a systematic survey of grammatical phenomena in the languages of the world. A **historical-comparative grammar** traces the evolution of the grammar of a proto-language into its daughter-languages by comparing these. A **contrastive grammar** compares grammatical phenomena of two languages, mostly with some practical application in mind. In contrast to comparative grammars, grammars devoted to just one language are sometimes called descriptive grammars; but of course such a grammar can also be prescriptive; see below.

A grammar of one language may take the synchronic or the diachronic perspective. A grammar that does the latter is a **historical grammar**. In contradistinction to a language history, a historical grammar is normally not

subdivided according to stages of the language, but may rather have the same structure as a synchronic grammar. A historical grammar usually defines one stage of a language – not seldom an ancient stage – as the end point of its scope and differs from a synchronic grammar in tracing the properties of that stage back to the relevant proto-language.

Focusing now on non-comparative synchronic grammars, we may distinguish between what has been known traditionally as **textbook** vs. **reference grammar**. A textbook (German *Sprachlehre*) aims at teaching the language. It arranges grammatical material in the sense of some didactic progression; and grammatical information may be intertwined with other didactic material such as lessons, cultural information and exercises. A reference grammar, on the other hand, gives a systematic and comprehensive overview of the grammatical system or even the whole system of a language, but is limited to that. Its purpose is to provide orderly access to information on the language system rather than to help the user learn the language. It may presuppose some familiarity with the language or with linguistics. There also used to be a distinction between linguists' vs. learners' ("scientific" vs. "pedagogical") grammars. This has lost somewhat in significance because it has been recognized that writing for colleagues is compatible with the requirement of user-friendliness (cf. also the notes on recent developments in 1.1).

The orientation of a grammar may be descriptive or prescriptive. A **descriptive grammar** is based on data that are independent from the analyst and describes these. Its purpose is to ascertain and systematize the facts in some empirical domain. A **prescriptive grammar** represents what the author considers correct. Its purpose is to orient the reader towards the norm of the language.

While the distinction between descriptive and prescriptive linguistics is clear in principle, there are some factors that tend to blur it in practice. Firstly, no grammar (descriptive or other) can account for all the variation that occurs in its object area. It will exclude at least speech errors, false starts, hesitation phenomena etc.; and it may exclude youth language or archaic ritual language from its scope. This delimitation amounts to the identification of a norm. This does not necessarily coincide with the highest norm in the entire speech community; it simply means

the identification of a prototypical manifestation of the variety under study and the marginalization of other varieties. The exclusion of some facts from the variety being described is similar to the decisions made in a prescriptive grammar. Secondly, most users of a grammar, even of a descriptive grammar, require reliable information on what is grammatical in the language or, at least, what is normally said. They want to be given rules. To satisfy this need, the author has to distinguish manifestations of the norm from deviations from it and has to formulate rules for grammatical constructions. This, again, is what a prescriptive grammarian does.

In this article we will concentrate on descriptive reference grammars. However, a good portion of the following metagrammatocigraphy applies to other types of grammar as well.

#### 1.4. Data and variation

Data are the empirical basis of a grammar (see Lehmann 2004 a). There are different kinds of them, and they play different roles in grammars. This is determined, first and foremost, by the orientation of the grammar in the sense of 1.3. In a descriptive grammar, data function as the empirical basis and the object of research. Therefore the data must exist independently from the production of the grammar. A descriptive grammar is therefore based on a corpus of data. In a prescriptive grammar, on the other hand, data have the function of illustrating the norm. They may be taken from a corpus of texts that is conceived as representing the norm – for instance the works of classical writers; but if the author considers herself master of the norm, then she may simply make up illustrative examples, which are then no data in the strict sense. Some descriptive grammars, it is true, are also based on material produced by the author of the grammar. In this case, however, the data are not independent from the scientist, which means they are not reliable data by scientific standards.

The **corpus** on which a descriptive grammar is based may essentially be of two kinds. It either exists prior to the grammar as a body of texts available in the speech community; or it is collected by the linguist in fieldwork (in the broadest sense). In morphology as opposed to syntax and discourse structure, many data may be systematically elicited in informant work, rather than drawn from texts (see Art. 168).

The object of a grammatical description is a certain language, more precisely, some particular **variety** of a language. The first task of the linguist is to delimit this variety in synchronic and diachronic terms. However narrowly she may delimit it, her empirical data will necessarily be only a sample of the object of description. With a corpus collected in a small village of the Mexican state of Quintana Roo, one may publish a grammar of the speech of those male peasants of that village who were born between 1940 and 1950; or one may publish a grammar of Yucatec Maya. It commonly corresponds both to the purpose of the author and to the interest of the readership to situate the variety described at a rather high level of generality. It is, then, the responsibility of the author to guarantee that her data actually represent her object, i.e. that they are valid. This presupposes some homogeneity of the object. The description of Yucatec Maya based on that sample will be valid to the extent that an observationally similar description could have been produced with data from Valladolid (Yucatan). *Mutatis mutandis*, the same goes for a grammar of French or Latin.

The actual sources of the examples – informants or texts – have to be identified (the former, if they agree to it). If their representativity could be an issue, it has to be demonstrated that they do represent the object of description. For instance, many a linguistic description is based on a corpus of narrative texts, and it is by no means to be taken for granted that such a corpus does truly represent the language in question. Again, young fieldworkers tend to cooperate with young informants, whose speech usually represents just one sociolect in the community. Therefore, the sociolinguistic coordinates of the data have to be indicated.

It is essential that the data sources be made entirely explicit. The usage observed in traditional grammars of giving examples without indication of the source is no longer acceptable. To put it differently, an example without indication of its source will be regarded as made up by the author of the description and insofar unreliable (no matter whether she is a native speaker or not).

### 1.5. Topicality

Scientists are all but resigned to seeing new theories and analyses pop up every day only to be superseded by even newer theories and analyses, so that most of scientific work be-

comes irrelevant not only to the general public but even to the scientific community itself. Things are different for a grammar (and likewise for a dictionary). First of all, for most languages of the world the publication of a grammar is an event that is not likely to be repeated within a generation. For all those languages that are threatened by extinction, a comprehensive description is the only chance for them to be revived and the only way to protect them from falling into oblivion after their extinction. Secondly, a grammar is a scientific product of potential importance for the general public, since it may and must be the basis of the development of a norm and the composition of primers. All of this heightens the responsibility of the grammarian. The main virtue of the grammarian is (alas) not originality and inspiration but consistency and reliability.

To the extent that grammarians have become aware of these conditions, they have refrained from tying their work to some fashionable model and from concentrating too much on trendy issues. The grammarian has the unenviable task of keeping herself informed on progress in general comparative linguistics but free of ephemeral trends. Fortunately this task has been facilitated in the past decades by comprehensive and up-to-date surveys of general comparative grammar such as Shopen (1985, ed.) and Payne (1997).

### 1.6. Concepts and terms

At least two issues arise with respect to the use of concepts and terms in grammatical descriptions: (a) What is the appropriate attitude towards conservatism and neology? (b) What degree of familiarity with concepts and terms should be presupposed?

#### 1.6.1. Conservatism and neology

Originality is a value in science that does not include terminology. Established terminology must be used. Most phenomena in the languages of the world, whether or not they happen to be known to the analyst or to the particular philological or linguistic school that she has been raised in, are actually known in general linguistics and have a term to them which, at least sometimes, is appropriate and well-established. For instance, the semantic relation often called *purposive* had been called ‘*destinative*’ by Finno-Ugricists half a century earlier. It is among the duties of the analyst to get informed about estab-

lished terminology before she coins her own terms. Sometimes, it is true, established terminology is inappropriate. For instance, the term ‘possessive classifier’ was deemed inappropriate even by its coiner, but for want of a better term continues to be used. It is most important that distinct concepts be designated by distinct terms and that these be unequivocal; questions of beauty of terms have less priority.

Secondly, the question regularly arises whether the particular phenomenon in the language being described is simply an instance of something known from other languages. Here, a middle course between two extremes must be steered (cf. 1.2). One is **particularism**. From the concept of the ‘langue’ which is both historically and systematically unique, and of the language sign which is arbitrary and unlike anything to be found in other languages, this position deduces that every category of the language must be named by a term of its own in order to avoid misleading identification with something known from other languages. This position has been cherished to the extreme in Russian and American structuralism. It renders a description unusable and blocks comparison of the language with other languages. The other extreme is **universalism**. It assumes a universal grammar and subsumes every phenomenon of the language under pre-established concepts and terms. For instance, the concept of ‘case’ is stretched to cover (non-morphological) syntactic properties of noun phrases; and then sometimes what used to be called case is renamed ‘surface case’. This position is well-represented in generative grammar, but not alien to some brands of functionalism either.

The appropriate middle course is characterized by two propositions: (a) descriptive concepts belong to the level of linguistic types; (b) descriptive concepts are prototypical in nature. Take the concept of ‘dative’ as an example. For it to be a typological concept means that it is not a universal property of human languages, that some languages have a dative and others don’t, that ‘dative’ is a type of which there are concrete instantiations in the languages and that there are partial functional equivalents to the dative both inside a language that has it and across languages. Being a typological concept for a (value of) a grammatical category, i.e. for a (kind of) linguistic sign, ‘dative’ is characterized both by its meaning or function

(roughly ‘case whose basic meaning (Grundbedeutung) is the recipient and whose generic meaning (Gesamtbedeutung) is a participant related indirectly to the situation’) and by its structure or form (which for the dative is the same as for ‘case’, i.e. a grammatical morpheme bound to a nominal expression that marks the latter’s syntactic or semantic function in the clause). The prototypicality of such concepts relates to both sides of the linguistic sign. For the meaning or function, it necessitates the Jakobsonian distinction between basic meaning and generic meaning, as just exemplified for ‘dative’. For the structure or form, it may involve the identification of a focal instance on a scale of grammaticalization, with some range of variation at both sides. For instance, a case is typically marked by an agglutinative affix, but may be marked by a highly general adposition, on the one hand (as in Japanese or Hawaiian), or by more fusional morphological means, on the other hand (as in Sanskrit).

Of course, such definitions are not always available for grammatical concepts, and often it is the task of the analyst to decide whether an established term covers her particular phenomenon. Extensions of traditional usage are allowed and necessary, because in the last analysis, most of these terms were originally conceived for Greek and Latin and must necessarily be extended in their meaning if they are to be applied to any other language at all. If this were not so, we could never speak of conjugation or of passive formation in English. What is essential here is that such extensions be controlled, i.e. that a general definition (at the typological level) of the kind alluded to be provided, from which it may be deduced that the term is indeed applicable to the particular phenomenon under analysis. If terms are used in this way, then neology in a grammatical description may be kept to a minimum.

The most strongly grammaticalized categories are, at the same time, the most deeply entrenched in a particular language system and consequently the most arbitrary and language-specific ones. Here the question of whether such a language-specific phenomenon should be subsumed under a category known in typology can become a real issue. An example is provided by the German verb category whose values are traditionally called ‘subjunctive I’ or ‘present subjunctive’ (e.g. *singest*) and ‘subjunctive II / past subjunctive’ (e.g. *sängest*). These terms are actually

not helpful as far as the functions of these categories are concerned. If German happened to be an exotic language first described in the second half of the twentieth century, then these categories would probably be called ‘hearsay evidential’ and ‘irrealis’, respectively. Consider the two sets of cross-reference markers in Mayan languages as another example. Members of the first set precede the finite verb, members of the second set (displayed in Tab. 170.9 below) follow it. Mayan linguistics, which originates in American structuralism, calls these two paradigms by purely arbitrary labels ‘set A’ and ‘set B’, in order to avoid any functional implications. The two sets do combine heterogeneous functions. In Yucatec Maya, set A cross-references the transitive subject and the subject of an intransitive verb in one aspect-mood category, while set B cross-references the subject of a nominal clause, the transitive object and the subject of an intransitive verb in another aspect-mood category. Relating this to certain ergativity-splits, one may say that set A cross-references the subject, while set B cross-references the absolute. However, this is odd, because a transitive verb would then be flanked by a subject and an absolute index. Moreover, set A also precedes nouns, cross-referencing the possessor. This function can certainly not be subsumed under the denominator ‘subject’. In other words, concepts provided (so far) by typology do not help here; language-specific concepts and terms have to be coined. The only question that remains is whether they must be as empty as ‘set A/B’.

#### 1.6.2. Definitions

In principle, the explicit introduction of concepts and terms of general linguistics, especially of analytic concepts such as ‘infix’ or ‘completive aspect’, is not the task of a grammatical description, but of genres such as terminological dictionaries and studies in general-comparative grammar. If one could presuppose a systematic methodological organization of linguistic science, a descriptive grammar should only have to categorize its phenomena by the concepts of general-comparative grammar and designate them by the corresponding terms. However, such rigor is not practicable. On the one hand, the extent to which familiarity with established concepts and terms may be presupposed obviously depends on the intended readership of the description. If non-specialists are addressed, technical terminology either has to be intro-

duced explicitly or must not be used at all. On the other hand, established usage may be insufficient with respect to the language being described. Then modifications to it must be made explicit. Still, a grammarian must sometimes be reminded that she is not writing an introduction to linguistics. Specific suggestions on morphological terminology are found in Art. 169, section 3.9.2.

#### 1.7. Description and argumentation

A grammar is not the kind of treatise that has a point. The data are not used to argue for a particular analysis, the analyses are not arranged in a way to convey a certain theoretical insight or to demonstrate a certain method. All of this is necessary and justifiable in other kinds of scientific treatises. A grammar is, in the first place, a systematic encyclopedia of the grammatical functions and structures of the language.

Consider the French clitic personal pronouns as an example. The grammarian, of course, has a position on the issue of whether they are affixes or clitics. In an article devoted to the subject, it would be appropriate to assemble the evidence for her thesis and argue for it. In a grammar, these pronouns will just be called clitic pronouns (or personal affixes), their various properties will be described, both those like their fixed relative order which are more affix-like and those like their alternative prethematic or postthematic position which are more clitic-like. It is not necessary to turn the description into a persuasory discourse, as is done in some grammars like Cole (1982) and McGregor (1990). Neither is a grammar the place for a linguist to teach linguistic methodology or to demonstrate that she has applied it. What matters is that her analytical decisions be transparent. This is chiefly done by means of examples that exhibit the phenomenon some descriptive statement attributes to the language. It is normally also not necessary to back statements on obligatory rules with ungrammatical examples of constructions that violate them. What is important is that it be made explicit under which conditions the rule applies. True, there will always be cases where the grammarian has not been able to ascertain these conditions. In order not to mislead the reader, she should admit this, for instance by modifying a descriptive statement by the phrase ‘under unclear conditions’.

Several grammaticographic treatments (e.g. Comrie & Smith 1977; Noonan to app.)

require that the analyst note those properties that the language does not possess. In a comparative grammar, such statements are natural by-products of the comparison. In a monolingual grammar, such statements involve an implicit comparison. Moreover, there is no limit to the things that a language does not have. A feasible form of complying with the requirement is to limit such negative statements to such features that might be expected on genetic, areal or typological grounds and to relate them to the description of some feature the language does possess. Thus, in the section on counting of an onomasiological grammar of German it would be mentioned that the nouns *Stück* ‘piece’ and *Mann* ‘man’ can replace the counted noun generically and anaphorically under certain conditions. This may naturally be rounded off by a statement to the effect that this is the closest to numeral classifiers that the language can muster.

## 2. Grammaticographic problems in morphology

### 2.1. Tasks of a morphological description

Basically, the core task of the morphological part of a grammar is to describe:

- A. (i) the internal structure of word forms (grammatical words) (**morphological structure**), and
- (ii) the form and the meaning of grammatical items (non-lexical **morphemes**).

These aspects of a morphological description overlap to the extent that (a) the grammatical items of a language are **bound morphemes**, and (b) the internal structure of word forms in this language can be described in terms of these morphemes (typically as a construction of a lexical stem and a chain of bound grammatical items). Since the domain covered by these two conditions is fairly large in many languages, the usual practice of grammarians is not to separate tasks A(i) and A(ii) from one another for descriptive purposes. As a result, those sub-domains where one of these conditions does not hold commonly come out in the periphery of a morphological description (as represented by sections on, e.g., prepositions or compounding).

Generally, preference is given to what may be termed **word-oriented morphology** (as in A(i)), that is, the word is taken as a starting

point for the description. Grammatical items thus fall into two strictly distinct groups which are described quite differently, viz. those which are constituent elements of word structure and those which are words themselves. Within this approach, the very existence of a bound grammatical item constitutes a descriptive statement about word structure. An alternative approach, which may be termed **morpheme-oriented morphology**, would be to start from the morpheme, so that the class of words as well as word classes emerge as a result of the distributional analysis of morphemes. No grammar takes this approach exclusively. However, many grammars contain sections which consist of an annotated list of grammatical formatives. For instance, McGregor (1990), in the relevant sections of his chapter 3, has various lists of grammatical items for each of which he provides the significans, a functional label and a list of functions fulfilled in various constructions. Each item of the latter list contains a reference to other sections where the item is treated in its relevant context.

The word-oriented approach implies a distinction between the morphological description proper (as outlined in A) and two “classificatory” tasks:

- B. (i) delimiting the class of linguistic units to be referred to as **words** (both in the sense of ‘word form’ and of ‘lexeme’; cf. Art. 26), and
- (ii) a classification of lexemes into **parts of speech** (cf. Art. 70) and further grammatically relevant sub-classes, at least inasmuch as morphological patterns vary by word class.

Quite commonly, classificatory solutions are just incorporated into the general structural outline of a grammar, rather than constituting a subject of description in its own right; in other words, the classes of linguistic units dealt with in morphology are presented as established a priori. There are obvious theoretical grounds for this well-established tradition:

First, these tasks are scarcely solved by purely morphological criteria; syntax and semantics play quite a role in an appropriate classification of linguistic units. In this sense, this classification can be regarded as “external” with respect to morphology proper (tasks A). Secondly, this way of presentation overcomes a certain circularity which is apparently inherent in grammatical reasoning,

whereby relevant classes are distinguished on the basis of distribution and associated morphological categories, the latter being, in their turn, defined with respect to and in terms of pre-established grammatical classes. However this problem may be solved in the process of grammatical analysis, the usual descriptive practice is to take both the level of the word and the word classes for granted.

On the other hand, classificatory solutions adopted in a given descriptive grammar do constitute autonomous descriptive claims and are even argued for in case they deviate from a “traditional” one, be it a language-specific or the general-linguistic tradition. In addition, minor (sub-)classifications which are relatively or exclusively language-specific (as, e.g., inflectional classes) are more likely to be presented as a subject of morphological description than the major ones. That is, many grammarians take the classificatory tasks to have (at least to some extent) a default solution which can be applied without special discussion. This default solution is in fact very close to the most traditional parts of speech. By accepting the default solution, a grammarian virtually refrains from a language-specific classification of linguistic units; that is, it is not considered a task of a particular descriptive grammar, but rather a universal ready-made structural template for describing languages.

However, the impact of classificatory solutions on the quality of grammatical description can hardly be overestimated:

- (a) They play a major role in structuring the morphological description and thus determine the general outline of the grammar (cf. 2.2).
- (b) They strongly affect the descriptive solutions (tasks A), insofar as **morphological patterns** and **categories** can be defined only with respect to the word classes; once a classification is established, lots of further descriptive solutions are predetermined.
- (c) They serve as a basis for the **morphology-syntax interface**, inasmuch as syntactic constructions often make reference to a word class and/or a further morphological sub-class of the elements involved.
- (d) Last not least, these solutions establish the **grammar-lexicon** interface, in the sense that they determine a system of grammatical indices to be provided for items listed in a lexicon (cf. 3.1).

This means that if a language does not fit the default classificatory solution, the consequences of adopting this solution can turn out harmful for the whole enterprise. On the other hand, the default solution is, in a sense, the most user-friendly one, in that it facilitates obtaining information from the grammar, since the reader is faced with a familiar structural outline: the key words (commonly constituting chapter headings, cf. 2.2) tend to be exactly those he expects to find. Thus, a grammarian normally has to weigh the default solution against its descriptive cost in order to find an appropriate compromise. This problem, although often not pronounced as such, seems to be fundamental for descriptive morphology.

A classification into word classes as represented in the morphological part of a grammar is often far more detailed than is needed for morphology proper. The usual descriptive practice is to introduce all more or less grammatically relevant classifications of words by structuring a chapter where the morphological patterns associated with some of these classes are described. In this sense, a description of the internal structure of word forms and a grammar-oriented classification of lexemes are not distinguished in the practice of grammarians, although these – obviously overlapping – descriptive domains presumably never coincide exactly. This practice has the advantage of highlighting the correlation between classifications based on distinct parameters (morphological, syntactic, semantic), the descriptive significance of which is beyond doubt.

To sum up: the practice of descriptive morphology combines three distinct domains of language structure: the word structure, the grammatical items, and the grammar-oriented classification of lexemes, the latter constituting the skeleton of a morphological description. As will be seen below, this combination, however well-motivated, involves some descriptive problems which could be avoided otherwise.

## 2.2. Structuring a morphological description

The well-established tradition of structuring morphological descriptions is based on a simple and elegant idea of converting grammar-oriented classifications of linguistic units into the section headings of a grammar. This idea makes it possible to combine two descriptive domains (word classes and word structure) in

- 
1. Open word classes with rich morphology:
    - 1.1. (Optional:) a general overview of morphological structure (e.g., in terms of a general structural template comprising slots for constituent morphemes).
    - 1.2. Further (functional) sub-classification, if relevant for morphology (e.g., transitive vs. intransitive verbs).
    - 1.3. Inflectional paradigm:
      - 1.3.1. The structure of the paradigm(s), i.e., the set of grammatical categories associated with the word class under description and their possible values.
      - 1.3.2. Further (formal) sub-classification into inflection classes (if any) and how the items of the paradigm are constructed for each inflection class.
      - 1.3.3. Exceptions, defective paradigms etc.
      - 1.3.4. (Optional:) semantics and usage of inflectional categories can be picked up as a special issue.
    - 1.4. Word formation (sections organized first by the category of the output of a process, second by the category of the base).
      - 1.4.1. Compounding
      - 1.4.2. Derivation:
        - 1.4.2.1. Regular (productive) derivation; the internal structure is close to a semantically ordered morphemic lexicon (a morphological item + its meaning and distribution).
        - 1.4.2.2. Irregular (non-productive) derivation.
      - 1.4.3. Reduplication, conversion, etc.
    - 1.5. (Optional:) some important semantic and/or derivational sub-classes (e.g., reflexive verbs).
  2. Open word classes with little or no morphology:
    - 2.1. Semantic sub-classes (e.g., spatial adverbs, temporal adverbs, etc.).
    - 2.2. Correlated structural features (if any) and word formation (often including etymology, grammaticalization notes, etc.).
  3. Closed classes:
    - 3.1. A general semantic and syntactic overview.
    - 3.2. (Optional:) exhaustive listing of items with their meanings and distribution, ordered according to functional (e.g., interrogative pronouns, indefinite pronouns etc.) or syntactic properties (e.g. postpositions with dative, postpositions with genitive etc.).
      - 3.2.1. If the items are inflected, the paradigm is given for each.
      - 3.2.2. (Optional:) distinctive structural features of each group (if any) and word formation techniques (e.g., conversion, grammaticalization, etc.).
- 

Tab. 170.1: Types of morphological chapters with respect to their internal structure

a rather natural fashion, inasmuch as it can be assumed that different word (sub-)classes are associated with distinct morphological patterns; hence, the morphological information is appropriately distributed over these (sub-)classes. The internal structure of chapters varies according to the word class. Roughly, they fall into three groups: open classes with rich morphology, open classes with no or little morphology, and closed classes. A schematic representation of how the respective chapters are most commonly structured is given in Tab. 170.1.

As shown by this scheme, the morphological parts of grammars contain rather heterogeneous chapters: some are morphological in the word-oriented sense, since they deal with the internal structure of words (cf. the dis-

tinction introduced in 2.1); others are morphological in the morpheme-oriented sense, since they describe the form and the functions of closed-class (= grammatical) items; and, finally, there are chapters which are concerned almost exclusively with (non-morphological) sub-classifications of a word class. This may be seen as a direct consequence of the fact that the three descriptive domains combined in morphology (cf. 2.1) remain essentially distinct. Not only do the respective chapters differ in their internal structure, they are bound to take distinct descriptive approaches. In some cases, a semantic classification seems to appear in a grammar just because there is nothing morphological to say about a class. The well-established tradition of providing a semantic classification of

adverbs independently of its grammatical significance is a case in point. Also, it seems that if a language has no nominal morphology, the chapter on nouns is much more likely to contain their semantic sub-classification. In fact, a chapter subdivision is the form of introducing a comprehensive classification, but it seems bizarre to leave a section thus generated empty.

A more consequential drawback of the structuring scheme outlined above is that it provides no natural way of accounting for morphological items shared by distinct classes (e.g., for person/number paradigms attached to both verbs and nouns, as, for example, in Yucatec Maya). Another instance of the same problem is represented by grammatical items which function both as a free morpheme (hence, have to be described in a separate chapter of the closed-class type) and as a bound morpheme (hence, a constituent element of some morphological pattern related to a certain word class); cf., e.g., locative prepositions vs. applicative suffixes on verbs in Rwanda (Art. 141, section 3.5.1). Various ways to overcome this drawback of the traditional approach have been applied in different grammatical descriptions. The options are:

- (a) Introduce **generalized word classes** comprising several word classes sharing some morphological categories (e.g., “nominals”). This solution has a limited domain of applicability because of its hierarchical nature. If the distribution of morphological patterns in a given language does not follow any hierarchical classification (e.g., a sub-class of nouns has an adjectival paradigm, as in Russian), this strategy does not work.
- (b) Select a “locus” for the description of an item – a morphological category, a single morpheme, or whatever – (for example, the noun for the category of case), and refer to this description in all other relevant chapters (in this case, for instance in the section on pronouns). This solution may work if the semantic and formal properties of the item in question do not vary by word class. If comparable items are even slightly different, a reference will not suffice.
- (c) Link the descriptions of similar items by cross-references. This is a solution widely

applied, obviously necessary, but not sufficient, since it does not provide a way to describe the item with several instantiations as a whole.

- (d) Separate word formation (for all word classes) from the main body of description and treat it in a special chapter. This compromise solution can overcome the drawbacks of the word-oriented approach at least for derivational morphemes.
- (e) Separate morphological semantics from the main body of description and treat it in a special chapter. This is a very strong, but rarely applied solution, presumably because of its “ambitious” flavor.
- (f) Create a **morphemicon** as part of the morphological description as described below (3.5).

The last three structuring solutions tend to be applied in one or another form in recent descriptive grammars. They are, in fact, a concession to the morpheme-oriented approach to grammar. In any case, an appropriate structure for a grammatical description can be achieved only by means of an appropriate combination of various solutions.

### 2.3. Identification of grammatical items

The task of describing the morphological structure of a word form (cf. A(i) in 2.2) is almost never addressed directly in descriptive grammars. Instead, it is decomposed into several distinct sub-tasks. The most widely applied decomposition is based on the distinction between **inflection** and **word formation** (in particular, **derivation**; see Art. 38). Its controversial theoretical status notwithstanding, this distinction is widely applied in morphological descriptions. Not only are these types of phenomena often rigidly distinguished (cf. 2.2); these sub-domains of morphology are commonly described in quite disparate fashion:

- (a) To describe the inflection of a language means to describe each word class in terms of the associated **paradigms**, that is, the inflectional patterns are considered properties of pre-established word classes. To describe the derivation is to describe the semantic impact and distribution of each single **derivational morpheme**; hence, a word sub-class to which some derivational morpheme applies is

- viewed as a property of that morpheme (this includes the issues of productivity, regularity, etc.).
- (b) Inflectional forms of different words can be identified as representing the same item of a paradigm independently of their formal similarity; the identity of a paradigm is grasped in terms of certain semantic-functional labels (even if these only imply an approximation of the respective meanings), not in terms of formal identity. In contrast with this, the derivational sub-task involves a semasiological description of particular morphemes, which usually presupposes formal similarity of its various allomorphs.
- (c) The semantics of inflectional items is assumed by default to be independent of the semantic properties of the word; they are allowed to have a few, usually context-dependent functions. Conversely, a derivational morpheme can have a variety of meanings depending on individual stems it is applied to.

These properties can be traced back to the concepts of “prototypical” inflection and derivation and the role they are assumed to play in the grammar and, in particular, in the grammar-lexicon interface (Art. 36). The distinction between inflection and derivation thus determines the choice of a descriptive perspective. That is, there is a considerable difference in how inflectional and derivational items are treated in descriptive grammars, from the identity of an item up to the types of information to be provided.

In particular, the concept of inflection implies a more function-oriented identification procedure, that is, both formal identity of items even in case of clearly related meanings, and an obvious formal contrast are very easily disregarded, as far as the inflection is concerned (hence, multiple non-distinctions within paradigms; see Plank (1992) for an overview of this problem, and Zaliznyak (1967: 19–34, 129–148) for case paradigms). Here is an example: Russian is an aspect-dominated language in the sense that the grammatical category of tense works differently for perfective and imperfective verbs. Tense allows for two different ways of description, schematically represented in Tab. 170.2 by the tense forms of the verb pair *delat'* ‘make/do (impfv.)’ and *sdelat'* ‘make/do (prfv.)’:

(a) form-based description

+Past	<i>dela-l</i> , <i>s dela-l</i>
–Past	<i>dela-et</i> , <i>s dela-et</i>
Imperf. Future	<i>budet delat'</i>

(b) function-based description

	imperfective verbs	perfective verbs
Past	<i>dela-l</i>	<i>s dela-l</i>
Present	<i>dela-et</i>	—
Future	<i>budet delat'</i>	<i>s dela-et</i>

Tab. 170.2: Russian tense: alternate paradigms

Basically, the two accounts differ in how they treat forms like *s dela-et* ‘PFV:make-FUT/NPAST.3.SG (will make, will have made)’. Variant (a) highlights the formal identity between this item and the Imperfective Non-Past *dela-et*. They are identified as instantiating the same item of the tense paradigm and get the same functional label (Non-Past) (see Comrie 1978: 66 f.). Yet these forms are clearly distinct semantically: the form of perfective verbs can refer only to the future, while with imperfective verbs it is employed basically for reference to the present (although the future meaning is not excluded). Thus, the descriptive cost of this solution is a dependency of the range of functions of an inflectional item on the verb class.

By contrast, variant (b) disregards the formal identity between the items under discussion. Instead, it highlights their functional distinction and, accordingly, the functional similarity between the Perfective Non-Past (Future) of the perfective verbs and the Imperfective (analytical) Future. The latter pair of forms are assigned the same position in the paradigm and get the same functional label (Future), despite the obvious formal contrast. The dependency on the aspectual meaning of the verb stem is thereby built into the structure of the paradigm, instead of constituting an independent descriptive statement. The descriptive cost of this solution is (i) the formal heterogeneity of the Future and (ii) the unmotivated formal identity of the Perfective Future and the Imperfective Present (which would show up recurrently in the description of multiple formal variants of this marking, only one of which is represented in Tab. 170.2).

Variant (b) is apparently preferred in existent descriptions of Russian (even though

some of them do suggest variant (a), most often the formal identity between the Imperfective Present and the Perfective Future is not mentioned). In other words, the paradigm is defined in such a way as to simplify the semantic description at the cost of somewhat artificial formal complexity.

Formal identity of this type is never neglected as easily for derivational formatives; to continue with Russian examples, its verbal suffix *-s'a/-s'* expresses an extremely wide variety of meanings in combination with different verb stems (middle, reflexive, reciprocal, anticausative, dispersive, to mention only some of them). Yet these are commonly treated as instances of the same linguistic unit.

In view of this difference in descriptive perspective, the well-known problem of delimiting inflection and derivation gets a new kind of relevance: how a grammatical item is identified and described depends on its affiliation with inflection vs. derivation, at least above a certain degree of discrepancy between form and meaning. It might be the case that these two sub-domains of morphology are in fact delimited in descriptive grammars depending on which approach seems more “suitable” for a given item. On the other hand, the choice of one or another perspective can affect the very properties of the items identified (as they would be described in the grammar), inasmuch as the prototypical features of inflection vs. derivation are embodied in the identification procedure.

### 3. Structure of a grammar

#### 3.1. Comprehensive presentation of a language

No grammar can be complete. Given limitations of every kind, the author has to set priorities. These follow from the main purpose of the grammar and from external conditions. Needless to say, on the basis of available publications, one may decide to produce a partial description. However, the aim may be to produce a comprehensive presentation of the language. Assume a field structured in a hierarchical fashion. For an account of it to be comprehensive means that it is balanced in terms of the amount of detail provided for each of the sections at a given level of the hierarchy. In this sense, a grammar such as McQuown (1990) (of Totonac), which comprises 64 pages on phonology, 107 pages on

morphology and 11 pages on syntax is unbalanced and therefore not comprehensive.

A comprehensive account of a language is articulated on the three levels shown in Tab. 170.3.

level 1	the <b>documentation</b> of the language, which is a corpus representing (analyzed) primary data;
level 2	the <b>description</b> of the language, whose object are the data of level 1;
level 3	a <b>methodological reflection</b> on the description, whose object is the account of level 2.

Tab. 170.3: Levels of presentation of a language

The levels are, thus, in a meta-relation to each other. The reflection of level 3 has the function of accounting for the purpose of the description, the conditions under which it was carried out, including the achievements of previous scholarship, its theoretical and methodological prerequisites, the many choices and decisions that the author has made; and as mentioned in 1.4, it reflects on the character and limitations of the data base. Most of this part commonly takes the form of an introduction to the description. Some of it may be relegated to an appendix.

The corpus on which the grammar is based (level 1) will normally not be reproduced in full. Good modern presentations of a language contain an appendix that presents some specimina of representative texts (see Mosel 2002, § 6.2 for the scientific importance of such a collection). These are provided in the form of video or audio recordings and rendered in the canonical trilinear representation as explained in Art. 169 (more on the relationship between the corpus and examples in the running description in 4.5).

The descriptive part of a comprehensive account of a language (level 2) consists of two subparts, the system of the language and the setting of the language. Here we may be brief on the latter. The **setting of the language** comprises an explanation of relevant glossonyms, the genetic affiliation of the language and its dialects, its ethnographic situation (i.e. the situation of its speech-community), its cultural situation (including, importantly, its written tradition) and its sociolinguistic situation, i.e. its internal stratification, its status in the speech community and its areal relations. This part of the description provides a referential background for paramet-

ricization by diachronic, diatopic, diastratic and diaphasic variables that will prove necessary in the account of the language system. It is, at the same time, the one part of a comprehensive presentation of a language that is most likely to take a historical perspective even if the rest is purely synchronic.

The **language system** has two main parts, the expressive subsystems and the significative subsystems (corresponding to Martinet's second and first articulation, respectively). The primary expressive subsystem is the phonology (with the phonetics), the secondary one is normally the writing system. (This is, of course, also true for [ancient] corpus languages.) The significative subsystems are not alternative, but jointly exhaustive: the grammar and the lexicon. They are significative because their units embody a mapping of meanings onto expressions. The significative subsystems are articulated in terms of levels of complexity: At the level of the word form and below, we have morphology, which – in the form of inflection – is an essential part of the grammar and also – in the form of word formation – of the lexicon (see Art. 36). At higher levels, we have the syntax as part of the grammar, and phraseology as its counterpart inside the lexicon.

Since there is, in principle, no borderline between grammar<sub>1</sub> and lexicon, there is also overlap between a grammar<sub>2</sub> and a dictionary of a language. The morphology describes the word-formation patterns of the language. Their products are nevertheless listed in the dictionary. Only in the limiting case of a word-formation process that applies completely regularly and productively – e.g. formation of certain verbal nouns – may one renounce to representing each of the products in a lexical entry of its own. Contrariwise, the dictionary will contain a couple of entries that result from a word-formation process that is no longer productive in the language. Mentioning it in the morphology would entail repeating the same list of items that is in the dictionary. In such cases, redundancy is preferred to parsimony. The drawbacks of redundancy are that it induces inconsistency and that it may be uneconomical. They are outweighed by the advantages, which include user-friendliness and theoretical soundness (the mental grammar and lexicon are redundant in the same way). Moreover, we are not talking about a literal repetition of material in two parts of the description, since the ordering principles and, conse-

quently, the ways of accessing the information are entirely different between dictionary and grammar<sub>2</sub>.

Another important relationship between the grammar and the dictionary is that the former introduces and defines the terms that appear in the cells on grammatical information of the microstructure of a dictionary entry. This concerns such categories as noun class, gender, countability, *akitionsart*, inflection class etc. Technically speaking, this kind of specification in a dictionary entry determines whether the item can be used in a certain construction described in the grammar.

### 3.2. Onomasiology and semasiology

As we said in section 1.2, the fundamental problem of grammaticography is to provide a common format for descriptive grammars while at the same time taking care not to obliterate the individuality of the language being described by forcing it into a Procrustean bed. The general task of a language is to provide a mapping between meanings and expressions. The meanings have an extra-linguistic substrate (i.e. one independent either of language altogether or at least of particular languages) in cognition and social interaction (communication). The expressions have an extra-linguistic substrate in phonetics and semiotics. It is the mapping itself, achieved in the grammar and in the lexicon, that is proper to each language. There can therefore be a universal system of cognitive and communicative domains coded by languages and a universal system of expression techniques and sounds used by languages. There cannot be a universal system of grammar.

A system of lexical and grammatical description that provides comparability of the object language with other languages can therefore be either based on a system of cognitive and communicative domains and then describe how the language in question manifests each facet of these in its expressions; or else it can be based on a system of expressive (structural) devices and then describe what each of them is used for in the object language. The former approach is traditionally called **onomasiological** (or synthetic, more recently 'functional'), the latter **semasiological** (or analytic, more recently 'structural') (cf. first Gabelentz 1891/1901: 84–104 and, more recently, Lehmann 1980, <sup>2</sup>2002: § 1.2.1 and Mosel 2003: §7). Each of these approaches is in itself coherent and capable of providing a complete description. Many grammars stick

to one of them. For instance, Jespersen (1937) is a purely semasiological grammar, while Givón (1993) is a purely onomasiological grammar. However, each of these approaches is one-sided, as the onomasiological approach corresponds to the viewpoint of the speaker, while the semasiological approach corresponds to the viewpoint of the hearer. Thus, an onomasiological grammar answers questions of the type “how can I express such and such a thought, or fulfill such and such a communicative function, in this language?”, while a semasiological grammar answers questions of the type “what does such and such an expression of this language mean?”. Since grammars, just as dictionaries, are generally meant to serve both the speaker and the hearer, the ideal grammar<sub>2</sub> consists of two parts, an onomasiological and a semasiological one. It should be clear that such an arrangement is also maximally user-friendly, because whatever question a user may pose to a grammar is posed either in the speaker or in the hearer perspective. Moreover, most of the descriptive problems reviewed in 2 resolve themselves if the twofold approach is chosen.

A couple of published grammars approach this ideal to some extent. Despite appearances and declarations of the author, Gabelentz (1881), a grammar of Classical Chinese, is relatively far off his own mark because the “synthetic grammar” is not really onomasiological, but just a construction-based syntax. McGregor (1990), a grammar of Gooniyandi, organizes its chapters in the familiar bottom-up fashion appropriate for a semasiological account. However, some syntactic chapters and part of a final chapter on semantics take an onomasiological perspective on the items introduced in lower-level sections. The chapter on morphology in Haspelmath (1993), a grammar of Lezgian, contains sections on noun morphology and verb morphology, each of which is subdivided into a form-based and a function-based subsection; but otherwise the grammar mixes the two approaches. Lehmann (2002) is not a grammar, but just a description of possession in Yucatec Maya. It is subdivided into a chapter that introduces the relevant structures in a semasiological perspective, and three chapters that take the opposite perspective. It must be said that to this day, most grammars mix the two approaches in uncontrolled ways (cf. Lehmann 2004 b).

It is nowadays standard to tie the distinction between morphology and syntax to the distinction between levels of grammatical complexity. However, it has often been associated with the distinction between an onomasiological and a semasiological grammar. From traditional grammars of the nineteenth century to our day, many a grammarian has said that her morphology deals with the system and structure of linguistic forms, while her syntax deals with their use. It might thus appear that the morphology of such a grammar is semasiological, while the syntax is onomasiological. However, such systematicity is only apparent. Most traditional morphologies contain chapters on functional categories such as person and tense – and to this extent they are onomasiological. And the typical traditional syntax has a chapter on the use of the cases, thus starting off from a structural concept introduced in the morphology and tracing its functions – and to this extent it is semasiological.

Actually, the canonical level-dependent distinction between morphology and syntax can only be made in a semasiological grammar. A semasiological grammar is organized according to the hierarchy of structural complexity of linguistic units. For all those languages which possess the word form as a level of grammatical structure, the section of semasiological grammar devoted to this and lower levels will be the morphology. In the last quarter of the twentieth century, there have also been functional (i.e. onomasiological) grammars that presupposed a distinction between morphology and syntax; e.g. Comrie & Smith (1977) and the series of grammars based on it. Since this distinction cannot be made in an onomasiological grammar, it introduces inconsistency into it. For instance, the grammar of many languages manifests the concept of definiteness, but some do so at the level of morphology, while others do so at the level of syntax.

In the twentieth century, many structural (thus, semasiological) grammars were published. Contrariwise, there are as yet few purely functional grammars. A theory of cognitive and communicative domains of language started to be developed only in the last quarter of the twentieth century, and for most of these domains we still lack both a solid foundation and an internal structure adequate for linguistic description. The present handbook tries to partly compensate for this shortcoming by organizing those of its

chapters that are devoted to grammatical categories and operations (XII and XIV) according to an onomasiological perspective. It is inevitable that the articles of those chapters break the boundaries between morphology and syntax. The handbook on syntax in this series might have done the same with equal or greater justification.

### 3.3. Structure of a semasiological grammar

Much of the language system is structured in terms of hierarchies of levels of complexity. This concerns, first of all, the system of grammar<sub>1</sub> as mentioned in 3.1; but it may also include, with some imprecision, the distinction between expressive and significative systems. This yields the well-known series ‘phonology – morphology – syntax – discourse’, and inside grammar the sequence ‘stem – word form – phrase/syntagm – clause – (complex) sentence’. The history of linguistics has, *cum grano salis*, followed this progression. Syntax was a step-child of linguistic description up to the end of the nineteenth century, and even the 20<sup>th</sup> century has seen many “grammars” that actually boil down to a phonology plus a morphology. Discourse has come to be studied systematically only in the last third of the 20<sup>th</sup> century. The progression is deeply entrenched in the awareness of linguists, who think it must shape the organization of their grammar. Almost all grammars work essentially bottom-up (cf. Mosel 2003: § 5.2). This seems to correspond to a didactic progression which starts from elementary units and proceeds stepwise to complex units. A top-down progression, as it is prescribed in Comrie & Smith (1977) and followed in *Lingua Descriptive Studies*, seems unnatural because in the treatment of a given unit (e.g. the complex sentence) one is forced to appeal to constituent concepts (the clause, in this case) which have not yet been introduced.

However, the picture must be modified slightly. The bottom-up approach corresponds to analytical thinking, which combines elementary units according to rules and aims at constructing a complex whole in a compositional fashion. The top-down approach corresponds to holistic thinking, which starts from a whole and understands its parts in terms of their function in the whole. As is well-known, the two approaches do not exclude, but complement each other. To give an example: A cleft-sentence is best understood if one knows what contrastive focus is and what it entails for semanto-syntac-

tic structure. One can then identify the structural constituents in terms of their role in the complex construction. Contrariwise, it will be hard to construct the purport of a cleft-sentence in a bottom-up fashion, starting from a copular predication over an empty subject that combines with something that looks like a relative clause. More precisely, although the chapter on the cleft-sentence does presuppose the notions of copula clause, relative clause and complement clause, it is not the case that the grammar can build on one of these and expand it into a cleft-sentence. The external grammatical relations of a given unit – the subordinate clause in this example – are not treated in the chapter dealing with this unit, but instead in the chapter of a higher unit whose internal relations they are and of which the given unit is a constituent. This responds to the principle – last put forth in construction grammar – that the formation of complex constructions is goal-directed.

The same goes for the morphology. The semasiological description of word-formation does not start from a certain derivational suffix, combine it with bases of different categories and then look what the category of the result is. Instead, there is a section on stems of a certain category, e.g. the adjective. The category has certain elementary members, i.e. adjectival roots. Next there are possibilities of forming adjective stems by various formal processes, e.g. by suffixal derivation. One of them is our derivational suffix. Finally, it is seen that it may combine with bases of various categories to yield the result at stake.

The general principles of such an arrangement of a semasiological description may be formulated as follows.

- (a) The description works bottom-up through the **hierarchy of grammatical levels**.
- (b) For the grammatical unit of each of these levels, the categories into which it is articulated are identified. For each of the potentially complex categories, its internal syntagmatic structure is analyzed: First, a set of constructions according to kinds of syntagmatic relation is enumerated. For each of these constructions, the nature and distribution of its elements is set forth. Finally, given a certain construction of elements of two categories, one of the categories may comprise a grammatical (in particular morphological) para-

- digm of elements. Such paradigms are discussed as part of the description of the particular construction.
- (c) Only the internal syntagmatic structure of a given unit is part of the treatment of that unit. Any structural phenomena which concern the relation of a given element to its context are treated at the point where the including construction – the one which provides the context – is treated.

To give a final example: At the level of the word form, categories such as ‘finite verb form’ are identified. There are kinds of finite verb forms according to their internal structure, e.g. periphrastic forms and synthetic forms of different kinds. One of these is the synthetic form that consists of a tensed stem and a personal ending. In the latter position, there is a paradigm of morphemes whose internal structure is treated now. It may be seen that this approach combines a bottom-up progression to ever more complex constructions with a top-down analysis of each of these constructions.

#### 3.4. Structure of an onomasiological grammar

The cognitive and communicative domains that are coded in language comprise concepts and operations a subset of which manifest themselves in grammatical structure. These are such concepts as the addressee or directed motion and such operations as abstraction or making a question. Such concepts and operations are assembled in **functional domains**. They are by now well-researched in functional typology, so that they can be tentatively enumerated. Since they provide the highest-level subdivision of an onomasiological grammar, this will be done here in Tab. 170.4.

Onomasiological description proceeds from very general cognitive and communicative functions as indicated in the central column of Tab. 170.4 through more specific subareas such as those of the right-hand column down to the functional categories and processes of the language under description. These are finally mapped onto the structural devices and grammatical formatives introduced in the semasiology. For instance, in a grammar of German, the chapter on possession treats, among other things, possessive pronouns, the genitive attribute and its equivalents, the possessive dative, possessive predictions with

*haben* and *sein* etc. Most of these structural devices recur also in other chapters of the onomasiological description; but this chapter is where the question is answered how German expresses an attribution of possession as in Turkish *vakt-im var* ‘time-POSS.1.SG EXIST (I have time)’.

#### 3.5. Additional parts of a grammar

The subdivision of a grammar follows mainly from the overall organization of the presentation of a language as outlined above. Since the term ‘grammar’ is not seldom used in a broad sense almost equivalent to ‘language description’, it bears repeating that, conceptually, the data corpus, on the one hand, and the methodological reflection, on the other hand, are not part of the grammar; and neither are the account of the setting of the language, the lexicon, the phonetics, phonology and orthography. What remains is simply the morphology and the syntax. Of the morphology, inflection is properly included in grammar, while word formation could, in principle, be treated either in the dictionary or in the grammar. For practical reasons, it is always included in the grammar (if it is treated at all).

An important challenge taken up in many modern grammars written on a typological background is to present an epitome of what the language is like and how it works, i.e. to present the language as instantiating a **linguistic type**. From what has been said in 1.7 it follows that the overall presentation style of a reference grammar does not by itself fulfill this task. It is therefore widespread practice to provide an additional chapter – usually at the beginning of the grammar – that characterizes the language at the typological level. While the main body of a grammar is descriptive, this chapter provides the opportunity to compare the language with other languages and to answer the Humboldtian (1836: 417) question “auf welche Art [sie] die hauptsächlichen Fragen löst, welche aller Spracherzeugung als Aufgaben vorliegen”.

Furthermore, a description of a language system contains a number of lists. The most important of these is, of course, the list of significative units (signs) of the language. It comprises the lexemes (including phraseologisms) in their citation form plus the morphemes, i.e. the roots and the grammatical and derivational formatives, including the affixes. The question is in which part of the

domain	basic functions	representative concepts and operations
apprehension & nomination	an entity is grasped by categorizing and individuating it; it is named by a label or a descriptive expression	categorization, types of concepts, empathy
concept modification	a concept is enriched, or an object is identified	attribution, apposition, relativization
reference	a representation is related to and delimited within the universe of discourse	determination, deixis, reference tracking
possession	the relation of an entity to another one is established or inheres in one of them	possession in reference, possessive predication, external possessors
spatial orientation	an entity is localized in space statically or dynamically	reference points, local relations, spatial and gestalt properties of objects
quantification	the extent of the involvement of a set of entities in a predication is delimited	quantification in reference and in predication; counting, ordering
predication	information is attributed to a referent	existence, situation, characterization
participation	a situation is articulated into an immaterial center and a set of participants and circumstances related to it and to each other	control & affectedness, central vs. peripheral roles, alignment of fundamental relations
temporal orientation	a situation is designed with respect to its internal temporal structure and limits and temporally related to another situation	situation types, aspectuality, temporal relations
illocution, modality, evidentiality	a proposition is rendered relative to speaker, hearer and reality	speech acts, obligation, volition, possibility, toning, evidentiality
contrast	a concept or proposition is assessed qualitatively by comparison with similar ones	negation, comparison, gradation, intensification
nexion	a situation is expanded into a complex one, or several situations are linked together	speech reproduction, complementation, interpropositional relations
communicative dynamism	a proposition is articulated in foreground and background	discourse structure, functional sentence perspective (topicalization, focusing, emphasis)

Tab. 170.4: Functional domains

overall description and in which way these lists are presented.

In the ideal situation, there is a **dictionary** of the language beside the grammar. If so, then the dictionary contains entries for all of these items. It then properly includes a **morphemicon** (inventory of morphemes), i.e.

morphemes are lemmata just like stems or citation forms of words. For the sake of user-friendliness, not only morphemes, but also variants are conceded an entry, the latter reducing to a reference to the main entry. A dictionary entry refers to the grammar in two ways: (a) implicitly by the grammatical cat-

ategorization and other grammatical information provided in the microstructure of each entry (see 3.1); (b) explicitly in the form of references to the relevant sections of the grammar for such entries which are grammatical formatives. In such a situation, the grammar contains no morphemic. Instead, each grammatical morpheme is introduced in the semasiological description as part of its construction and paradigm. The distribution of information on individual grammatical items between dictionary and grammar is then a question of fine-tuning. There must be a certain amount of overlap as regards generic properties. Details on meaning and function, especially idiosyncratic properties, are provided in the lexicon, details on distribution and conditions of variation are provided in the grammar.

If there is no extra dictionary, then a number of second-best solutions are available. Of these, the relative best is a **glossary** or vocabulary appended to the grammar that contains all the lexemes and morphemes that are mentioned in the grammar, including the examples, each coupled with its meaning or function. Again, variants are listed, too; and for the grammatical items there are references to all the places where they are discussed. The absolute minimum requirement is an alphabetical **index of grammatical items** at the end of the work, again with the relevant references.

Among the indices there is also a **subject index**. While this is nothing special of books on grammar, it does assume a special role here. If the grammar is bipartite and both the semasiology and the onomasiology are organized hierarchically according to established conceptions, then the user can retrieve the information by systematic search through the table of contents if this is sufficiently detailed. Even then the grammar probably needs an index, as there are liable to be at least some concepts and terms that are language-specific or appear in unwonted contexts. But as long as grammars are not organized in such a way, an index of functional terms is necessary in a structural grammar to compensate for the lack of the opposite perspective; and vice versa for a functional grammar.

Another kind of list that is often attached to a grammar – and not seldom also to a dictionary – is the list of inflection paradigms (mostly of the specimen type; see 4.3) of various word classes and their subclasses. If the language has a relatively complicated

inflectional morphology, as e.g. Ancient Greek conjugation, then it makes sense not to interrupt its systematic description by pages filled with conjugation tables and instead to relegate these to an appendix.

Among the various lists that a grammar shares with other genres, such as lists of tables, figures, primary text sources, bibliographical references and the like, the list of **abbreviations** is of special interest. Among all the abbreviations used in a grammar, the (abbreviations of) grammatical category labels form a distinct subset. Whether or not other kinds of abbreviations are listed, the latter must certainly be listed, and they may constitute a list of its own. A separate list of grammatical category labels also helps the author to be consistent.

#### 4. Descriptive devices

##### 4.1. Representation of a single grammatical item

###### 4.1.1. Representation of the significans

Basically, one representation type is chosen for the whole grammatical description (i.e. excepting the phonetics and phonology). The primary choice is between a standard **orthographic representation** and a (technical) linguistic representation. The former is generally preferred because it makes the description more accessible to the speech community and other non-linguists interested in the language. Sometimes, standard orthography is supplemented by additional diacritics, e.g. to mark word stress.

There are, however, situations where a linguistic representation is necessary or preferable, if a language has no orthography or if using traditional orthography would complicate comprehension of the grammar and is not in accordance with its descriptive tasks (cf., e.g., recent Chinese grammars). In this case, authors choose a basic representation trying to obtain a compromise between its theoretical ambitions, convenience for descriptive purposes and user-friendliness (especially in a learner's grammar, the latter includes easy rules of reading). Sometimes printing (publishing) facilities prove to be a limiting factor. If the representation is based on linguistic principles, it is either a **broad phonetic transcription** or a **phonemic** or a **morphophonemic representation**. Use of a phonetic transcription as the basic representation in a grammar is very rare and essentially lim-

ited to situations in which a grammar is published before a satisfactory phonological analysis of the language is available. In general, a morphophonemic representation is to be preferred to a phonemic one, but lower-level representations are, of course, necessary in those sections that deal with lower level variation. In phonetic representations, symbols are chosen from the International Phonetic Alphabet. At higher levels and if a certain phonetic distinction is irrelevant in the language, usually less technical letter symbols are preferred.

The basic representation is used throughout a grammar; yet, a description of morphology often involves more or less significant deviations from this representation. In particular, an orthographic representation can prove insufficient for an adequate and sufficiently detailed morphemic segmentation, either in a language with intricate morphophonology or in case the traditional orthography obscures certain morphological phenomena. Hence, some additional elements can be incorporated into representations of some forms or even phrasal examples in order to provide a transparent morphemic segmentation.

For example, Russian orthography makes use of graphemes like *я*, *ю* which stand for combinations /ja/, /ju/ after a vowel or for single phonemes /a/, /u/ after a palatal consonant (very loosely, in the latter environment these graphemes are employed to signify the palatalization of the preceding consonant, in contrast with their plain equivalents *a*, *y* which occur after non-palatal consonants). This spelling convention makes the orthographic representation unsuitable for morphemic segmentation and for the representation of single morphemes, as is necessary, i.a., in combination with interlinear morphemic glossing (Art. 169). In a strictly orthography-based Russian grammar one would find numerous items like an “ending -*a/-*я*”, a representation which suggests allomorphy, but in fact just reflects the fact that the ending /a/ occurs both after non-palatal and palatal consonants (in particular, after /j/). In order to avoid such misleading representations, a spelling convention can be employed which renders /j/ between vowels as well as palatal consonants in a more transparent fashion (Zaliznyak 1967: 201–294). As an example, consider the following forms, where morphemic borders are indicated as imposed by the orthographic representation:*

	student	father-in-law	hero
sg.nom	студент	тесь	герой
sg.gen	студент-а	тест-я	repo-я
sg.dat	студент-у	тест-ю	repo-ю

Tab. 170.5: Russian declension in standard orthography

The phonemic representation is:

	student	father-in-law	hero
sg.nom	stud'ent	t'es't'	g'iroj
sg.gen	stud'ent-a	t'es't'-a	g'iroj-a
sg.dat	stud'ent-u	t'es't'-u	g'iroj-u

Tab. 170.6: Russian declension in phonemic representation

It is now clear that the endings do not vary between these words; yet the transcription cannot help but contain deviations from the orthography which are of no relevance here and can only be misleading (cf., e.g., /i/ for *e* in the last column). Now consider the same forms in an alternative spelling (the deviations from the established convention are shown by capital letters):

	student	father-in-law	hero
sg.nom	студент	тесь	герой
sg.gen	студент-а	тестЬ-а	героЙ-А
sg.dat	студент-у	тестЬ-у	героЙ-У

Tab. 170.7: Russian declension in adapted orthography

The orthographic representation is brought closer to the phonemic one in order to show that the three nouns follow the same inflection pattern, yet the deviations are reduced to a necessary minimum. The symbols reflecting a distinct representation type are underscored either by capitalizing or by using symbols from another alphabet (continuing the Russian example above, the symbols /l/ and /j/ might be used instead of *Л* and *Й*). Such mixed representations may be used as a compromise between the incompatible requirements of using a standard orthography and providing an adequate morphological analysis.

Another side of the same problem is the representation of a morpheme abstracted from its morphological environments; that is, technically speaking, a unitary representation

of a set of allomorphs. Several techniques (which may also be combined) are in use:

- (a) Regular phonemic alternations are ignored (a technique normally favored by the use of the orthographic representation).
- (b) For more or less regular morphonemic phenomena, special symbols for morphemes are employed, commonly, a capitalized phoneme symbol (for instance to abstract from vowel harmony phenomena).
- (c) The allomorphs are just listed, commonly separated by a slash.

There seems to be no established means to represent the significans of non-segmental items (e.g. metaphony).

#### 4.1.2. Representation of the significatum

The meaning of a grammatical morpheme may be rendered in several ways, among them:

- by using an appropriate functional (mnemonic) term, e.g. ‘iterative’, ‘ablative case’, etc., which may then be referred to by means of a conventional abbreviation, e.g. ITER, ABL etc., where upper case identifies (abbreviations of) grammatical category labels;
- by translating it into the metalanguage (e.g. ‘constantly’, ‘from’).

The choice of a name for a category may be determined by a language-specific terminological tradition or by a desire to keep the terminology as transparent and “speaking” as possible (see 1.6.1).

Semantic information in morphology almost never reduces to naming categories appropriately (although, in some cases, a reference to the syntactic part of grammar is considered sufficient). The semantics of derivational morphemes is commonly described in terms of processes and operations. This is appropriate wherever operations apply in a regular and productive way, yet it can be misleading in the case of fossilized patterns. In the latter case, restrictions on the distribution should be carefully described (sometimes, just a list of relevant items seems the best solution, to be sure, with comments on their semantic motivation). The semantics of inflectional items is commonly described in terms of grammatical features structuring the respective paradigm (cf. 4.2).

To visualize the range of meaning of a polysemous (multifunctional) morpheme, a **semantic map** is a useful descriptive device (see Haspelmath 2003). This is a language-independent n-dimensional (mostly with  $n = 2$ ) arrangement of monosemic grammatical functions by their similarity. A semantic map of a particular case, e.g. the Latin dative, will cover a contiguous area that comprises a subset of case functions (recipient, addressee, experiencer etc.) appearing on the map. A similar map can be drawn for the dative of another language, e.g. Turkish, and then the semantic expansion of the dative in the two languages may be easily compared. Such a semantic space is also the locus for diachronic semantic variation.

#### 4.2. Category and feature

Rules of syntax may refer to a morphological category or **grammatical feature** that a constituent bears. The formulation of such rules often presupposes the specification of such categories in a parameter-value format, where a grammatical category constitutes a parameter, and its subcategories, the values (s. Art. 28). Thus, the inflectional information of a Latin verb form may be represented as follows:

*cantaveritis*

2	person
m	number
u	tense
m	anteriority
m	mood
u	voice

Fig. 170.1: Inflectional information of a Latin verb form

Another task of a morphological description is to provide an appropriate system of **inflection classes** to be picked up in a lexicon (cf. 2.2 and Art. 65). The following duties await the analyst here:

- (a) The inflection classes have to be brought into a hierarchy. Wherever a language has some complexity in this area, major inflection classes usually have minor subclasses, and these have a few totally irregular members. For instance, one of the major conjugation classes in Latin is the consonantal conjugation. One of its subclasses is constituted by a perfect formation which involves lengthening of the

- root vowel. This, in turn, has *emō* ‘buy’ as a member, which is slightly irregular in having an epenthetic consonant in the perfect participle *emptum* ‘bought’.
- (b) The inflection classes have to be named by appropriate terms. Since the conditioning factor of an inflection class is often the stem final, this is suitable as a label of the class, as in the Latin *a* conjugation. The concept of a particular inflection class reflects its position in the taxonomy, e.g. ‘consonantal conjugation with lengthening perfect’. This, of course, leads to cumbersome terms which may be abbreviated by numbers and letters, e.g. ‘3e’ (third conjugation, subclass e). The abbreviations are needed, among other things, in the section of grammatical information of the lexical entry of such a verb.
- (c) For didactic purposes, those inflected forms of the paradigm are identified from which one can deduce the entire inflectional paradigm. For this reason, Latin nouns are often quoted not only in the nominative, but also the genitive, e.g. *gēnus, gentis* ‘stem’. For German, one even needs the nominative plural in addition (*Hahn, Hahns, Hähne* ‘rooster’). For many semiregular Latin verbs, if one knows the first person singular present indicative active, the first person singular perfect indicative active and the perfect participle, then one can conjugate the verb correctly through all of its categories. Traditional grammars and dictionaries therefore specify the forms of these categories for semiregular verbs, e.g. *emō – ēmi – emptum* ‘buy’. Instead of such a set of word forms, it is also possible to mention a set of allomorphs appearing in such word forms. For instance, a German apophony class may be identified by *i – a – u* (as in *singe – sang – gesungen*).

#### 4.3. Paradigms

Inflectional morphology is commonly represented in terms of paradigms arranged as one- or multidimensional tables. For each word class, the set of tables has the same structure in terms of number of columns and rows and their category labels. There are two linguistic forms of such tabular representations. In the traditional paradigm, a plain example word is represented in the forms of all the values of the categories, as in Tab. 170.8.

	number	singular	plural
person			
1		<i>laud-o</i>	<i>laud-amus</i>
2		<i>laud-as</i>	<i>laud-atis</i>
3		<i>laud-at</i>	<i>laud-ant</i>

Tab. 170.8: Specimen paradigm of Latin personal endings

This kind of representation is the original linguistic sense of the Greek term *parádeigma* ‘example’. It is still used in textbooks or if the morphology is so fusional that inflectional markers are hard to represent separately.

This leads us to the second form of representing a paradigm, in which the cells of the table are occupied just by inflectional morphemes, as in Tab. 170.9.

	number	singular	plural
person			
1		<i>-en</i>	<i>-o'n</i>
2		<i>-ech</i>	<i>-e'x</i>
3		–	<i>-o'b</i>

Tab. 170.9: Marker paradigm of Yucatec personal suffixes

This kind of representation is more abstract and proper to linguistic science. It is viable for agglutinative morphology. The **specimen paradigm** is, in a sense, a definition *per ostensionem* that appeals to intuitive understanding. The **marker paradigm** only works to the extent that the information that it presupposes is made explicit in the grammar. This concerns, in particular, syntagmatic information on the way the markers combine with the stem (see 4.4). If complex morphophonemics are involved, it is prudent and user-friendly to complement the rules by a specimen paradigm.

For a given morphological category, the order in which its values are enumerated is fixed by tradition in many cases; e.g. ‘1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> person’, ‘nominative, genitive, dative, accusative, ablative, vocative case’. A properly linguistic criterion for their arrangement is **syncretism** (Art. 66). Especially for semantically motivated syncretism, it is convenient if the syncretistic category values are adjacent in the table. Tab. 170.10 and Tab. 170.11 are two examples to illustrate ways of displaying syncretistic paradigms:

case \ number	singular	plural	dual
vocative	<i>deva</i>	<i>devās</i>	
nominative	<i>devas</i>	<i>devān</i>	<i>devau</i>
accusative	<i>devam</i>		
instrumental	<i>devena</i>	<i>devais</i>	
dative	<i>devāya</i>	<i>devebhyaṣ</i>	<i>devābhyaṁ</i>
ablative	<i>devāt</i>		
genitive	<i>devasya</i>	<i>devānām</i>	
locative	<i>deve</i>	<i>devesu</i>	<i>devayos</i>

Tab. 170.10: Declension of Old Indic *a*-stems (paradigm of *deva-* ‘god’)

The tables representing paradigms are at most two-dimensional. For agglutinative morphology (e.g. the Turkish case paradigm), even a one-dimensional table suffices to represent the set of values that each morphological category can take. If two morphological categories are cumulated in a morpheme, as e.g. person and number in many conjugation paradigms, a two-dimensional table becomes necessary. If more than two categories are fused in a morpheme, or if inflection classes are involved in addition, a suitable combination of two-dimensional tables is chosen, as also shown in Tab. 170.11.

Although such complex paradigms are not necessary for agglutinative morphology, it is prudent at least to give a couple of examples of morphologically complex word forms to render the formulas concrete.

#### 4.4. Syntagms

The accurate statement of the combination of a morphological marker (of a certain category) with its host may require specification of the category of the base, the conditions for the combination and specification of the resulting category. This may take the form of a

rule or **instruction** how to construct a certain morphological form. E.g., concerning the Latin supine II like *dictū* ‘to be said’: in order to make the supine II of a verb, form its passive participle, convert this into a stem of the *u*-declension and form the ablative singular of this stem.

Syntagmatic relations are commonly analyzed in terms of constituency or dependency and accordingly visualized by **tree diagrams** or **dependency stellmatia**, each illustrated here with an example:

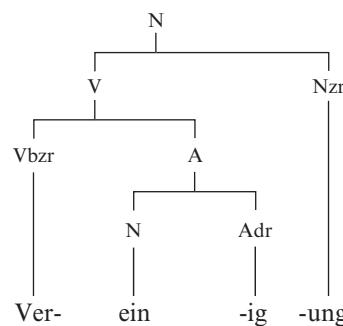


Fig. 170.2: Constituent structure of complex stem



Fig. 170.3: Dependency structure of complex stem

The **valence class** of a verb or some other relational word is, above all, a syntactic property. However, in many languages, valence classes coincide with inflection classes, so that specification of a valence class may be relevant for morphology, too. This may be done in the form of a construction formula, e.g. for a trivalent verb: [NP]<sub>Sbj</sub> — [NP]<sub>dir.Obj</sub> [AdvP]. Such valence classes may also con-

tense	number	gender person	present	past		
			m/f/n	m	f	n
sg.	1		<i>čítáju</i>			
	2		<i>čítáeš</i>	<i>čítál</i>	<i>čítála</i>	<i>čítálo</i>
	3		<i>čítáet</i>			
pl.	1		<i>čítáem</i>			
	2		<i>čítáete</i>			
	3		<i>čítájut</i>	<i>čítáli</i>		

Tab. 170.11: Russian conjugation

position category	5	4	3	2	1	0
	negation	subject	tense/aspect	relative	object	stem

Fig. 170.4: Swahili verb prefixes

rank category	0	1	2	3	4
	stem	ADN/NUM*	relational <sup>0-2</sup>	modal <sup>0-2</sup>	(associating/complementizing)

Fig. 170.5: Kayardild nominal suffixes

tribute to an account of the function of a given case.

There are two basic types of **sequential order** of elements, scope order and template order. **Scope order** presupposes a binary construction consisting of a head and a dependent one of which has scope over the other, and specifies the position of the dependent as against the head. Thus, the position of an adnominal relative clause is determined with respect to its head nominal, and the position of the complement NP with respect to an adposition. At the morphological level, this convention is replaced by specifying the position of bound morphemes with respect to their hosts, or of grammatical formatives with respect to a lexical root or stem. In a syntagm such as Fig. 170.3, the relative order of the suffixes *-ig* and *-ung* is a contingent consequence of the fact that each of them has scope order with respect to the operand preceding it. This may be summarized in a formula like [ [ X ]<sub>N/V</sub> -ig ]<sub>A</sub>.

**Template order** is the relative order of elements which do not form a construction with each other. It is typical of clitics which attach to whatever their host is, in a fixed order that does not reflect any semantic scope. Therefore template order cannot be specified in the same straightforward way as scope order. Instead, a schema is set up consisting of a fixed sequence of slots, a **slot** being a syntagmatic position for an element. For instance, the order of the prefixes on the Swahili verb may be described by the template in Fig. 170.4.

The interpretation of such a template is straightforward: Each of the five prethematic slots is reserved for an element out of the paradigm designated by the slot label. If a slot remains unoccupied, this does not concern the order of the other elements.

Some morphological systems are so complex that different ranks of complex stems may be assumed such that affixes may be added at each rank. The suffix positions of the Kayardild noun may be schematized as in Fig. 170.5 (according to Evans 1995: 122).

At the first rank, the stem may be followed by any number of adnominal case and number suffixes. At rank 2, up to two relational cases, and at rank 3, up to 2 modal cases may follow. At rank 4, there may be either an associating or a complementizing suffix.

Templates are necessary when the sequential position of elements is not in consonance with their semantic scope. Structuralist accounts of the middle of the 20<sup>th</sup> century often give templates even for constructions which are decomposable into binary constructions in each of which scope order obtains. An example would be ‘stem – past tense – personal ending’ for the German regular finite verb. This is helpful in visualizing a complex form as a whole, but should be preceded by a separate account of each of the elementary constructions involved.

#### 4.5. Illustrative examples

Abstract or formal statements should be supplemented by example expressions from the language being described. One of the most conspicuous differences between older and modern linguistic descriptions is the growing extent to which this principle is heeded. One might suppose this to be a trivial issue. Actually, however, descriptions that contain no examples are very often not intelligible and, to this extent, worthless. Since most linguistic statements fail in one or another respect and simply do not do justice to the real complex-

ity of the language, an example can teach more than an ever so neatly formalized rule.

Above a certain degree of syntagmatic complexity, such an example should be presented in the canonical trilinear representation (cf. Art. 169), i.e. the text should be provided with an interlinear morphemic gloss and an idiomatic translation. While this was practically never done in traditional descriptions and even in structural linguistics, modern grammatical descriptions, especially those written in a functional or typological framework, often abound in analyzed illustrative examples. In this respect, there has been a real progress in descriptive linguistics.

There is a certain amount of trade-off between the corpus specimina as presented in the appendix and the examples in the running descriptive text. First of all, extensive illustration in the description can be reduced and be replaced by references to the corpus. Generally, such segments from the corpus are copied as illustrative examples in the running text which represent frequent patterns and do not involve any additional complications irrelevant to the issue at hand. Second, not all of the levels of representation figuring in the edition of the text corpus need be repeated in an example of the running text. For many purposes and especially for simple examples, some representation of the significans (orthographic or morphological, i.e. phonological with morpheme boundaries marked) coupled with a translation as a shorthand semantic representation will be sufficient. Thirdly, wherever uniformity of a set of examples is required, for instance in a conjugation paradigm, examples will not (entirely) be drawn from the corpus, but will rather complement it.

Normally, the illustration of a morphological phenomenon does not require a whole example sentence. Very often, a word form or, at most, a phrase will do. Where a sentence appears to be needed to adequately exemplify a grammatical phenomenon, this is a hint at the syntactic rather than morphological nature of the phenomenon. Also, while a corpus sentence may appear to be a more natural piece of language use, it may also distract, by its complexity, from the morphological phenomenon that is at stake. The situation is, of course, different in syntax. There the author bears heightened responsibility if she simplifies corpus sentences for illustrative purposes.

## 5. References

- Cherubim, Dieter (1973), "Grammatikographie". In: Althaus, Hans Peter & Henne, Helmut & Wiegand, Herbert Ernst (eds.), *Lexikon der germanistischen Linguistik*. Tübingen: Niemeyer, 768–778
- Cole, Peter (1982), *Imbabura Quechua*. Amsterdam: North-Holland (Lingua Descriptive Studies, 5)
- Comrie, Bernard (1978), *Aspect: An Introduction to Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press
- Comrie, Bernard & Smith, Norval (1977), "Lingua Descriptive Studies: Questionnaire". *Lingua* 42, 1–72
- Evans, Nicholas (1995), *A Grammar of Kayardild*. Berlin, New York: Mouton De Gruyter
- Gabelentz, Georg von der (1881), *Chinesische Grammatik: mit Ausschluß des niederen Stils und der heutigen Umgangssprache*. Leipzig: Weigel
- Gabelentz, Georg von der (1891), *Die Sprachwissenschaft: Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig: Weigel Nachf. [2. überarbeitete Aufl.: Leipzig: Tauchnitz, 1901]
- Givón, Talmy (1993), *English Grammar: A Function-Based Introduction*, Vol. 1–2. Amsterdam & Philadelphia: J. Benjamins
- Haspelmath, Martin (1993), *A Grammar of Lezgian*. Berlin: Mouton de Gruyter (Mouton Grammar Library, 9)
- Haspelmath, Martin (2003), "The Geometry of Grammatical Meaning: Semantic Maps and Cross-linguistic Comparison". In: Tomasello, Michael (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, Vol. II. Hillsdale/NJ: L. Erlbaum, 211–242
- Humboldt, Wilhelm von (1836), "Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechtes". Reimpr.: Humboldt, Wilhelm von (1963), *Schriften zur Sprachphilosophie*. [= *Werke in fünf Bänden*, hrsg. v. A. Flitner und K. Giel, Bd. III]. Darmstadt: Wissenschaftliche Buchgesellschaft, 368–756
- Jespersen, Otto (1937), *Analytic Syntax*. Copenhagen: E. Munksgaard
- Lehmann, Christian (1980), "Aufbau einer Grammatik zwischen Sprachtypologie und Universalistik". In: Brettschneider, Gunter & Lehmann, Christian (eds.), *Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 145), 29–37

- Lehmann, Christian (2002), *Possession in Yucatec Maya*. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 10)
- Lehmann, Christian (2004 a), “Data in Linguistics”. *Linguistic Review* 21.3/4: 175–210
- Lehmann, Christian (2004 b), “Funktionale Grammatikographie”. In: Premper, Waldfried (ed.), *Dimensionen und Kontinua: Beiträge zu Hansjakob Seilers Universalienforschung*. Bochum: N. Brockmeyer (Diversitas Linguarum, 4), 147–166
- McGregor, William B. (1990), *A Functional Grammar of Gooniyandi*. Amsterdam & Philadelphia: J. Benjamins (SLCS, 22)
- Mosel, Ulrike (2002), “Analytic and Synthetic Language Description”. In: Ezawa, Kennosuke & Kürschner, Wilfried & Rensch, Karl H. & Ringmacher, Manfred (eds.), *Linguistik jenseits des Strukturalismus. Akten des II. Ost-West-Kolloquiums, Berlin 1998*. Tübingen: G. Narr, 199–208
- Mosel, Ulrike (2003), *Grammaticography – the Art and Craft of Writing Grammars*. Kiel: Seminar für Sprachwissenschaft der Universität
- McQuown, Norman A. (1990), *Gramática de la lengua totonaca (Coatepec, Sierra Norte de Puebla)*.
- México, D.F.: Universidad Nacional Autónoma de México (Colección Lingüística Indígena, 4)
- Noonan, Michael to app., “Grammar Writing for a Grammar-Reading Audience”. In: Payne, Thomas & Weber, David (eds.), *Grammar Writing*. Amsterdam & Philadelphia: J. Benjamins (Typological Studies in Language)
- Payne, Thomas (1997), *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press
- Plank, Frans (1992, ed.), *Paradigms: The Economy of Inflection*. Berlin & New York: Mouton de Gruyter (Empirical Approaches to Language Typology, 9)
- Shopen, Timothy (1985, ed.), *Language Typology and Syntactic Description, Vol. 1–3*. Cambridge etc.: Cambridge University Press
- Zaliznyak, Andrey (1967), *Russkoe imennoe slovoizmenenie* [Russian nominal inflection]. Moscow: Nauka

*Christian Lehmann, Erfurt (Germany)  
Elena Maslova, Stanford (U.S.A.)*

## 171. Lexicography

1. Morphology in dictionaries
2. Morphological models
3. Morphology and lexicography
4. References

### 1. Morphology in dictionaries

#### 1.1. Monolingual dictionaries

We begin our discussion of the contribution of morphology to **lexicography** with a review of the treatment of inflection (cf. Mugdan 1989–1990 a) and word formation in **monolingual dictionaries** intended for native speakers. The questions that need to be asked are: What kinds of morphological information do dictionaries contain? How do they present them?

##### 1.1.1. English

Our starting point is the second edition of the *Oxford English Dictionary* (Simpson & Weiner 1989, eds.). Each “main word”, i.e. headword or primary entry, begins with its “identification”, which includes:

“The inflections, i.e. plural of substantives, and principal parts of verbs, when other than the ordinary -s, -ed.” (Simpson & Weiner 1989, eds.: xxvii)

Irregular inflections only are indicated for nouns and verbs. In the area of word formation, the OED distinguishes “combinations”, i.e. compound words, and “derivatives”, i.e. words formed by suffixation. Combinations are treated as main words if they have a specialized meaning which cannot be deduced from the meanings of their elements and therefore require a full definition, or if they have multiple senses. Otherwise they are treated as run-ons, at the end of the article for the main word constituting the first element of the combination. The category of derivatives does not include words derived by prefixation, since, because of the alphabetical ordering of the dictionary, these must be treated as main words in their appropriate place. Derivatives by suffixation are also treated as main words, unless they are of infrequent occurrence or have only one or a small number of meanings directly relatable

to the meaning of the main word from which they are derived; in which case they are treated in the article for the main word in a separate paragraph at the end. Also given main word status are prefixes, suffixes and combining forms that are especially productive in the formation of derived and compound words.

The practice of the OED in the treatment of morphological information is followed in broad terms by other British English dictionaries, though with some significant variations. In the treatment of inflections, *Collins English Dictionary* (Hanks 1979, ed.; Treffry 1998, ed.), *Chambers English Dictionary* (Schwarz et al. 1988, eds.) and the *Concise Oxford Dictionary* (Allen 1990, ed.; Pearsall 1999, ed.) give not only those formed irregularly but also regular formations which are likely to cause confusion or difficulty, e.g. where the final consonant of the root doubles with the addition of the inflectional suffix (*dragging*) or where a letter is dropped from the root (*arguing*) or where the addition of a suffix changes the pronunciation of the root (*younger*). The inflections of adjectives are treated as well as those of nouns and verbs, and the *Concise Oxford* at least indicates which two-syllable adjectives regularly take the *-er/-est* comparative and superlative suffixes. These dictionaries also enter irregularly inflected forms of words as headwords, with a cross-reference to the base or citation form. The treatment of compounds and derivatives is, however, variable. *Chambers* treats all compounds and suffixed derivatives as run-ons; prefixed derivatives are headwords, except that derivatives with productive prefixes, like *re-*, *sub-*, *un-*, are listed without definition at the bottom of the page unless they require definition by virtue of a specialized or particularly complex meaning. *Collins* enters all compounds as headwords and all suffixed derivatives, unless the meaning of the derivative is clear from the root, when it is treated as a run-on under the root. It also, like *Chambers*, has lists of prefixed derivatives at the bottom of the page. The *Concise Oxford* has prefixed derivatives only as headwords, and suffixed derivatives are treated in the same way as in *Collins*. With compounds, however, the *Concise Oxford* uses a spelling rather than a meaning criterion: if the compound is written solid it is treated as a headword, otherwise it is treated as a run-on under the first element.

The treatment of inflection is particularly detailed in *Webster's Third New International*

*Dictionary* (Gove 1961, ed.). All inflectional forms are indicated in the dictionary, whether irregular or regular. Regular inflections are given only by suffix, but where any inflected form of a word is irregular all the forms are given in full. If an inflected form would fall in the alphabetical sequence "more than five inches" from its root, then it is entered as a headword and cross-referenced to the root. These practices also mean that all adjectives which form their comparative and superlative by suffixation have that indicated. In addition to the inflectional information in the dictionary entries, the front matter to *Webster's Third* contains two pages of detailed spelling rules for forming the plurals of nouns. It also contains two pages of detailed rules for writing compounds. All types of compound – solid, hyphenated and open – are entered as headwords, as are suffixed derivatives where they need separate definition.

### 1.1.2. French

English has a relatively limited inflectional morphology, with a small number of irregularities, affecting only the plural morpheme in the noun and the past tense and past participle morphemes in the verb (cf. Art. 119). It is, therefore, practicable to indicate, where necessary, the irregular forms in the dictionary entries themselves. French, however, shows a more complex inflectional morphology, especially in the verb (cf. Art. 121). The *Grand Larousse* (Guilbert et al. 1971, eds.) contains a "Tableau des Conjugaisons", covering several pages, to which the verb entries in the dictionary are cross-referenced. For example, *acquérir* is indicated as belonging to conjugation 13. This dictionary also contains in its front matter extensive discussion of the word formation processes of prefixation, suffixation and compounding, from both a diachronic and a synchronic perspective. An alternative solution to the inflectional question is found in *Le Petit Robert* (Rey 1973, ed.; cf. Drivaud 1997, ed.). For nouns and adjectives, the feminine form and the plural form (when it is not the regular *-s*) are given near the beginning of the entry for the word (e.g. *journal*, *aux*; *jumeau*, *elle*). For verbs, a number of strategies are employed: there is an appendix containing the "Conjugaisons des Verbes Français"; regular *-er* and *-ir* verbs have their inflections given only in the table of conjugations; verbs with minor irregularities, such as the inclusion of *-e-* in *bougeons* (1<sup>st</sup> person plural present of *bouger*) or the *ç* in *plaçons*,

have these irregularities given in the dictionary entry for the verb; major irregular verbs are referred to a prototype, e.g. “PRODUIRE, conjug. *conduire*”, where the forms for that group of irregular verbs are given, as well as in the table of conjugations. *Le Petit Robert* includes derivational prefixes and suffixes as headwords and also contains a table of suffixes as an appendix.

### 1.1.3. German

In German, the inflectional forms are yet more diverse and varied (cf. Art. 120). In the *Deutsches Wörterbuch* (Wahrig et al. 1980–1984, eds.), the front matter contains numbered tables for the declension of nouns and pronouns, for the declension of adjectives and indefinite pronouns, and for the conjugations of verbs, which are used for cross-reference from the dictionary entries. Each noun entered in the dictionary is marked for gender, and the genitive singular and the plural inflections are given; uncountable nouns (i.e. with no plural inflection) are marked as such, as are nouns occurring only in the plural form; for the adjectival and nominal declension of nouns the user is referred to the declension tables. A similar cross-reference is made for adjectives, where three tables are devoted to regular patterns and five to deviations from these. In the case of verbs, so-called “weak” verbs are regarded as the regular category, and there is no cross-reference to a conjugation table, though one is provided. In the case of “strong” or irregular verbs, the entry for the verb contains a number which cross-refers to a conjugation table. Irregular inflectional forms are entered as headwords in the dictionary, with an indication of their form and the citation form to which they belong, e.g. “brachte <1, 3 Sg Ind Prät von> bringen”. Clearly, in the case of German, and to some extent of French, an important question is how and where to show the inflectional forms of words, whether in the dictionary entries themselves or in the tables outside of the main body of the dictionary with cross-references in the entries. It is a question to which we shall return (cf. 3.2).

### 1.2. Learners' dictionaries

**Learners' dictionaries** are a particularly English phenomenon, though a French and a German learners' dictionary will be referred to below. They aim to be a comprehensive and detailed guide to usage for, mainly, ad-

vanced foreign learners. Care is, therefore, taken to ensure that all necessary information, especially grammatical information, is included (Jackson 1988: 174–191). The first in the marketplace, in 1948, was A. S. Hornby's *Oxford Advanced Learner's Dictionary*, now in a much revised fifth edition (Crowther<sup>5</sup> 1995, ed.). In this dictionary, irregular inflections are shown and where the addition of an inflectional suffix causes spelling changes, e.g. the doubling of a stem-final consonant. All words with irregular inflections are listed as headwords, with cross-reference to the stem form. There is a table of irregular verbs as an appendix. All adjectives forming the comparative and superlative in -er/-est are said to be marked for this, though the dictionary (4<sup>th</sup> edition, cf. Cowie<sup>4</sup> 1989, ed.) is less than consistent in practice (Bolinger 1990: 135). In the area of word formation, suffixed derivatives are entered as run-ons after a distinctive triangle symbol, if the meaning is clear from the root word; otherwise a derivative is entered as a separate headword, as it is if the spelling connection is not clear, e.g. with *satisfaction* from *satisfy*. Compounds are usually listed as run-ons under the first element and are preceded by a distinctive square symbol; however, if the first element of a compound is not a headword, or if the compound is a loanword or has an idiomatic meaning, then it is entered separately. Affixes and combining forms are listed as headwords and accompanied by explanations and examples.

The second advanced learners' dictionary was the *Longman Dictionary of Contemporary English* in 1978, now in a third edition (Summers<sup>3</sup> 1995, ed.). It has the same information on inflections as its Oxford competitor, except that it contains no information about adjective inflections in the entries apart from irregular ones. In word formation, suffixed derivatives are entered as run-ons if the meaning is derivable from the root, but compounds are always entered as separate headwords. The dictionary (2<sup>nd</sup> edition, 1987) contains a ten-page appendix on word formation, including alphabetical listings of “word beginnings” and “word endings”. Common affixes and combining forms are listed as headwords and cross-referenced to the appendix.

The third English advanced learners' dictionary was the *Collins Cobuild English Language Dictionary* (Sinclair<sup>2</sup> 1995), which, in its construction – based on a large computer

corpus – and in its presentation of information, marked a significant departure from conventional lexicographic practice. Inflectional forms are given in full, at the beginning of the entry, for all words that have inflections, both regular and irregular. Each entry is for a single orthographic word; there are no separate entries for a word form belonging to more than one word class. There is one entry for *play*, for example, with the inflectional forms *plays*, *playing*, *played*; the form *plays* indicates both ‘noun plural’ and ‘verb 3<sup>rd</sup> person singular present tense’. In word formation, a suffixed derivative is entered as a run-on, after a distinctive diamond symbol, only if the meaning is directly derivable from the stem; otherwise, if the derivative is very common or has a different meaning from the stem, it has a separate entry. Compounds always have separate headword status, as do common prefixes and suffixes (including combining forms).

A fourth advanced learners’ dictionary in English is the *Cambridge International Dictionary of English* (Proctor 1995, ed.). For inflection the general principle is that, in the case of nouns and verbs, only irregular inflections are given in the dictionary, but regular inflections are described in the front matter and boxed “language portraits” in the body of the dictionary add further information. In the case of adjectives, all those that can form comparative and superlative with the suffixes *-er/-est* are so marked in the dictionary, as are ungradable adjectives. One of the features of the *Cambridge International* is the multiple entry of headwords with more than one core sense. Derivatives are entered, each on a separate line, below the sense to which they relate. Additionally, there is a “language portrait”, entitled “Combining Forms”, which describes the main derivational prefixes and suffixes of English. Compound words all have headword status.

English learners’ dictionaries both reflect and vary the diversity of practice in the treatment of morphological information that was evident in the monolingual dictionaries of English. The same is true of the French learner’s dictionary, *Dictionnaire du Français Contemporain* (Dubois et al. 1971, eds.). It contains in its front matter tables of “Conjugaisons des Verbes” and “Suffixes et Préfixes”. In the dictionary entries, if a verb has an irregular conjugation, it is cross-referenced by number to the table of conjugations, e.g. “*sortir* (conj. 28)”. An interesting

feature of this dictionary are the boxed articles in the body of the dictionary on derivational prefixes, e.g. *archi-*, *hypo-*, *in-*.

A first German advanced learners’ dictionary appeared in 1993, *Langenscheidts Großwörterbuch Deutsch als Fremdsprache* (Götz et al. 1993, eds.). The inflections of all nouns (genitive and plural) and of all verbs (simple past and perfect tenses) are given; additionally the “3<sup>rd</sup> person singular present tense” is given for irregular verbs where it differs from the stem of the infinitive headword. This dictionary has paid particular attention to aiding the learner understand the word formation processes of German. A larger number of prefixes and suffixes are entered as headwords, with a label indicating how productive they still are, and with some typical examples. Common compounds are listed within the entries for words, after a *K*- (for “Komposita”) symbol where the word is the first element of a compound, and *-K* where it may be the second element.

### 1.3. Bilingual dictionaries

In **bilingual dictionaries**, inflectional information is normally given for headwords. In a French/English bilingual dictionary, for example, inflections are given for French words in the French-English side of the dictionary, and for English words in the English-French side; though this does not invariably apply. In the *Collins-Robert French/English Dictionary* (Atkins et al. 1987, eds.), the irregular or unobvious plurals of nouns are indicated in this way. However, for gender variants, the feminine and masculine forms of a French noun are given on the English-French side if they could both be the translation of an English noun, e.g. “*singer .. n chanteur m, -euse f*”.

The feminine forms of French adjectives are also given on the English-French side, if they are not formed regularly. In the case of verb inflections, the *Collins-Robert* follows the respective monolingual traditions: French verbs with irregular inflections, on the French-English side of the dictionary, are given a number, which refers the user to the tables of verb conjugations in the appendix; the inflections of irregular English verbs are given at the entry for the verb in the English-French side of the dictionary, though there is also an appendix of irregular English verbs. Inflected forms of nouns and verbs that cannot be easily related to the root form are en-

tered as headwords, with a cross-reference to the citation form.

The separate monolingual traditions are also followed in the *Collins German Dictionary* (Terrell et al. 1980, eds.). The genitive and plural inflections are given for all German noun headwords, on the German-English side, except for certain regular noun inflections, indicated in a table in the front matter. For English, the irregular plurals, and only these, are given for nouns on the English-German side. Similarly, the irregular, or “strong” verbs have their inflections given in the entries for the simple verb. This means that verbs which are derived by prefix or compounding from a simple irregular verb are cross-referenced to that verb for the inflectional information. German verbs that have regular inflections apart from not forming the past participle with *ge-* are marked with an asterisk, e.g. *bedienen\**, *hantieren\**. Irregular inflectional forms are listed as headwords and cross-referenced to the citation form.

In bilingual dictionaries, the need to provide the information that a user requires at the most convenient place vies with the pressing need to keep the dictionary within manageable bounds of both volume and cost. Inflectional information would most conveniently, from the user's point of view, be provided at the point of look-up, e.g. for German verbs on the English-German side of the dictionary. But providing information there would mean either increasing the volume of the dictionary or leaving something else out. The user must therefore look in two places to discover, for example, the past tense form of the German translation of *speak*: in the English-German side to discover that the equivalent in German is *sprechen*, and in the German-English side to discover that the past tense form is *sprach*. How to present inflectional – and word formation – information in an accessible way is an important question, to which we shall return below (in 3).

## 2. Morphological models

### 2.1. Dictionary entries

Before we consider which of the **models of morphology** is most used in describing morphological data in dictionaries, it will be useful to consider the nature of a **dictionary entry**. A dictionary entry consists of a headword and an accompanying article that gives

a lexical description of the headword. The lexical description includes phonological, morphological, syntactic, lexicological, semantic and pragmatic information (Jackson 1988: 35–47; but cf. Dubois & Dubois 1971: 90).

The headwords of a dictionary usually include the “simple” words (i.e. the root forms that have not undergone any word formation processes) of the vocabulary in their citation form (see 3.1), some compound words, and some derived words. Compounds and derivatives are usually given separate headword status if their meaning has diverged from that of the root so that it can no longer be deduced directly from the form of the word itself. Otherwise, compounds and derivatives are treated as run-ons under the simple word, i.e. as part of the morphological/lexicological description of the simple word.

Headwords may also include inflected forms of words (for cross-reference purposes), affixes and combining forms, abbreviations, and so on. This raises the question whether dictionaries take the word or the morpheme as their basic unit of description. The vast majority of headwords in dictionaries, for European languages at least, are words, in their citation form, though what counts as a dictionary word is not always a simple matter to decide (Dubois & Dubois 1971: 61–65). Other items that appear as headwords (prefixes, suffixes, and so on) are included because they are deemed to represent useful reference material for dictionary users; they are sometimes included additionally or instead in an appendix to the dictionary. Whether the word is the appropriate basic unit for highly inflecting languages like Swahili is another question (Bwenge 1989).

Some of the information contained in the article that accompanies a headword may be morphological. This includes the inflectional forms of the headword, usually if these are not formed on the regular pattern for the word class concerned (Jackson 1988: 148). It also includes the words that have been formed from the headword by the word formation processes of compounding and derivation, where the resulting forms do not require separate definition. In the case of derivation, however, it is only suffixed derivatives that are included in the article; words derived by prefixation have separate headword status. This discrepancy arises from the exigencies of the alphabetical ordering of dictionary entries and the assumption that a user

cannot be expected to make a morphological analysis of a word, i.e. recognise the presence of a derivational prefix, before looking it up.

## 2.2. Word and paradigm

Dictionaries for European languages operate with words as their basic unit (Zgusta 1971: 159). The implicit working model of lexicographers tends therefore to be the **word-and-paradigm** model of morphology (Zgusta 1971: 119). This is most obvious in the French monolingual dictionaries reviewed in 1.1.2, where the inflections of verbs are indicated in the dictionary entry merely by a number, which refers the user to a paradigm in the table of conjugations at the front or back of the dictionary. But it is equally the model implicit in the indication of irregular inflections in the dictionary entry itself. First of all, a regular paradigm of inflections is assumed – the “declension” of nouns, the “conjugation” of verbs – which is either listed in the front or back matter of the dictionary or assumed to be in the general knowledge of the dictionary user (Hoenigswald 1967: 109). Secondly, the inflectional forms are given in the articles for irregularly inflected words that enable the user to construct the appropriate paradigms for each word. For example, the “principal parts” of verbs in German and English are given (3<sup>rd</sup> person singular present tense, past tense, past participle), from which the paradigms for the various tenses can be worked out (Lyons 1977: 515).

What the word-and-paradigm model does not always make explicit, because of the assumption of a regular paradigm, is that there may be gaps in the paradigm even for some regularly inflected words (Hoenigswald 1967: 109). Not all verbs in English have the full paradigm forms: stative verbs (e.g. *know*) do not usually form progressive tenses; modal verbs (e.g. *can*) do not have non-finite forms. There are nouns that have no plural forms (German *Armut*, English *poverty*), usually marked in learners’ dictionaries as “uncountable”.

The word-and-paradigm model assumes a language whose major word classes inflect and where a dictionary user can either be referred to an appropriate table or be given some principal parts from which to construct the paradigms. Where a language does not have a series of inflected forms, the paradigm model is irrelevant, e.g. for Burmese (Zgusta 1971: 162). The model really only has something to say about the inflectional aspects of

morphology; it does not treat the forms that arise by regular or irregular derivational processes (Hoenigswald 1967: 106).

## 2.3. Word formation

One reason why the word-and-paradigm model does not apply well to derivation and other types of word formation is that it is essentially a static model. **Word formation** is conceived in terms of processes. General monolingual synchronic dictionaries are static descriptions; they record the products of word formation processes, rather than the processes themselves. However, there are ways in which dictionaries can indicate the process nature of word formation. Firstly, the process by which a compound or derivative comes into being may be indicated in the etymology of the word (see Art. 174). Secondly, dictionaries often list affixes and combining forms as headwords, or in an appendix, with an implication that they continue to be productive in the processes of forming new words. With compounding, that is more difficult. There is much that the practice of lexicography can learn from the linguistic investigation of morphology (Jackson 1988: 250) and *vice versa*; it is to this we now turn.

# 3. Morphology and lexicography

## 3.1. Citation forms

The form of a word that is entered as the headword is known as the “citation form” (Lyons 1977: 513; Jackson 1988: 9) or the “canonical form” (Zgusta 1971: 119; Landau 1989: 76). For words that undergo inflection, the **citation form** represents the paradigm (Landau 1989: 76; Zgusta 1971: 119). In that sense they may be regarded as “theoretical units” (Dubois & Dubois 1971: 62), since a headword (e.g. *biscuit*) stands for all the possible forms of the word (i.e. singular *biscuit* and plural *biscuits*).

In English dictionaries, the citation form is the stem to which the inflectional suffixes are added (Lyons 1977: 513). So, for regular verbs in English, the citation form is the infinitive without *to* (e.g. *play*), to which the inflectional suffixes are added (e.g. *plays*, *played*, *playing*). In German and French dictionaries, the citation form is also the infinitive (e.g. *jouer*, *spielen*); but in these languages it is not the stem to which the inflectional suffixes are added. The infinitive suffix, *-er* and *-en* respectively, has to be

removed to produce the stem forms (*jou-*, *spiel-*). In *Der Sprach Brockhaus* (see Brockhaus<sup>8</sup> 1972), however, verbs are entered in the 1<sup>st</sup> person singular present tense form (e.g. "ich spiele"). In the case of Swahili, it is by no means clear what the citation form should be (Bwenge 1989: 5 f.). Nouns in Swahili, for example, whether singular or plural, always have a prefix that marks number, but which also acts as a gender/class marker: should the headword be the stem form, without prefixes, or the singular form? The former would make it difficult for a user of the dictionary to locate a word unless they knew which noun class it belonged to. The latter would group together in the dictionary all the nouns belonging to the same class, and it assumes that the user will then recognise the root and be able to deduce the plural form.

What the citation forms should be in any particular instance is not always immediately obvious. The choice will be informed both by the insights of morphology and by the practical demands of the dictionary. This is why it is important, for example, for inflected forms, if they are not readily relatable to their stem, to be entered as headwords and cross-referenced to the citation form. Such a practice is especially necessary in bilingual dictionaries, where the users may not be familiar with the word that they are looking up (Landau 1989: 77).

### 3.2. Inflectional variants

It is well-established that, as part of the idiosyncratic information about words that dictionaries contain, inflections are included, especially if they are irregular (Jackson 1985: 54). Idiosyncratic information is not predictable from the general rules of grammar; in the case of inflections, the declension of a noun or adjective or the conjugation of a verb is not predictable from the phonological form of a word (Bauer 1983: 192). The requirement on a dictionary is that it should be possible to construct the whole paradigm of a word from the information given in the entry (Zgusta 1971: 121). This is usually interpreted to mean that regular inflections need not be given, since the dictionary user is assumed to know them; *Webster's Third* (Gove 1961, ed.) and *Collins Cobuild* (Sinclair<sup>2</sup> 1995, ed.) are exceptions to this generalisation. For irregular inflections, enough information needs to be given for the complete paradigm to be deduced, as the principal parts of "strong" verbs in German allow the para-

digm for all the tenses to be constructed. However, this is not always or not always obviously the case.

In German dictionaries it is customary to give for nouns the genitive singular inflection and the plural inflection (e.g. *Blick* -(e)s, -e; *Genosse* -n, -n). Unless a user knows the way in which the singular paradigm for German nouns works, he will not be able to conclude that in the case of *Blick* the dative singular form is *Blick*, but in the case of *Genosse* it is *Genossen*. And if he knows that, he hardly needs the information that is given in the dictionary (Wiegand & Kučera 1982: 309 f.).

What information does a dictionary user need to have about inflections? There are two different look-up questions in respect of inflections that a user might come to a dictionary with (Mugdan 1983: 180). The user may come with an inflected word form and want to find the corresponding headword or citation form (lemma). Alternatively, the user may come with a citation form and want to know what the inflected forms of the lemma are (cf. Bergenholz & Mugdan 1982: 29 f.). To satisfy the first query, a dictionary needs to contain a grammatical description that includes inflectional paradigms, as well as entering irregularly inflected forms as headwords with cross-reference to the citation form (Mugdan 1983: 182 f.; cf. Cowie 1983: 101). To satisfy the second query, again the grammatical description needs to be included in the dictionary, and, if possible, an indication of irregular inflections needs to be given at the entry for the citation form (Mugdan 1983: 187 f.). There appears to be a strong argument for including a brief grammatical description in the dictionary, to which dictionary entries can be cross-referenced (cf. Bergenholz & Mugdan 1982: 19). This is customary practice for the conjugations of verbs in French dictionaries (see 1.1.2), but because the complexities are too great to be able to handle in the entries themselves rather than for any principled reason. The inclusion of such a grammatical sketch in bilingual dictionaries may be a solution to the problem of where to indicate the inflectional information in these dictionaries (cf. 1.3; Mugdan 1983: 189).

A particular problem – though it is probably not unique – is posed by the adjective inflection in English. First of all, there is an alternation in forming the comparative and superlative forms between an inflectional suffix (-er, -est) and *more/most* before the adjec-

tive (Jackson 1985: 54 f.). The inflectional suffix is usual with words of one syllable and with some two-syllable words; the periphrastic form is obligatory with three-syllable words and longer. Dictionaries need to indicate, at least for two-syllable words, where *-er/-est* is allowed and where not, as in the *Cambridge International* (Proctor 1995, ed.) (see 1.2). But the inflectional suffix may not be possible even for some one-syllable words, but not for any morphological reason (Bolinger 1990: 135): e.g. *souer* or *wronger* do not “sound right”. Further linguistic investigation and guidance would greatly aid lexicographers in instances like these.

### 3.3. Derivatives

**Derivation**, as a word formation process, is of great interest to dictionary compilers. They are keen to include in their works all new words that have become lexicalised (accepted into the language) since the previous edition of a dictionary. Because the number of derivatives can greatly enhance the calculation of the number of “references” included, they may be used to inflate the claimed size of a dictionary (Landau 1989: 78; Mugdan 1984: 241; Jackson 1998: 27). In their function as a record of the vocabulary of a language, dictionaries are naturally much concerned with word formation processes. Dictionaries need clear policies and practical guidance on several matters relating to the treatment of derivatives and compounds.

One of the matters concerns the relationship between derivation and inflection, “a key point in the linguistic theory of word formation” (Ilson 1985: 163). The sharp boundary that is implied by the separate treatment of inflection and derivation in this article does not exist in reality (Zgusta 1971: 129); it is in many respects considered to be controversial (Lyons 1977: 521). The point can be illustrated from the present and past participles in a number of languages (Dubois & Dubois 1971: 63; Landau 1989: 91). The participles are usually considered to be inflectional forms of the verb; so they do not merit an entry in a dictionary unless they are formed irregularly. However, some participles are used more commonly than others as adjectives and, in this function, may develop a meaning that differs in some respect from the meaning of the verb to which the participle is related, e.g. French *étonnant*, English *surprising*. Such participles would warrant being entered as separate headwords, which implies

a lexical derivation rather than a grammatical inflection. From the perspective of derivation, many derivational processes are highly regular and predictable, e.g. *un-* as a negative-forming prefix in English, or the *-ish* suffix added to adjectives, neither of which involve a change of word class of the item to which they are affixed. To that extent they are like inflections (Zgusta 1971: 127), and a dictionary of contemporary Polish did not intend to include regular derivations (Saloni et al. 1990: 7). However, derivational rules tend to be less productive than inflectional ones, and they sometimes lead to unpredictable semantic effects (Lyons 1977: 524). A lexicographer has to decide at what point to count an item as a different word and therefore deserving a separate entry in the dictionary.

How then are derivatives treated in dictionaries? Prefixed derivatives, because of the alphabetical ordering of dictionary entries, are separated from the stem to which they are related, “so that the user loses the chance to grasp with a quick glance the combinatory potentialities of the given stem” (Kahane & Kahane 1967: 257). Suffixed derivatives are treated in two ways, and dictionaries differ in how individual derivatives are handled. The first way is to include a derivative in the article for its stem as a run-on entry, in bold type but usually with no, or minimal, definition. Derivatives whose meaning is directly related to that of the stem are normally handled in this way. The second way is to accord a derivative separate entry status as a headword, thus confirming the derivative as a different word. Derivatives whose meaning differs from that of the stem or whose use requires more extensive explanation are normally handled in this way (Landau 1989: 78). How a dictionary treats individual derivatives may depend on practical considerations of economy and retrievability and the weight given to these by the lexicographer (Cowie 1983: 102). The inconsistency in the treatment of prefixed as against suffixed derivatives and the implications this has for the view that a user has of derivational processes needs further attention from linguists and lexicographers. (Cf. the comments on the treatment of word formation in learners’ dictionaries by Prcic (1999).

### 3.4. Affixes and combining forms

Should the elements of word formation processes, though they are not words themselves, be accorded headword status in the dictio-

nary? The practice of doing this is variable (Bergenholtz & Mugdan 1982: 31), but their inclusion is seen as highly desirable (Benson et al. 1986: 236; Mugdan 1984: 300). The arguments for inclusion include the fact that they are meaning-bearing items (Müller 1982: 156), and that their inclusion may enable the dictionary user to decode new derivatives that are not yet lexicalised (Bergenholtz & Mugdan 1982: 32; Müller 1982: 167). For example, the inclusion of the combining form *-ectomy* will enable a user who knows the meaning of *tonsil* to decode the derivative *tonsilectomy* (Benson et al. 1986: 236).

The information included for affixes and combining forms needs to indicate not only their meaning, but also how they form new words (Benson et al. 1986: 236, cf. Bauer 1983: 193). Arguably, this kind of explanation may be more satisfactorily handled in an appendix, where all word formation processes can be considered together, as in some French dictionaries (see 1.1.2) and learners' dictionaries (see 1.2). To provide adequate points of access for users, it may be necessary to provide a three-fold treatment of derivatives: the inclusion of lexicalised derivatives as words; the inclusion of affixes and combining forms as headwords; and the treatment of word formation processes in an appendix.

Lexicographers usually make a distinction between affixes and combining forms, though some dictionaries operate just with the labels "prefix" and "suffix", or more unusually just with "combining form" (Proctor 1995, ed.). It is not certain that this is a distinction that all linguists would make (Ilson 1985: 164). There needs to be some clarification for lexicographers of the distinctions between inflectional affix, derivational affix, and combining form. They probably form a continuum from the obviously inflectional affix (like the English plural *-s*), through the inflectional affixes like English present participle *-ing*, through highly productive derivational affixes like the agentive *-er*, to the more obviously lexical *re-* ("again") and combining forms like *-ology* ("study of") (Ilson 1985: 164).

### 3.5. Compounds

In contrast to derivatives, one of the crucial issues in the treatment of **compounds** in dictionaries is deciding when a word + word or root + root combination constitutes a word or lexical item in its own right and therefore warrants entry in a dictionary. On the one hand this issue is connected with the general

problem of the definition of a lexical unit, when this is not a simple word nor apparently a syntactic phrase (Dubois & Dubois 1971: 30). On the other hand, it is connected with the criteria that might be applied to decide that a combination should be regarded as a compound. Spelling is no guide in English (Cowie 1983: 103; cf. Mugdan 1984: 285 for German), though some dictionaries use that as a criterion, requiring a compound to be written "solid" or at least "hyphenated", but not "open". More acceptable and reliable criteria include the requirement that a compound should have a semantic unity or form a semantic unit (Dubois & Dubois 1971: 30) or have a "unity of signification" (Cowie 1983: 104), as well as the requirement that a compound should be stable (Zgusta 1971: 135) or institutionalised (Lyons 1977: 535). This last criterion excludes occasional or nonce compounds, which may be readily produced in some languages but are bound to the context in which they are coined. Only if they subsequently become generally accepted by use – institutionalised or lexicalised – do they merit inclusion in a dictionary; but the lexicographer is faced with deciding when that point has been reached (Lyons 1977: 536), and it may in any case be indeterminable (Mugdan 1984: 245).

The other important issue for lexicography is how to treat compounds in the dictionary. Should they all be accorded headword status, or can they be included as run-ons under one or other of the constituent roots? From the user's perspective, a strict alphabetical ordering of headwords is probably ideal (Mugdan 1984: 290). But economy usually dictates that the second course is followed as far as possible. In that case, are there any criteria for deciding which element of the compound it should be entered under? If, for example, it is considered desirable to indicate the relationship between a compound like *fruit-picking* and the construction from which it derives (e.g. "They picked fruit"), then this compound needs to be entered under the verb *pick* rather than the noun *fruit* (Cowie 1983: 99). But compounds are usually entered under the first lexical constituent. Not only may this be arbitrary, but it also implies a relationship between compound and the item (Lyons 1977: 537), when the compound may have developed a specialised meaning that is no longer relatable to either of its constituents; though that may be a reason for entering the compound as a separate headword. Another approach to this question is to take

the practical perspective of the dictionary user: where would a user expect to find a particular compound? It may be necessary to enter a compound in more than one place, or to take account of the strategies that users employ. It would appear that strategies differ from language to language (Bogaards 1990: 79 ff.); French speakers, for example, take more account of a constituent's or word's frequency, while Dutch speakers are more sensitive to a word's class, tending to look for an item under a noun element before an adjective or verb element.

### 3.6. Conclusion

That morphological information should be included in dictionaries is without dispute. Such information includes at least the inflectional properties of headwords (Lyons 1977: 517), and additionally derivational features (Saloni et al. 1990: 5). Another proposal specifies as the morphological information about lexical items: structure in terms of morphemes or alternating phonological structures; irregular morphological structures linked to particular morpho-syntactic features; partial similarities to other words (in derivatives and compounds); cliticizing properties (Hudson 1988: 311). These are the areas where linguists and lexicographers can fruitfully interact in assembling the information that needs to be included in a dictionary.

More particularly lexicographical are the considerations of how to present this information, both to be true to the data and its description, and to take account of the practical needs of the dictionary user. How much of the information should be included in the dictionary entries? How extensive should the cross-referencing be? What is the place of grammatical descriptions in the front or back matter of the dictionary? Perhaps there is an argument for every dictionary to contain a brief grammatical sketch to which the entries can cross-refer (Bergenholtz & Mugdan 1982; Mugdan 1989–1990 b) or even for there to be a reference book that is dictionary and grammar combined (Lemmens & Wekker 1991).

## 4. References

- Allen, Robert E. (81990, ed.), *The Concise Oxford Dictionary of Current English*. Oxford: Clarendon Press
- Atkins, Beryl T. et al. (21987, eds.), *Collins-Robert French-English English-French Dictionary*. Glasgow: Collins
- Bauer, Laurie (1983), *English Word Formation*. Cambridge: Cambridge University Press
- Benson, Morton & Benson, Evelyn & Ilson, Robert (1986), *Lexicographic Description of English*. Amsterdam: John Benjamins
- Bergenholtz, Henning & Mugdan, Joachim (1982), "Grammatik im Wörterbuch – Probleme und Aufgaben". In: Wiegand (ed.), 17–36
- Bogaards, Paul (1990), "Où cherche-t-on dans le dictionnaire?". *International Journal of Lexicography* 3.2, 79–102
- Bolinger, Dwight (1990), "Review of Oxford Advanced Learner's Dictionary, Fourth Edition". *International Journal of Lexicography* 3.2, 133–145
- Brockhaus (81972) = *Der Sprach Brockhaus*. Wiesbaden: F.A.Brockhaus
- Bwenge, C. (1989), "Lexicographical Treatment of Affixational Morphology: A Case Study of Four Swahili Dictionaries". In: James, Gregory (ed.), *Lexicographers and their Works*. University of Exeter (Exeter Linguistic Studies 14), 5–17
- Cowie, Anthony P. (1983), "On Specifying Grammatical Form and Function". In: Hartmann, Reinhard R. K. (ed.), *Lexicography: Principles and Practice*. London: Academic Press, 99–107
- Cowie, Anthony P. (41989, ed.), *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press
- Crowther, Jonathan (51995, ed.), *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press
- Drivaud, Marie-Hélène (1997, ed.), *Le Robert Collège*. Paris: Dictionnaires Le Robert
- Dubois, Jean et al. (1971, eds.), *Dictionnaire du Français Contemporain*. Paris: Librairie Larousse
- Dubois, Jean & Dubois, Claude (1971), *Introduction à la Lexicographie: Le Dictionnaire*. Paris: Librairie Larousse
- Götz, Dieter & Haensch, Günther & Wellmann, Hans (1993, eds.), *Langenscheidts Großwörterbuch Deutsch als Fremdsprache*. Berlin: Langenscheidt
- Gove, Philip B. (1961, ed.), *Webster's Third New International Dictionary*. Springfield, Mass: G. & C. Merriam Co.
- Guilbert, Louis & Lagane, René & Niobey, Georges (1971, eds.), *Grand Larousse de la Langue Française*. Paris: Librairie Larousse
- Hanks, Patrick (1979, ed.), *Collins English Dictionary*. London, Glasgow: Collins
- Hausmann, Franz Joseph & Reichmann, Oskar & Wiegand, Herbert Ernst & Zgusta, Ladislav (1989–1990, eds.), *Wörterbücher, Dictionaries, Dic-*

- tionnaires – Ein internationales Handbuch zur Lexikographie*. Berlin, New York: Walter de Gruyter
- Hoenigswald, Henry M. (1967), "Lexicography and Grammar". In: Householder, Fred W. & Saporta, Sol (eds.), *Problems in Lexicography*. Bloomington/IN: Indiana University; The Hague: Mouton, 103–110
- Hudson, Richard (1988), "The Linguistic Foundations for Lexical Research and Dictionary Design". *International Journal of Lexicography* 1.4, 287–312
- Ilson, Robert (1985), "The Linguistic Significance of Some Lexicographic Conventions". *Applied Linguistics* 6.2, 162–172
- Jackson, Howard (1985), "Grammar in the Dictionary". In: Ilson, Robert (ed.), *Dictionaries, Lexicography and Language Learning*. Oxford: Pergamon Press (British Council ELT Documents 120), 53–59
- Jackson, Howard (1988), *Words and their Meaning*. Harlow: Longman
- Jackson, Howard (1998), "How Many Words in YOUR Dictionary?". *English Today* 55, Vol. 14.3, 27–28
- Kahane, Henry & Kahane, Renée (1967), "Problems in Modern Greek Lexicography". In: Householder, Fred W. & Saporta, Sol (eds.), *Problems in Lexicography*. Bloomington: Indiana University; The Hague: Mouton & Co.
- Landau, Sidney I. (1989), *Dictionaries, The Art and Craft of Lexicography*. Cambridge: Cambridge University Press
- Lemmens, Marcel & Wekker, Herman (1991), "On the Relationship between Lexis and Grammar in English Learners' Dictionaries". *International Journal of Lexicography* 4.1, 1–14
- Lyons, John (1977), *Semantics*, Vol. II. Cambridge: Cambridge University Press
- Mugdan, Joachim (1983), "Grammatik im Wörterbuch: Flexion". In: Wiegand, Herbert Ernst (ed.), *Studien zur neuhochdeutschen Lexikographie III*. Hildesheim etc.: Olms (Germanistische Linguistik 1–4.82), 179–237
- Mugdan, Joachim (1984), "Grammatik im Wörterbuch: Wortbildung". In: Wiegand, Herbert Ernst (ed.), *Studien zur neuhochdeutschen Lexikographie IV*. Hildesheim etc.: Olms (Germanistische Linguistik 1–3.83), 237–308
- Mugdan, Joachim (1989–1990 a), "Information on Inflectional Morphology in the General Monolingual Dictionary". In: Hausmann et al. (eds.), 518–525
- Mugdan, Joachim (1989–1990 b), "Grundzüge der Konzeption einer Wörterbuchgrammatik". In: Hausmann et al. (eds.), 732–749
- Müller, Wolfgang (1982), "Wortbildung und Lexikographie". In: Wiegand (ed.), 153–188
- Pearsall, Judy (101999, ed.), *The Concise Oxford Dictionary*. Oxford: Oxford University Press
- Prcic, Tvrtnko (1999), "The Treatment of Affixes in the 'Big Four' EFL Dictionaries". *International Journal of Lexicography* 12.4, 263–79
- Proctor, Paul (1995, ed.), *Cambridge International Dictionary of English*. Cambridge: Cambridge University Press
- Rey, Alain (1973, ed.), *Le Petit Robert: Dictionnaire Alphabetique et Analogique de la Langue Française*. Paris: Dictionnaires Le Robert
- Saloni, Zygmunt & Szpakowicz, Stanislaw & Swidzinski, Marek (1990), "The Design of a Universal Basic Dictionary of Contemporary Polish". *International Journal of Lexicography* 3.1, 1–22
- Schwarz, Catherine et al. (1988, eds.), *Chambers English Dictionary*. Edinburgh: W & R Chambers
- Sinclair, John (21995, ed.), *Collins Cobuild English Language Dictionary*. London & Glasgow: Collins [1987]
- Simpson, John A. & Weiner, Edmund S. C. (21989, eds.), *Oxford English Dictionary*. Oxford: Clarendon Press
- Summers, Della (31995, ed.) *Longman Dictionary of Contemporary English*. Harlow: Longman [21987; 11978]
- Terrell, Peter et al. (1980, eds.), *Collins/Pons German Dictionary*. London, Glasgow: Collins; Stuttgart: Ernst Klett
- Treffry, Diana (41998, ed.), *Collins English Dictionary*. Glasgow: HarperCollins
- Wahrig, Gerhard & Krämer, Hildegard & Zimmermann, Harald (1980–1984, eds.), *Brockhaus Wahrig Deutsches Wörterbuch*. Wiesbaden: F. A. Brockhaus; Stuttgart: Deutsche Verlags-Anstalt
- Wiegand, Herbert Ernst (1982, ed.), *Studien zur neuhochdeutschen Lexikographie II*. Hildesheim etc.: Olms (Germanistische Linguistik 3–6/80)
- Wiegand, Herbert Ernst & Kučera, Antonin (1982), "Brockhaus-Wahrig Deutsches Wörterbuch auf dem Prüfstand der praktischen Lexikologie". In: Wiegand (ed.), 285–373
- Zgusta, Ladislav (1971), *Manual of Lexicography*. Prague: Academia

Howard Jackson, Birmingham  
(Great Britain)

## 172. Computational linguistics

1. Computational linguistics
2. Models of morphology in computational linguistics
3. Morphology in language technology
4. Morphology learning
5. Morphology tools for linguists
6. References

### 1. Computational linguistics

It is useful to distinguish in Computational Linguistics between applications and modules. Applications are geared toward a specific user-oriented goal (e.g., automatic translation, or dialogue with an information system), whereas modules are necessary in a wide range of applications. For example, syntactic parsing as a module is useful in the development of both Machine Translation systems and Dialogue Systems. Modules can be characterised by a transformation between different linguistic representation levels, e.g. from text to speech, or from a string of words to a tree structure representing the sentence's structure. The goal of much work in Computational Linguistics is to design efficient and accurate computational models for such transformations. In developing a syntactic parser, for example, we may build a model using a grammar, a lexicon and a heuristic search procedure which together transform strings of words (the input representation) into labeled trees representing the syntactic structure (the output representation).

In this article, I will treat morphology as one such module, which can operate in two directions. In morphological analysis, a complex word form is transformed into a string of morphemes with a characterisation of the structural relations between the different morphemes (possibly represented in a tree structure), and its stem or citation form. In

morphological synthesis (or generation), a stem or root form of a morphological paradigm and a set of grammatical features is taken as input, and the corresponding complex word form is generated.

At a general level, all morphological processing modules will need a lexical database associating morphemes with linguistic information, a model of the combination possibilities of morphemes and the effects thereof on spelling and pronunciation (mostly in the form of rules, and a search procedure using this information to actually generate or analyze words.

In this article, I will briefly discuss the kind of models that have been proposed for designing the morphological module in Computational Linguistics applications (cf. 2), go into the different application areas where morphological modules are used (cf. 3), and give an overview of tools which could help morphologists in their research (cf. 5). I will also provide pointers to some recent work on “learning of morphology” which seems to attract the attention of both linguistic and computational morphology (cf. 4). There are several good introductions to the field of Computational Linguistics in general (Jurafsky & Martin 2000; Manning & Schütze 1999; Allen 1995).

### 2. Models of morphology in computational linguistics

As Computational Morphology focuses on developing models that achieve one of the morphological mappings discussed earlier (segmenting a string into its parts and disambiguating the parts in morphological analysis, construction of a string on the basis of lexical and morphological information in

morphological analysis	
<i>surprisingly</i>	(( <i>surprise</i> )Verb <i>ing</i> )Adjective <i>ly</i> )Adverb
<i>took</i>	( <i>take</i> )Verb-Past
morphological synthesis	
<i>establish</i> (past participle)	<i>established</i>
<i>clean</i> (superlative)	<i>cleanest</i>

Tab. 172.1: Morphological analysis vs. synthesis

morphological synthesis), theoretical linguistic distinctions like inflection versus derivation (cf. Art. 38) do not play an important role. Formalisms developed for modeling morphology focus on the concrete construction processes involved (concatenation, Ab-laut, root-and-template interleaving, suppletion, etc.; cf. Art. 53), on the spelling and phonological changes these processes produce (cf. Art. 35, 44), and on the constraints under which all this occurs (morphotactics; cf. Art. 42). Issues of productivity (cf. Art. 33) will be reflected in Computational Morphology by a particular choice of the division of labour between lexical storage and rule-based processing. We will only briefly describe the main types of models here. More technical and broader overviews of the field of Computational Morphology can be found elsewhere (Sproat 1992; 2000).

### 2.1. The role of lexical databases

In any Computational Morphology model, there is a trade-off between the contents of the lexical database and the size of the rule set. In languages like English with a relatively poor morphology, it is feasible to construct a lexical database with complete paradigms of complex word forms rather than morphemes (a full form lexicon). The construction can be done semi-automatically (by morphological synthesis, which is easier than analysis), leading to an analysis by synthesis approach. All inflectional and most of the derivational processes can then be solved by lexical retrieval rather than computation. However, for most languages this is not a feasible solution. For example, compounding is extremely productive in languages like German and Dutch, and for languages like Turkish and Finnish, inflection and derivation cannot be solved by this “pre-computation” and lexical storage approach. Moreover, the approach fails for words (stems) not contained in the lexicon.

For implementation, lexical databases tend to be represented as **letter tries** (Fredkin 1960). A trie is a tree data structure with nodes representing a position in a word, and arcs leaving a node representing letters that can follow the corresponding position in the word. The root node represents the start position, and individual words are paths in the trie. Leaf nodes represent the lexical information associated with the word the path of which ends at that node. This way the redundancy in the spelling of words is used to com-

press the lexicon into a data structure that is small and that can be searched efficiently.

Fig. 172.1 is an example of the trie structure representing the five words *ask asking away be became*. Note how the redundancy is removed.

### 2.2. Finite-state morphology

The most influential model developed for morphological analysis and synthesis to date is **finite-state morphology** (FSM), also called two-level morphology. It was developed by Koskenniemi (Koskenniemi 1983; 1984) for Finnish and other languages, inspired by unpublished work by Martin Kay and Ron Kaplan, published only much later (Kaplan & Kay 1994). The model consists of a trie lexicon structure and a set of rules implemented as a finite-state transducer, and can perform both analysis and synthesis.

The most important ideas in this approach were that all or most morphological phenomena can be described with regular expressions, and that the morphological mappings can be described with only two levels: a lexical representation, and a surface (spelling or speech) representation. Traditional linguistic descriptions like generative phonology (Chomsky & Halle 1968) made use of an ordered series of rewrite rules with intermediate representations to transform lexical representations into surface forms, where each rule works on the output of the previous rule, thereby giving rise to complex rule interactions. In addition, the expressive power of the formalism used to describe each rule (context-sensitive grammar) was obviously higher than needed, and in a computational perspective, this type of rule only works in one direction (from lexical to surface). A further advantage of finite-state approaches is that the computational machinery involved is language-independent, and not *ad hoc* for a specific language.

In finite-state morphology, each rule is implemented as a bi-directional Finite-State Transducer linking the lexical and surface representation, making it useful for both analysis and synthesis, and all rules are applied in parallel. The base unit of a two-level rule or constraint is a pair of symbols, one from the alphabet of the lexical representation, one from the alphabet of the surface representation. E.g. the pair *n:m* (a lexical *n* corresponding to a surface *m*), or *V:0* (a lexical vowel deleted at the surface level). See Fig. 172.2 (a), which represents an assimila-

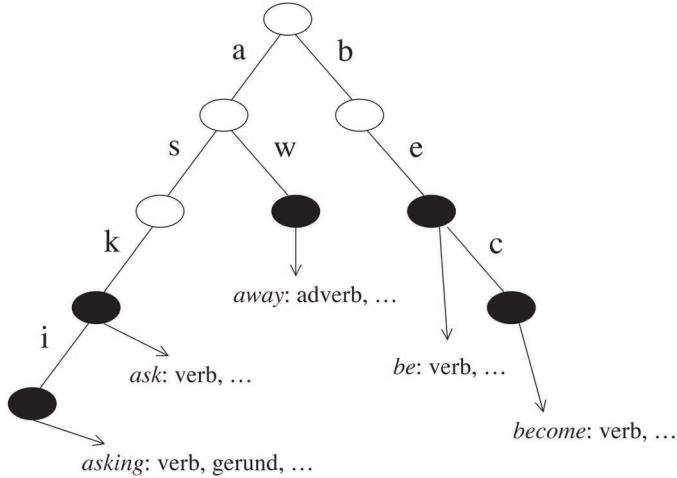


Fig. 172.1: Example of a trie data structure containing the lexical items  
ask asking away be became

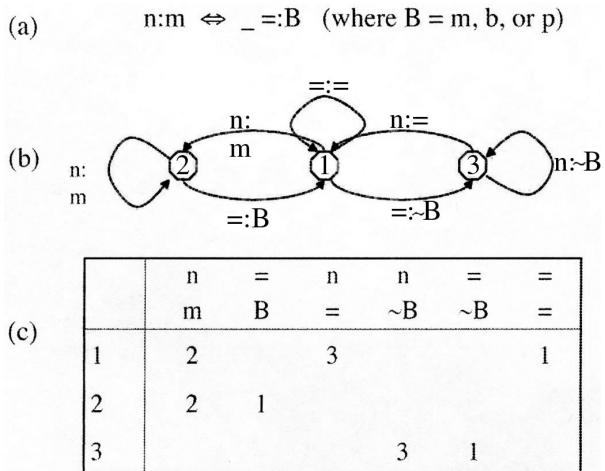


Fig. 172.2: Three representations of a finite-state morphology rule; two-level rule (a), transition network (b), and transition table (c)

tion rule: a lexical  $n$  has to be realised as a surface  $m$  when followed by a bilabial consonant ( $B = m, b$ , or  $p$ , the symbol  $=$  can refer to any symbol). Fig. 172.2 (b) and (c) show the corresponding Finite-State Transducers as a network (transitions show how states are allowed to change) and as a transition table. The automatic compilation of two-level rules into finite-state transducers is not trivial, but can be done. The strength of the approach is that two-level rules can refer to the lexical as well as to the surface representations for defining their context.

A finite-state morphology model consists of a set of rules of this type. In morphological synthesis, the rules get a series of lexical symbols as input and produce all surface forms allowed by the rules, in morphological analysis, they get a surface form as input, and produce all possible lexical representations (constrained by the lexicon). The lexicon in finite-state morphology consists of a list of stem forms, and a number of affix lexicons (lists of inflectional and derivational suffixes). Lexicon entries contain associated lexical information, including a list of pointers to contin-

uation lexica (lexica which can follow this lexicon entry).

The model has been applied to many languages and many morpho-phonological phenomena, and has proven to be an elegant and practical approach to Computational Morphology. Many overviews, tutorial material, extensions, and references are available (Gazdar 1985; Ritchie et. al. 1992; Antworth 1990). An extension for morphologically motivated phenomena like Umlaut is presented in Trost (1991). Despite its success, finite-state morphology is not without problems; the lexicon system leads to redundant representations, and some processes like reduplication and root-and-template morphology cannot be described without complicating the formalism.

The success of finite-state approaches in Computational Morphology no doubt contributed to the investigation of the applicability of finite-state methods to syntactic analysis and in applications such as shallow semantic interpretation (Abney 1996; Karttunen et al. 1996; Roche & Schabes 1997, eds.). Together with statistical approaches, finite-state methods dominate Computational Linguistics today.

### 2.3. Hierarchical lexicons

Another popular approach to computational morphology makes use of the concept of taxonomies and inheritance of properties to represent morphological knowledge. The basic insight here is that irregular words normally deviate only in a few characteristics from regular words. E.g., an irregular verb like *run* is just like a regular verb like *work* (compare *run, runs, running to work, works, working*), except that the past tense and past participle are formed in a different way. This “elsewhere condition” type of reasoning, so common in linguistic morphology, can be modeled elegantly with inheritance networks (Daelemans et al. 1992). Although there are many general-purpose knowledge representation and programming languages (of the frame-based or object-oriented type) that allow the implementation of this type of reasoning, the most important exponent of this approach is DATR (Evans & Gazdar 1996), a special-purpose programming language for lexical knowledge representation. DATR was guided in its design by formal adequacy goals (explicit declarative semantics and explicit theory of inference), and notational adequacy goals (it should be expressive enough

to describe all relevant generalizations). Another concern was efficient implementation. DATR fragments have been developed for the morphology of various languages.

Both approaches are complementary as the finite-state approach is limited in the mechanisms it allows for handling morphotactics, which is what most DATR work focuses on. Not all models can be assigned to one of these approaches; one successful morphological analysis module in the context of speech synthesis, DECOMP (Allen et al. 1987), uses ad hoc weights to improve disambiguation. I will return to the problem of morphological ambiguity resolution in the section on morphology learning (cf. 4).

### 3. Morphology in language technology

Morphological analysis is a basic component in many language technology applications. In Word Processing applications, for example, the accuracy of hyphenation, spelling checking, and grammar checking is highly dependent on the presence of some form of morphological analysis. Morphological processing can also be considered a core component in any language processing system, from speech synthesis over information retrieval to machine translation.

Depending on the sophistication of the morphological processing involved, different terms have been used in language technology. The term **stemming** or **suffix stripping** is used to refer to a sort of “poor man’s” morphological analysis involving simple rules (often without a lexicon) to reduce an inflectional form of a word to its stem. A classical example of this approach, adapted to many languages is the Porter stemmer (Porter 1980). This approach is of course only feasible with some accuracy for languages like English which have a simple morphology. The term **lemmatization** refers to a more advanced process in which a complex word form is reduced to its lemma (or citation form) and the possible morphosyntactic classes it can have as retrieved from the lexicon and deduced from the rules. In this case the structure of the word is not further analyzed. So whereas full morphological analysis would analyze *optimalizations* as “((((*optimal*)Adjective -ize) Verb -ation)Noun -s)Noun-plural”, a lemmatizer would output “*optimalizationNoun-plural*”, depending on the contents of the lexicon, and the definition of the rules, of course.

### 3.1. Word processing

**Automatic hyphenation** is the process of splitting words at the end of lines in order to minimize whitespace when right justification is used. Although the process is conventional, it is mostly based on linguistic units like syllables and morphemes. In languages like English where hyphenation is morpheme-based, morphological segmentation is required. However, given the poor morphology of English, a good computational solution is to use a set of splitting patterns found in a dictionary rather than full morphological analysis (Liang 1983). In a language like Dutch, however, where hyphenation is based on syllabification and morphological structure, the situation is much more complex. It is fairly easy to implement the basic syllabification rules for Dutch (based on the maximal onset principle and a language-specific rule avoiding syllables ending in a short vowel). It is sufficient to collect a list of possible syllable onsets. Interestingly, these rules are overridden by morphological rules. In compounds and some derivations, e.g. with the suffix *-achtig* (transl. *-ly*), the morphological boundary overrides the syllable boundary, giving rise to oppositions like *groe-nig* (*groen* + *ig*, ‘greeny’) versus *groen-achtig* (*groen* + *achtig*, ‘greeny’) where in the first case the maximal onset principle is preserved, and in the second case the morphological boundary has precedence over the syllable boundary in hyphenation. A fairly accurate morphological analysis is required to solve this problem in principle, in this case splitting patterns will necessarily be error-prone because of the high incidence of new compounds in Dutch, as can be witnessed every day in Dutch newspapers (Daelemans 1988).

**Spelling Checking** is based on a very simple principle to detect errors. Given a list of words of the language, every word encountered in a text which does not belong to the list is a spelling error. In languages with a productive morphology, this leads to an annoying “overkill”; the software continuously flags correct words as errors. Especially compounding is a problem. Extremely productive in languages like German and Dutch, this morphological process is responsible for the creation of many complex words which are used *ad hoc* in a text, and never gain enough frequency to warrant their inclusion in the word list. Morphological analysis is the only way to solve this problem. Similar arguments hold for languages with an extensive deriva-

tional or inflectional morphology like Finnish or Turkish, where storage of all word forms in a word list is impossible. Kukich (1992) describes the role of natural language processing in spelling correction.

In **Grammar Checkers**, software that checks the grammatical correctness of sentences in a text, morphological analysis is essential to be able to detect agreement errors, e.g. between subject and verb in many languages, and between modifiers and nouns in languages like German and French.

Processes very similar to morphological analysis are also necessary in word processing of languages like Chinese, and Japanese and Korean when written with Chinese characters, where word boundaries are not marked by spaces or other typographical means. A further good example is Vietnamese: it is written in Latin characters, but white space indicates syllable boundaries rather than word boundaries.

### 3.2. Module in larger systems

Morphological analysis is a necessary component in complete Natural Language Processing systems (for speech recognition and synthesis, language understanding, language generation, language translation, information retrieval, etc.), mostly as a means to increase the **lexical coverage** of such a system.

For example, a syntactic parser needs lexical information about every word in a sentence to be parsed. In case words are not in the lexicon, morphological analysis helps extract useful lexical information from these unknown word forms (often complex forms of known words), increasing the “virtual” coverage of the lexicon. In speech processing, morphological analysis is important as well. In Speech Recognition as a means to keep the recognizer lexicon small, in Speech Synthesis as a means to increase lexical coverage and solve ambiguous pronunciations like *th* in *nothing* versus *ant hill*.

With the availability of the WWW, **Information Retrieval** has become a ubiquitous technology. In search engines, keywords can be input to retrieve a number of documents containing them, ranked according to relevance according to statistical or heuristic measures. The reliability of an information retrieval engine is measured in terms of **recall** (how many of the documents relevant for my query did I get) and **precision** (how many of the documents returned by the system actu-

ally were relevant). Morphological processing is one way of increasing the recall of search engines. By expanding a keyword to all its morphologically related forms (by morphological synthesis) on the basis of its stem (found by morphological analysis), a wider range of (possibly relevant) documents is found, increasing recall.

#### 4. Morphology learning

The last decade, statistical approaches have started dominating the field of computational linguistics (Manning & Schütze 1999). The field has evolved from **deductive** to **inductive**. Recently, machine learning methods have been added to the tools of inductive computational linguistics. Machine Learning is a subfield of Artificial Intelligence concerned with the design of algorithms that learn from examples (Mitchell 1997). When used for building a linguistic model explaining some set of data, machine learning algorithms and linguists share the same task and purpose, which makes the approach potentially interesting for linguistics.

Machine Learning algorithms can be **supervised**, in which case they get examples of available input and required output, or **unsupervised**, in which case they only get examples of available input, and have to figure out useful groupings or clusters of the data. The goal of a learning approach is to use the examples to find useful generalizations about the input-output mapping to be learned. In an example of supervised machine learning applied to morphological analysis (van den Bosch & Daelemans 1999), a morphological analysis system is induced based on a large set of examples of complex words and their corresponding morphological analysis from the CELEX lexical database (Baayen et al. 1993, eds.). It can reconstruct the analyses it used for training, and apply the same systematicity to previously unseen complex words with high accuracy. The approach has been applied to English, German, and Dutch, and shows that even complex spelling changes can be handled in this classification-based way. Daelemans et al. (1997) is an example of a similar approach for morphological synthesis. In this case, the output of the learning is a rule system for diminutive formation in Dutch which is very similar to linguistic solutions proposed for the problem.

Supervised learning methods solve a problem which approaches like finite-state trans-

ducer leave basically unsolved: the disambiguation problem. Even in languages with a minimally complex morphology, a morphological analysis system can lead to many possible analyses (different segmentations into morphemes, different assignments of grammatical classes to morphemes), many of which are spurious. A supervised machine learning approach implicitly uses frequency information from the data it was trained on to make probabilistic disambiguation decisions.

Especially in **unsupervised** learning of morphology, a lot of progress has been achieved the last few years. Starting from a list of complex words, several unsupervised learning techniques have been experimented with to automatically extract information like lists of stems and affixes, and of how they can be combined (Goldsmith 2001; Kazakov & Manandhar 2001; Yarowsky & Wicentowski 2000). An exciting new approach in between supervised and unsupervised learning would start from a bilingual corpus, semi-automatically aligned. Suppose a morphological analyzer exists for German and not for Dutch. Once a bilingual aligned corpus German – Dutch is collected, the analyses generated for German can be projected to the Dutch part of the corpus and used to induce a morphological analyzer for Dutch (Yarowsky & Ngai 2001).

#### 5. Morphology tools for linguists

Some of the approaches used in Computational Morphology are at the heart of what constitutes a linguistic approach to morphology (finding the right generalizations and morpheme inventories for describing the morphology of a language, and looking for formalisms which allow describing them in sufficient detail). It is therefore not surprising that many Computational Morphology researchers have tried to build reusable tools that will be of interest to linguists as well. In this section a (probably incomplete) overview of the most important ones is given (cf. also Art. 168).

Finite-state morphology is a well-developed field which has spun off many useful tools. The Xerox Research Center in Europe (XRCE) has developed tools for morphological analysis in many languages, based on finite-state technology and is also active in research in this area. Evan Antworth of SIL

provides a useful two-level morphology software package called PC-KIMMO. Finite-state tools for inflectional morphological analysis and synthesis of English implemented using widely-available unix utilities are available from the University of Sussex (Minnen et al. 2001).

Also many implementations for DATR have been developed. Dafydd Gibbon provides an implementation called Zdatr, and the developers of DATR (Roger Evans and Gerald Gazdar) maintain a web portal with tutorial information, references, and references to implementations. Simpler application-oriented tools like Porter stemmers are widely available as well.

Many implementations of Machine Learning algorithms are available as well, but are generally not directly usable for linguistic research. Systematic evaluation is only just starting (Maxwell 2002). BOAS is a user-friendly environment for the development of morphological analyzers which makes use of Machine Learning (Oflazer et al. 2001). Interactive demonstrations of the supervised learning approach of van den Bosch & Daelemans (1999) are available as well via the demonstrations section of their website.

## 6. References

- Abney, Steven (1996), "Partial Parsing via Finite-State Cascades". In: Carroll, John & Briscoe, Ted (eds.), *Proceedings of the ESSLLI '96 Robust Parsing Workshop*. Prague: ESSLLI, 8–15
- Allen, John & Hunnicutt, M. Sharon & Klatt, Denis (1987), *From Text to Speech: The MITalk System*. Cambridge: Cambridge University Press
- Allen James (1995), *Natural Language Understanding*. Redwood City/CA: The Benjamin/Cummings Publishing Company
- Antworth, Evan L. (1990), *PC-KIMMO: A Two-Level Processor for Morphological Analysis*. Dallas/TX: Summer Institute of Linguistics (Occasional Publications in Academic Computing, 16)
- Baayen, R. H. & Piepenbrock, P. & Rijn, H. van (1993, eds.), *The CELEX Lexical Database (CD-ROM)*. Philadelphia/PA: Linguistic Data Consortium, University of Pennsylvania
- Bosch, A. van den & Daelemans, W. (1999), "Memory-Based Morphological Analysis". *Proceedings of the 37th Annual Meeting of the ACL*. San Francisco/CA: Morgan Kaufmann, 285–292
- Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York/NY: Harper & Row
- Daelemans, Walter M. P. (1988), "Automatic Hyphenation: Linguistics versus Engineering". In: Heyvaert, F. J. & Steurs, Frieda (eds.), *Worlds behind Words*. Leuven: Leuven University Press, 347–364
- Daelemans, W. & Berck, P. & Gillis, S. (1997), "Data Mining as a Method for Linguistic Analysis: Dutch Diminutives". *Folia Linguistica XXXI*.1–2, 57–75
- Daelemans, Walter & Smedt, Koenraad De & Gazdar, Gerald (1992), "Inheritance in Natural Language Processing". *Computational Linguistics* 18.2, 205–218
- Evans, Roger & Gazdar, Gerald (1996), "DATR: A Language for Lexical Knowledge Representation". *Computational Linguistics* 22.2, 167–216
- Fredkin, E. (1960), "Trie Memory". *Communications of the ACM*, 490–499
- Gazdar, Gerald (1985), Finite State Morphology. *Linguistics* 23, 597–607
- Goldsmith, John (2001), "Unsupervised Learning of the Morphology of a Natural Language". *Computational Linguistics* 27.2, 153–198
- Jurafsky, Daniel & Martin, James H. (2000), *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey: Prentice Hall, Englewood Cliffs
- Kaplan, Ron & Kay, Martin (1994), "Regular Models of Phonological Rule Systems". *Computational Linguistics* 20.3, 331–378
- Karttunen, L. & Chanod, J. & Grefenstette, G. & Schiller, A. (1996), "Regular Expressions for Language Engineering". *Natural Language Engineering* 2.4, 305–338
- Kazakov, Dimitar & Manandhar, Suresh (2001), "Unsupervised Learning of Word Segmentation Rules with Genetic Algorithms and Inductive Logic Programming". *Machine Learning* 43, 121–162
- Koskenniemi, Kimmo (1983), "Two-Level Model for Morphological Analysis". *Proceedings of the 8th International Joint Conference on Artificial Intelligence*. Los Alamos/CA: Morgan Kaufmann, 683–685
- Koskenniemi, Kimmo (1984), "A General Computational Model for Word-Form Recognition and Production". *Proceedings 10th International Conference on Computational Linguistics*. Stanford/CA: ICCL
- Kukich, Karen (1992), "Techniques for Automatically Correcting Words in Text". *ACM Computing Surveys* 24.4, 377–439

- Liang, Franklin Mark (1983), *Word Hy-phen-a-tion by Com-put-er*. Ph.D. dissertation, Stanford University
- Manning, Christopher D. & Schütze, Hinrich (1999), *Foundations of Statistical Natural Language Processing*. Cambridge/MA: MIT Press
- Maxwell, Mike (2002), "Resources for Morphology Learning and Evaluation". In: Gonzalez Rodriguez, Manuel & Suarez Araujo, Carmen Paz (eds.), *LREC 2002: Third International Conference on Language Resources and Evaluation, Vol. III*. Paris: ELRA, 967–974
- Minnen, Guido & Carroll, John & Pearce, David (2001), "Applied Morphological Processing of English". *Natural Language Engineering* 7.3, 207–223
- Mitchell, Tom M. (1997), *Machine Learning*. New York/NY: McGraw-Hill
- Oflazer, Kemal & Nirenburg, Sergei & McShan, Marjorie (2001), "Bootstrapping Morphological Analyzers by Combining Human Elicitation and Machine Learning". *Computational Linguistics* 27.1, 59–86
- Porter, M. F. (1980), "An Algorithm for Suffix Stripping". *Program* 14.3, 130–137
- Ritchie, Graeme D. & Russell, Graham J. & Black, Alan W. & Pulman, Stephen G. (1992), *Computational Morphology: Practical Mechanisms for the English Lexicon*. Cambridge/MA: MIT Press
- Roche, Emmanuel & Schabes, Yves (1997, eds.), *Finite-State Language Processing*. Cambridge/MA: MIT Press
- Sproat, Richard W. (1992), *Morphology and Computation*. Cambridge/MA: MIT Press
- Sproat, Richard (2000), *Lexical Analysis*. In: Dale, Robert & Moisl, Hermann & Somers, Harold (eds.), *Handbook of Natural Language Processing*. New York, Basel: Marcel Dekker, 37–57
- Trost, Harald (1991), "X2morph: A Morphological Component Based on Augmented Two-Level Morphology". *Proceedings of the 12<sup>th</sup> International Joint Conference on Artificial Intelligence*. San Francisco/CA: Morgan Kaufmann, 1024–1030
- Yarowsky, D. & Ngai, G. (2001), "Inducing Multilingual pos Taggers and np Bracketers via Robust Projection across Aligned Corpora". *Proceedings of NAACL-2001*. San Francisco/CA: Morgan Kaufmann, 200–207
- Yarowsky, D. & Wicentowski, R. (2000), "Minimally Supervised Morphological Analysis by Multimodal Alignment". *Proceedings of ACL-2000*. San Francisco/CA: Morgan Kaufmann, 207–216

*Walter Daelemans, Antwerpen (Belgium)*

## XXI. Morphologie und Nachbardisziplinen

### Morphology and related fields

#### 173. Namenkunde

1. Eigennamen und ihre Funktionen
2. Interne onymische Morphologie
3. Externe onymische Morphologie
4. Zitierte Literatur

##### 1. Eigennamen und ihre Funktionen

**Eigennamen** (auch *Propria*, *Onyme*) werden unter die Substantive subsumiert und erfüllen spezifische referentielle Funktionen. Im Gegensatz zu den **Appellativen** (Gattungsbezeichnungen) wie z. B. *Mensch* oder *Stadt*, die eine ganze Klasse von Gegenständen bezeichnen, referieren Eigennamen prototypischerweise nur auf ein einziges Denotat (Monoreferentialität), z. B. *Goethe* oder *Frankfurt*. Im Falle mehrfacher Referenz werden feste Erweiterungen vorgenommen: Bei Familiennamen kommt der Rufname hinzu (*Fischer* → *Joschka Fischer*), bzw. bei Rufnamen der Familienname, bei Toponymen weitere Spezifikationen (*Frankfurt (am) Main* vs. *Frankfurt (an der) Oder*). Während Appellative ihr semantisches Potential nutzen, um die Referenz zu leisten, kommt Eigennamen in der Regel keine Semantik zu (s. jedoch Christoph 1987).

Diese idealtypische Generalisierung ist bei genauerem Hinsehen einzuschränken: So können Spitznamen durchaus ein motiviertes semantisches Potential besitzen (z. B. *Schlaffi*

für einen wenig dynamischen Menschen), doch müssen sie dies nicht (man kann theoretisch auch jeden anderen Gegenstand *Schlaffi* nennen). In manchen Sprachen wie z. B. dem Deutschen müssen Rufnamen den Sexus des Referenten anzeigen (*Christian* vs. *Christiane/Christine*). Auch sog. Gattungseigennamen (*Teutoburger Wald*, *Goethestraße*) enthalten semantische Strukturen.

Zwar können Eigennamen lexikalische Rest- oder Scheinstrukturen enthalten – was aus ihrer diachronen Abkunft aus Appellativen oder aus volksetymologischen Umformungen resultiert –, doch unterstützen diese keineswegs die Identifikation: Ein Lehrer kann *Fischer* heißen und eine Großstadt *Düsseldorf*. Potentielle semantische Strukturen werden also beim Eigennamen neutralisiert. Doch kann es, gerade bei eher außergewöhnlichen Eigennamen, zu Resemantisierungen kommen, etwa wenn ein Anwalt *Mörder* heißt; doch geschieht dies seltener im umgekehrten Fall, also wenn ein Mörder den gewöhnlichen Namen *Richter* trägt. Auch in Wortspielen, Witzen und der Literatur werden transparente Strukturen (re)motiviert. In der Regel gilt jedoch: Eigennamen individualisieren, Appellative generalisieren und charakterisieren (s. Abb. 173.1).

Der gestrichelte Pfeil deutet den potentiellen Restbestand lexikalisch-semantischer

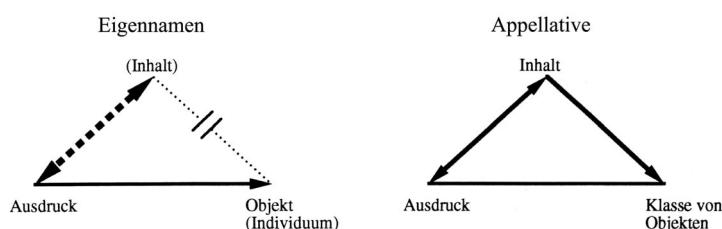


Abb. 173.1: Die unterschiedlichen Referenzleistungen von Eigennamen und Appellativen

Strukturen an, der jedoch unmotiviert bleibt und nicht zur Referenzleistung beiträgt (s. die unterbrochene gepunktete Linie). Die Identifikation des Objekts verläuft direkt vom Ausdruck zum Objekt (Direktreferenz; s. den durchgezogenen Pfeil). Damit können Eigennamen zwar über transparente, doch nicht über motivierte Strukturen verfügen (vom genuinen Gattungseigennamen abgesehen). Eigennamen benennen, identifizieren bzw. etikettieren ein Objekt. Ihre Funktion entfalten sie nicht auf der Basis semantischer Merkmale.

Monoreferenz kann auch mit anderen Mitteln hergestellt werden: Zum einen durch definite Beschreibungen mithilfe appellativischer Ausdrücke (z. B. *Deutschlands größtes Bankenzentrum für Frankfurt*), wobei diese häufig onymische Bestandteile enthalten; zum anderen durch Indikatoren, d. h. Deiktika und Proformen, die jedoch auf einen Ko- bzw. Kontext verweisen und damit einen solchen zur Voraussetzung haben (*diese Stadt hier* für ‘Frankfurt’; *er* für ‘Goethe’) (Werner 1974). Eigennamen haben diesen Verfahren gegenüber den Vorteil der Eindeutigkeit, des kurzen Ausdrucks und der Ko(n)textunabhängigkeit. Daher leisten sich alle Sprachen den “Luxus” von Eigennamen; dabei erhalten nur relevante Objekte einen Namen, in aller Regel Personen, der menschliche Lebensraum und vom Menschen verfertigte Objekte. Gemäß der Beschaffenheit des Denotats unterscheidet man zwischen Anthroponymen (Rufnamen, Familiennamen, Pseudonymen, Spitznamen etc.), Toponymen (Städte-, Straßen-, Flur-, Gewässernamen etc.), Ergonymen (Produktnamen, Bücher-, Filmtitel etc.), Praxonymen (Kriege, Abkommen etc.) und Phänonymen (Wetterhochs und -tiefs, Tai-fune etc.). Dabei nimmt die Prototypik von Eigennamen nach hinten hin ab: Konkrete, distinkte Objekte wie Personen und Örtlichkeiten bilden die Objektgruppe, die am ehesten onymisch realisiert wird. Da man immer wieder mit neuen Personen, Örtlichkeiten etc. in Kontakt kommt, ist man wie bei keiner anderen Wortart lebenslang mit dem Erwerb neuer Einheiten befaßt: Eigennamen bilden kein geschlossenes Inventar. Auch wenn sie in der Regel nicht bzw. nur rudimentär in Wörterbüchern verzeichnet werden, so gehören Eigennamen dem Sprachsystem im Allgemeinen und dem Lexikon im Besonderen an. Auf der anderen Seite gehören Eigennamen zu den Wörtern, die man am ehesten wieder

## XXI. Morphologie und Nachbardisziplinen

vergibt, vor allem dann, wenn man sie nicht häufig gebraucht.

Die Onomastik hat sich wie kaum eine andere linguistische Disziplin bisher in fast ausschließlicher Weise mit diachronen Gesichtspunkten befaßt (Fleischer 1970). In der Regel geht es um Fragen der Etymologie (s. Art. 174), da die meisten Namen appellativischen Ursprungs sind, sich jedoch im Laufe der Zeit zu mehr oder weniger opaken Ausdrücken entwickelt haben. Daß es dabei auch zur Herausbildung interner morphologischer Strukturen kommen kann, ist erst in jüngster Zeit ins Blickfeld geraten. Prinzipiell gilt es, zwei Domänen onymischer Morphologie zu unterscheiden: interne Morphologie, die die Binnenstruktur des Onyms selbst betrifft bzw. die Markierung der Proprialität (s. 2), und externe Morphologie, die die spezifisch onymische Flexion und Wortbildung zum Gegenstand hat, z. B. die Pluralbildung von Eigennamen oder die Bildung von Adjektiven aus Eigennamen (s. 3).

### 2. Interne onymische Morphologie

Da Eigennamen nur Benennungsfunktion haben, sollte erwartbar sein, daß interne Morphologie afunktional ist. Dennoch lassen sich verschiedene morphologische Strukturen erkennen, die nicht nur Relikte einstiger appellativischer Morphologie darstellen. Diese morphologischen Strukturen tragen jedoch nicht dazu bei, das Referenzobjekt zu identifizieren, sondern sie markieren die Proprialität des Wortes, also seinen Eigennamenstatus. Dabei handelt es sich gemäß Dressler (2000) um extragrammatische Morphologie. Die Erfassung interner morphologischer Strukturen ist bisher sowohl von der Onomastik als auch von der Morphologie vernachlässigt worden.

#### 2.1. Der Eigename im Spannungsfeld zwischen Motivierbarkeit, Transparenz und Opakheit

Das “Dilemma” von Eigennamen besteht darin, daß sie zwar eine grundlegend andere Funktion als Appellative ausüben, ihre diachrone Quelle jedoch in aller Regel in der Appellativik besteht. Dies schließt nicht aus, daß es auch Phantasienamen gibt (z. B. Spitznamen, Pseudonyme, Tiernamen) oder arbiträre Zahlen- und/oder Buchstabenkombinationen (z. B. Autokennzeichen, Identitätsnummern), doch ist dies ein wegen seiner ho-

hen kognitiven Kosten realiter selten verwendet. Namentyp: Eine arbiträre und womöglich lange Zahlen- oder Buchstabenfolge stößt schnell an die Grenzen der Memorierbarkeit. Der prototypische Eigenname rekrutiert sich in seiner Entstehungsphase aus motivierter appellativischer Lexik (*Eichstatt*, *Homburg* < (zer) hohen Burg, *Becker*); dabei kann es sich auch um Fremdlexik handeln (*Köln* < lat. *Colonia Agrippinensis* ‘Kolonie des Agrrippa’). Um der Verwechslungsgefahr mit einem Appellativ zu entgehen, sollte der Eigenname idealerweise möglichst weit entfernt vom Appellativ sein. Dem kommen gerade opake Namen wie *Köln*, *Tolsterglope*, *Golodkowski* nahe. Doch haben solche Namen, vor allem wenn sie mit Länge gekoppelt sind, hohe kognitive Kosten: Ein Familienname wie *Becker* ist leichter memorierbar als *Golodkowski*. Ist es ein weitgehend opaker Name, so bildet man Merkhilfen, d. h. man schafft sich sekundär transparente Strukturen.

Weder totale Opakheit noch potentielle Motivierbarkeit können als propria Ideale betrachtet werden. Üblicherweise bewegen sich die Eigennamen zwischen diesen beiden Polen, was sich wie folgt skizzieren läßt:

Man geht davon aus, daß zu germanischer Zeit die zweigliedrigen Rufnamen motiviert waren im Sinne sog. Heilwünsche, z. B. *Theoderich* bzw. *Dietrich* ‘im Volke mächtig’. Dabei herrschten semantisch, morphologisch, syllabisch und euphonisch gesteuerte Abfolgeprinzipien und Restriktionen (Seibicke 1985; Greule 1996). Eine solche motivierte Rufnamenvergabe findet sich heute z. B. in China.

Viele Eigennamen enthalten volltransparente Strukturen, ohne daß Motivierbarkeit, ein sinnvoller Bezug zum Objekt, möglich wäre: Familiennamen wie *Beckenbauer*, *Neumann*, ebenso Städtenamen wie *Mannheim*, *Würzburg* bestehen aus bekannten lexikalischen Bausteinen, ohne damit eine Identifikation des Referenzobjekts vollziehen zu können. Das schwedische Familiennamensystem hat diese Strategie zum Prinzip erhoben: Namentkomposita wie *Sjöberg* ‘Seeberg’, *Sjögren* ‘Seeast’, *Lindström* ‘Lindenstrom’ etc. gehören nicht nur zu den häufigsten Namentypen, sondern sie werden bei dem in Schweden einfachen Familiennamenwechsel sogar explizit empfohlen (Nübling 1997).

Partielle Transparenz besteht, wenn appellativische Bestandteile leichte formale Veränderungen erfahren: *Schmitt/Schmidt/Schmitz*

zu *Schmied*, *Eichstatt* zu *Eiche* und *Stadt/Stätte*. Hierzu gehören auch Abweichungen in der Akzentposition (*Ebersfelde*), Apokopen (*Ulm*), Erweiterungen (*Schilfa*, *Oberhausen*), Abweichungen in der Graphie (<*Duisburg*, *Becker*, *Mueller*>) etc. Solche formal leicht dissozierten und damit noch identifizierbaren Morphe bezeichnet man als Pseudomorphe (Nübling 2000: 130) oder Paläomorphe (Tournier 1985: 87 f.).

Auch Semitransparenz bildet eine vielgenutzte Kompromißlösung zwischen motivierten und opaken Strukturen: Hier ist mindestens ein Glied opak und ein Glied transparent (meist das letzte): *Klinsmann*, *Kullmann* bzw. *Hamburg*, *Homburg* etc.

Vor dem Endstadium totaler Opakheit sind solche Eigennamen anzusiedeln, die noch morphologische Strukturen erkennen lassen, doch deren Einzelmorphe opak sind. So interpretiert man das Toponym *Tolsterglope* als Kompositum (als *Tolster-* + *-glope*), ohne daß der Sprecher einen dieser Teile an bekannte Lexik oder Onymik anschließen könnte.

Eigennamen sind morphologisch erstarrt. Morphologische Operationen, wie sie bei Appellativkomposita möglich sind, sind bei Onymen nicht mehr gegeben, wie etwa Koordinierbarkeit: \**Morgen fahre ich erst nach Ham- und dann nach Homburg*. Das Erstglied solcher zweigliedrigen Komposita steht in der Regel der Appellativik ferner und weist eine höhere morphologische Bandbreite auf, während das Letzтglied eher an die Appellativik anschließt, einem begrenzteren Inventar angehört und eher Rückschlüsse auf den Eigennamentyp zuläßt. Ähnlich wie in der Wortbildung scheint es auch hier das Letzтglied zu sein, das den kategorialen Status markiert. Häufig sind gerade bei deutschen Toponymen Erweiterungen zur Dreigliedrigkeit, wobei dann meist ein transparentes, oft sogar motiviertes Erstglied präponiert wird: *Oberjosbach* – *Niederjosbach*. Prinzipiell scheint im Deutschen die Peripherie eines morphologisch komplexen Toponyms eher der Appellativik nahezustehen, während das Zentrum (-*jos-*) opak ist.

Auf einer Skala von (potentieller) Motivierbarkeit (*Koch*, *Neustadt*) bis hin zu totaler Opakheit (*Stratz*, *Tolsterglope*) sind es diese drei dazwischenliegenden Verfahren der vollen, der partiellen und der Semitransparenz, die am meisten genutzt werden. Hierdurch entgehen die Eigennamen potentieller

Verwechslungsgefahr mit den Appellativen, nutzen jedoch gleichzeitig deren Bekanntheit, was der Memorierbarkeit zugute kommt.

## 2.2. Integration proprialen Materials

Eine häufig genutzte Technik der Eigennamenbildung besteht in der Nutzung bereits bestehender Eigennamen: Toponyme verwenden anderweitiges toponymisches oder anthroponymisches Material; entsprechend nutzen auch Anthroponyme anderes anthroponymisches oder toponymisches Material: *Rhein* → *Rheinhausen, Rheinfelden* etc.; *Ludwig* → *Ludwigsburg, Ludwigshafen*; *Peter* → *Petersen, Peters*; *Hessen* → *Hesse, Hess* etc. (s. 3.2). Bisher wurde noch nicht hinreichend untersucht, in welche Richtung diese onymischen Wander- und Entlehnungsbewegungen verlaufen.

## 2.3. Spezifisch onymische Morphe

Von besonderem Interesse sind **spezifisch onymische** (auch: propriae oder onomastische) **Morphe**. Diese haben keinen synchronen Bezug zur Appellativik. Indem sie ausschließlich Propriätät markieren, erfüllen sie extrem kategoriale Funktionen. Daher werden diese Verfahren der internen Morphologie zugerechnet; allerdings ergeben sich hier Übergänge zur externen onymischen Morphologie (zu den Techniken s. 3.3).

In manchen Sprachen sind Rufnamen produktiv aus Appellativen ableitbar, d. h. sie bilden kein festes Inventar (Onomastikon), aus dem geschöpft wird. Im Kinyarwanda (s. Art. 141), einer Klassensprache, besteht ein Appellativ üblicherweise aus einem Präfix, einem Präfix und einem Stamm, während der Eigename durch die Abwesenheit des Präfixes gekennzeichnet wird: *umuhuungu* bedeutet 'Junge', segmentierbar als *u-mu-huungu*; *Muhuungu* dagegen ist ein Eigenname; ebenso *amabuye* 'Steine' (*a-ma-buye*) vs. *Mabuye* als Eigenname. Eine umgekehrte Markierung findet sich in einer anderen Klassensprache, dem Zulu: Hier erhält der Eigenname ein sekundäres Präfix, *u-*, wobei der anlautende Vokal des ursprünglichen (appellativen) Suffixes schwindet: *intombi* 'Mädchen' vs. *uNtombi* als Eigenname (Kuhn & Serzisko 1982).

Onymische Affixe zeigen nicht nur Propriätät an, sondern sie können auch Auskunft über die Art des Eigennamens und damit – indirekt – des Referenzobjekts erteilen. Die polnischen Familiennamen machen besonders intensiven Gebrauch von onymi-

schen Suffixen: Zu den häufigsten zählen *-ska* (für weibliche Träger) und *-ski* (für männliche Träger), die ursprünglich aus Beinamen von Adligen hervorgingen und deren Besitz anzeigen (z. B. *Tarnów* → *Tarnowska, Tarnowski*). Damit hatten sie ein Toponym als Basis. Später breitete sich dieses Suffix auf Toponyme aus, die die Herkunft (nicht mehr den Besitz) der betreffenden Person bezeichneten, und schließlich entwickelte es sich zum häufigsten aller polnischen Familiennamensuffixe, indem es sich ab dem 16. Jh. auch mit Rufnamen und Appellativen verband, z. B. *Wiśniewski* aus *wiśnia* 'Sauerkirsche' (Szczepaniak 2002).

Auch das Schwedische kultiviert mit *-son* ein Familiennamensuffix, dessen appellative Herkunft zwar noch ersichtlich ist, doch das sich formal davon dissoziiert hat (keine Homophonie mehr mit dem Simplex *son* 'Sohn'). Suffigiertes *-son* in *Gustavsson* bedeutete ursprünglich 'Sohn von' und ist als Onym längst auch auf Frauen beziehbar. Mit dieser Desemantisierung ist es zu einem Indikator für Familiennamen geworden. In jüngster Zeit wird es sogar produktiv gemacht, indem es auch in Verbindung ganz anderer Erstglieder auftreten kann (*Lindson, Balson*).

Das Deutsche liefert gerade bei den Toponymen viele Beispiele für onymische Morphe: Bei Eigennamen auf *-ach, -itz, -a, -ow* [o:], *-wang(en), -holt*, in der Schweiz *-ikon*, kann es sich nur um Siedlungsnamen handeln (allenfalls um sekundäre Familiennamen). Das bekannteste toponymische Suffix dürfte *-ing(en)* sein (*Tübingen, Gundelfingen, Freising*), das sich aus einem Zugehörigkeitssuffix entwickelt hat und ursprünglich mit (männlichen) Rufnamen verband: *Gundolf + -ingen* (> *Gundelfingen*) bedeutete ursprünglich 'die Leute/die Siedlung des Gundolf'.

Daß onymische Suffixe morphologischen Status besitzen, erweist sich bei volksetymologischen Prozessen: *Rellingen* ist eine sekundäre Neubildung aus *in Reinlage, Ording* aus (*in*) *Urden, Heida* und *Wiesa* (aus *Heide, Wiese*) wurden sekundär neugebildet nach Toponymen wie *Vechta* (s. 2.4).

Bei der Entlehnung und Integration von Toponymen aus fremdsprachlichem Material ins Deutsche sind auffällige Strukturbildungen beobachtet worden: Sowohl bei Entlehnungen aus dem (Gallo-) Romanischen am Südwestrand Deutschlands als auch aus dem Slawischen im Osten Deutschlands kommt es zu strukturellen Übereinstimmungen, besonders bzgl. des Letztglieds: So enden viele To-

ponyme trotz unterschiedlicher Etyma auf *-atsch* (z. B. *Fillatsch* im Westen, *Trebatsch* im Osten), *-itsch*, *-an*, *-aun*, *-un* etc. Dies läßt auf eine gewisse “Gestaltung des onymischen Systems des Deutschen” (Eichler 2001:165) schließen. Auffällig ist auch, daß sich viele Ortsnamen trotz ihrer fremdsprachlichen Herkunft in zwei Glieder segmentieren lassen, d.h. daß sekundär pseudomorphologische Strukturen mit meist bekannten onymischen Morphemen entstehen (z. B. *-en* und *-in* wie in *Wilthen* und *Berlin*).

#### 2.4. Sekundäre Herstellung von Transparenz und onymischer Morphologie: Volksetymologie

In ihrer theoretischen Tragweite unterschätzt wurden bisher die gerade im onymischen Bereich ausgesprochen häufig auftretenden Prozesse sog. Volksetymologie. Unzutreffend ist der alternative Terminus der *sekundären Motivierung*. Wenn – meist opake – Eigennamen volksetymologisch umgeformt werden, so entstehen dabei keineswegs adäquatere Beschreibungen des Referenzobjekts, sondern es entstehen nur Strukturen, die an Bekanntes anschließen und – hierin dürfte der Sinn liegen – die Memorierbarkeit erleichtern. Beispiele aus der deutschen Toponymik sind *Hiddensee* < dän. *Hiddens-o*, eigentlich ‘Hiddens-Insel’, *Sauerland* < *Suderland*, *Venusberg* < *Vennsberg*, *Meerholz* < *Mērold* etc., aus der Anthroponymik *Augstein* < *Augustin*, *Wohlrabe* < *Wallraff* etc. (Bach 1953; Leys 1966; Olschansky 1996; Vennemann 1999). In keinem der (beliebig zu vermehrenden) Beispiele entsteht erhöhte referenzsemantische Adäquatheit im Sinne motivierter Strukturen, sondern sie verharren auf der Stufe der formalen Transparenz. Koch (1963) geht so weit, gerade den Mangel an Sinn als bestes Indiz volksetymologischer Umbildung zu bewerten: “Eben das Widersinnige und Ungeheimte eines ON. [Ortsnamens] ist das beste Kennzeichen der Volksetymologie. Deren Wesen muß also woanders liegen als beim Sinngeben oder Sinngebenwollen” (Koch 1963:165). Auch er sieht die Funktion in mnemotechnischer Erleichterung. Zu diesem Komplex besteht noch großer Forschungsbedarf. Stichprobenhafte Untersuchungen zeigen, daß die opake Vorlage meist mehrere Silben umfaßt (was die Anforderungen an die Memorierbarkeit und damit den Bedarf an Herstellung vermehrter Transparenz erhöht), und daß es meist die Silbengrenzen sind, die

die späteren (sekundären) morphologischen Zäsuren vorgeben.

Wie das Beispiel *Venusberg* < *Vennsberg* zeigt, können auch sekundäre onymische Strukturen entstehen. Damit werden sämtliche Verfahren der Markierung von Propriät durch volksetymologische Umformungen produziert. Dies läßt den Schluß zu, daß es sich um funktionale Strukturen handelt.

### 3. Externe onymische Morphologie

Eigennamen können Basis für unterschiedliche morphologische Prozesse sein. Solche Prozesse können (1) Flexionsformen ein und desselben Namens (*Goethe-s* Genitiv Singular) hervorbringen (s. 3.1) oder durch die Wortbildungstechniken der Derivation (Personename *Leverkus* → Ortsname *Leverkus-en*) und Komposition (*Rhein-hausen*) zu (2) Eigennamen einer andern Kategorie (tsch. *Hrabal-ová* movierter Familienname) und (3) Appellativen einer gleichen oder andern Wortart (*Benz-in*, *boykott-ieren*) führen (s. 3.2). Ein Sonderfall von (1) sind onymische Pluraliatantum (*die Azoren*), aus denen kein Singular zurückgebildet werden kann (\**eine Azore*). Lexikalische Techniken der Bedeutungs-Konversion wie bei *Pegnitz* (Flußname → Ortsname) zu (2) oder *Manchester* (Ortsname → Stoffbezeichnung) zu (3) werden ebenfalls berücksichtigt, auch wenn hier keine explizit morphologischen Mittel eingesetzt werden. Für die explizite deonymische Morphologie spielen solche Vorgänge eine Rolle, wenn die zu Appellativen umgedeuteten Eigennamen Basen morphologischer Prozesse sind: Personename *Krösus* → Appellativ *Krösus* ‘reicher Mann’ (Singular) → *Krö-suss-e* (Plural). Ob “Personengruppennamen” wirklich onymisch sind, ist umstritten (s. 1). Der propriale Status von Ethnonymen (*Kamerun-er* ‘Bewohner des Landes *Kamerun*’) ist fragwürdiger als der von Kollektivbezeichnungen für die ‘Mitglieder einer Familie’ ((*die*) *Buddenbrook-s*).

#### 3.1. Flexion

Eigennamen teilen als Unterklasse der Substantive deren grammatische Kategorien. Invariant ist die Person (immer 3.), variant sind Numerus, Kasus und Genus. Plural und oblique Kasus werden in der Regel morphologisch overt symbolisiert. Genus ist inhärentes Merkmal, kann aber ausdrucksseitig z. B. in typischen Endungen reflektiert sein (*Michael-*

*a* ‘feminin’ zu *Michael*; *Claudi-us/Claudi-a* ‘maskulin/feminin’). Morphosyntaktische Auswirkungen haben die namensubstantivischen grammatischen Kategorien insofern, als sie die Deklination ihrer Begleiter bestimmen (*ewig-es Rom*).

Numerus als eine stark semantische Flexionskategorie zum Ausdruck natürlicher Ein- bzw. Mehrzahligkeit (s. Art. 100) kollidiert bei Eigennamen in pluralischer Form mit dem für Propria konstitutiven Faktum der Monoreferentialität. Dennoch bleibt ein Plural wie (*die*) *Mann-s* (zum Familiennamen *Mann*) ebenso Eigename (‘Kollektiv, das von den zu dieser Familie *Mann* gehörenden Personen gebildet wird’) wie ein geographisches Namen-Pluraletantum *die Azor-en*. In Redeweisen wie *die zwei Deutschlands* (BRD, DDR) dagegen handelt es sich um appellativischen Gebrauch.

Im Deutschen kommt Kasus von Eigennamen fast nur in Form des attributiven Genitivs vor (*Goethe-s; Polen-s*). Vom Verb regierte syntaktische Kasus sind anders als etwa im Polnischen (*Golęb-owil-a* Dativ/Akkusativ Singular) archaisch (*Goethe-n* Dativ/Akkusativ Singular). Es gibt die Tendenz, Eigennamen möglichst in ihrer Nennform und dadurch unflektiert zu gebrauchen, d. h. – den Kasus betreffend – den Nominativ auch da stehen zu lassen, wo syntaktisch ein anderer Kasus gefordert wäre: (*in einem Bericht*) *des “Neuer Tag”* statt *des “Neuen Tags”*.

Nicht eindeutig der Kasus- oder Numerusflexion zuzuordnen sind wegen der Homonymie der Suffixe die Bildungen auf *-s* vom Typ *Fischer-s Fritz* ‘Fritz Fischer’. Die Paraphrasen ‘Fritz aus der Familie der *Fischer-s*’ oder umgangssprachlich ‘den *Fischer-s* ihr Fritz’ sprechen für pluralisches *-s*, die syntaktische Konstruktion und die Paraphrase ‘(Vater) *Fischer-s* Sohn Fritz’ dagegen für genitivisches *-s*. Möglicherweise ist aber dieses Suffix bei Eigennamen weder numeralisch noch kasuell, sondern konstituiert eine eigene (Misch-)Kategorie ‘kollektiv-genealogisch’ o. ä., wobei ‘kollektiv’ nur Affinität zu ‘pluralisch’ und ‘genealogisch’ ebenfalls nur Affinität zu ‘genitivisch’ (‘herkunftsmäßig’) aufweist.

Eigennamen mit flexivischer Kennzeichnung weichen gegenüber der Flexion von Appellativen oft ab. Das wird besonders deutlich, wenn es zu Eigennamen homonyme Appellative gibt (von denen jene z. T. herkommen): vgl. proprial/appellativisch *die Mann-s / die Männ-er; die Herz-en-s* (mit phonotaktisch bedingtem *-en-* zwischen den Sibilan-

ten) / *die Herz-en; des Karl Rabe / des Rabe-n*. Zusätzlich fällt die allomorphe Armut der Namenflexive gegenüber den appellativischen Flexiven auf: proprial *die Mann(-s), Rabe(-s), Zahn(-s)* vs. appellativisch *die Männ-er, Rabe-n, Zähn-e*.

Das *s*-Suffix ist hier wie in andern Spezialbereichen des Wortschatzes, wo die formale Identität des Lexems möglichst unangetastet bleiben soll (etwa bei Fremdwörtern), das probate Mittel. Es erhöht die Silbenzahl nicht, führt nicht zu möglichen Hiatusproblemen zwischen Stammauslaut und Suffix, ist nicht mit konkomitanten Änderungen des Basismorphems verbunden (wie etwa *-er* mit Umlaut des Stammvokals bei *Männ-er*) usw. Entsprechenden Appellativen nachgebildete Plurale wie *die Liebermänn-er* sind spielerischer Natur.

Daran, daß pluralische Formen wie *Mann-s (kommen)* oder genitivische Formen wie (*Thomas*) *Mann-s (Werk)* artikellost gebraucht werden können bzw. müssen und bei Artikelgebrauch ein nichtsymbolisierter Plural *die Mann-Ø (kommen)* möglich bzw. ein nichtsymbolisierter Genitiv *des Thomas Mann-Ø* zwingend ist, sieht man, daß im Deutschen bei Phrasen mit Namen ebenso ‘Monoflexion’ (Admoni 1982) anzutreffen ist wie bei Phrasen aus Artikel und Adjektiv (*d-er groß-e* vs. (*ein-Ø*) *groß-er*).

Genus (s. Art. 98) kann ausdrucksseitig mit (Pseudo-)Suffixen korrelieren, die zu bestimmten, vom Genus mitkonstituierten flexivischen Klassen passen, z. B. *Ann-e* ‘feminin’ (vgl. appellativisch *Nicht-e, Kann-e* ‘feminin’; s. 2 zur ‘internen’ onymischen Morphologie). Genus ist ausdrucksseitig auch mit derivativischen Suffixen verbunden, die Genus determinieren wie bei *Lott-chen/Fritz-chen* ‘neutrumb’ (vgl. appellativisch *Töchter-chen/Söhn-chen, Lämp-chen/Stühl-chen* ‘neutrumb’) oder Sexus symbolisieren und mit dem passenden Genus einhergehen wie bei tsch. *Hrabal-ová*, dt. archaisch (*Luise*) *Müller-in*, dt. *Michael-a*, alle ‘feminin’ (s. 3.2 zur ‘externen’ onymischen Wortbildungsmorphologie).

### 3.2. Wortbildung

Eigennamen unterliegen als Basen für Wortbildungsprozesse keinen prinzipiell anderen Beschränkungen als die appellativischen Substantive, d. h. sie können explizit und implizit abgeleitet sowie zu Kompositionen verwendet werden. Die Wortbildungsprodukte können wiederum Eigennamen sein oder Appellative

	Derivation	Konversion	Komposition
Eigennamen	(Hermann) Bayreuth-er Personenname	Bayreuth Eigenname Synonym für die Wagner-Festspiele	Neu-Bayreuth Eigenname Synonym für die Nachkriegs-Festspiele
Appellative	(allelein) Bayreuth-er Einwohnerbezeichnung  Bayreuth-er (Studenten) bayreuth-ische (Lande) Relationsadjektive	(es gibt inzwischen viele) Bayreuth-s 'Festspiele mit Merkmalen, die ursprünglich nur die Bayreuther Festspiele aufwiesen'	Bayreuth-Besucher 'Besucher der Stadt oder der Festspiele'

Tab. 173.1: Der Ortsname *Bayreuth* als Basis für die Bildung von anderen Eigennamen und Appellativen

werden. Beispielbasis für die folgenden deonymischen Bildungen ist der Ortsname *Bayreuth*:

Derivationen von Eigennamen aus Eigennamen dienen u. a. den Zwecken der Movie rung (tsch. *Hrabal-ová*), der Personennamenbildung aus Ortsnamen (*Oppenheim-er*), der Ortsnamenbildung aus Personennamen (*Leverkus-en*), der Namen-Diminution (Kosenname *Ton-i* aus *Anton*; *Änn-chen* aus *Ann-e*). Die derivativen Mittel zur deonymischen Bildung von neuen Eigennamen sind spezifisch proprial bei *Hrabal-ová*, *Ton-i*, *Leverkus-en*, homomorph mit solchen zur Bildung von appellativischen Substantiven bei *Oppenheim-er*, *Änn-chen* (vgl. *Städt-er*, *Töchterchen*). Derivierte Eigennamen können zu Appellativen konvertiert werden (*Hein-i* 'Klein Heinrich' → 'Mensch mit seltsamem Ver halten').

Ziel der meisten deonymischen Derivationen (s. Art. 89) sind Appellative. Eine große Gruppe machen hier die Einwohnerbezeichnungen aus. Im paradigmatischen Verhältnis der Ableitungen zu ihrer toponymischen Basis lässt sich auch deren mögliche morphologische Binnengliederung erkennen.

- |                  |                         |
|------------------|-------------------------|
| (1) Ortsname     | → Einwohnerbezeichnung  |
| <i>Bayreuth</i>  | → <i>Bayreuth-er</i>    |
| <i>Veron-a</i>   | → <i>Veron-es-e</i>     |
| <i>Damask-us</i> | → <i>Damasz-en-e(r)</i> |
| Ländername       | → Einwohnerbezeichnung  |
| <i>Schweiz</i>   | → <i>Schweiz-er</i>     |
| <i>Israel</i>    | → <i>Israel-i(-s)</i>   |
| <i>Chin-a</i>    | → <i>Chin-es-e(-n)</i>  |
| <i>Marokk-o</i>  | → <i>Marokk-an-er</i>   |

Einwohnerbezeichnungen müssen nicht de (top)onymisch gebildet sein. Sie können auch

primär vorliegen und ihrerseits Basis für die Ableitung von Orts- und Ländernamen sein: Einwohnerbezeichnung [*bei den*] *Franke-n* (Dativ Plural) → Orts-/Ländername *Frank-en* (vgl. auch den Typ [*der*] *Grieche-n Land* → *Griechen-land*). Die hier angezeigten Ableitungsrichtungen entsprechen der diachronischen Entwicklung, sind aber auch synchronisch gerechtfertigt, wenn man die längere Form als Wortbildungsprodukt annimmt, die kürzere als Wortbildungsbasis. Eine rein auf Proportionalanalogie begründete Morphologie sieht aber auch synchronisch rückbildungende deonymische Ableitungen kürzerer Ethnonyme aus längeren Ländernamen vor:

- |                         |                      |
|-------------------------|----------------------|
| (2) <i>Afghan-istan</i> | : <i>Afghan-en</i> = |
| <i>Usbek-istan</i>      | : x                  |
|                         | x = <i>Usbek-en</i>  |
| <i>Finn-land</i>        | : <i>Finn-en</i> =   |
| <i>Lett-land</i>        | : x                  |
|                         | x = <i>Lett-en</i>   |

Eine große, kaum blockierte Gruppe deonymisch abgeleiteter Appellative sind die einwohnerbezeichnenden Substantive (*Bayreuth-er* Singular/Plural, *Chin-es-e(n)*) und die aus den Ländernamen abgeleiteten Adjektive (*Bayreuth-er/bayreuth-isch, chin-es-isch*). Die Beziehung zwischen adjektivischem und (letztem) ethnonymschem Suffix ist hochgradig regulär (-er, -e, -i, wird durch -isch ersetzt vice versa): s. zu obigen Einwohnerbezeichnungen die Adjektive *bayreuth-isch, veron-es-isch, damasz-en-isch, chin-es-isch, marokk-an-isch, israel-isch* (*schweiz-er-isch* ist eine der wenigen Ausnahmen). Die Basen für die Ableitung der Bewohnerbezeichnungen und der Adjektive aus Toponymen sind also in der Regel identisch und können intern noch ein-

mal aus toponymischem Stamm und einer Endung bestehen: Der intern strukturierte Derivationsstamm *chin-es* (zu *Chin-a*) ist Basis für *Chin-es-e* und *chin-es-isch* (Fuhrhop 1998: 141 ff.). Allomorphie des toponymischen Stamms kommt vor (*Damask-* vs. *damasz-*).

Eine andere Gruppe von deonymischen Derivaten bilden aus Personennamen abgeleitete appellativische Substantive wie *Benz-in* (zu *Benz*) oder Verben wie *mendel-n* (auf Basis der Grundform *Mendel*), *röntg-en* (auf Basis des Stamms *Röntg* von *Röntg-en*). Detponymisch ist etwa *finnland-is-ier-en*.

Konversionen (s. Art. 90) aus Eigennamen können intra-onymisch vor sich gehen, so wenn aus Personennamen Firmen- und Markennamen gebildet werden (*Ford*®), aus Ortsnamen Vereinsnamen (*Liverpool*), aus Flussnamen Ortsnamen (*Pegnitz*), und sie können zu Appellativen werden wie der Markenname *Uhu*® für ‘Flüssigklebstoff’. Oft ist schwer zu entscheiden, ob ein Name direkt konvertiert wurde oder das Ergebnis einer Rückbildung aus Komposita mit diesem Namen ist: *Diesel* für (und aus?) appellativisch *Diesel-motor*, *Diesel-kraftstoff*.

Auch Komposition (s. Art. 87) unter Beteiligung von Eigennamen kann entweder neue Eigennamen hervorbringen oder Appellative. Beispiele für Komposita aus mehreren Namen sind kopulative Bildungen wie *Leuthäuser-Schnarrenberger* (Familienname + Familienname), *Castrop-Rauxel* (Ortsname + Ortsname), *Frankenpfalz* (Ländername + Ländername; oberfränkisch-oberpfälzisches Grenzgebiet) und determinative Bildungen wie *Karl-Rüdiger* (in einer Reihe mit *Wolf-/Heinz-Rüdiger* usw.) zur progressiven Differenzierung von *Rüdiger* oder wie *Müller-Lüdenscheid* (Personenname + Ortsname) oder *Baden-Baden* (Ortsname + Ländername) zur regressiven Differenzierung der vielen Personen namens *Müller* bzw. mehrerer Orte namens *Baden*. Beispiele für Komposition unter Beteiligung von Appellativen sind determinative Bildungen wie *Ober-/Unterammergau* (Adjektiv + Ortsname → Ortsname), *Klein-hans* (Adjektiv + Familienname → Familienname), *Karl-s-ruhe* (Personenname + Fuge + Substantiv → Ortsname) zur progressiven Differenzierung eines Eigennamens (*Klein-hans*) bzw. eines Appellativen durch einen Eigennamen (*Karl-s-ruhe*) oder Bildungen wie *Davos-Dorf*, *Davos-Platz* zur regressiven Differenzierung. Aus ursprünglich regressiv-determinierend erscheinenden Reihenbildungen

wie *Hans-dieter*, *Hans-veit*, *Hans-jürgen* entwickelt sich eine Art Halbpräfix *Hans-*. Dreigliedrige weibliche Familiennamenkomposita wie *Greiner-Petter-Memm* sind kopulativ aus den unmittelbaren Konstituenten *Greiner-Petter* und *Memm* aufgebaut, wobei *Greiner-Petter* selber eine Struktur aufweist, in der *Petter* den im Thüringischen häufigen Familiennamen *Greiner* regressiv determiniert. Beispiele für Komposition unter Eigennamenbeteiligung, die Appellative hervorbringt, sind etwa *Otto-Motor* mit der Struktur Personename + Substantiv oder *Grüß-August* mit der Struktur Verb + Personename, wobei *August* schon vorher zum Appellativ konvertiert gewesen sein kann (Vorname → ‘unbeholfener Mensch’).

### 3.3. Morphologische Typen deonymischer Prozesse

Im Prinzip unterscheiden sich deonymische Prozesse morphologisch nicht von deappellativischen. Hier wie dort sind davon die Stammposition (*Mailand/Mailänd-er* wie *Land/Länd-er*) und die Suffixposition betroffen, an der es additive, modulatorische und – mit den selben Einschränkungen wie in der appellativischen Morphologie – subtraktive Prozesse gibt.

Was den Sprachbau betrifft, sind die additiven und subtraktiven Prozesse dem agglutinierenden Typ und dem Grundform-Prinzip zuzuordnen, die modulatorischen dem alternierenden Typ und dem Stamm-Prinzip (Harnisch 2001). Die in Tab. 173.2 zusammengestellten Prozesse operieren mit morphologischen Einheiten, wenn auch z. B. das *Jo-* von *Joseph* allenfalls sekundär morphologisch interpretiert sein kann: s. die scherhaft proportionanalalogische Reihe *Sepp/Jo-seph*, *Hansl/Jo-hann*, *Kurt/x*; *x* = *Jo-kurt*. Bei *-seph/Sepp* oder *Urs-/Usch-* (über *Ursch* mit Sibilantenassimilation) liegt Stammallomorphie vor.

Andere Prozesse an Eigennamen operieren mit rein lautlichen Mitteln. Die lautliche Entstellung der Urformen von Eigennamen ist ein automatischer Prozeß, der mit den semantischen Besonderheiten und den besonderen Gebrauchsbedingungen der Eigennamen zu tun hat (s. 1). Motiviert dagegen sind lautliche Manipulationen an Eigennamen, wie sie v. a. bei der Herstellung von Koseformen zu Personennamen vorgenommen werden. Neben den Vorteil, den gekürzte Formen bei häufigem Gebrauch im kommunikativen Nahbereich haben, tritt das Prinzip der laut-

additiv	modulatorisch	subtraktiv
<i>Hans</i>	<i>Ägypt-en</i>	<i>Mark-us</i>
<i>Hans-i</i>	<i>Ägypt-er</i>	<i>Marc</i>
additiv-modulatorisch	additiv-subtraktiv	modulatorisch-subtraktiv
<i>Chin-a</i>	<i>Jo-seph</i>	<i>Urs-ul-a</i>
<i>Chin-es-e</i>	<i>Sepp-i</i>	<i>Usch-i</i>

Tab. 173.2: Suffixmorphologische Prozeßtypologie

lichen Verringerung als Ikon der semantischen Verniedlichung. Das hindert nicht, daß morphologische Substanz, die von der Lautsymbolik her geeignet ist, ‘Kleinheit’ anzudeuten, wie das *-i*, wieder hinzutritt: *Joseph* → *Sepp* → *Sepp-i*. Ideale Grundeinheit rein langer Verkürzungen ohne Rücksicht auf eine eventuell vorliegende interne morphologische Struktur sind offene Silben, die entweder ohnehin vorliegen wie bei *Fe.li.ci.tas* → *Fe.li* (Silbentilgung), *Lie.se.lo.tte* → *Li.lo* (Wahl auseinanderliegender betonter Silben), *E.li.sa.beth* → *Li.li* (Reduplikation) oder erst hergestellt werden wie bei *Hans-Joachim* → \**Hansjo* → *Hajo*). Offensichtlich werden Zweisilbler mit vokalischem Auslaut angestrebt, s. auch *Fe.li.ci.tas* → *Zi.ta*, wo das *s* lautlich und nicht als \*-s-Suffix getilgt wird. Wo ein typischer Namenausgang durch die Kürzung nicht schon angeboten wird wie bei *Rü.di.ger* → *Rü.di*, wird er hergestellt: *Ru.dolf* → *Ru.di*.

Diese vokalischen Wortausgänge werden dann morphologisch reanalyisiert und sekundär motiviert (s. 2.4), hier als typisch weibliches Vornamenssuffix *-a* (*Zita* → *Zit-a* wie *Paul-a* zu *Paul*), als typisch diminuierendes Kosesuffix *-i* (*Feli/Rudi* → *Fel-i/Rud-i* wie *Hans-i* zu *Hans*), als untypisch weibliches oder typisch männliches, aber immer reihenbildendes Suffix *-o* (*Lilo* → *Lil-o* in der Reihe [*Ca.ro.lin* →] *Caro* → *Car-o* usw.; *Hajo* → *Haj-o* in der Reihe *Brun-o*, *Ud-o*, usw.). Daß solche (pseudo)suffixischen Muster wahrgenommen werden, zeigen Namenmoden wie die gleichzeitige Beliebtheit von Namen auf *-ian* (*Christ-ian*, *Sebast-ian*, *Flor-ian* usw.).

#### 4. Zitierte Literatur

- Admoni, Wladimir (1982), *Der deutsche Sprachbau*. München: Beck  
 Bach, Adolf (1953), *Deutsche Namenkunde*. Heidelberg: Winter

Christoph, Ernst-Michael (1987), “Studien zur Semantik von Eigennamen”. *Namenkundliche Informationen*, Beiheft 10

Doleschal, Ursula & Thornton, Anna M. (2000, Hrsg.), *Extragrammatical and Marginal Morphology*. München: Lincom

Dressler, Wolfgang U. (2000), “Extragrammatical vs. Marginal Morphology”. In: Doleschal & Thornton (Hrsg.), 1–10

Eichler, Ernst (2001), “Germania Romana und Germania Slavica im toponymischen Vergleich”. In: Bentzinger, Rudolf & Nübling, Damaris & Steffens, Rudolf (Hrsg.), *Sprachgeschichte – Dialektologie – Onomastik – Volkskunde*. Stuttgart: Steiner, 159–167

Eichler, Ernst & Hilty, Gerold & Löffler, Heinrich & Steger, Hugo & Zgusta, Ladislav (1995–1996, Hrsg.), *Namenforschung: Ein internationales Handbuch zur Onomastik*. Bd. 1–3. Berlin, New York: Walter de Gruyter

Fleischer, Wolfgang (1970), “Onomastische Strukturen in der deutschen Sprache der Gegenwart”. In: *Onomastica Slavogermanica V, Abhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Phil.-hist. Klasse* 62.2, 35–44

Fuhrhop, Nanna (1998), *Grenzfälle morphologischer Einheiten*. Tübingen: Staufenburg

Greule, Albrecht (1996), “Morphologie und Wortbildung der Vornamen: Germanisch”. In: Eichler et al. (Hrsg.), Bd. II, 1183–1187

Harnisch, Rüdiger (2001), *Grundform- und Stammprinzip in der Substantivmorphologie des Deutschen: Synchronische und diachronische Untersuchung eines typologischen Parameters*. Heidelberg: Winter

Koch, Max (1963), “Volksetymologie und ihre Zusammenhänge”. *Beiträge zur Namenforschung* 14, 162–168

Kuhn, Wilfried & Serzisko, Fritz (1982), “Eigennamen im Rahmen der Dimension der Apprehen-

- sion". In: Seiler, Hansjakob & Lehmann, Christian (Hrsg.), *Apprehension. Teil 1: Bereich und Ordnung der Phänomene*. Tübingen: Narr, 277–293
- Leys, Odo (1966), "Der Eigename in seinem formalen Verhältnis zum Appellativ". *Beiträge zur Namenforschung*, Neue Folge 1, 113–123
- Nübling, Damaris (1997), "Deutsch-schwedische Divergenzen in Entstehung und Struktur der Familiennamen: Ein Beitrag zur kontrastiven Onomastik". *Beiträge zur Namenforschung* 32.2, 141–173
- Nübling, Damaris (2000), "The Semiotic and Morphological Structure of German Toponyms: Different Strategies for Indicating Proprietary". In: Dolleschal & Thornton (Hrsg.), 127–137
- Olschansky, Heike (1996), *Volksetymologie*. Tübingen: Niemeyer
- Seibicke, Wilfried (1985), "Überblick über Geschichte und Typen der deutschen Personennamen". In: Besch, Werner & Betten, Anne & Reichmann, Oskar & Sonderegger, Stefan (Hrsg.), *Sprachgeschichte: Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung, Bd. II*. Berlin, New York: Walter de Gruyter, 2148–2163

## XXI. Morphologie und Nachbardisziplinen

Szczepaniak, Renata (2002), *Onymische Suffixe als Signal der Proprialität: Das Polnische als Paradebeispiel* [Vortrag, gehalten am 20.08.2002 beim 21. Internationalen Kongress für Namenforschung in Uppsala; wird veröffentlicht]

Tournier, Jean (1985), *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris: Champion-Slatkine

Vennemann, Theo (1999), "Volksetymologie und Ortsnamenforschung: Begriffsbestimmung und Anwendung auf ausgewählte, überwiegend bayerische Toponyme". *Beiträge zur Namenforschung, Neue Folge* 34, 269–322

Werner, Otmar (1974), "Wie kann man mit einem begrenzten Inventar über unbegrenzt viele Gegenstände sprechen?" In: Heilmann, Luigi (Hrsg.), *Proceedings of the 11th International Congress of Linguistics 1972, Bd. II*. Bologna: Il Mulino, 173–187

Rüdiger Harnisch, Passau (Deutschland)  
Damaris Nübling, Mainz (Deutschland)

## 174. Etymologie

1. Was ist Etymologie?
2. Etymologie und Wortbildungslehre
3. Etymologie und Wortgeschichte
4. Die Wortprägung
5. Theoretische Voraussetzungen
6. Erschließungsmethoden
7. Zitierte Literatur

### 1. Was ist Etymologie?

Die Etymologie ist der Teil der Sprachwissenschaft, in dem Entstehung (und Geschichte) der Wörter einer Sprache untersucht werden. Im strengen Sinn beschreibt die Etymologie nur die **Entstehung der Wörter**; doch verlangt eine verwertbare Beschreibung der Entstehung eine Ergänzung durch die Behandlung der darauffolgenden **Wortgeschichte** mindestens in dem Umfang, in dem der Zugang zu der Bildung des Wortes von der betrachteten Sprachstufe aus beeinträchtigt wird. Soll also etwa die Etymologie des neuhighdeutschen Wortes *Kiefer* (= Nadelbaum) beschrieben werden, so ist einerseits auf die Bildung des althochdeutschen Wortes *kien-foraha* einzugehen: Es ist ein Kompositum aus *foraha*

'Föhre' und *kien* 'Kienspan' ('die Föhre, aus der Kienspäne vor allem als Beleuchtungsmittel gewonnen werden'); dann müssen aber, um diese Erklärung der Bildung mit dem heutigen Wort zu verbinden, die besondere Lautentwicklung, die morphologische Verdunkelung und die Veränderung von der Benennung einer besonderen Art Föhre zu einem undurchsichtigen Pflanzennamen beschrieben werden, also die Wortgeschichte (oder zumindest ein Teil von ihr). In einem solchen Fall wird das Material zur Beurteilung der Etymologie, die **Vorform**, aus der Überlieferung der betreffenden Sprache selbst gewonnen, in anderen Fällen wird die Beurteilungsgrundlage durch den Vergleich von Wörtern gleichen Ursprungs in verwandten Sprachen (im Rahmen einer **Wortgleichung**) erst erschlossen. Während sich die traditionelle etymologische Wissenschaft fast ausschließlich um die Wortentstehung gekümmert hat, wurde vor allem in der romanistischen Etymologie nicht ohne Grund die Wichtigkeit der Geschichte des Wortes betont; denn die tiefgreifenden Entwicklungen der französischen Sprache lassen bei früh ent-

standenen Wörtern eine bloße Herkunfts-Etymologie in der Tat als unbefriedigend erscheinen. Der Gegensatz zwischen den beiden Interessenschwerpunkten wurde dabei schlagwortartig ausgedrückt durch **étyologie-origine** und **étyologie-histoire du mot** (Baldinger 1959 [1977: 219]). Diese Umorientierung führte in der romanistischen Sprachwissenschaft dazu, daß nun umgekehrt die Wortentstehung vernachlässigt wurde (indem z. B. bei Erbwörtern mit lateinischer Vorform das französische Wort lediglich bis zu dieser zurückgeführt wurde und die Erklärung der Entstehung der lateinischen Etymologie überlassen blieb). Eine befriedigende Etymologie wird aber (in der oben geschilderten Weise) immer beide Aspekte miteinander verbinden müssen. Dabei ist jedoch zu berücksichtigen, daß zur Erklärung der Entstehung (also der Wortbildung) das Quellenmaterial ergiebig genug sein muß. Deshalb hat z. B. Kluge (1911 [1977: 106]) bloßen Wortgleichungen, die die Erschließung der Bildung nicht erlauben (weil keine Basis erkennbar ist), einen „etymologischen Gehalt“ abgesprochen, sie also von der Etymologie im engeren Sinn ausgeschieden, selbst wenn die Zahl der an der Gleichung beteiligten Sprachen ein hohes Alter des betreffenden Wortes sichert. Es ist jedoch zweckmäßiger, auch in solchen Fällen von einer Etymologie zu sprechen, dabei aber zu vermerken, daß der Bildungsvorgang nicht erschlossen werden kann. In den neuesten Auflagen des etymologischen Wörterbuchs von Kluge (231995) wird dies mit dem Vermerk *Herkunft dunkel* getan. Allerdings können die Wörter einer Sprache auch auf andere Weise als durch Wortbildung entstehen. Solche Möglichkeiten, die für eine Behandlung im Rahmen der Morphologie nicht einschlägig sind, sind die Entlehnung, die Urschöpfung (z. B. Lautnachahmung), die Bedeutungs-Übertragung und -Verschiebung, das Erstarren syntaktischer Fügungen und die Wortkürzung. In der Regel wird auch die Entstehung unselbstständiger Morpheme (etwa der Suffixe) der Etymologie zugerechnet, doch könnte dieser Bereich zweckmäßiger in einer historischen Wortbildungslehre (und einer historischen Flexionslehre) behandelt werden.

## 2. Etymologie und Wortbildungslehre

Damit berührt sich ein Teil der Etymologie (nämlich die Etymologie derjenigen Wörter, die als in der betreffenden Sprache oder einer

Vorstufe zu ihr gebildet erwiesen werden können) sehr eng mit der Wortbildungslehre (der verschiedenen Stadien dieser Sprache). Eine naheliegende, aber nicht ganz richtige Abgrenzung weist dabei die Etymologie der historischen, die Wortbildungslehre der synchronisch beschreibenden Sprachwissenschaft zu. Dieser Standpunkt verkennt, daß die Etymologie einen wichtigen synchronischen Aspekt in sich schließt, nämlich die Frage, wie die Sprecher einer Sprache neue Benennungen finden, und welche Möglichkeiten sie dabei auswählen. Zur Beantwortung dieser Frage bietet die Wortbildungslehre die allgemeinen Muster, die zur Gewinnung eines neuen Wortes herangezogen werden können (und dieser Aspekt ist zugleich auch ein etymologischer); doch sucht die Etymologie darüber hinaus die nicht nur mit dem Bildungstyp, sondern auch mit dem speziellen Wort verbundenen Fragen zu klären: die des Benennungsmotivs (s. 4.), die der Abgrenzung im Wortfeld oder gegenüber sonstigen Konkurrenten, die der Auswahl aus verschiedenen Bildungsmöglichkeiten und ähnlichem. Die Wortbildungslehre hebt das an dem Bildungsvorgang Typische heraus und erfaßt auf der semantischen Seite in der Regel nur die **systematische Bedeutung** (die durch die Bedeutung der Elemente und die Funktion der Bildungsregeln voll erschließbar ist). Die Etymologie hebt vor allem auf der Bedeutungsseite heraus, wie mit Hilfe dieser systematischen Bildung ein Zugriff auf das mit dem Wort speziell Gemeinte möglich ist. Deshalb ist es durchaus sinnvoll, nach der Etymologie eines ganz neuen Wortes zu fragen, z. B. nach der Etymologie von *Freischwinger* (womit Stühle bezeichnet werden, bei denen Sitzfläche und Lehne auf einem elastischen Gestell ruhen, weshalb sie schwingen können). In vielen Fällen ist nämlich auch bei neuen Wörtern die Etymologie nicht klar (während die Wortbildung je nach der Stellung, die man der Semantik in ihrem Rahmen einräumt, gegebenenfalls als eindeutig aufgefaßt werden kann). Solche Wörter sind etwa *Geisterfahrer* oder *Radler(maß)*. So gesehen beschreibt die Etymologie synchronisch den Vorgang der Wortentstehung, zumindest bei den Wörtern, die gebildet werden, um etwas Bestimmtes zu bezeichnen (die **Wortprägungen**); voll systematisch gebildete und verwendete Wörter, wie etwa Abstrakta zu Verben und Adjektiven, sind an sich kein Gegenstand der Etymologie; ihre Bildung wird durch die Wortbildungslehre voll erklärt. Sie

können allerdings im Gebrauch Besonderheiten annehmen, die zu Lexikalisierungen führen (**Entfaltung** nach Seibold 1981: 56 f.); für diesen Vorgang ist dann wieder die Etymologie zuständig. Da auch die Prägung von Wörtern nicht notwendigerweise zum normalen Sprachgebrauch gehört und die meisten Sprecher neue Wörter als bereits gebildete vorfinden, sind für sie die Bildungsvorgänge auch ‘Geschichte’, deshalb ist von der Einstellung der Sprecher her gesehen die Einordnung der Etymologie in die historische Sprachwissenschaft gerechtfertigt.

### 3. Etymologie und Wortgeschichte

Von der Etymologie zu verlangen ist, wie ausgeführt, so viel von der Wortgeschichte, daß Form und Bedeutung des Wortes zum Zeitpunkt der Bildung mit seiner Form und Bedeutung in der zu untersuchenden Sprache eindeutig in Zusammenhang gebracht werden können. Im Rahmen einer Wortgeschichte kann natürlich noch viel mehr relevant sein, etwa das Auftreten und die Verwendung des Wortes in bestimmten literarischen Bereichen, seine Häufigkeit, die Herausbildung bestimmter syntaktischer Verwendungen usw. Es ist also durchaus möglich, eine von der Etymologie unterscheidbare Wortgeschichte als Bestandteil der historischen Sprachwissenschaft zu betreiben. Allerdings sind von diesen wortgeschichtlichen Feststellungen die meisten in irgendeiner Weise auch für die Etymologie und ganz besonders für die Erschließung der Etymologie relevant, so daß es unzweckmäßig wäre, in zwei verschiedenen Bereichen der historischen Sprachwissenschaft fast das gleiche zu tun. In der Regel wird deshalb zwischen Wortgeschichte und Etymologie kein systematischer Unterschied gemacht. Mit dem Aspekt ‘Wortgeschichte’ gehört die Etymologie selbstverständlich voll in die historische Sprachwissenschaft.

### 4. Die Wortprägung

Sucht ein Sprecher einer Sprache eine Benennung für eine neue Sache (oder eine neue Benennung für eine bereits benannte Sache), so hat er bereits zwei Vorgaben für den Bildungsvorgang: Er sucht erstens ein Wort, also eine Benennung für eine Klasse ähnlicher Gegenstände (Vorgänge, Eigenschaften usw.), nicht einen Eigennamen (für etwas Individuelles), nicht eine Typenbezeichnung

(die Austauschbarkeit der bezeichneten Einheiten voraussetzt) und nicht einen Fach-Terminus (der nur strikt nach Definition verwendet werden darf). Zweitens hat der Sprecher den ungefähren Bedeutungsumfang vorgegeben: Er will etwas Bestimmtes benennen. Das heißt nicht, daß die Bedeutung bereits festliegt – sie wird erst durch den Gebrauch (durch die Abgrenzung vom Gebrauch teilweise konkurrierender Wörter) im einzelnen bestimmt; aber eine ungefähre Vorgabe (sei es in Form von Prototypen oder in Form von Klassen) ist vorauszusetzen. Dieser Bedeutungsumfang wird nun bei der hier zu betrachtenden Art der Wortgewinnung erfaßt durch eine Wortbildung, die als solche zunächst eine systematische Bedeutung hat. Diese systematische Bedeutung kann den geplanten Bedeutungsumfang mehr oder weniger deutlich wiedergeben: In einem Fall wie *Holzstuhl* fallen beide mehr oder weniger zusammen: Ein *Holzstuhl* ist ein Stuhl aus Holz, und jeder Stuhl aus Holz kann *Holzstuhl* genannt werden. Bei *Freischwinger* dagegen ist lediglich eine charakteristische Eigenschaft genannt: daß er schwingt, wenn man sich hineinsetzt (und es ist durchaus nicht gesagt, daß nach der Absicht des Wortbildners jeder Stuhl, der schwingt, wenn man sich hineinsetzt, *Freischwinger* genannt werden soll, und schon gar nicht ist alles, was frei schwingt, z. B. ein Pendel oder eine Schaukel, ein *Freischwinger*; nur ein Stuhl, der frei schwingt, wird so genannt). Dieses zur Benennung herangezogene Merkmal ist das **Benennungsmotiv** (weniger genau sagt man auch **Bezeichnungsmotiv**). Natürlich hat auch *Holzstuhl* ein Benennungsmotiv (‘besteht aus Holz’), doch ist diese Feststellung hier trivial. Das Benennungsmotiv muß durch die systematische Bedeutung erfaßt werden, und dafür gibt es meist verschiedene Möglichkeiten. Bei *Säugetier* oder *Säuger* (als Benennung einer Tierklasse) ist das Benennungsmotiv ‘säugt seine Jungen’; das wird im Deutschen durch den Rückgriff auf das Verb *säugen* erfaßt, im Französischen und Englischen durch die Erwähnung der Zitzen (lat.  *mammae*, deshalb frz. *mammifère*, engl. *mammal*), im Russischen durch die Feststellung der Ernährung der Jungen mit Milch (russ. *mlekopitajuščee*) – das Benennungsmotiv ist gleich, die **Benennungsmittel** sind verschieden. Bei der Erschließung des Benennungsmotivs sind nun zumindest bei neueren Wörter (wie *Geisterfahrer*) die Benennungsmittel ohne weiteres zu erkennen (*Geister* + *Fahrer*), während die

Präzisierung des Benennungsmotivs häufig Kenntnis der benannten Sache wie auch einen Zugriff durch historische Belege erfordert. Was die Kenntnis der bezeichneten Sache anbelangt, so wird bei der Wortprägung in der Regel nach dem auffälligsten oder am charakteristischsten erscheinenden Merkmal benannt, wobei vor allem die Gebrauchsmöglichkeit eine große Rolle spielt (die Preiselbeere z. B. wird regional als *Steinbeere* bezeichnet, weil sie als Heilmittel gegen Steinleiden (Gallensteine, Nierensteine) galt; ohne Kenntnis dieses Sachverhalts ist das Benennungsmotiv nicht erschließbar). Das bedeutet gleichzeitig, daß das Erfassen des Benennungsmotivs beim Erwerb des Wortes zugleich etwas über die bezeichnete Sache verrät, und das ist einer der Gründe des starken Interesses auch von Nicht-Sprachwissenschaftlern an der Etymologie.

Während sich nun der Gebrauch des Wortes nach den notwendigen Abgrenzungen auf den ursprünglich vorgegebenen Bedeutungs-umfang einpendelt (**Gebrauchsbedeutung**), hat die systematische Bedeutung der Bildung (**Bildungsbedeutung**) beim Gebrauch an sich keine Funktion mehr; sie kann vernachlässigt werden (obwohl das die Sprecher in der Regel nicht tun) und wird bei zunehmender Lexikalisierung des Wortes mehr oder weniger undurchsichtig (die Bildungsbedeutung und damit das Benennungsmotiv etwa von *Junggeselle* ist einem nicht historisch geschulten Sprecher des Deutschen nicht mehr zugänglich; s. Art. 150). Solange sie den Sprechern noch durchschaubar ist, kann sie aber den Gebrauch steuern. Im Normalfall ist diese Steuerung (**Verdeutlichung** nach Seibold 1981: 223–230) nicht zu erkennen, weil sie naheliegende Folgen hat; erkennbar wird sie erst, wenn die Steuerung infolge falscher Anschlüsse durch die Sprecher einen Verlauf nimmt, der durch die Herkunft des Wortes nicht gerechtfertigt ist. Dies ist besonders häufig beim falschen Anschluß von Fremdwörtern der Fall (z. B. *irritieren*, ursprünglich ‘reizen, erregen’, durch falschen Anschluß an *irre* in Richtung auf ‘unsicher machen, beunruhigen, irre machen’ gesteuert). Das Aktivieren der Bildungsbedeutung bei den Sprechern (das für den normalen Sprachgebrauch nicht notwendig ist) kann eine Reihe von Funktionen haben: Es erleichtert das Verständnis des Wortes, unterstützt die Organisation des Wortschatzes beim einzelnen Sprecher, hält aktive Muster für weitere Wortbildungen bereit, ermöglicht Wortspiele usw. Die Verfüg-

barkeit über die Bildungsbedeutung wird häufig “Motivation/Motivierung/Motiviertheit” (s. Art. 150) oder auch “Durchsichtigkeit” genannt, doch sind diese Begriffe nicht sehr präzise gefaßt: Sie beziehen sich in erster Linie darauf, wie nahe die Bedeutung eines Wortes der vorauszusetzenden systematischen Bedeutung steht, wobei aber nicht unterschieden wird, ob die Zurückdrängung der systematischen Bedeutung auf nachträglicher Lexikalisierung oder auf einem schwer oder gar nicht erkennbaren Benennungsmotiv beruht.

## 5. Theoretische Voraussetzungen

Auf die für die synchronische oder diachronische Betrachtung von Sprache allgemein vorauszusetzenden Theorien soll hier nicht eingegangen werden, doch ist eine für die Etymologie speziell wichtige Konzeption besonders zu erwähnen: Die etymologische Untersuchung geht immer von einem Wort einer bestimmten Sprache aus, das etymologisiert werden soll; und im Falle eines nach Wortbildungsregeln gebildeten Wortes ist das Ergebnis der Untersuchung ein Bildungsvorgang, der unter Umständen zeitlich sehr weit zurückliegt, und dessen unmittelbares Ergebnis unter Umständen sowohl lautlich wie auch semantisch von dem Wort der zu untersuchenden Sprache stark verschieden sein kann. Unter diesen Umständen ist es die Frage, ob das Wort der zu untersuchenden Sprache und die gefundene oder erschlossene Bildung der früheren Sprachstufe als das gleiche Wort aufgefaßt werden können (genauer: ob und inwieweit das Wort der späteren Sprache als Nachfolger des Wortes der früheren Sprache aufgefaßt werden kann). Von den Erfordernissen der Etymologie her gesehen können zwei Wörter dann miteinander identifiziert werden, wenn ihre jeweilige Überlieferung auf denselben Bildungsvorgang oder dieselbe Lexikalisierung einer ursprünglich systematischen Bildung zurückführt (**Herkunftsgleichheit**). Eine gewisse Beeinträchtigung der Gleichheit, die aber durch die Beschreibung ausgeglichen werden kann, besteht bei Bedeutungs-Übertragungen und -Verschiebungen, bei Beeinflussungen durch andere Wörter (**Attraktion**) und ähnlichem. Keine Gleichheit im etymologischen Sinn können dagegen parallele Bildungen zu verschiedenen Zeiten und/oder an verschiedenen Orten beanspruchen, selbst wenn sie das glei-

che Sprachmaterial auf gleiche Weise verwenden (**materielle Gleichheit**). Andernfalls würde materielle Gleichheit von Wörtern in verwandten Sprachen eine Rekonstruktion in der gemeinsamen Grundsprache erfordern, was in solchen Fällen zu Ungereimtheiten führt. Dieselbe Schwierigkeit liegt vor, wenn in der Überlieferung dieser Wörter Entlehnungen (s. Art. 152) aufgetreten sind (**Verbreitungsgleichheit**). So haben etwa die nordgermanischen Sprachen materiell gleiche Wörter für die Briefmarke (dän. *frimærke*, schwed. *frimärke*, norw. *frimerke*, isl. *fímerki*), die selbstverständlich auf sekundärer Verbreitung beruhen und nicht die Erschließung eines urnordischen \**frija-mark-ja-* (*n*) ‘Briefmarke’ erlauben, denn Briefmarken gibt es auch in den nordischen Ländern erst seit der Mitte des 19. Jahrhunderts. Ein anderer Fall, bei dem die oberflächliche Gleichheit bei der Etymologie aufzugeben ist, ist die **Homonymie**, besonders wenn eine zusätzliche **Attraktion** der Bedeutungen eingetreten ist. So ist etwa das Wort *Spieß* in der Bedeutung ‘Jagdspieß’ etymologisch ein anderes Wort als in der Bedeutung ‘Bratspieß’ (mhd. *spiez* und *spiz*). Ähnlichkeit in Form und Bedeutung, die für den Sprachgebrauch durchaus eine Identifizierung als gleiches Wort zulassen (**Gebrauchsgleichheit**), sind für die Etymologie erst auf ihre historische Haltbarkeit hin zu prüfen. Selbstverständlich ist die Bestimmung als herkunftsgleich oder materiell gleich nicht (oder zumindest nicht immer) unmittelbar aus dem Befund ablesbar, sondern Ergebnis der etymologischen Analyse. (Ausführlicher zu solchen Fragen: Seibold 1981: 26–35).

## 6. Erschließungsmethoden

Für die Erschließung und Beschreibung der Etymologie sind im Prinzip dieselben Methoden verwendbar wie für die Wortbildungsslehre, die (Wort-) Semantik, die Lautgeschichte, die Morphologie-Geschichte usw., sofern den besonderen Erfordernissen der Etymologie (Benennungsmotive usw.) entsprochen wird. Besonderheiten ergeben sich, wenn für das betreffende Wort eine Basis nicht oder nicht unmittelbar erschlossen werden kann. Terminologisch wäre die Möglichkeit der Erschließung einer Basis als **Anschluß** zu bezeichnen. Liegen verwandte Wörter vor, die auf eine entsprechende Basis zurückführbar sind, die Basis selbst aber nicht, so ist

dies eine (etymologische) **Verknüpfung** (man spricht auch von **Wurzeletymologie**, doch ist dieser Ausdruck mehrdeutig und etwas irreführend). Am ungünstigsten ist die Beleghlage, wenn nur Wörter der gleichen Struktur verglichen werden können, aber weder eine Basis noch aus der gleichen Basis gebildete andersartige Wörter vorliegen (**reine Wortgleichung**). In diesem Fall sind in morphologisch einigermaßen durchschaubaren Fällen etymologische Hypothesen möglich, die aber in der Regel nicht ausreichend zu sichern sind (eine Sicherung wäre z. B. durch gewichtige Parallelen möglich).

Eine Bemerkung ist noch am Platz zu der Analyse in [Wurzel + Affixe] gegenüber der Analyse in [Basis gegenüber Bildung]. In den modernen Sprachen ist es angemessen, bei der Bildung eines Wortes seine Basis, sein Grundwort, aufzuweisen und dann den Bildungsvorgang zu beschreiben. In der älteren Etymologie wurde statt dessen vielfach mit dem Konzept der **Wurzel** operiert, d. h. einem abstrakten Morphem, das allen Bildungen der betreffenden Wortfamilie gemeinsam ist und eine entsprechend abstrakte Bedeutung hat. An diese Wurzel konnten **Primärsuffixe** treten, um **Stämme** zu bilden, an diese wiederum **Sekundärsuffixe**, die Sekundärstämme hervorriefen. Als Wort konnten sowohl die Wurzel (**Wurzelverb**, **Wurzelnomen**) wie auch ein Stamm verwendet werden. Diese Analyse geht auf die indischen Grammatiker zurück, die auf diese Weise das Sanskrit und das Vedische grammatisch beschrieben (vgl. Art. 5). Nun wird eine solche Analyse durch die Struktur des ältesten Indischen tatsächlich nahegelegt, weil diese Sprache z. B. bei den primären Verben (die bei weitem den Hauptbestand der altindischen Verben ausmachen) kein festes Paradigma kennt, sondern nach Bedarf Verbalstämme, wie Präsens, Perfekt, Aorist usw. und auch nominale Ableitungen bildet, so daß eine Grundform (als die wir im Deutschen etwa das Präsens oder den Infinitiv ansehen) nicht zu existieren schien und dadurch das abstrakte Konzept der Wurzel nahegelegt wurde. Diese Analyse-Art ist aber schon bei dem mit dem Altindischen nahe verwandten Altgriechischen nicht mehr zweckmäßig, weil dort die Paradigmen wesentlich stärker ausgebildet sind, und bei den modernen Sprachen, wie Deutsch, Englisch oder Französisch, sind die Paradigmen so fest, daß keinerlei Anlaß besteht, mit abstrakten Wurzeln statt mit konkreten Wörtern zu operieren. Die Gefahr bei Wurzel-

Analysen besteht hauptsächlich darin, daß sie durch ihren abstrakten Charakter leichter an irgendwelche hypothetischen Annahmen anzupassen sind, während konkrete Wörter mit erheblich größerer Präzision (besonders im Bereich der Semantik) bei der Analyse eingesetzt werden können.

Die etymologische Forschung hat viel mit frühen, ja, wo dies möglich ist, sogar mit erschlossenen Sprachzuständen zu tun und setzt deshalb eine umfassende Kenntnis der Sprachgeschichte und der Verwandtschafts- und Entlehnungsverhältnisse der betrachteten Sprache voraus. Für die früheren Sprachstadien und ganz besonders für die erschlossenen Sprachstufen ist sie die wichtigste und vielfach einzige Quelle der Erschließung der Wortbildungslehre (und anderer Teile der Sprachbeschreibung). Dies ist nicht nur wichtig für die Beschreibung von unproduktiv gewordenen Affixen und dergleichen, sondern auch für Wortbildungsmittel, die in den lebenden Sprachen der betreffenden Sprachfamilie gar nicht mehr faßbar sind, so etwa für das Deutsche die Vriddhi-Bildung (ein Bildungstyp für Zugehörigkeitsbildungen, der durch Dehnung oder sonstige Verstärkung des Vokals der ersten Silbe charakterisiert ist; Beispiele wären *Huhn* zu *Hahn* oder *Schwager* zu *Schwieger*), Konsonanten-Geminierung bei emphatischen Bildungen, Wortbildung durch Akzentversetzung und ähnliches. Auf der anderen Seite bietet die präzise Erfassung eines Wortbildungstyps im Rahmen der Wortbildungslehre der Etymologie nicht

nur ein Mittel der Untersuchung und Beschreibung, sondern auch der Bewertung: eine Bildung, deren Merkmale nicht denen ihres Bildungstyps folgen, ist in besonderem Maße erklärbungsbedürftig und möglicherweise falsch eingeordnet. Auf diese Weise sind Etymologie und Wortbildungslehre (auch allgemeiner: und Morphologie) gegenseitig von einander abhängig und können sich gegenseitig erhellen.

## 7. Zitierte Literatur

Baldinger, Kurt (1959), "L'étymologie hier et aujourd'hui". *Cahiers de l'Association Internationale des Études Françaises* 11, 233–264 [Nachdruck in: Schmitt (1977, Hrsg.), 213–246]

Kluge, Friedrich (1911), "Aufgabe und Methode der etymologischen Forschung". *Neue Jahrbücher für das Klassische Altertum, Geschichte und deutsche Literatur* 14, 365–376 [Nachdruck in: Schmitt (1977, Hrsg.), 103–119]

Kluge (231995) = KLUGE: *Etymologisches Wörterbuch der deutschen Sprache*. Berlin, New York: de Gruyter [231995, Nachdruck 1999, 23. erw. Aufl., bearbeitet von Elmar Seibold; 242002]

Schmitt, Rüdiger (1977, Hrsg.), *Etymologie*. Darmstadt: Wissenschaftliche Buchgesellschaft

Seibold, Elmar (1981), *Etymologie: Eine Einführung am Beispiel der deutschen Sprache*. München: C. H. Beck

*Elmar Seibold, Andechs (Deutschland)*

## 175. Schriftsysteme

1. Sprachsystem und Schriftsystem
2. Morphologische Aspekte in logographischen Schriftsystemen
3. Morphologische Aspekte in phonographischen Schriftsystemen
4. Zum Parameter der "Tiefe"
5. Zitierte Literatur

### 1. Sprachsystem und Schriftsystem

Der überkommenen Auffassung zufolge wäre eine breite Transkription die ideale Schrift – wenn die Schriftzeichen direkt die kleinsten lautsprachlichen Einheiten abbilden, d. h. *grosso modo* den Phonemen entsprechen.

Dementsprechend wurde die Geschichte der Schrift als Fortschreiten hin auf das Ideal einer reinen Alphabetschrift gesehen (die klassische Darstellung ist Gelb 1963). Schriften, in denen die Schriftzeichen Wörter, Silben, Morpheme, nur Konsonanten etc. repräsentieren, haben dieser Auffassung zufolge den höchsten Entwicklungsstand noch nicht erreicht. Umgekehrt gelten Alphabetschriften, in denen vom phonographischen Ideal (ein Schriftzeichen – ein Laut bzw. Phonem) abgewichen wird, als in beklagenswerter Weise defizient (vgl. z. B. Saussure 1916).

Diese Sichtweise ist im letzten Viertel des 20. Jahrhunderts zunehmend kritisiert und

revidiert worden. Wiewohl der Schrift ihr phonographischer Charakter nicht abgesprochen werden kann, ist dieser doch nicht auf eine linear-eindimensionale Abbildung der Phonemebene zu reduzieren. Zudem ist die Abbildungsfunktion der Schrift nicht ihre einzige, übernimmt sie doch gleichzeitig auch Modellfunktion für die Analyse (Olson 1993) und Entwicklung der gesprochenen Sprache (Koch & Oesterreicher 1994). Zu unterscheiden ist zwischen einem semantischen Aspekt (Schriftzeichenfolgen repräsentieren Bedeutungen), einem phonographischen Aspekt (Schriftzeichenfolgen repräsentieren die lautive Seite der Sprache), und einem grammatischen Aspekt (Schriftzeichenfolgen repräsentieren die grammatische Organisation von Äußerungen) (vgl. Maas 1992; Günther 1995). Als Schriftsystem einer Sprache wird die Gesamtheit der die schriftliche Ausdrucksseite der Sprache betreffenden Mittel bezeichnet (Eisenberg 1996 a: 1368 f.), von denen einige in der Lautsprache nicht originär sind (z. B. Abkürzungen, abstrakte Symbole, Groß-/Kleinschreibung, Determinative). Moderne Schriftsysteme weisen stets semantische, phonographische und grammatische Aspekte gleichzeitig auf. So sind die arabischen Ziffern, wo immer sie in Texten verwendet werden, grundsätzlich semantisch bestimmt; in den modernen Alphabetschriften sind die Wortkörper durch die Buchstaben phonographisch, die Interpunktions- und der Gebrauch von Großbuchstaben dagegen grammatisch bestimmt. Im folgenden soll anhand exemplarischer Beispiele die Rolle der Morphologie als Brücke zwischen phonographischem und grammatischem Aspekt in Schriftsystemen dargestellt werden.

## 2. Morphologische Aspekte in logographischen Schriftsystemen

Im Hinblick auf die Grundbezugsebene lassen sich drei Schrifttypen unterscheiden: In **logographischen** Schriften beziehen sich die Schriftzeichen in erster Linie auf Bedeutungsträger (z. B. Chinesisch); in **Silbenschriften** bilden die Schriftzeichen Silben ab (z. B. die japanischen Kana); in **Alphabetschriften** sind die Schriftzeichen auf Phoneme bezogen (Eisenberg 1996 a: 1372).

Der Terminus "logographisch" besagt wörtlich, daß die einzelnen Schriftzeichen dazu dienen, Wörter zu bezeichnen. In den meisten logographischen Schriften wird aber

eher davon die Rede sein, daß Schriftzeichen Morpheme bezeichnen. In isolierenden Sprachen wie dem Chinesischen wird damit im Prinzip jedes Morphem durch ein Schriftzeichen repräsentiert. In flektierenden Sprachen wird dagegen in der Regel nur die Wortbasis durch ein Logogramm repräsentiert. So wird in der sumerischen Schrift anfänglich ein Logogramm für alle möglichen grammatischen Formen eines Lexems verwendet; wenn aber später "syllabisch ausgedrückte grammatische Elemente obligatorisch zu den Logogrammen treten, wird deren Funktion auf den Ausdruck der Wortbasis eingeengt" (Krebernik & Nissen 1994: 279). Das Zitat kennzeichnet eine Veränderung, die überall beobachtet werden kann, wo von einer semantischen hin zu einer phonographischen Schreibweise übergegangen wird: Die ausschließliche Verwendung von Zeichen für Inhaltswörter wird ergänzt durch silbische Elemente; solche Schriftsysteme werden häufig als "morpho-syllabisch" bezeichnet. Denn die Schriftzeichen für Flexionselemente repräsentieren nicht in direkter Weise Flexionsmorpheme (ein Zeichen pro Morphem), sondern indirekt durch Kennzeichnung der Aussprache.

Es gibt verschiedene Verfahren, die den Übergang zu einem morphosyllabischen und damit jedenfalls teilweise phonographischen System kennzeichnen. Der bekannteste Fall sind sog. Rebusschreibungen: Ein Wortzeichen wird nach seinem Lautwert, nicht seiner Bedeutung gelesen, besonders bei Homonymen. So wird im Sumerischen das (ursprünglich piktographische) Schriftzeichen für *sar* 'Pflanze' dann auch benutzt für ein homophones Verb mit der Bedeutung 'schreiben' (Cooper 1996: 42; mit weiteren Beispielen). In der späteren Zeit werden logographische Zeichen auch für gleich oder ähnlich lautende Flexionselemente verwendet, z. B. dient dann das Logogramm für *ga* 'Milch' auch dazu, als Silbenzeichen den Auslaut der Form *duga zu du* 'sagen' zu schreiben (Krebernik & Nissen 1994: 285).

Ein weiterer uneigentlicher Gebrauch logographischer Schriftzeichen sind semantische Determinative (auch Komplemente genannt). Hier wird einem Logogramm für einen homonymen Ausdruck ein weiteres Zeichen aus dem gleichen Bedeutungsfeld beigegeben, das die Lesart indiziert; in einer logographischen Schrift für das Deutsche würde man z. B. dem Wortzeichen für *Schloß* das Zeichen für *Schlüssel* beigegeben, um die Lesart

*Türschloß* zu spezifizieren. In der ägyptischen Hieroglyphenschrift stehen einmal etablierte Determinative auch bei nicht ambigen Zeichen, um sie inhaltlich näher zu spezifizieren; so kann dem Wort *'fd.t'* ‘Kasten’ ein semantisches Determinativ hinzugefügt werden, das ihn als aus Holz, Stein oder Metall gefertigt kennzeichnet (Ritner 1996: 77); es wird so ein umfangreiches semantisches Klassifikationssystem aufgebaut, das in der gesprochenen altägyptischen Sprache keine Entsprechungen hat (Schenkel 1994: 291). Es läßt sich hier durchaus von einer morphologischen Kennzeichnung semantischer Klassen in der Schreibung sprechen, die es im Mündlichen nicht gibt.

Eine systematische Aufteilung der logographischen und der phonographischen Ebene bietet das japanische Schriftsystem (vgl. Stalp 1996). Es besteht aus drei funktional und formal differenzierbaren Schriftzeichensets. Das eine, *Kanji* genannt, ist eine heute auf knapp 2000 beschränkte Menge von Logogrammen, denen durch vielfache Homonymie ca. 4000 Bedeutungen entsprechen (Stalp 1996: 1415 f.). *Kanji* stellen praktisch immer Basismorpheme dar. Daneben stehen zwei homologe, funktional differenzierte Syllabare. Das eine, *Katakana* genannt, dient der Schreibung von Namen, Fremdwörtern, Dialektismen etc., also rein phonographischen Zwecken. Das andere, *Hiragana* genannt, dient der Repräsentation grammatischer Aspekte (Hilfswörter, Verbalsuffixe, Konjunktionen) und repräsentiert mithin (auch) die Flexionsmorphologie des Japanischen. Eine so eindeutige funktionale Aufteilung der Schriftzeichen ist in den Schriften der Welt eher selten.

### 3. Morphologische Aspekte in phonographischen Schriftsystemen

#### 3.1. Der Parameter der Tiefe

Die Abbildung der Phoneme auf die Buchstaben gilt als Endstufe der phonographischen Entwicklung. In der reinsten Form verkörpert wohl das Altgriechische der vorklassisch-klassischen Periode dieses Prinzip: Hier scheint es keinen Bezug auf die Bedeutungsseite zu geben. Dies wird besonders deutlich durch das Prinzip der *scriptio continua*: Quasi den mündlichen Sprachstrom direkt abbildend werden im Gegensatz zu semitischen Vorbildern keine Wortgrenzen markiert (vgl. Ludwig in Vorb.). Die Schriftgeschichte zeigt

indes, daß dies nicht der Endpunkt der Entwicklung ist: Wenn sie länger in Gebrauch sind, entwickeln sich in allen alphabetischen Systemen Auszeichnungsmethoden, die mehr als die Darstellung der Lautseite der Sprache zum Ziel haben. Durch typographische Mittel, Einführung der Worttrennung durch das Spatium, Groß-/Kleinschreibung, Interpunktion u. a. m. wird in der schriftlichen Form die “grammatische Artikulation des Textes” repräsentiert (Maas 1992): Absätze organisieren den Textfluß, initiale Majuskel und Schlußpunkt gliedern den Satz als die maximale syntaktische Einheit aus, das Spatium markiert minimale syntaktische Einheiten (Wörter), etc.

Auch die Repräsentation morphologischer Aspekte gehört zu diesen Mitteln. Neuerdings wird hier mit dem Konzept der **Tiefe eines Schriftsystems** gearbeitet. Wohl im Anschluß an die Konzeption einer Tiefenphonologie (Chomsky & Halle 1968) gelten Schriftsysteme als um so flacher, je direkter die Buchstabenfolgen die Oberflächenphonologie einer Sprache reflektieren. So werden im modernen Neuhochdeutschen morphonemische Phänomene wie die Auslautverhärtung in der Schrift nicht ausgedrückt, die noch im Mittelhochdeutschen und Frühneuhochdeutschen auch im Schriftbild erschien, vgl. das Wort *Tag*:

- (1) (a) mhd. [tak] [tages]  
 <tac> <tages>
- (b) nhd. [ta:k] [ta:ges]  
 <tag> <tages>

Es wird davon gesprochen, daß das Mittelhochdeutsche eine flachere Graphie als das Neuhochdeutsche aufweist.

Die systematischsten neueren Untersuchungen zum Parameter der Tiefe von Schriftsystemen in linguistischer Hinsicht stammen von Trudel Meisenburg. Abb. 175.1 kennzeichnet den Ansatz.

Unterschieden werden zunächst zwei durch den Parameter der Tiefe verbundenen Ebenen, eine (oberflächen-)phonologische und eine morphosemantische Ebene. Eine dritte (phonetische) Ebene ist zwar ebenfalls anzusetzen, aber ohne große praktische Relevanz (Meisenburg 1996 a: 25). Außerdem ist zu unterscheiden, ob die Tiefe des jeweiligen Systems in diachroner oder synchroner Hinsicht betrachtet wird. Für die verschiedenen Ebenen werden Beispiele aus dem Lateinischen gegeben (<v> signalisiert im Lateini-

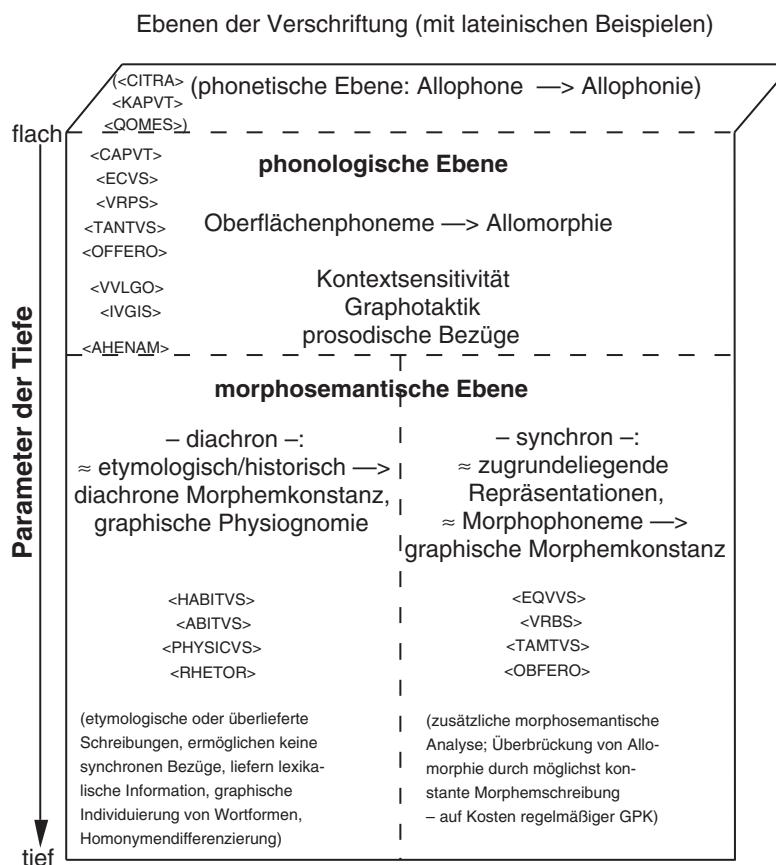


Abb. 175.1: Der Parameter der Tiefe (aus Meisenburg 1998:45)

schen sowohl /v/ als auch /u/). Da gibt es zunächst in <citra, kapvt, comes> die Wiedergabe phonetisch bedingter allophonischer Variation durch auch unterschiedliche Schriftzeichen: In den älteren Zeugnissen steht für /k/ vor vorderen Vokalen <c>, vor hinteren <q>, vor Konsonanten <k> – eine Praxis, die aber schon sehr früh aufgegeben wird: <k> verschwindet, und <q> wird nur noch vor <v> verwendet (Meisenburg 1996 a: 39). Auch oberflächenphonologische Variation wird anfänglich berücksichtigt, z. B. *urbs* ‘Stadt’ Nom. Sg. <vrps> vs. Gen. Sg. <vrbis> wegen Entstimmung des zugrundeliegenden /b/ vor /s/ in der Nominativform (/urbs/ -[urps]), später nicht mehr; ähnlich die anderen Fälle. Entgegen dem Eindruck der vereinheitlichten Klassikerausgaben gibt es in klassischer und nachklassischer Zeit Schreiberunterschiede bezüglich der “tieferen” Schreibung, z. B. <obfero> trotz der vorherrschenden Schreibung nach der Aus-

sprache [offero] wegen der morphologischen Struktur *ob* + *fero*, oder, wiewohl selten, <tamtus> trotz der Dissimilation [tantus] wegen des Bestandteils *tam* (vgl. ausführlich Meisenburg 1996 a: 35–47).

Die bisher angeführten Beispiele betreffen Schreibungen, die synchron nachvollziehbar sind. Andere Mechanismen beziehen sich u. a. auf die diachrone Ebene (vgl. ausführlich zur lateinischen Graphie Meisenburg 1996 a: 31–47). Dazu gehört z. B. die Kennzeichnung (ursprünglich) fremder Wörter durch fremde Buchstaben(kombinationen), etwa wenn aus dem Griechischen entlehnte Wörter mit <ph> und <y> erscheinen wie in <physicus>, obgleich wohl nur Gebildete dem <y> auch den entsprechenden, dem Lateinischen fremden Laut [y] zuordnen konnten; die Normalaussprache dürfte [fiskus] gewesen sein (Meisenburg 1996 a: 46). Es findet sich auch Homonymendifferenzierung, z. B. [abitus]

<habitus> vs. <abitus> wegen der Ableitungsstruktur *ab-ire*.

Im folgenden soll der Parameter der Tiefe in Anschluß an Meisenburg (1996 a; 1998) exemplarisch an einigen Schriftsystemen erläutert werden. Die Anordnung soll die zunehmende Tiefe der Systeme spiegeln.

### 3.2. Spanisch

Das spanische Schriftsystem wird von Meisenburg diachron (1996 a) und synchron (1996 b) als nahezu idealtypisch flaches System gekennzeichnet. Die Elemente eines relativ einfachen Phonemsystems werden meist direkt durch lateinische Buchstaben repräsentiert. Allomorphien, ohnehin eher selten und unsystematisch, werden durchgängig schriftlich repräsentiert, ebenso Neutralisierungen oder Assimilationen. Systematisch phonologiebezogen ist auch die graphische Markierung des relativ freien und teilweise distinktiven Wortakzents. Die synchrone Flachheit des spanischen Schriftsystems wird lediglich durchbrochen bei der Akzentmarkierung homophoner Funktionswörter. Eine geringfügige diachrone Tiefe erhält das System durch aus dem Lateinischen stammende Alternationen, die keine morphophonemischen Entsprechungen haben, z. B. <tocar>, <toques> trotz der gleichen Aussprache [to'kar], [to'kes].

### 3.3. Koreanisch

Das moderne koreanische Schriftsystem *Han'gul*, im 15. Jahrhundert entwickelt, ist in vieler Hinsicht einzig unter den alphabetischen Schriftsystemen (vgl. King 1996). Die Basis des koreanischen Schriftsystems ist eine quasi ikonische Abbildung der artikulatorischen Geste bei den Konsonanten; die genaue Abbildung der Lautstruktur ist Ziel dieser Schriftform (vgl. Günther 1988: 54–58). Anders als in Lateinschriften werden (ähnlich wie in indischen Schriften) die alphabetischen Symbole zu Silben repräsentierenden Schriftzeichen zusammengefaßt, deren Form und Größe an chinesische Schriftzeichen erinnert.

Trotz der Zielsetzung einer möglichst genauen Darstellung der Lautebene gab es von Anfang an auch Bestrebungen, morphologische Aspekte zu berücksichtigen (King 1996: 223). Aufgrund der Zusammenfassung einzelner Phoneme zu Silben ergeben sich bei flektierten Formen nämlich unterschiedliche Schriftzeichen. So lautet der Lokativ des Wortes *saram* ‘Mensch’ *saram-ege* ‘Mensch-LOK’. Ersteres ergibt die Silbenfolge *sa-ram*,

letzteres *sa-ra-me-ge*. Die zweite Silbe ist also in der Grundform *ram*, in der flektierten Form *ra*. Dies aber ergibt unterschiedliche Zeichen: Zwar sind in beiden Silbenzeichen die Elemente für /r/ und /a/ enthalten, doch die Verwendung des Zeichens für /m/ hängt ab von der Syllabifizierung, so daß das komplexe Zeichen unterschiedlich aussehen kann (vgl. Abb. 175.2). In der modernen koreanischen Orthographie gilt die morphemische Schreibweise.

(a) *saram* ‘Mensch’

사	람
sa	ra
	m

(b) *saram-ege* ‘Mensch’, Lokativ

morphemische Schreibweise	silbische Schreibweise
사 람	사 라
ra e	ra me
ge	ge

Abb. 175.2: Schreibweisen des Wortes *saram-ege* in *Hang'ul*

Das koreanische System ist also, betrachtet man die Komponenten der einzelnen Schriftzeichen, wie das Spanische sehr flach; betrachtet man die Komposition der Silbenzeichen, so ergibt sich eine gewisse morphologische Tiefe, die aber die phonologische Flachheit nicht tangiert (s. 4).

### 3.4. Deutsch

Das deutsche Schriftsystem ist durch das Prinzip der Morphemkonstanz bestimmt (Günther 1988: 86–94), das in 3.1 schon am Beispiel der deutschen Auslautverhärtung erläutert wurde: Mit wenigen Ausnahmen haben deutsche Wortstämme im Geschriebenen eine konstante Form ungeachtet der Variation in der phonetisch-phonologischen Realisation. Instruktiv ist das Morphem /kɔ:nig/. In der Standardsprache wird /ig/ im Auslaut [iç] ausgesprochen, also lautet die Basisform [kɔ:niç]. Die Ableitung mit Movierungssuffix /kɔ:nig + in/ wird [kɔ:nigin] ausgesprochen, das denominale Adjektiv /kɔ:nig + lic/ lautet [kɔ:nikliç]. Für das Stammorphem gibt es also drei phonetische Realisierungen, aber stets die gleiche Schreibung <könig>; synchron repräsentiert die Schreibung direkt die morphologische Ebene.

Auch in anderen Fällen signalisiert die geschriebene Form morphologische Struktur,

die an der lautsprachlichen Oberfläche nicht sichtbar ist. So gibt es im gesprochenen Nhd. keine konsonantischen Geminate; wo sie aus morphologischen Gründen auftauchen würden, werden sie im Gesprochenen vereinfacht, in der Schrift jedoch nicht, vgl. *Tief-flug* [ti:flu:k], *verraten* [ve:ra:tn], etc. In der Flexion gibt es zwar auch im Geschriebenen Geminatenvereinfachungen, vgl. <er rät> /re:t + t/, <du feixt> /feiks + st/, doch bleibt auch hier die Stammkonstanz erhalten.

Ein etwas komplexerer Fall ist der Umlaut. Bei vokalischer Alternation gibt es im Nhd. (anders als im Englischen, s. 3.6) keine graphische Morphemkonstanz, vgl. *singe* – *sang* – *gesungen*; *ziehen* – *Zug*, etc. Die Vokale [ɛ], [ø] und [y] und die entsprechenden Langvokale alternieren in einer Reihe von Fällen, z. B. der nominalen Pluralbildung, mit [a], [o] und [u], vgl. /bart/ – /bərtə/, /korb/ – /kɔrbə/, /hu:t/ – /hy:tə/, teilweise auch in Derivationen, vgl. /alt/ – /eltlich/, /to:t/ – /tɔ:tn/, /vu:t/ – /vy:tɔri/. Die umgelautete Form wird mit dem Basisvokal plus Trema als Diakritikon geschrieben: <bärte>, <körbe>, <hüte>, <ältlich>, <töten>, <wüterich>. Im heutigen Deutsch ist der Umlaut morphologisch, nicht mehr phonologisch geregelt (s. Art. 147; Wurzel 1984). Außerdem müssen <ä>, <ö> und <ü> in der nhd. Orthographie als Grapheme gelten, weil sie auch ohne Umlaubasis vorkommen, vgl. z. B. <lärche, dünn, könig>, wozu es keine \*<larch, dunn, kon> gibt. Natürlich sind diese Schreibungen diachron als Umlaut erklärbar; dies ändert nichts am graphematischen Status der Umlautbuchstaben. In der seit 1998 gültigen neuen Rechtschreibung werden in einigen Fällen diachron erklärbare Fälle synchron umgedeutet, so ist z. B. früher <bendel> (zu ahd. *bentil*) nunmehr <bändel> zu schreiben, weil es synchron zu <band> gestellt wird, d. h. der Bezug der Umlautbuchstaben auf die Basisvokale wird als morphologisches Schreibprinzip angesehen.

In die diachrone Dimension gehört die Morphemkonstanz u. a. bei unproduktiven Ableitungen, z. B. das <h> bei <draht> (historisch zu *drehen*), da das Dehnungs-*h* nur vor <l, m, n, r> auftritt (vgl. Eisenberg 1996 b). Hierher gehört auch die Verteilung von <f> und <v> zur Bezeichnung von /f/ im Anlaut im indigenen Wortschatz: Dem unmarkierten Fall der <f>-Schreibung stehen eine Handvoll <v>-Schreibungen gegenüber, die jedoch großteils sehr häufige Wörter

bezeichnen (vgl. Günther 1988: 94–97). Diachrone Tiefe durch Homonymendifferenzierung (*Saite* – *Seite*, *Mohr* – *Moor*, etc.) ist offenbar im 17. Jahrhundert erheblich stärker ausgeprägt gewesen als heute. Ein orthographisches Problem bietet schließlich die Fremdwortschreibung: Fremde Wörter werden zunächst grundsätzlich in der Originalschreibung wiedergegeben, wodurch diachrone Tiefe erreicht wird; die allmähliche (und unsystematische) phonologische und morphologische Angleichung an das deutsche Schriftsystem führt hier zu Inkonsistenzen (z. B. ist die eingedeutschte Form *Frisör* heute wieder veraltet; das Duden-Rechtschreibwörterbuch (2000) verzeichnet es als Hauptlemma *Fotografie* mit dem Verweis „auch *Fotographie*“, dagegen als Hauptlemma *Lithographie* mit dem Verweis „auch *Lithografie*“; etc.).

Daß im heutigen Deutsch in der Tat die konstante Morphemschreibung systematisch dem phonographischen Oberflächenbezug vorgeordnet ist, belegt das Kuriosum der Schreibung von <fünf, elf, zwölf, doof>. Phonologisch ist stets von stimmhaftem Auslaut auszugehen, denn es heißt standardsprachlich z. B. im Plural <doofe> [do:və]. Da jedoch die Grundformen mit <f> geschrieben werden (wohl, weil diese Stämme fast nur so vorkommen), wird die Schreibung auch in den anderen Formen beibehalten, obgleich <f> im Deutschen stets /f/, nie /v/ bezeichnet (vgl. Günther 1988: 96 f.). Insgesamt ist das Deutsche tiefer als z. B. Spanisch oder Koreanisch, aber flacher als die beiden im folgenden besprochenen Schriftsysteme.

### 3.5. Französisch

In noch stärkerem Ausmaß als im Deutschen ist im Französischen graphematische Morphemkonstanz gewahrt. Dies betrifft insbesondere nicht hörbare, aber geschriebene Morphologie, etwa das Plural -*s*, vgl. <enfant – enfants> [afɑ̃] ‘Kind(er)’, <le – les> [lə; le] oder den Genusmarker -*e*, vgl. <petit – petite> [pə'ti – pə'tit]. Verkompliziert wird das Ganze durch das Phänomen der Liaison; vor vokalisch anlautenden Folgewörtern wird z. B. das Plural -*s* hörbar, vgl. *les Etats Unis* [leze:'tazy:'ni] ‘USA’. Da die Wortgrenze überschritten wird, wenn nur geschriebene Morpheme hörbar werden, spricht Eisenberg (1996 a: 1376) vom Wortformenbezug des Französischen. Wenn unterschiedlich lautende Allomorphe durch graphische Morphemkonstanz aufeinander bezogen bleiben, vgl. z. B. <champ – champêtre> [ʃã –

šā'peṭr] ‘Feld – ländlich’ oder <hasard – hasarder> [a'za:r – aza:r'de] ‘Zufall – aufs Spiel setzen’, spricht Catach (1996: 1448 f.) von ‘Morphogrammen’. Bezuglich diachroner Tiefe ist vor allem die morphemidentifizierende Distinktschreibung von Homonymen zu nennen, vgl. z. B. die verschiedenen Schreibungen der [köt] ausgesprochenen Morpheme in <conte> ‘Märchen’, <comte> ‘Graf’, <compte> ‘Rechnung’, wodurch zugleich derivationelle Zusammenhänge verdeutlicht werden, vgl. <conter> ‘erzählen’, <comtal> ‘gräflich’, <compter> ‘zählen’. Eine systematische Darstellung der diachronen Entwicklung der französischen Graphie bis heute bietet Meisenburg (1996 a: 68–206).

### 3.6. Englisch

Englisch ist für viele Autoren der Prototyp eines “tiefen” Systems; als “deeply morphologized” wird es in Stubbs (1996: 1442) bezeichnet. Die Auffassung, daß es sich um ein notorisch irreguläres System handelt, ist spätestens seit der Behauptung, es handele sich um ein “near to optimal system” (Chomsky & Halle 1968: 49) zumindest fragwürdig geworden. Aufgrund der englischen Sprachgeschichte ist insbesondere die diachrone Tiefe des englischen Schriftsystems wohl einzigartig (zur historischen Entwicklung vgl. Scragg 1974). Hier sind insbesondere zwei Aspekte zu nennen. Der eine ist die Zusammensetzung des Wortschatzes aus einem angelsächsischen und einem romanischen Teil mit unterschiedlichen Wortstrukturen und orthographischen Konventionen, vgl. z. B. die Schreibung des Auslauts [šən] in <freshen> [frēʃən] und <nation> [neiʃən]. Der andere ist der Umstand, daß durch den Drucker Caxton die englische Orthographie schon im 15. Jahrhundert weitgehend festgelegt wurde; die nachfolgenden gravierenden Lautveränderungen, insbesondere der “great English vowel shift”, wurden in der Schreibung nicht berücksichtigt. Ähnlich wie das Französische archiviert so das englische Schriftsystem in Paaren wie <sign – signature> oder <divine – divinity> morphologische Zusammenhänge, die in der gesprochenen Sprache nicht mehr erkennbar sind, vgl. [sa:n – signətʃər], [di've:n – di'venɪtɪ]. Ein weiterer morphologischer Aspekt ist die Worttrennung am Zeilenende, wo nicht wie in vielen anderen Alphabetschriften Schreibsilben abgetrennt werden (vgl. Eisenberg 1996 b), sondern Flexionssuffixe, sofern sie

Vokale enthalten, vgl. *read-ing, walk-ed, substitu-tion*.

Obgleich solche Beispiele morphologischen Bezug aufweisen, muß doch festgestellt werden, daß ihre Reichweite und Systematizität ungleich kleiner ist als z. B. im Deutschen oder auch im Russischen (dazu Meisenburg 1998). Es ist deshalb eher davon zu sprechen, daß die Tiefe des englischen Schriftsystems sich auf die Wortebene bezieht; “the invariant unit of English spelling is not the morpheme, but the word” (Stubbs 1996: 1442). In der Tat sind einzelne Autoren so weit gegangen, das englische als ein logographisches Schriftsystem zu bezeichnen – sicherlich ein Unfug, denn jenseits der Unrichtigkeit von G.B. Shaws Behauptung, das Englische sei so unregelmäßig, daß z. B. auch <ghoti> als [fiʃ] gelesen werden könne – [f] wie in <tough>, [i] wie im Plural <women>, [š] wie in <nation> –, gibt es auch im Englischen eine phonographische Grundebene; die Buchstabenfolge <tomb> kann niemals genauso wie <grave> (beides kann ‘Grab’ bedeuten) gelesen werden (Beispiel von Daniels 1996: 654).

### 4. Zum Parameter der “Tiefe”

Betrachtet man die vorstehenden Beispiele im Zusammenhang, so wird deutlich, daß die so einfache scheinende Dimension von morphologisch flach bis tief keineswegs homogen ist. Schon in Abb. 175.1 (Meisenburg 1998) werden zusätzlich die synchrone und die diachrone Ebene unterschieden. Letztere ist überdies höchst disparat: Fremdwortschreibung, Homonymendifferenzierung und etymologische Schreibung können nur a posteriori zusammenwirken. Ein einheitliches Gebiet bilden sie nicht; zudem ist morphologischer Bezug für keinen der Bereiche konstitutiv. “Verallgemeinerungen über den Ebenenbezug alphabetischer Schriftsysteme sind [...] noch ein Forschungsdesiderat” stellt dementsprechend Eisenberg (1996 a: 1376) fest, so daß die von ihm tentativ formulierten Tendenzen noch der Bestätigung harren:

- Sprachen mit wenig Allomorphie neigen eher zu flachen Schriftsystemen (Spanisch).
- Isolierende Sprachen und solche mit segmental schwer faßbarer Morphologie neigen zu Wortformbezug (Französisch, Englisch, Arabisch).

- (c) Sprachen mit segmentaler Morphologie neigen zu Morphembezug (Deutsch, Russisch).

Die Notwendigkeit weiterer Forschungen liegt auf der Hand. Naheliegende Umkehrschlüsse z. B. sind offenbar nicht zutreffend. So ist das serbokroatische Schriftsystem ähnlich flach wie das spanische, obgleich es erheblich mehr Allomorphie gibt, z. B. durch die Palatalisierung, die hier im Unterschied zu anderen slawischen Sprachen in der Schrift angezeigt wird (vgl. Comrie 1996), vgl. z. B. <ruka – ruci> ‘Arm’; <snaha – snasi> ‘Schwiegertochter’ (jeweils Nom./Dat. Sg.). Die phonographische Genauigkeit geht so weit, daß sogar Unterschiede zwischen serbischen und kroatischen Dialekten graphisch repräsentiert werden, vgl. serbisch <mleko>, kroatisch <mljeko> ‘Milch’. Wohlgemerkt gilt dies für die serbokroatische Schrift seit ihrer grundlegenden Reform im 19. Jahrhundert (Feldman & Barac-Cikola 1996).

Eine besondere Stellung nehmen die semitischen Schriftsysteme ein. Auch wenn es Ansätze zur Schreibung von (Lang-)Vokalen gibt, sind sie systematisch als Konsonantenalphabete zu kennzeichnen. Aufgrund der semitischen Morphologie mit (extrem vereinfacht!) drei Konsonanten als Wurzel und vocalischer Variation zur Kennzeichnung von Flexion und Derivation wird die morphologische Ebene in besonderer Weise ausgezeichnet (das Konsonantengerüst signalisiert morphologische Beziehungen); das System scheint morphologisch “tief”. Eine Kennzeichnung der arabischen Schriften als Morphemschriften aber ist unangemessen. So spricht Eisenberg (1996 a: 1376) von einer morphologischen Analyse, die aber erst auf der Ebene der Wortformen strukturbildend wirkt, und ganz explizit wehrt sich Bauer (1996: 1435) gegen die landläufige Vorstellung, “die arabische Schrift diene vor allem zur Schreibung von Wurzeln”. Die vielen Details, worin dies sich zeigt, können hier nicht ausgeführt werden; wesentlich ist das Ergebnis: “Die durch die Defektivschreibung des Arabischen entstehende Tiefe der Schrift ist [...] eher lexikalischer als morphologischer Natur” (Bauer 1996: 1455). Ähnlich war in 3.6 das englische Schriftsystem als weniger morphologisch und vielmehr lexikalisch tief gekennzeichnet worden.

Eine solche differenzierte Betrachtung des Parameters der Tiefe ist insbesondere deshalb wesentlich, weil er auch in anderen Wissen-

## XXI. Morphologie und Nachbardisziplinen

schaften eine Rolle spielt. In der psychologischen Leseftorschung etwa ist er weitgehend als reguläre vs. irreguläre Graphem-Phonem-Korrespondenzen interpretiert worden (z. B. Katz & Feldman 1996). Eine entwicklungspsychologische Interpretation besagt dann, daß Schriftsysteme um so leichter zu lernen seien, je flacher sie sind (Wimmer & Goswami 1994 zum Vergleich von Englisch und Deutsch), und eine kognitionspsychologische Deutung vermutet, daß der Erwachsene in flachen Systemen beim Lesen phonologisch rekodiert (Katz & Feldman 1996) oder in tiefen Systemen morphologisch zerlegt (Schreuder et al. 1989; s. Art. 162). Die oben gegebenen Hinweise lassen es als ratsam erscheinen, im jeweiligen Fall genau zu spezifizieren, was unter “flach” bzw. “tief” zu verstehen ist.

Schon Bolinger (1946) war die besondere Rolle von “visual morphemes” bewußt. Es scheint, als ob die Analyse von Schriftsystemen in besonderer Weise geeignet ist, die Rolle der Morphologie als Interface von Bedeutung und Form in der menschlichen Sprache und ihre unterschiedliche Ausprägung in den Sprachen zu kennzeichnen.

In den letzten Jahren ist die Erforschung der Schriftsysteme weiter vorangetrieben worden, insbesondere die Rolle der Schreibsilbe zwischen segmentaler und morphologischer Ebene steht im Zentrum der Aufmerksamkeit. Systematisch wird dies am Deutschen exemplifiziert in Eisenberg (1998: 286–311); einen guten Einblick in verschiedene Ansätze vermittelt der Sammelband von Neef et al. (2002, Hrsg.).

## 5. Zitierte Literatur

- Bauer, Thomas (1996), “Das arabische Schriftsystem”. In: Günther & Ludwig (Hrsg.), 1433–1436
- Bolinger, Dwight L. (1946), “Visual Morphemes”. *Language* 22, 333–340
- Catach, Nina (1996), “The French Writing System”. In: Günther & Ludwig (Hrsg.), 1445–1450
- Chomsky, Noam & Halle, Morris (1968), *The Sound Pattern of English*. New York: Harper & Row
- Comrie, Bernard (1996), “Languages of Eastern and Southern Europe”. In: Daniels & Bright (Hrsg.), 663–689
- Cooper, Jerrold S. (1996), “Sumerian and Akkadian”. In: Daniels & Bright (Hrsg.), 37–57

- Daniels, Peter T. (1996), "English". In: Daniels & Bright (Hrsg.), 651–655
- Daniels, Peter & Bright, William (1996, Hrsg.), *The World's Writing Systems*. Oxford: Blackwell
- Duden (2000), *Die deutsche Rechtschreibung*. 22., völlig neu bearbeitete und erweiterte Auflage. Mannheim et al.: Dudenverlag (Duden Band 1)
- Eisenberg, Peter (1996 a), "Sprachsystem und Schriftsystem". In: Günther & Ludwig (Hrsg.), 1368–1380
- Eisenberg, Peter (1996 b), "Das deutsche Schriftsystem". In: Günther & Ludwig (Hrsg.), 1451–1455
- Eisenberg, Peter (1998), *Grundriss der deutschen Grammatik I: Das Wort*. Stuttgart/Weimar: Metzler
- Feldman, Laurie B. & Barac-Cikoja, Dragana (1996), "Serbo-Croatian: A Bilingual Language". In: Daniels & Bright (Hrsg.), 769–776
- Gelb, I[gnaz] (1963), *A Study of Writing*. Chicago: University Press
- Günther, Hartmut (1988), *Schriftliche Sprache: Strukturen geschriebener Wörter und ihre Verarbeitung beim Lesen*. Tübingen: Niemeyer
- Günther, Hartmut (1995), "Die Schrift als Modell der Lautsprache". *Osnabrücker Beiträge zur Sprachtheorie* 51, 15–32
- Günther, Hartmut & Ludwig, Otto (1994, Hrsg.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung*. Berlin: de Gruyter (Handbücher zur Sprach- und Kommunikationsforschung 10.1)
- Günther, Hartmut & Ludwig, Otto (1996, Hrsg.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung*. Berlin: de Gruyter (Handbücher zur Sprach- und Kommunikationsforschung 10.2)
- Katz, Leonard & Feldman, Laurie B. (1996), "The Influence of an Alphabetic Writing System on the Reading Process". In: Günther & Ludwig (Hrsg.), 1094–1101
- King, Ross (1996), "Korean Writing". In: Daniels & Bright (Hrsg.), 218–227
- Koch, Peter & Oesterreicher, Wulf (1994), "Schriftlichkeit und Sprache". In: Günther & Ludwig (Hrsg.), 587–604
- Krebernik, Manfred & Nissen, Hans J. (1994), "Die sumerisch-akkadische Keilschrift". In: Günther & Ludwig (Hrsg.), 274–288

- Ludwig, Otto (in Vorb.), *Geschichte des Schreibens*. Buchmanuskript
- Maas, Utz (1992), *Grundzüge der deutschen Orthographie*. Tübingen: Niemeyer
- Meisenburg, Trudel (1998). "Zur Typologie von Alphabetschriftsystemen anhand des Parameters der Tiefe." *Linguistische Berichte* 173, 43–64
- Meisenburg, Trudel (1996 a), *Romanische Schriftsysteme im Vergleich*. Tübingen: Narr
- Meisenburg, Trudel (1996 b), "Das spanische Schriftsystem". In: Günther & Ludwig (Hrsg.), 1437–1441
- Neef, Martin & Neijt, Anneke & Sproat, Richard (2002, Hrsg.), *The Relation of Writing to Spoken Language*. Tübingen: Niemeyer
- Olson, David R. (1993), "How Writing Represents Speech". *Language & Communication* 13, 1–17
- Ritner, Robert K. (1996), "Egyptian Writing". In: Daniels & Bright (Hrsg.), 73–84
- Saussure, Ferdinand de (1916), *Cours de linguistique générale* [édition critique préparée par Tullio de Mauro; 1974]. Paris: Payot
- Schenkel, Wolfgang (1994), "Die ägyptische Hieroglyphenschrift und ihre Weiterentwicklungen". In: Günther & Ludwig (Hrsg.), 289–297
- Schreuder, Rob & Jarvela, Robert & Job, Raimo & Sandström, Göran (1989), "Morphologische Strukturen im mentalen Lexikon". In: Günther, Hartmut (Hrsg.), *Experimentelle Untersuchungen zur deutschen Flexionsmorphologie*. Hamburg: Buske, 27–40
- Scragg, Donald G. (1974), *A History of English Spelling*. Manchester: Manchester University Press
- Stalphi, Jürgen (1996), "Das japanische Schriftsystem". In: Günther & Ludwig (Hrsg.), 1413–1427
- Stubbs, Michael (1996), "The English Writing System". In: Günther & Ludwig (Hrsg.), 1441–1445
- Wimmer, Hans & Goswami, Usha (1994), "The Influence of Orthographic Consistency on Reading Development: Word Recognition in English and German Children". *Cognition* 51, 91–103
- Wurzel, Wolfgang U. (1984), *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie

*Hartmut Günther, Köln (Deutschland)*

## 176. Terminology in special languages

1. Introduction
2. What are 'special languages'?
3. The lexicon in special languages
4. Term formation: morphological trends
5. Standardisation in term formation
6. Appendix
7. References

### 1. Introduction

Special languages are heavily dependent for their development on creative word formation processes. While the specialised vocabulary encountered in special languages has the same range of morphological structures as the general lexicon, it exhibits far greater regularity as a result of the deliberate and often systematic techniques of term creation.

This article outlines the distinguishing characteristics of specialised designation (cf. 2) and briefly examines the nature of the lexicon in special languages (cf. 3). The semantic content of these designations and the motivations for their expression forms are explored, and a description is provided of the prevalent morphological variations encountered in English term formation (cf. 4). (For German and French term formation see Drozd & Seibicke 1973; Hoffman 1988; Kocourek 1982.) An appendix (cf. 6) presents guidelines for the creation of terminology in which certain preferences in practical term formation are motivated.

The linguistic examples given are drawn from a variety of specialised subject domains since our objective is to make statements which apply to special languages in general.

### 2. What are 'special languages'?

**Special languages** (or **sublanguages**) are semi-autonomous, complex semiotic systems based on and derived from general language. Their use presupposes special education and is restricted to communication among specialists in the same or closely related fields (cf. Sager et al. 1980).

The features of special languages implied in this definition, namely special reference, special speech acts, special message forms and special pragmatic situations belong to quite separate aspects of language. From the point of view of function and structure, special languages highlight the referential and

systematic nature of language, maximally exploit the communicative and classificatory use of language and aim at a clear distinction between their three favoured intentions of speech acts, i.e. the informative, evaluative, and directive. This functional restriction of language use permits deliberate planning of special languages by means of rules for the creation and use of specialised linguistic forms.

### 3. The lexicon in special languages

The lexicon of special languages has developed its own particular characteristics which set it apart from the general lexicon. These characteristics reflect and arise from the referential and classificatory needs of special languages.

#### 3.1. Special reference

A description of the particular referential needs of special language is fundamental to an account of the lexical trends in specialised designation. (For a detailed exposition of the theory of reference in special languages see Sager et al. 1980: 70–86.)

The lexicon of a special subject language reflects the organisational characteristics of the conceptual domain by tending to provide as many lexical units as there are distinct concepts conventionally established in the knowledge structure of the domain, and by restricting the reference of each such lexical unit to a well-defined entity within it. It is this enhanced rigour which characterises **special reference**, and distinguishes it from **general reference**, where "general knowledge" tends to have a less well-defined, less "disciplined" structure.

In addition to a large number of items which have the property of special reference (**terms**), the lexicon of a special language also contains items of general reference which are not usually specific to any discipline and whose referential properties are vague or generalised (**words**). The criterion which determines whether a particular lexical item is a term or word is thus the nature of its reference within the discipline concerned.

The creation of new lexical items in special languages is a conscious and deliberate process carried out by subject specialists and is

thus part of the scientific method. Consequently, the naming of concepts is as subject to procedural regulation and systematisation as scientific investigation and has its own special rules and recommendations (cf. BS 3669 1963 and ISO 704 1986). The standardisation of terminology is discussed further in 5 and 6.

### 3.2. Linguistic modes of designation

Designation in special languages aims to achieve transparency and consistency. The structure of a designation often reflects the major conceptual features it represents, and related terms often have a comparable expression form.

The most common mode of term formation is determination, which operates by means of compounding and by derivation. On the basis of the superordinate concept (hyperonym), it creates a more specific term whose designation is motivated by systematic criteria, e.g. according to material (*iron chain*), function (*safety chain*), location (*door chain*), etc. Consecutive stages of determination produce sequences of terms in hierarchy.

Determination by means of pre-modification can use any word category as the specifying element (*pre-edit*, *virtual storage*, *T-iron*, *four square*, *feed hole*, *preshrunk fabric*, *programming language*). When a concept is determined by more than one modifier, it may be necessary to make use of hyphens to clarify the relationship (*wide-angle lense*). In some cases, the determined element is omitted and the new concept is represented solely by the modifier (*floppy disk* → *floppy*).

Conjunction is a much less frequent technique of term formation, although it does occur in designations of chemical compounds (*nickel steel*) and of dual purpose devices (*lifting and forcing pump*). Disjunction, which joins the extensions of two or more concepts into a new hyperonym by presenting two alternatives as a single concept, is used to express an ‘either-or’ relationship (*glide-sail parachute*).

## 4. Term formation: morphological trends

The three main linguistic methods of designation are:

- (a) the use of existing resources;
- (b) the modification of existing resources;
- (c) the creation of new resources.

The following description of the usage of these various modes of designation in English special languages is not intended to be exhaustive but is rather indicative of the range of possibilities (cf. Sager 1996: 250–253 for further references on scientific and technological word formation).

### 4.1. Use of existing resources

The creation of new designations by narrowing or widening the meaning of existing linguistic resources is frequent in pre-scientific enquiry and in popular expositions of science. The two principal methods used are polysemy and redefinition.

Simile is a common method of new designation which highlights a distinguishing feature of a concept through comparison (*G-clamp*, *star wheel*). Special language also uses many metaphoric expressions based on form, function or position (*ear of a hammer*). Metaphors are also common in popular naming in biology (*bluebell*, *seahorse*).

Other transfers of meaning occurring in special reference include the use of proper names to designate objects and measurements (*machmeter*, *watt*), the form of an object for another object of that form (*square*, *bulb*), and the name of the material for the object made of it (*glass*, *iron*).

The process of redefinition usually entails reducing the extension of a general language word. For example, *speed* and *velocity* are used synonymously in general reference, but in special reference they designate two distinct concepts (cf. Sager et al. 1980: 255 f.).

### 4.2. Modification of existing resources

The principal method of designation in special reference is the modification of existing resources by means of affixation, compounding, conversion or compression.

#### 4.2.1. Affixation

Affixation is a particularly important method of designation in special languages because of its ability to contribute to precision of expression and systematic reference.

Affixes are far more numerous in special languages than in general language, largely because of the borrowing or adaptation of neoclassical words and methods of word formation, many of which are not required in general language. English freely assimilates word elements and words from other languages, and in science and technology, in particular, makes substantial use of Greek

and Latin prefixes, suffixes, and stems (cf. Nybakken 1959; Flood 1960; Hogben 1969).

**Prefixes** make an important contribution to the systematic structuring of special vocabularies. Contrasting sets of terms can be created easily by the negatives *un-*, *in-*, *dis-*, *a-* (*synchronous/asynchronous*, *organic/inorganic*), the privatives *un-*, *de-*, *dis-* (*stable/unstable*, *code/decode*), the prefixes of degree or size *super-*, *sur-*, *sub-*, *hyper-*, *ultra-*, *mini-* (*subsonic*, *hypersonic*, *supersonic*), the locatives *super-*, *sub-*, *inter-*, *trans-* (*amination*, *transamination*), the temporal or sequential *pre-*, *post-*, *re-* (*post-emphasis*, *pre-emphasis*), and the numerals *uni-*, *mono-*, *bi-*, *di-* (*monovalent*, *divalent*), etc. Such sets maintain terminological unity and appear to be transparent. However, prefixes are not always used systematically and some prefixed words take on special meanings in the process of terminologisation (*denature* in chemistry means to render ethanol unfit for human consumption).

In special languages, certain neoclassical stems are so frequently used that they can be considered prefixes from the functional and structural point of view (*auto*, *equi*, *ext(e)r*, *iso*, *macro*, *mega*, *meta*, *micro*, *pan*, *para*).

The noun-noun transformation **suffix** *-ing* is highly productive in the conversion of countable nouns to mass nouns in order to indicate an assembly of objects (*wire* – *wires* – *wiring*). Concrete nouns are also derived from verbs by means of the suffix *-er* and its neoclassical alternative *-or* (*scanner*, *collator*). In this way, terminological agreements can be established between the action and the instrument of the action.

The causative Latin-derived pattern *-fy*, *-fier*, *-fication* also contributes to systematicity by providing names for action verbs, agents or instruments and the corresponding processes (*amplify* – *amplifier* – *amplification*). Agent nouns can also be formed by *-ant* and often have matching verbs in *-ate* (*penetrate* – *penetrant*). Most process nouns end in *-ation* or *-ion* (*collation*). They can also end in *-ing* but this is less desirable in specialised designation because of the danger of confusion with the identical concrete nouns designating the product (*carving*, *fitting*).

Properties are designated by adjectives derived from nouns, and by nouns derived from adjectives. Adjectival derivations in *-al*, *-ic*, *-ive*, and *-ous* are particularly prevalent (*trinomial*, *concentric*, *coercive*, *contiguous*). Property nouns are derived from adjectives ending mainly in *-al*, *-ent*, *-ic*, *-ous* (*trinominality*, *absorbency*, *concentricity*, *contiguity*).

Terms designating measurable properties frequently end in *-ance* or *-ivity* (*capacitance*, *inductivity*). Measuring devices are commonly designated by suffix-like expressions which indicate the method of measurement, e.g. *-meter* (*spectrometer*), *-graph* (*thermograph*).

Certain suffixes are restricted to one field in which they are used exclusively or with a special meaning (*-ase*, *-ate*, *-ol*, *-ose* in chemistry; *-asis*, *-itis*, *-oma*, *-ectomy* in medicine). However, even within a single field suffixes may have more than one function (*-ine* and *-ic* in chemistry).

#### 4.2.2. Backformation

The process of **backformation**, or back-derivation, permits the formation of complex verb forms which correspond very closely in their range of reference to nominal concepts of processes. It thus achieves a great deal of concentration of information while maintaining terminological accuracy.

Backformation in special languages occurs most often in complex nouns and in technology rather than science where there is a greater need for verbs to describe processes. It is closely linked to the process of compound verb formation which often terminologises verbs by narrowing their range of reference. The first element may designate the instrument or agent (*steam-bend*), the purpose (*test-drive*), the method (*colour-code*), or the location (*factory-test*).

#### 4.2.3. Compounding

**Compounds** play a very important role in special language designation where their potential for systematicity and regularity is maximally exploited to create coherent terminological systems.

The vast majority of special language compounds are nouns. Noun compounds provide an ideal basis for the building of terminological systems as the nucleus indicates the category to which the concept belongs and the determinant the criterion for the subdivision of the category (cf. Art. 87).

Three major types of noun compounds can be identified: those which designate entities, those which designate properties, and those which designate operations and processes. Most noun compounds fall into the first of these groups which exhibits the greatest variety of relationships between determinant and nucleus. Thus, the determinant may compare the nucleus to another object (*daisy wheel*); specify the material of which the nucleus is

made (*steel band*); express an inherent property of the new concept which is not inherent in the nucleus (*pneumatic drill*); specify the use to which the nucleus is regularly put (*splitting axe*); specify the object of an instrument or agent noun (*card punch*); express the product which is regularly associated with the nucleus (*cotton mill*); specify the mode of operation of the nucleus (*tumble-drier, slide-rule*); refer to the whole of the nucleus as a part (*brake lining*); assert identity with the nucleus (*fuel oil*) (for further exemplification of these and other possible relationships see Sager et al. 1980: 267–271).

Eponymic noun compounds are also frequent in special languages. While having the advantage of precision through unique differentiation, these compounds lack systematic import (*Venn diagram, Eddy's theorem*).

Compounds with particles occur in all word categories and form a large number of specialised terms (by *-chamber, -effect, -product; on -line, -stream*). There are also many phrasal compounds linked by prepositions (*on the fly printer*).

Compounds are often made up of two, three or more elements, but five- or six-element compounds are rare. Multiple-element compounds occur in schedules for classificatory purposes but in other text types they are shortened because contextual reference resolves any possible ambiguity.

Compound adjectives are rarely ambiguous and are therefore frequently used in special language where there is a need for economy and transparency of expression. Such compounds are used to determine processes and operations, and to identify new entities (*heat-resistant, lead-free, voice-activated*).

Most compound verbs in special languages are formed by backformation (cf. 4.2.2) or conversion. A third major group are compounds with the particles *out, over, under* (*underwrite, output, overcoal*).

#### 4.2.4. Conversion and compression

**Conversion** occurs most frequently from verb to noun and vice versa (*a drill, to telex*), but there are also cases from adjective to noun (*a solvent*), from noun to adjective (*imitation, model*), and from adjective to verb (*to dry, to smooth*). The productivity of conversion is reduced in scientific English where there is a high proportion of terms derived from Latin and Greek word elements with identifiable noun endings which are unsuitable for conversion.

Lexical **compression** in special languages is achieved by abbreviating existing expression forms, omitting elements in compounds, creating letter symbols, combining letters and numbers into short designations, and by using pictograms (cf. Sager et al. 1980: 277–280).

#### 4.3. The use of new resources

Absolute invention in special language designation is rare since it is appropriate to reflect in new designations the relationships of new concepts to existing ones. In English, the creation of neologisms is often hard to distinguish from Greek and Latin borrowing.

The principal methods of creating special language neologisms are the combination of existing English and anglicised Greek or Latin word elements (*videodisc*); the use of Latin or Greek stems (*apex*); blending (*transceiver*) (cf. Art. 91); the use of eponyms (*Mach number*); the use of letters, singly or in compounds (*beta-ray*); the rare use of absolute invention (*quark*).

Borrowing from other languages is relatively infrequent in contemporary English, but is widespread in languages which import new technology, often together with English terminology. Of morphological interest are the methods by which foreign borrowings are adapted to the requirements of inflection or derivation in the target language. This process is relatively smooth in English because of its limited inflectional endings, its weak requirements for identifying word classes and its great openness to conversion. The spelling is anglicised if a word enters into general usage, but even this offers no problems with the dual English tradition of Germanic and Romance vocabulary (cf. Rey 1995: 3–85).

#### 5. Standardisation in term formation

Once the validity of a new concept is generally acknowledged, it becomes established within the specialist community which usually regularises it by agreeing on a suitable linguistic designation. **Standardisation** is a separate process by which users reach ‘public’ agreement on the adoption of a given term for use in specific circumstances. In addition to fixing meaning, standardisation may also involve a choice from among competing terms using the criteria of economy, precision or appropriateness (cf. Cabré 1998: 194 ff.).

The mandatory fixation of form and meaning of a group of terms is exceedingly rare and can be as limited in application as a single legal document (contracts with definition clauses), or cover the entire activities of a professional group (the language used for formal communication between aircraft pilots and air traffic controllers). Contrary to appearances and popular belief, the terminology of so-called "standard vocabularies" is only rarely made mandatory. In most cases it represents a set of terms recommended, for use as established in a sequence of object or procedural standards and collected in the first instance to preserve consistency of reference in an extended series of interrelated standards (cf. Sager 1996: 114–128).

## 6. Appendix

The "Recommendations for the Selection, Formation and Definition of Technical Terms" (BS 3669 1963) issued by the British Standards Institute are guided by three principles: **appropriateness**, to the circumstances of communication, **precision**, of expression as demanded by specialised topics of communication, and **economy**, of expression as suitable for discourse among specialists.

The most important recommendation is that terms should be created systematically with respect to their morphological, syntactic, semantic, and pragmatic characteristics. Terms should conform to the morphology, spelling, and pronunciation conventions of the language for which they are intended, and should not give rise to ambiguities in pronunciation or irrelevant associations. They should be capable of permitting the creation of derivatives so that the expression and content forms can be made consistent wherever possible. Latin or Greek roots are particularly useful for this purpose.

Terms should be as short and concise as the requirements of precision and systematicity will permit. As function, mode of operation, production, and even origin can vary, it is recommended that in a complex term the determined constituent should be the genus of the new concept. In running text or spoken language terms may be abbreviated provided that the context is clear.

A term should represent only one concept. Homonymy may occur when a general language designation becomes a term or when a term from one area is applied in another. Homonyms should be avoided when two

## XXI. Morphologie und Nachbardisziplinen

meanings are very similar, when they may occur in the same context or when a term from one special subject is applied in another with a different definition.

A single concept should preferably be expressed by a single term. The creation of synonyms should be avoided because they may create confusion about the identity of a concept and because they burden the memory.

Once a term has gained wide acceptance it should not be changed without compelling reasons and a strong certainty that the new term will become accepted as a full substitute.

## 7. References

- BS 3669 (1963) = *Recommendations for the Selection, Formation, and Definition of Technical Terms*. London: British Standards Institution
- Cabré, M. Teresa (1998), *Terminology: Theory, Methods and Applications*, ed. by Juan C. Sager. Amsterdam: Benjamins
- Drozd, Lubomír & Seibicke, Wilfried (1973), *Deutsche Fach- und Wissenschaftssprache*. Wiesbaden: Brandstetter
- Flood, Walter Edgar (1960), *Scientific Words: Their Structure and Meaning*. London: Oldbourne
- Hoffman, Lothar (1988), *Vom Fachwort zum Fachtext*. Tübingen: Narr
- Hogben, Lancelot (1969), *The Vocabulary of Science*. London: Heinemann
- ISO 704 (1986) = *Principles and Methods of Terminology*. Geneva: International Organization for Standardization
- Kocourek, Rostislav (1982), *La langue française de la technique et de la science*. Wiesbaden: Brandstetter
- Nybakken, Oscar E. (1959), *Greek and Latin in Scientific Terminology*. Ames, Iowa: Iowa State College Press
- Rey, Alain (1995), *Essays on Terminology*, transl. & ed. by Juan C. Sager. Amsterdam: Benjamins
- Sager, Juan C. (1996), *A Practical Course in Terminology Processing*. Amsterdam, Philadelphia: Benjamins [1990]
- Sager, J[uan] C. & Dungworth, D[avid] & McDonald, P[eter] F. (1980), *English Special Languages: Principles and Practice in Science and Technology*. Wiesbaden: Brandstetter

*Jeanette M. Pugh, Manchester  
(Great Britain)*  
*Juan C. Sager, Manchester  
(Great Britain)*

## 177. Sprachunterricht

1. Eingrenzung des komplexen Themas
2. Konzepte für das Lehren und Lernen einer Sprache
3. Konzepte pädagogischer Grammatiken für den Fremdsprachenunterricht
4. Grammatik im muttersprachlichen Unterricht
5. Zitierte Literatur
  
1. Eingrenzung des komplexen Themas

**Sprachunterricht** ist in der Praxis wie in der Theorie ein hochkomplexes Geschehen, sowohl was den Unterricht selbst, das Lehren und Lernen, als auch was den Unterrichtsgegenstand, die Sprache, als auch was die Unterrichtsziele, das Beherrschende der Sprache in Wort und Schrift, betrifft.

Praxis und Theorie des Sprachunterrichts, des muttersprachlichen wie des fremdsprachlichen, können zudem auf eine lange Geschichte zurückblicken. Sie ist faßbar bereits in den antiken Grammatiken (Art. 6), von denen Donat und Priscian eine überragende Rolle für die Vermittlung des Lateinischen ebenso wie für die allmähliche Herausbildung nationalsprachlicher Grammatiken und Lehrbücher zur Vermittlung der europäischen "Nationalsprachen" gespielt haben, den antiken Rhetoriken, von denen Cicero und Quintilian intensiv rezipiert wurden (s. Gwynn 1926), schließlich auch in den antiken literarischen Werken selbst, insofern an und mit ihnen im Lateinunterricht Sprache gelernt wurde. Die Ziele, die in diesem Unterricht verfolgt wurden, sind vielfältig: Die Lernenden sollten diese Texte verstehen und bei entsprechendem Lernerfolg selbst Texte (prosaische wie Briefe, Reden, Traktate, Verträge und poetische wie Epigramme, Epen, Oden, Dramen) nach den antiken Exempla verfassen können. Außerdem sollten sie, spätestens seit dem Beginn des Humanismus in Italien im 15. Jahrhundert, das Lateinische als *lingua franca* in Europa in Wissenschaft, Verwaltung, Politik möglichst vollkommen in Wort und Schrift, rezeptiv wie produktiv, zu beherrschen lernen.

Während die Komplexität des Unterrichtsgegenstandes "Sprache" keiner weiteren Erläuterung bedarf, ist die Komplexität des Unterrichtsgeschehens selbst wenigstens stichpunktartig anzudeuten: Unterricht zielt durch

Lehrverfahren vielfältiger und unterschiedlichster Art darauf, Lernprozesse anzustoßen, zu ermöglichen, zu steuern, zu sichern, zu evaluieren. Die über der Praxis des Sprachunterrichts operierenden Theorien rezipieren und integrieren in vielfältiger und zum Teil unterschiedlicher Weise Erkenntnisse und Konzepte verschiedener wissenschaftlicher Disziplinen, insbesondere der Pädagogik, der Psychologie, der Linguistik, der Literaturwissenschaft. Insofern es sich um Fremdsprachenunterricht handelt, integrieren sie zudem fremdsprachendidaktische Konzepte sowie Ergebnisse der Zweitsprachenerwerbsforschung (s. Art. 166) und der Fremdsprachenlehr- und -lernforschung.

Erst vor dem Hintergrund dieses komplexen Geflechts interdependent Einflußfaktoren läßt sich das Thema "Sprachunterricht" so eingrenzen, daß seine Behandlung in dem vorgegebenen Rahmen nicht als inadäquate Verkürzung oder Simplifizierung mißverstanden wird. Der Beitrag fokussiert die folgenden Aspekte:

- Konzepte für das Lehren und Lernen einer Sprache (insbesondere im fremdsprachlichen, aber auch im muttersprachlichen Unterricht) in Vergangenheit und Gegenwart in Auswahl und im Vergleich mit dem Schwerpunkt auf der Morphologie (s. 2),
- Konzepte pädagogischer Grammatiken für den Fremdsprachenunterricht (s. 3),
- Grammatik im muttersprachlichen Unterricht (s. 4).

### 2. Konzepte für das Lehren und Lernen einer Sprache

#### 2.1. Überblick

Überblickt man die Konzepte für das Lehren und Lernen einer Fremdsprache seit der Antike, so lassen sich immer wiederkehrende Grundpositionen feststellen, die einander zum Teil zu widersprechen scheinen:

- (a) monolinguale Konzepte vs. bilinguale Konzepte (s. 2.2),
- (b) Konzepte induktiven Lernens, das von Texten ausgeht und Regularitäten durch die Lernenden selbst entdecken läßt, vs. Konzepte deduktiven Lernens, das von den durch die Grammatik und die Leh-

renden zu vermittelnden Regeln und grammatischen Kategorien ausgeht, diese auswendig lernen und danach erst auf Beispielsätze und Texte anwenden läßt (s. 2.3),

## 2.2. Monolinguale vs. bilinguale Konzepte

### 2.2.1. Monolinguale Konzepte

**Monolinguale Konzepte** propagieren die möglichst ausschließliche Verwendung der Zielsprache im Unterricht und außerhalb des Unterrichts. Sie versprechen sich davon vor allem zweierlei: (a) das Zurückdrängen interlingualer, potentiell Fehler verursachender Interferenz der Muttersprache auf die zu erlernende Fremdsprache sowie (b) einen dem Erwerb der Muttersprache vergleichbaren direkten Erwerb der Fremdsprache.

Prominente Beispiele aus dem 20. Jahrhundert sind (a) die im Kontext der Reformpädagogik der zwanziger Jahre propagierte Direkte Methode, (b) die in den vierziger Jahren in den USA entwickelte "Army Method" für die schnelle Vermittlung von Grundfertigkeiten in ganz unterschiedlichen Sprachen an Angehörige der US-Army, die nach dem Zweiten Weltkrieg zur Audio-lingualen Methode mit dem Fokus auf dem Hören und Sprechen weiterentwickelt wurde, und (c) die in Frankreich am CREDIF ausgearbeitete Audio-visuelle Methode als Variante der Audio-lingualen Methode, die sich im Anfängerunterricht in extensiver Weise visueller Hilfen in Form von Bilderfolgen in Dia-Reihen zur "unmittelbaren" Bedeutungsvermittlung sowie zum Einprägen und Memorieren ganzer Dialoge anhand dieser Bilderserien bedient.

Angestoßen wurde die Entwicklung dieser Methoden aber bereits im 19. Jahrhundert, insbesondere durch die unter dem Pseudonym "Quousque tandem" publizierte grammatische Schrift des Anglisten Wilhelm Vietor mit dem Titel "Der Sprachunterricht muß umkehren", der eine radikale Abkehr von der im 19. Jahrhundert im Unterricht der klassischen Sprachen Latein und Griechisch praktizierten Lehr- und Lernverfahren forderte, die angeblich nur auf die Lektüre antiker Literatur und deren Übersetzung auf der Basis intensiven Grammatikunterrichts zielten.

Allerdings greift diese Auseinandersetzung, historisch betrachtet, zu kurz. Denn mit Sicherheit bis weit in das 18. Jahrhundert hinein zielte der Unterricht im Lateinischen auch auf die produktive Beherrschung der la-

teinischen Sprache in Schrift und Wort, war sie doch über Jahrhunderte hinweg die *lingua franca* Europas, die Sprache nicht nur der Kirche, sondern vor allem auch der Wissenschaften, in der sich die Vertreter aller Fakultäten mündlich wie schriftlich verständigten und ihre Werke publizierten. Das belegen die Unterrichtswerke selbst (beispielsweise schrieb Melanchthon für zwei seiner adeligen Privatschüler eine Grammatik des Griechischen für Lernzwecke in lateinischer Sprache), die überlieferten Unterrichtskonzepte (beispielsweise Comenius 1642) und die Schulordnungen für die Lateinschulen jener Zeit, in denen ausdrücklich gefordert wurde, daß sich die Schüler nicht nur in der Schule, sondern auch außerhalb der Schule des Lateinischen befleißigen sollten. In diesen Schulen wurde von Anfang an lateinisch gelesen, geschrieben, gehört und gesprochen; ja selbst die Grundlagen der Grammatik wurden mündlich im katechetischen Verfahren gelehrt und gelernt, bei dem die Schüler ihren Lehrer und der Lehrer seine Schüler nach festgelegten Mustern fragte:

"Partes orationis quot sunt? Octo. Quae? Nomen pronomen verbum adverbium participium coniunctio praepositio interiectio. Nomen quid est? Pars orationis cum casu corpus aut rem proprie communiter significans. Nomini quot accident? Sex. Quae? Qualitas comparatio genus numerus figura causus. ... Comparationis gradus quot sunt? Tres. Qui? Positivus, ut doctus, comparativus, ut doctior, superlativus, ut doctissimus ..." (Donati de partibus orationis ars minor, Keil 1864, Hrsg.: 355)

Der Unterschied zwischen diesem einsprachigen Konzept des Lateinlernens und den einsprachigen Konzepten zur Erlernung moderner Fremdsprachen ist freilich gravierend: In diesem Lateinunterricht wurde zunächst und vor allem das grammatische Kategoriensystem der acht Wortarten, meistens durch Beispiele illustriert, gelernt, im Grunde also eine Mischung von Primär- und Metasprache, bevor die Lektüre begann und die sprachlichen Muster durch dauernde mündliche Wiederholung eingeprägt wurden. Daß dieser Unterricht für die noch recht jungen Schüler häufig genug zur Qual wurde, belegen vor allem die heftigen Attacken bedeutender Humanisten (wie Erasmus, Comenius und vieler anderer), die das Nachplappern von nicht Verstandenen als sinn- und erfolglose Mühe verurteilten und statt dessen forderten, daß nur das gelernt werden könnte und sollte, was zuvor in seiner Bedeutung verstanden worden war.

Wie wurde und wird in einsprachigen Konzepten die Morphologie der Fremdsprache gelehrt und gelernt? Da auf die Muttersprache als Vergleichsbasis und als Erklärssprache zu verzichten ist, bleiben als Möglichkeiten vor allem (a) die direkte zielsprachliche Präsentation morphologischer Lerninhalte in Form von primärsprachlichen Beispielsätzen und tabellarischen Übersichten (Deklinations- und Konjugationsparadigma), (b) die metasprachliche Formulierung der jeweiligen Regularitäten in der Zielsprache, (c) das Reproduzieren und Variieren der jeweiligen Formen in zum Teil monotonen Übungen im Stile der "pattern drills" der audio-lingualen Methode zwecks unbewußter Automatisierung dieser Formen, (d) das Hoffen darauf, daß die "eingedrillten" Formen in der "freien" Sprachproduktion automatisch richtig verwendet werden – eine Hoffnung freilich, die sich nicht erfüllt, wenn nicht vielfältige Übungen zum produktiven Sprachgebrauch hinzutreten. Was demgegenüber bei strikter Handhabung des einsprachigen Prinzips als Möglichkeit ausscheidet, ist die Zuhilfenahme der Muttersprache als Vergleichsbasis und Kommunikationsmittel.

### 2.2.2. Bilinguale Konzepte

**Bilinguale Konzepte** propagieren den Einsatz der Muttersprache im Fremdsprachenunterricht in vielerlei Hinsicht, (a) zur Vermittlung der Bedeutung von Wörtern und Wortgruppen, (b) zur Formulierung grammatischer Regeln, (c) als Basis für den Vergleich der Zielsprache mit der Muttersprache, (d) zur Herüber- und Hinüberübersetzung sowohl zu Übungszwecken als auch als übergeordnetes Lernziel, (e) als Unterrichtssprache. Sie versprechen sich davon ein schnelleres und genaueres Verstehen der Bedeutung von Lexik und Grammatik, die Vermeidung von Interferenzfehlern durch Herausarbeitung der Unterschiede zwischen Zielsprache und Muttersprache und einen schnelleren Lernfortschritt.

Beispiele aus dem 20. Jahrhundert sind: (a) die von Butzkamm (1989) u. a. propagierte bilinguale Methode, bei der in Abkehr von dem zu engen Korsett der audio-lingualen Methode ein maßvoller und jeweils didaktisch genau zu begründender Einsatz der Muttersprache angeraten wird, (b) Suggestopädie und Superlearning, wobei in strenger Dogmatik der Ausgang vom muttersprachlichen Wort, Satzteil und Satz vorgeschrieben wird, denen jeweils die zielsprachlichen Wör-

ter, Formen und Strukturen unmittelbar zugeordnet werden, und (c) der Unterricht in Latein und Griechisch, in dem das Herüberübersetzen authentischer antiker Texte sowie deren Analyse und Interpretation das komplexe, erst allmählich zu erreichende Ziel ist, auf das hin jedoch alle didaktischen Überlegungen und methodischen Konzepte von Anfang an ausgerichtet sind.

Dieses didaktische Konzept des Lateinunterrichts wurde bereits im 19. Jahrhundert entwickelt und als generelles Konzept auf den Unterricht in den modernen Sprachen übertragen – was dann von Vietor und anderen (wie bereits gezeigt) heftig kritisiert wurde. Gleichwohl wird dieses Konzept der sogenannten Grammatik-Übersetzungsmethode bis in die Gegenwart zumindest in Teilespekten (z. B. muttersprachliche Erklärungen, Übersetzungsübungen) noch praktiziert. Dies gilt insbesondere für homogene Lernergruppen im Land der jeweiligen Muttersprache, auch wenn die Zielsetzung nicht auf Lesen, Verstehen und Übersetzen reduziert bleibt, sondern daneben und darüber hinaus auch Hören und Sprechen entwickelt werden sollen – mit häufig unbefriedigendem Erfolg, weil der Unterricht zu sehr durch ausschließlich deduktive Grammatikvermittlung und durch Übersetzungsübungen dominiert wird.

Ansätze des bilingualen Konzepts zur Vermittlung der Fremdsprache Latein sind bereits im frühen Mittelalter, wenn auch in sehr bescheidener Form, faßbar und weiten sich erst langsam in dem Maße aus, wie die *lingua vernacula* sich mit und an dem lateinischen Vorbild lexikalisch und grammatisch entwickelt; das vollzieht sich im romanisch-sprachigen Raum (Italien, Spanien, Frankreich) wesentlich früher und schneller als beispielsweise im deutschsprachigen Raum, bis auch hier aufgrund großer und zum Teil überspannter Anstrengung das Deutsche aufgrund extensiver Nutzung der Möglichkeiten der Wort(neu)bildung etwa Mitte des 17. Jahrhunderts einen Zustand erreicht hat, der die Sprachwissenschaftler wie Schottel und die Poeten und Poetiker wie Harsdörffer und Zesen nicht nur die Gleichrangigkeit, sondern sogar die Überlegenheit des Deutschen gegenüber dem Lateinischen und aller zeitgenössischen europäischen Sprachen konstatieren läßt.

Wie wurde und wird in diesen bilingualen Konzepten die Morphologie der Zielsprache gelehrt und gelernt? In den der antiken Grammatik in Gestalt der Ars Minor des Do-

nat verpflichteten zweisprachigen Grammatiken des Lateinischen seit dem 15. Jahrhundert sind folgende Verfahren zu beobachten: (a) Zunächst und vor allem wird das grammatische Kategoriensystem der *octo partes orationis* in Frage und Antwort gelehrt, wobei den lateinischen Beispielwörtern für die einzelnen Kategorien Übersetzungen dieser Beispielwörter in die jeweilige Muttersprache folgen. (b) Die metasprachlichen Fragen nach den Kategorien und die Antworten und Definitionen werden in einem weiteren Schritt ebenfalls in die jeweilige Muttersprache übersetzt, wobei die muttersprachliche Formulierung der zielsprachlichen lateinischen nachfolgen oder vorangehen kann. (c) Die Formen des Nomens und des Verbs erscheinen in systematisch geordneter und übersichtlicher tabellarischer Form, manchmal in Form kurzer Sätze in der Zielsprache, wiederum gefolgt von der Übersetzung in die Muttersprache. (d) Das Erlernen des kompletten morphologischen Systems dient (in Verbindung mit dem Erlernen eines Grundwortschatzes anhand von zweisprachigen Vokabellisten oder Wörterbüchern) zunächst dem Übersetzen lateinischer Texte ins Deutsche, sodann in zunehmendem Maße auch dem Verfassen lateinischer Texte; dabei wird implizit und kontinuierlich das fremdsprachliche morphologische System geübt. (e) Der Einbezug der Muttersprache in den Sprachunterricht dient vorrangig dem Erwerb der Zielsprache Latein, darüber hinaus aber auch der Entwicklung und Ausbildung der muttersprachlichen Kompetenz. (f) Gelegentlich wird die Einführung in die Grammatik der Muttersprache auch dem Lateinunterricht vorgeschaltet und verstanden als Propädeutikum des Lateinlernens.

Im gegenwärtigen Lateinunterricht finden sich viele dieser Merkmale wieder; ein wesentlicher Unterschied besteht freilich darin, daß von Anfang an mit lateinischen Texten gearbeitet wird, die es zu verstehen und zu übersetzen gilt; anhand dieser Texte erfolgt die Herausarbeitung morphologischer und syntaktischer Regularitäten, die in Form von Tabellen und Regeln sowie anhand von Beispielsätzen in grammatischen Übersichten in den Lehrbüchern und in eigenen Grammatiken des Lateinischen, in denen die Meta- und Erklärsprache die Muttersprache ist, zusammengestellt sind.

Analoge Übersichten über die Morphologie und Syntax finden sich in Lehrbüchern und Grammatiken der modernen Fremdspra-

chen; diese Übersichten kommen entweder ganz ohne Regelformulierungen aus oder bieten solche Formulierungen in möglichst einfacher Form in der Zielsprache oder in der Muttersprache. Sofern sie zweisprachig sind, kontrastieren sie darüber hinaus die morphologischen Systeme der Muttersprache und der Zielsprache, um das Verstehen zu erleichtern und um mögliche Interferenzfehler zu verhindern.

### 2.3. Konzepte induktiven vs. deduktiven Lernens

#### 2.3.1. Konzepte induktiven Lernens

**Konzepte induktiven Lernens** im Sprachunterricht zielen auf einen Lernprozeß, in dessen Verlauf die Lernenden selbst sprachliche Regularitäten entdecken und internalisieren und dabei zugleich die kognitiven Fähigkeiten des Vergleichens, Abstrahierens und Generalisierens entwickeln. Diese Konzepte werden des Weiteren mit dem Hinweis darauf begründet, daß die selbst entdeckten Regularitäten besser verstanden und behalten werden als vorformulierte Regeln, die auswendig zu lernen sind. Erkenntnisse der Gehirnforschung, nach denen Behaltensleistungen nach Differenziertheit und Dauer von der Intensität und Tiefe der kognitiven Verarbeitung von Informationen abhängen, tragen zur Fundierung dieser Konzepte bei: Je mehr Gehirnregionen gleichzeitig und parallel intensiv arbeiten, desto wahrscheinlicher ist es, daß die jeweilig bearbeiteten Sachverhalte aufgrund stabiler molekularer Veränderungen im sog. Langzeitgedächtnis gespeichert werden können.

Dabei ist jedoch zu beachten, daß viele Regularitäten einer Sprache aufgrund ihrer Kompliziertheit, Komplexität und Anzahl der "Ausnahmen" auf diesem Wege von den Lernenden selbst nicht oder nicht in einem überschaubaren und vertretbaren Zeitraum zu entdecken sind. Beispiele solcher komplexen und komplizierten Regularitäten sind: Gebrauch des definiten, des indefiniten und des Null-Artikels im Deutschen, der von der Linguistik mittels 55 Regeln – die Ausnahmen nicht mitgerechnet – beschrieben worden ist; Deklination der Nominal-Gruppe im Deutschen. Daneben gibt es aber eine Vielzahl von Regularitäten, die aufgrund ihrer Reichweite (keine oder nur ganz wenige Ausnahmen) und internen, vergleichsweise einfachen Struktur von den Lernenden in einem methodischen Drei-Schritt selbst entdeckt

werden können: (a) Sammeln/Ordnen (b) Vergleichen/Systematisieren und (c) Beschreiben.

Die Propagierung und Realisierung dieser Konzepte induktiven Lernens stehen in einem unlösbaren Zusammenhang mit grundlegenden pädagogischen Konzepten, welche die Lernenden, ihre Lernprozesse und Lernstrategien konsequent fokussieren, denen erst in einem zweiten Schritt geeignete Lehrprozesse, welche die Lernprozesse ermöglichen, fördern, und sichern sollen, zuzuordnen sind.

Beispiele für solche Konzepte finden sich überwiegend erst im 20. Jahrhundert, sind aber ansatzweise und indirekt bereits in früheren Jahrhunderten nachweisbar, zum Beispiel in der von Comenius propagierten und praktizierten Trias *exempla – pracepta – exercitia*, bei der die Beispiele den Regeln immer vorangehen sollen:

“Per Exempla, Praecepta, Exercitia, docentur et discuntur omnia ... Exemplum semper praecedat, Praeceptum semper sequatur, Imitatio semper urgeatur ... Per Exempla facilius discitur, quam per Praecepta. Facilius tamen per utrumque junctim. Sed ut Exempla praecedant ...” (Comenius 1642: 102 ff.)

Der Prozeß der Ermittlung der Regularitäten (“pracepta”) auf der Basis konkreten sprachlichen Materials (“exempla”) wird in Konzepten induktiven Lernens also von jedem Lernenden selbst vollzogen. Diese Konzepte haben zusammen mit anderen Einflußfaktoren etwa seit den siebziger Jahren des 20. Jahrhunderts zur Entwicklung von Fremdsprachen-Lehrwerken und von pädagogischen Grammatiken geführt, in denen selbständiges und selbsttätiges Lernen im Mittelpunkt aller didaktischen Überlegungen und methodischen Konkretisierungen steht.

Wie wurde und wird in Konzepten induktiven Lernens die Morphologie der Zielsprache gelernt? Die Lernenden selbst analysieren, spontan oder durch entsprechende Aufgabenstellungen dazu aufgefordert, das präsentierte sprachliche Material mit dem Ziel, Regularitäten der Formenbildung und des Gebrauchs dieser Formen zu entdecken und in Paradigmata und Regeln (zunächst in der Muttersprache, dann zunehmend auch in der Zielsprache) abzubilden. Diese Analyse läßt sich in die drei bereits erwähnten Analysestufen gliedern: (a) Sammeln/Ordnen (b) Vergleichen/Systematisieren und (c) Beschreiben. Sie operiert auf und vor dem Hintergrund sowohl des muttersprachlichen Kön-

nens und des sich allmählich parallel herauskristallisierenden Wissens, als auch des jeweils erworbenen zielsprachlichen Wissens und Könnens. Der Analyse müssen sodann vielfältige Übungen im Erkennen und im Gebrauch der jeweiligen Formen folgen.

Zu beachten ist, daß die einzelnen morphologischen Subsysteme (z. B. das Tempusystem) weder in einem einzigen Zugriff noch *in toto* behandelt und internalisiert werden können, sondern in mehreren, aufeinander aufbauenden Zugriffen und in einem Grad von “Vollständigkeit”, der didaktisch (auf der Ebene der Lernzielermittlung und -beschreibung) und nicht primär linguistisch (im Sinne des Vollständigkeitspostulats) zu definieren ist.

Hochkomplexe Strukturen mit einem hohen Maß an Irregularitäten jedoch sind mit Konzepten induktiven Lernens allein nicht zu bewältigen.

### 2.3.2. Konzepte deduktiven Lernens

**Konzepte deduktiven Lernens** im Sprachunterricht zielen auf einen Lernprozeß, in dessen Verlauf die Lernenden die in z. T. komplexen und zudem kompliziert formulierten Regeln generalisierten Regularitäten einer Sprache lesen und auf ausgewählte, meist kontextlose Beispielwörter und -sätze anwenden. Die Lernenden werden dabei von Anfang an mit abstrakter linguistischer Metasprache in komprimierter Form konfrontiert, welche das Verstehen erschwert, wenn nicht gar verhindert, weil sie vor dem konkreten Gegenstand, der Sprache, die abstrakte und komprimierte Beschreibung dieses Gegenstands kennenzulernen und verstehen sollen. Ein solches Verfahren läßt sich pädagogisch und lernpsychologisch schwerlich begründen. Gleichwohl scheint es sich immer noch einer gewissen Beliebtheit zu erfreuen. Über die Gründe dafür läßt sich nur spekulieren: Sicherheit seitens der Lehrperson aufgrund ihres überlegenen Wissens, scheinbare Zeitersparnis, auch in anderen Fächern praktizierte Vermittlung von Fachwissen, Verankerung in der Tradition der Vermittlung des Lateinischen und Griechischen.

Wie wurde und wird in Konzepten deduktiven Lernens die Morphologie der Zielsprache gelehrt und gelernt? Lehrbücher und Lehrpersonen präsentieren die jeweiligen Formen der zu erlernenden Sprache in enger Anlehnung an die linguistische Beschreibung dieser Formen in linguistischen Grammatiken in Form von Paradigmata, Regeln und

Ausnahmen von den Regeln in der Muttersprache der Lernenden; bei Einbeziehung des kontrastiven Aspekts erfolgt darüber hinaus ein Vergleich mit den jeweiligen Formen der Muttersprache, insbesondere dann, wenn systematische Unterschiede negativen Transfer möglich oder wahrscheinlich erscheinen lassen. Als Beleg für die jeweiligen Regeln und Ausnahmen folgen sodann Beispielwörter und Beispielsätze. Die Lernenden müssen als erstes die schwierigen metasprachlichen Regelformulierungen verstehen, bevor sie sie auf primärsprachliche Wörter, Sätze und Texte anwenden. Während dieses Verfahren bei der Rezeption von Texten in schriftlicher Form sinnvoll sein kann, kann es bei der schriftlichen und allemal bei der mündlichen Produktion dazu (ver)führen, Sätze durch parallele Anwendung einer Vielzahl von Einzelregeln konstruieren zu wollen, was die Konzentration auf den Inhalt der Aussage erschwert und die Flüssigkeit der Äußerung erheblich beeinträchtigen kann. Deshalb kommen auch deduktive Konzepte ohne viele und vielfältige Übungen, die auf "Automatisierung" der zu erlernenden Formen zielen, nicht aus.

### 3. Konzepte pädagogischer Grammatiken für den Fremdsprachenunterricht

Konzepte pädagogischer Grammatiken für den Fremdsprachenunterricht basieren nicht alleine auf linguistischen Beschreibungen der jeweiligen Zielsprache, sondern darüber hinaus auf lernpsychologischen Erkenntnissen allgemein sowie Erkenntnissen der Kognitionswissenschaft (*Cognitive Sciences*) und

der Zweitsprachenerwerbsforschung (s. Art. 166) sowie zunehmend auch der Fremdsprachenlehr- und -lernforschung. Da dem Fremdsprachenunterricht aufgegeben ist, das Verstehen und den Gebrauch der Fremdsprache bei den Lernenden zu entwickeln, führt das zu Funktionsbestimmungen und Realisierungsformen von Grammatik, die sich von Funktionsbestimmungen und Realisierungsformen rein linguistischer Grammatiken zum Teil deutlich unterscheiden.

Autoren pädagogischer Grammatiken für den Fremdsprachenunterricht bedienen sich deshalb auch morphologischer Theorien und der auf diesen Theorien basierenden Darstellungen der Morphologie in eklektischer und souveräner Weise: Sie sichten den Formenbestand einer Sprache und die den Formen, ihren Funktionen und ihrem Gebrauch innewohnenden Regularitäten. Sie analysieren diese Regularitäten im Hinblick auf ihre Reichweite (den Grad der ihnen innewohnenden Irregularitäten) und ihren Komplexitätsgrad. Die Ergebnisse dieser Analysen sind eine wichtige Basis für die Auswahl, Progression und Präsentation unter lernpsychologischen Gesichtspunkten.

Der Wert einzelner morphologischer Theorien, wie beispielsweise des Strukturalismus, bestimmt sich danach, ob und gegebenenfalls bis zu welchem Grad die auf ihnen basierenden Beschreibungen den Anforderungen des Lehr- und Lernprozesses entsprechen. Abstrakte Darstellungen sind von weit geringerem Wert als Darstellungen, die anhand von kontextualisierten Beispielen die jeweiligen Regularitäten besser verstehbar und behaltbar machen. Metasprachlich möglichst einfache Regelformulierungen mit Fokussierung

Linguistische Grammatik	Lerner-Grammatik
Totalität (Ausnahmen von der "Regel" besonders wichtig)	Auswahl
Abstraktheit (der Beschreibung/Darstellung)	Konkretheit/Anschaulichkeit (der Abbildung/Darstellung)
Kürze (der Darstellung)	Ausführlichkeit (der Darstellung der als wichtig erkannten Elemente)
Keine lernpsychologischen Vorgaben/ Rücksichten	Lernpsychologische Kategorien (Verstehbarkeit, Behaltbarkeit, Anwendbarkeit)

Tab. 177.1: Wesentliche Unterschiede zwischen einer linguistischen und einer Lerner-Grammatik (Schmidt 1992: 163)

des Wesentlichen sind von weit größerem Wert als in zum Teil eigenwilliger Metasprache gefaßte Regelformulierungen mit Fokussierung aller nur denkbaren Ausnahmen und Ausnahmen von den Ausnahmen.

Pädagogische Grammatiken für den Fremdsprachenunterricht lassen sich deshalb auch nicht einfach aus linguistischen Grammatiken durch Kürzung und Vereinfachung deduzieren, sondern sind Sprachbeschreibungen *sui generis* (s. Art. 11).

Zentrale Funktionen solcher Grammatiken sind:

- (a) Sie sollen die Einsicht der Lernenden in die Regularitäten grammatischer Formen, Strukturen und Funktionen fördern.
- (b) Sie sollen das Einprägen, das Behalten und die Abrufbarkeit dieser Regularitäten unterstützen.
- (c) Sie sollen die möglichst fehlerlose Beherrschung der Fremdsprache entwickeln helfen.
- (d) Sie sind damit "nur" Instrument, nicht Selbstzweck.

#### 4. Grammatik im muttersprachlichen Unterricht

Grammatik im muttersprachlichen Unterricht unterscheidet sich prinzipiell ebenso von rein linguistischen Grammatiken wie Grammatik im Fremdsprachenunterricht, insofern auch bei dieser Grammatik lernpsychologische Erkenntnisse sowie pädagogische und didaktische Konzepte eine entscheidende Rolle spielen.

Grammatik im muttersprachlichen Unterricht setzt – im Unterschied zu Grammatik im Fremdsprachenunterricht – die je nach Altersstufe und individuellen Lernbiographien unterschiedlich weit fortgeschrittene Beherrschung der Muttersprache voraus. Vor diesem Hintergrund dient sie vor allem der kognitiven Auseinandersetzung mit den Formen und Strukturen der eigenen Sprache sowie deren Funktionen und Leistungsfähigkeit in Wort und Schrift. Damit lassen sich Zielsetzungen auf unterschiedlichen Ebenen verbinden, die zudem noch nach Altersstufen variieren können, z. B.: Unterstützung beim Erlernen der Orthographie (auf der Primarstufe) mittels spielerischem und entdeckendem Umgang mit der eigenen Sprache; Einsicht in grundlegende Regularitäten und Funktionsweisen der eigenen Sprache zwecks

Verbesserung der mündlichen und Entwicklung der schriftlichen Kompetenz sowie zwecks Förderung der Fähigkeit zur Textanalyse und -interpretation (auf der Sekundarstufe I); Entwicklung der Fähigkeit zu Analyse, Generalisierung und Abstraktion ("formale" Bildung); Kennenlernen und Verstehen linguistischer Kategorien und Terminologie(n) (und damit in gewisser Weise Vorbereitung und Entlastung des nachfolgenden oder parallel einsetzenden Unterrichts in einer ersten Fremdsprache).

Welchen Stellenwert hat Morphologie in diesen Zielhorizonten? Die Behandlung morphologischer Themen im muttersprachlichen Unterricht dient vor allem dazu, die Morphologie der eigenen Sprache zu entdecken und die metasprachliche Beschreibung morphologischer Sachverhalte zu verstehen. Das impliziert zunächst und vor allem das (Kennens-)Lernen grundlegender Kategorien, insbesondere der Wortarten und ihrer Akzidentien (z. B. Verb; die Tempora wie Präsens, Perfekt usw., die Genera wie Aktiv und Passiv, die Modi wie Indikativ, Konjunktiv usw.); unterschiedliche linguistische Verfahren, Modelle und Terminologien stellen dabei ein großes Problem dar, insofern sie zu erheblichen Konfusionen in den Köpfen der Lernenden führen können, wenn sie beispielsweise auf der Primarstufe mit anderen Begriffen konfrontiert werden als auf der Sekundarstufe I, oder wenn zeitlich parallel im muttersprachlichen Unterricht andere Begriffe für dieselbe Sache verwendet werden als im fremdsprachlichen Unterricht.

Des weiteren war und ist Grammatik im muttersprachlichen Unterricht immer beeinflußt von Entwicklungen in der Sprachwissenschaft: Während bis weit ins 20. Jahrhundert hinein Verfahren, Modelle und Kategorien der lateinischen Grammatik die Grundlage grammatischer Darstellungen in Sprachbüchern gewesen sind, haben etwa seit der Mitte des 20. Jahrhunderts Entdeckungsprozeduren, Modelle, Kategorien und Begrifflichkeiten des Strukturalismus verstärkt in Sprachbüchern und damit auch im Sprachunterricht Platz gegriffen. Exemplarisch verwiesen sei auf den "Sprachspiegel" von Hans Glinz (1969, Hrsg.), in dem Entdeckungs- und Analyseprozeduren des Strukturalismus zu von den Schülern handhabbaren operationalen Werkzeugen für den entdeckenden Umgang mit der eigenen Sprache gemacht werden: Der Permutationstest wird zur Umstellprobe, mit deren Hilfe Satzglieder ent-

deckt und abgegrenzt werden sollen; der Deletionstest wird zur Weglaßprobe, mit deren Hilfe Kernsätze entdeckt werden sollen; die Infinitivierung wird dazu verwendet, das Subjekt (oder die Nominalphrase des Knotens Satz im strukturalistischen Syntax-Modell) von dem Rest des Satzes (oder der Verbalphrase des strukturalistischen Modells) zu trennen, und so fort. Zugleich werden allerdings auch neue Begrifflichkeiten eingeführt (z. B. Grundgröße für Subjekt, Gleichgröße für Prädikatsnomen, Anteilgröße für Genitivobjekt, Zuwendgröße für Dativobjekt usw.), die sich weder im muttersprachlichen Deutschunterricht generell noch gar im Fremdsprachenunterricht durchsetzen konnten.

Neuere Entwicklungen innerhalb der Linguistik, insbesondere die Erweiterung des Analysegegenstands über Morphem, Syntagma und Satz hinaus auf Text, Kommunikation, sprachliches Handeln, Spracherwerb und die damit einhergehende Untergliederung der Linguistik in weitere Teildisziplinen (Textlinguistik, Pragmalinguistik, Psycholinguistik u. a.) haben sich ebenfalls auf den muttersprachlichen Unterricht und den Teilbereich "Reflexion über Sprache", innerhalb dessen Grammatikunterricht wiederum einen Teilbereich darstellt, ausgewirkt. In ihrem Buch "Der andere Grammatikunterricht" fragen Boettcher & Sitta "... nach Begründungen, Prinzipien und Realisierungsmöglichkeiten von Grammatikunterricht im Rahmen des Lernfelds 'Reflexion über (sprachliche) Kommunikation' im Kontext eines insgesamt schüler- und situationsbezogenen Deutschunterrichts" (Boettcher & Sitta 1978: 5).

Welche Bedeutung kommt in einem solchen Konzept noch der Vermittlung der Einsicht in die Morphem-, Phrasen- und Satzstrukturen der Muttersprache zu? Generell läßt sich feststellen, daß die Reflexion über diese Strukturen instrumentalisiert wird für die Lösung von sprachlichen und kommuni-

## XXI. Morphologie und Nachbardisziplinen

kativen Problemen in für die Schüler relevanten Situationen außerhalb der Schule, aber auch innerhalb der Schule und innerhalb einzelner Fächer. Es bleibt aber zu fragen, ob und in welcher Weise in diesem anderen Grammatikunterricht eine auch situationsunabhängige systematische Vermittlung der o. g. Strukturen erfolgen kann oder darf, welche die Schüler kennengelernt haben müssen, wenn sie mit ihnen arbeiten sollen.

### 5. Zitierte Literatur

- Boettcher, Wolfgang & Sitta, Horst (1978), *Der andere Grammatikunterricht*. München usw.: Urban & Schwarzenberg
- Butzkamm, Wolfgang (1989), *Psycholinguistik des Fremdsprachenunterrichts. Natürliche Künstlichkeit: Von der Muttersprache zur Fremdsprache*. Tübingen: Francke
- Comenius, Johann Amos (1642), *Didacticorum Operum Pars II, (Text II: Novissima Linguarum Methodus, Fundamentis Didacticis solide superstructa &c.)* [Konstanz: Archiv für Fremdsprachenvermittlung 1978]
- Glinz, Hans (1969, Hrsg.), *Deutscher Sprachspiegel: Sprachgestaltung und Sprachbetrachtung*. Düsseldorf: Schwann
- Gwynn, Aubrey (1926), *Roman Education from Cicero to Quintilian*. Oxford: Oxford University Press [New York: Teachers College Press o. J. (Classics in Education 29)]
- Keil, Heinrich (1864, Hrsg.), *Probi Donati Servii Qui feruntur De arte grammatica libri*. Leipzig: o. V. [Hildesheim: Olms 1961]
- Schmidt, Reiner (1992), "Das Konzept einer Lern(er)-Grammatik für DaF und seine linguistischen und lernpsychologischen Grundlagen". In: Leirbukt, Oddleif & Lindemann, Beate (Hrsg.), *Psycholinguistische und didaktische Aspekte des Fremdsprachenlernens*. Tübingen: Narr, 159–177

Reiner Schmidt, Bielefeld (Deutschland)

## Verzeichnis der Abkürzungen / List of Abbreviations

### 1. Deutsch

Ø	Null	INCH	Inchoativ
1	1. Person	IND	Indikativ
2	2. Person	INDEF	indefinit
3	3. Person	INF	Infinitiv
ABL	Ablativ	INSTR	Instrumental
ABS	Absolutiv (Kasus/Funktion des ergativischen Systems)	INT	Interrogativ
ADJR	Adjektiv(alis)ator	INTS	Intensiv
ADVR	Adverbalisator	IO	indirektes Objekt
AG	Agens	IPFV	Imperfektiv
AKK	Akkusativ	IRR	Irrealis
AKT	Aktiv	KAUS	Kausativ
ALL	Allativ	KLN	Nominalklasse n
AOR	Aorist	KMPR	Komparativ
APPL	Applikativ	KOMIT	Komitativ
ART	Artikel	KOMP	Kompletiv
ASRT	Assertiv	KOND	Konditional
AUG	Augmentativ	KONJ	Konjunktiv
AUX	Auxiliar	KONN	Konnektor
DAT	Dativ	KONSEK	Konsekutiv
DEF	definit	KOP	Kopula
DEM	Demonstrativ	LIG	Ligatur
DIR	Direktional	LOK	Lokativ
DISTR	Distributiv	M	maskulinum
DO	direktes Objekt	MED	medial (nicht-finaler Satz)
DU	Dual	MEDPASS	Mediopassiv
DYN	dynamisch	N	neutrum
ERG	Ergativ	NEG	negativ
EXIST	Existential	NOM	Nominativ
F	femininum	NPRT	Nicht-Präteritum
FAKT	Faktitiv	NR	Nominalisator
FOK	Fokus	OBJ	Objekt
FUT	Futur	OBL	Oblik
GEN	Genitiv	OBLG	Obligativ
GER	Gerundium	OPT	Optativ
HABIT	Habitualis	PART	Partizip
IMP	Imperativ	PASS	Passiv
IMPF	Imperfekt	PAT	Patiens
INAKT	inaktiv	PF	Perfekt
		PFV	Perfektiv
		PL	Plural
		PLUP	Plusquamperfekt
		POSS	Possessiv (pronominales Element)

POT	Potentialis	ANIM	animate
PRÄS	Präsens	AOR	aorist
PRÄT	Präteritum	APASS	antipassive
PRIV	Privativ	APPL	applicative
PROG	Progressiv	ART	article
PROLAT	Prolativ	ASRT	assertive
PROZ	Prozessiv	ASS	(as)sociative
RDP	Reduplikation	AT	actor topic
REFL	reflexiv	AUG	augmentative
RELL	Relational(isator)	AUX	auxiliary
REZ	reziprok	AVERS	aversive (abessive)
S	intransitives Subjekt	BEN	benefactive
SG	Singular	CAUS	causative
STAT	Stativ	CL	classifier
SUP	Superlativ	CLN	noun class n
TOP	Topik	CMPR	comparative
VEN	Venitiv	COLL	collective
VOK	Vokativ	COMIT	comitative
VR	Verbalisator	COMPL	completive
ZIRK	Zirkumstantiell	COND	conditional
		CONJ	conjunctive
		CONN	connector
		CONSEC	consecutive
		CONSTR	construct
		CONT	continuous
		COP	copula
Ø	zero	D1	deictic of 1 <sup>st</sup> person
1	1 <sup>st</sup> person	D2	deictic of 2 <sup>nd</sup> person
2	2 <sup>nd</sup> person	D3	deictic of 3 <sup>rd</sup> person
3	3 <sup>rd</sup> person	DAT	dative
ABL	ablative (separative)	DECL	declarative
ABS	absolutive (case/function of ergative system)	DEF	definite
ABSL	absolute (free form of noun)	DEM	demonstrative
ABSTR	abstract	DES	desiderative
ACAUS	anticausative (deagentive)	DET	determiner
ACC	accusative	DETR	detransitivizer
ACNNR	action nominalizer	DIM	diminutive
ACR	actor	DIR	directional
ACT	active	DIST	distal (remote)
ADESS	adessive	DISTR	distributive
ADJR	adjectiv(al)izer	DON	donative
ADM	admonitive	DR	direct
ADVR	adverbializer	DS	different subject
AG	agentive	DU	dual
AGNR	agent nominalizer	DUB	dubitative
AGR	agreement	DUR	durative
AGT	agent	ELAT	elative
ALL	allative	EMPH	emphasizer/emphatic
ANA	anaphoric	EQT	equative
AND	andative	ERG	ergative
		EVID	evidential

## 2. English

Ø	zero	D1	deictic of 1 <sup>st</sup> person
1	1 <sup>st</sup> person	D2	deictic of 2 <sup>nd</sup> person
2	2 <sup>nd</sup> person	D3	deictic of 3 <sup>rd</sup> person
3	3 <sup>rd</sup> person	DAT	dative
ABL	ablative (separative)	DECL	declarative
ABS	absolutive (case/function of ergative system)	DEF	definite
ABSL	absolute (free form of noun)	DEM	demonstrative
ABSTR	abstract	DES	desiderative
ACAUS	anticausative (deagentive)	DET	determiner
ACC	accusative	DETR	detransitivizer
ACNNR	action nominalizer	DIM	diminutive
ACR	actor	DIR	directional
ACT	active	DIST	distal (remote)
ADESS	adessive	DISTR	distributive
ADJR	adjectiv(al)izer	DON	donative
ADM	admonitive	DR	direct
ADVR	adverbializer	DS	different subject
AG	agentive	DU	dual
AGNR	agent nominalizer	DUB	dubitative
AGR	agreement	DUR	durative
AGT	agent	ELAT	elative
ALL	allative	EMPH	emphasizer/emphatic
ANA	anaphoric	EQT	equative
AND	andative	ERG	ergative
		EVID	evidential

EXIST	exist(ential)	N	neuter
EXPER	experiential	NARR	narrative
F	feminine	NEG	negative
FACT	factitive	NFIN	non-finite
FIN	finite	NFUT	non-future
FOC	focus	NHUM	non-human
FUT	future	NOM	nominative
GEN	genitive	NPST	non-past
GER	gerund (conjunctive participle)	NR	nominalizer
GNR	generic	OBJ	object
HABIT	habitual (customary)	OBL	oblique
HEST	hesternal	OBLG	obligative
HODPST	hodiernal past	OBV	obviative
HON	honorific	OPT	optative
HORT	(ad-/co-)hortative	PART	participle
HUM	human	PASS	passive
ILL	illative	PAST	past
IMM	immediate	PAT	patient
IMP	imperative	PATNR	patient nominalizer
IMPF	imperfect	PE	plural exclusive
IMPR	impersonal	PF	perfect
INACT	inactive	PFV	perfective
INAN	inanimate	PI	plural inclusive
INCH	inchoative	PL	plural
INCOMPL	incomplete	PLUP	pluperfect
IND	indicative	PNCT	punctual
INDEF	indefinite	POSS	possessive (pronominal element)
INDEP	independent	POT	potential
INESS	inessive	PRES	present
INF	infinitive	PRET	preterite
INFR	inferential	PRIV	privative
INGR	ingressive (inceptive)	PROC	processive
INSTNR	instrument nominalizer	PROG	progressive
INSTR	instrumental	PROH	prohibitive
INT	interrogative	PROPR	proprietary
INTS	intensive (reinforcement marker)	PROX	proximal (local deixis)
INV	inversive	PRTV	partitive
INVIS	invisible	PRX	proximate (pronominal)
IO	indirect object	PT	patient topic
IPFV	imperfective	PURP	purposive
IPS	impersonal passive	QUOT	quotative
IRR	irrealis	REC	reciprocal
ITER	iterative	RECPST	recent past (immediate past)
JUSS	jussive	RDP	reduplication
LOC	locative	REFL	reflexive
LOG	logophoric	REL	relative
LT	locative topic	RELL	relational(izer)
M	masculine	REM	remote
MANNR	manner nominalizer	REMPST	remote past
MED	medial (non-final clause)	REP	repetitive
MID	middle (medio-passive)	RES	resultative
		RLS	realis

RPRT	reportative	SUP	superlative
S	intransitive subject	SUPESS	superessive
SBJ	subject	TOP	topic
SENS	sensory evidential	TRNSL	translative
SEQ	sequential	TRL	trial
SG	singular (restricted)	TRR	transitivizer
SPEC	specific	UGR	undergoer
SR	subordinator	VEN	venitive
SS	same subject	VOC	vocative
STAT	stative	VOL	volitive
SUBESS	subessive	VR	verbalizer
SUBJ	subjunctive	VIS	visible

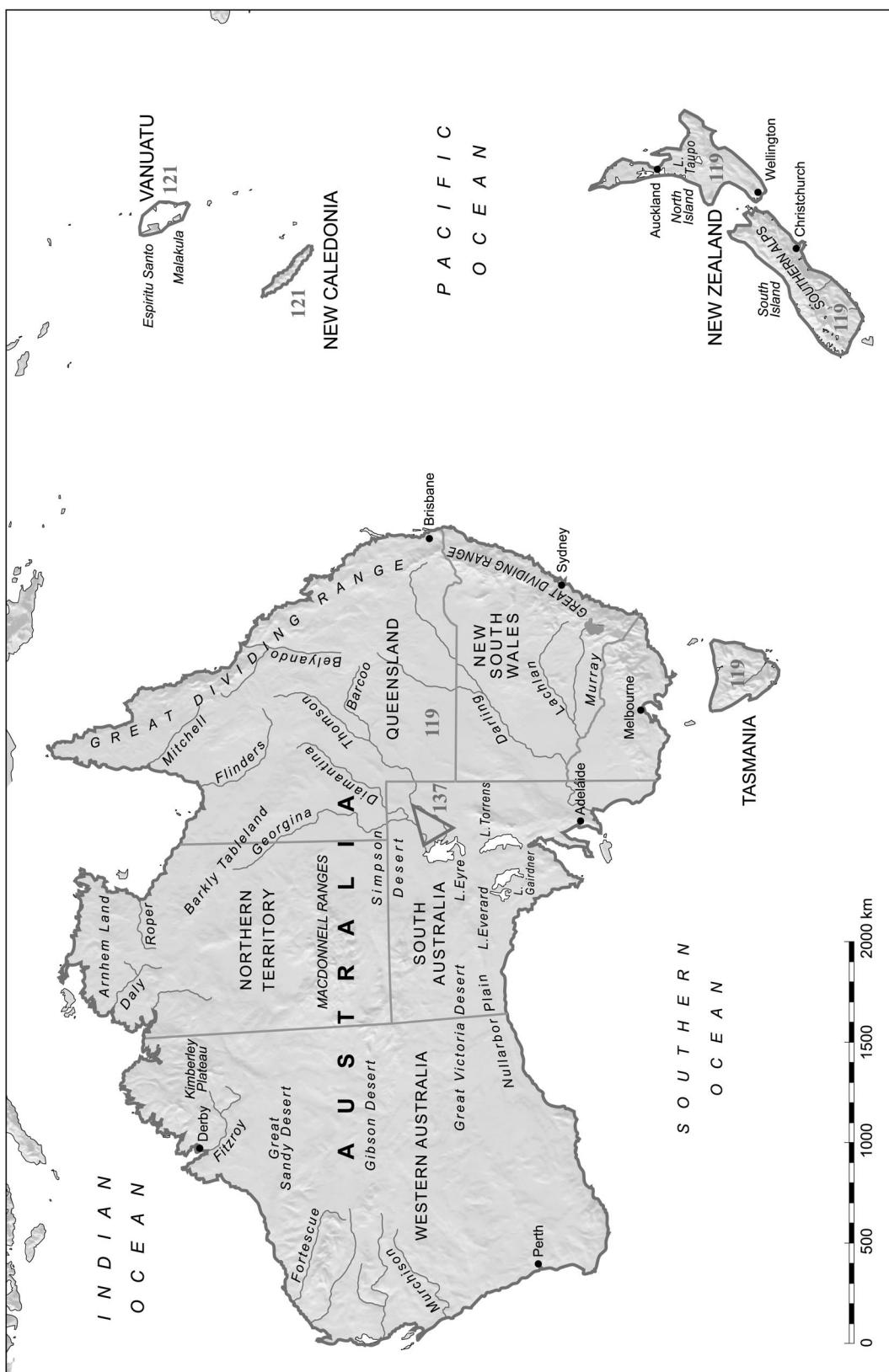
# Sprachenkarten/Language maps

## 1. Afrika/Africa



119. English (Indo-European: Germanic)  
121. Français (Indo-européen: Roman)  
139. Turkana (Nilo-Saharan)  
140. Twi (Kwa)  
141. Kinyarwanda (Bantu)

## 2. Australien/Australia

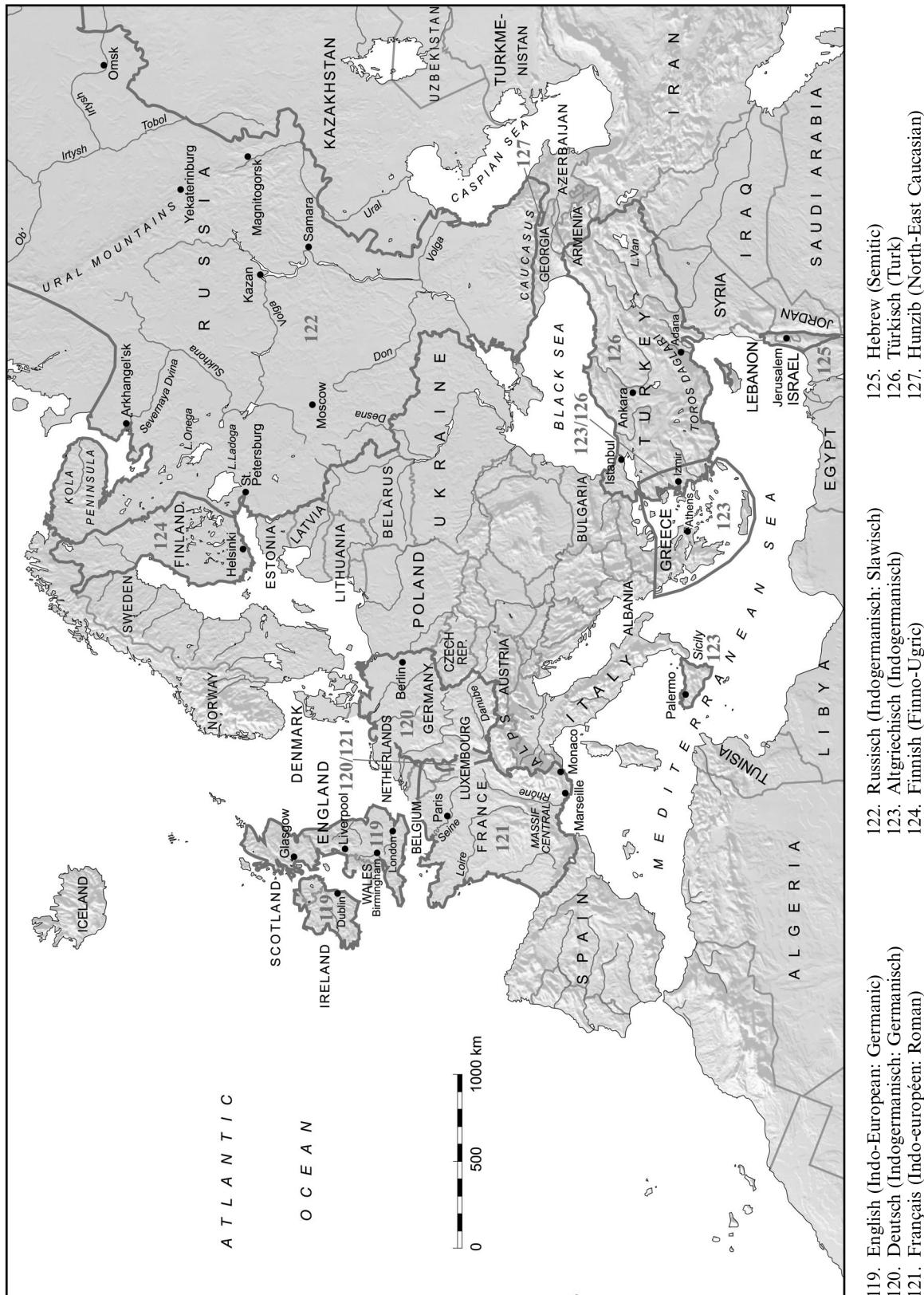


119. English (Indo-European: Germanic)

121. Français (Indo-européen: Roman)

137. Diyari (Pama-Nyungan)

## 3. Europa und Nahost/Europe and Near East



## 4. Nordasien/North Asia



122. Russisch (Indogermanisch: Slawisch)

128. Ketisch (Jenisieisch)

## 5. Nordamerika/North America



119. English (Indo-European: Germanic)  
 121. Français (Indo-européen: Roman)  
 129. West Greenlandic (Eskimo)

130. Koyukon (Athapaskan)  
 131. Montagnais/Innu-aimun (Algonquian)

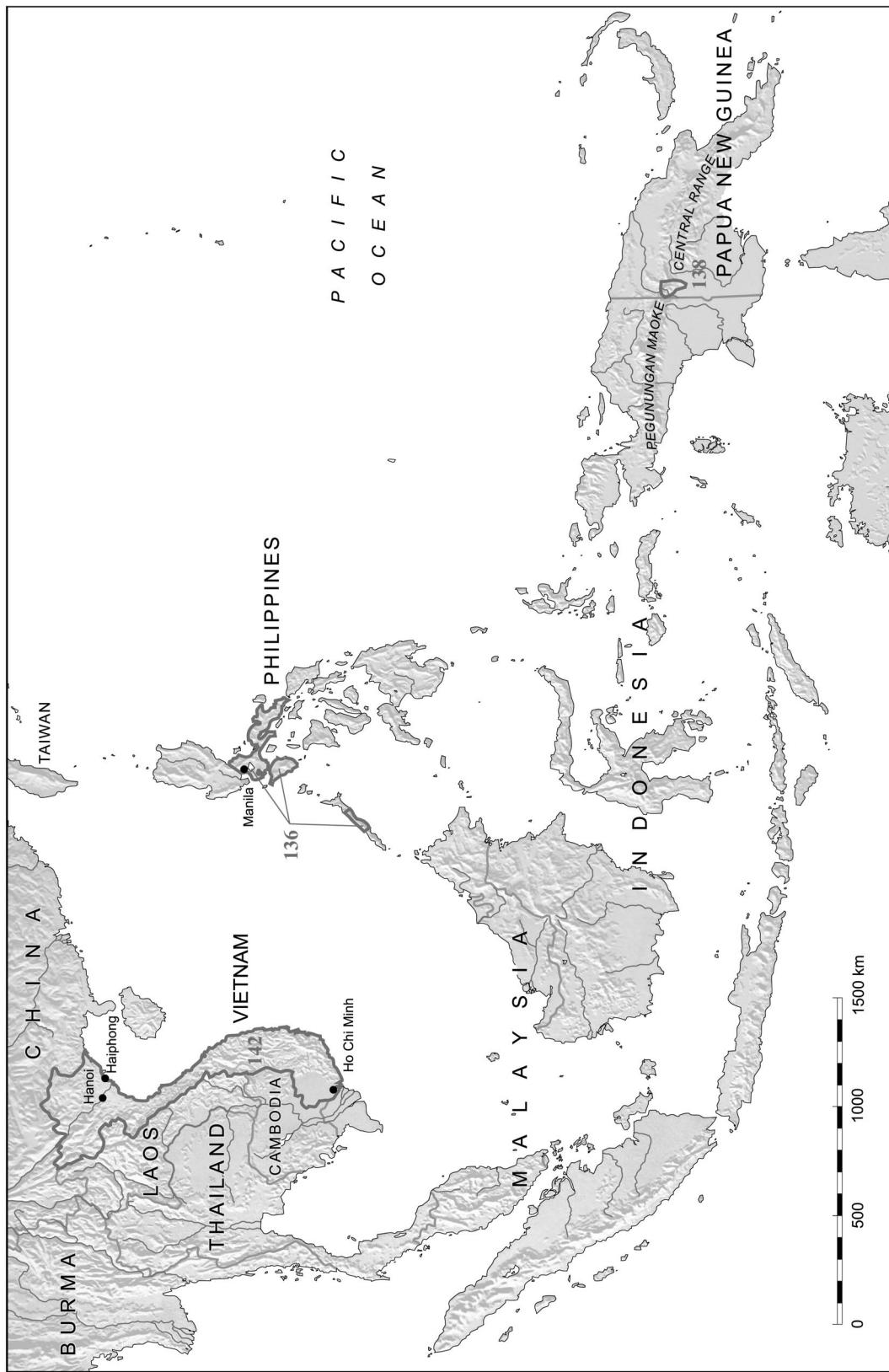
## 6. Südamerika/South America



119. English (Indo-European: Germanic)  
 121. Français (Indo-européen: Roman)  
 132. Guarani (Tupi-Guarani)

134. Quechua (Quechua)  
 135. Yagua (Peba-Yaguan)

## 7. Südostasien/South East Asia



136. Tagalog (Austronesian)

138. Wambon (Awyu)

142. Vietnamese (Viet-Muong)

## Namenregister / Index of names

### A

Aaseb, Ivar 146  
Abbi, Anvita 546, 561  
Abbott, Miriam 1097  
Abercrombie, David 97  
Abhyankar, Kahinatha Vasudev 42  
Abney, Steven 1896  
Abondolo, Daniel 1080  
Abraham, Werner 1274  
Achmánova, Ol'ga S. 133  
Ackema, Peter 409  
Ackrill, John L. 211  
Acson, Veneeta 585  
Adams, Catherine 1820  
Adams, G. Brendan 88  
Adams, Karen 1016, 1475  
Adams, Valerie 960, 155, 1255  
Adelaar, Willem F. H. 164, 1206, 1453 f., 1458, 1462, 1654  
Adelung, Johann Christoph 91, 113, 668, 698, 1267, 1682, 1690  
Admoni, Vladimir G. 611, 665, 668, 677, 681 f., 1270, 1906  
Adodurov, Vasilij 1300  
Adontz, Nicolaus 59  
Adrados, Francisco Rodriguez 122  
Aelfric, Abbot of Eynsham 86, 696  
Aerts, Willem Joan 656  
Aeschylus 1316  
Agha, Asif 993  
Agricola, Mikael 1328  
Ahlquist, Anders 88  
Aichinger, Carl Friedrich 698  
Aikhenvald, Alexandra Y. 1024, 1031 f., 1034–1042, 1070, 1197  
Aissen, Judith L. 324, 1135  
Aitchison, Jean 206, 386, 1771–1774, 1781, 1783–1785  
Ajnewtegin, M. V. 315  
Akamatsu, Tsutomu 493  
Akmajian, Adrian 818, 1005  
Aksakov, Konstantin S. 608  
Aksu-Koç, Ayhan 1364, 1800  
Al, Bernard P. F. 163  
Al-Astarabadi, Rađi 67, 71 f.  
Aldama y Guevara, Joseph 1433  
Alerten, D. J. 6

Alexander der Große 55  
Al-Farra?, Abu Zakariyya 67, 72  
Algeo, John 954, 957 f.  
Ali, Abdul Sahib Mehdi 1750  
Allan, Keith 1016, 1018, 1060, 1067  
Allen, Barbara J. 924 f.  
Allen, James 1893  
Allen, John 1896  
Allen, Margaret Reece 189, 199, 351–353, 356–358, 411, 541, 907, 948  
Allen, Robert E. 1883  
Allen, Shanley E. M. 1821  
Allen, William Sidney 93, 218, 437  
Allerton, D. J. 696  
Allières, Jacques 488  
Allport, D. A. 1790  
Allwood, Jens 1195  
Almazan, M. 1822  
Almeida, Matthew 767  
Al-Mubarrad, Ibn Yazid 68, 70, 703  
Alonso, Dámaso 1582  
Alpher, Barry 719, 1034  
Alsdorf-Bollée, Annegret 1708  
Alsina, Alex 325  
Althaus, Hans Peter 435  
Altmann, Gabriel 1239–1243  
Altner, Günter 16  
Ameka, Felix 719, 1092  
Amith, Jonathan D. 711  
Ammann, Hermann 667  
Amritavalli, A. 859  
Anceaux, Johannes C. 164, 166, 803  
Andersen, Elaine 750  
Andersen, Henning 273, 383  
Andersen, Karl Gustaf 1620–1623  
Andersen, Paul Kent 1086, 1149, 1159, 1212, 1216  
Andersen, Roger 1654, 1806–1808  
Anderson, Arthur J. O. 1452  
Anderson, John M. 507, 812 f., 1596  
Anderson, Stephen R. 6, 20, 27, 32 f., 153, 197, 199–202, 205, 207, 233, 235, 261, 268, 316, 320 f., 325 f., 328 f., 337, 339, 341, 346 f., 350 f., 353, 358, 361, 365 f., 373, 395 f., 406, 410, 412, 429, 437, 439, 442, 453, 456, 460, 476, 489, 499 f., 502, 508, 526, 542, 546, 549, 571, 580, 584 f., 587, 589, 625, 627 f., 716, 722, 735, 739, 742 f., 753, 784, 814, 842, 851, 854, 877, 882, 976, 983, 986, 988, 991–993, 1005, 1047, 1049, 1252, 1655, 1732, 1772, 1790, 1796 f., 1801  
Andrade, Manuel J. 560 f.  
André, Jacques 1218  
Andreev, Nikolaj D. 133  
Andrews, Avery D. 877  
Andrews, Edna 279  
Andrews, J. Richard 621 f., 1440, 1447  
Anshen, Frank 206, 346, 348, 852  
Antinucci, Francesco 1807  
Anttila, Raimo 100, 121, 281, 856, 1086, 1662, 1667  
Antworth, Evan L. 1833, 1896, 1898  
Anward, Jan 710  
Aoki, Haruo 998  
Aoun, Joseph 1265  
Aoyma, Takashi 988  
Aphek, Edna 1346  
Apollonius Dyscolus 55, 64 f.  
Appel, Elsbeth 1271  
Appleyard, D. L. 809  
Apresjan, Jurij D. 514, 1546  
Arce-Arenales, Manuel 1162  
Archaimbault, Sylvie 697  
Archangeli, Diana 346  
Arens, Hans 7, 91–93, 674  
Arensen, Jonathan E. 1100  
Ariel, Mira 983  
Ariel, Shlomo 1345, 1347 f.  
Aristarchus 55  
Aristophanes 53, 1316  
Aristoteles 2, 53 f., 57, 77, 83, 211 f., 264, 459, 600, 695, 732, 734, 1168, 1172, 1182, 1326  
Arlotto, Anthony 1664  
Armagost, James 1005  
Armon-Lotem, Sharon 1796  
Arnason, Kristján 1607  
Arnauld, Antoine 698  
Arndt, Walter W. 1269

Arnim, Ioannes von 54f.  
 Arnott, Donald Whitehorn 762,  
 1134  
 Aronoff, Mark 31, 165, 189,  
 196–198, 200–203, 205f.,  
 230, 249, 252f., 290, 308f.,  
 341, 346, 348, 350, 364, 408,  
 410, 412, 414, 419, 423, 430,  
 433, 445, 477, 479, 541, 546,  
 583, 596, 633, 657, 834, 838,  
 841f., 848, 852, 854, 856,  
 868, 870, 873, 878–880, 882,  
 887, 899, 932, 1255f., 1259–  
 1262, 1637, 1639f., 1642f.,  
 1780, 1790f.  
 Aronson, Howard I. 597  
 Arrivé, Michel 1291  
 Arutjunowa, N. D. 680f.  
 Asdahl Holmberg, Märta 890,  
 1277  
 Asher, R. E. 738  
 Ashton, Ethel O. 762, 1120,  
 1123  
 Askedal, John Ole 1268, 1273  
 Asporius (Asperius, Asper) 78  
 Atkins, Beryl T. 1885  
 Atkins, John 178  
 Atlla, Catherine 1409  
 Augst, Gerhard 420, 905,  
 1267f., 1272, 1622f., 1686,  
 1689, 1692  
 Augustinus 76  
 Auroux, Sylvain 39, 693f.,  
 697f., 701  
 Austerlitz, Robert 1217  
 Austin, Peter 987, 1078, 1088,  
 1158f., 1202–1205, 1491–  
 1495, 1497f.  
 Averroes 77  
 Avicenna 77  
 Avram, Andrei 493  
 Axelrod, Melissa 917–919, 1402  
 Axelsson, Monica 1413  
 Ayto, John 1674  
 Azzopardi-Alexander, Marie 760

**B**

Baarle, Peter van 823  
 Baayen, Rolf Harald 164f., 206,  
 299, 346, 1639, 1773, 1779,  
 1898  
 Babby, Leonard H. 1157  
 Bach, Adolf 1905  
 Bach, Emmon 201, 727, 851  
 Bacher, Wilhelm 74, 703  
 Back, Otto 1563  
 Bacon, Francis 1816  
 Bacon, Roger 92  
 Baddeley, Alan D. 1822  
 Badecker, William 364, 1784,  
 1817, 1819

Bader, Françoise 1325  
 Bader, Thomas 1481  
 Bahner, Werner 104  
 Bai, Juyi 1737  
 Bailey, Charles-James N. 289  
 Bailey, Nathalie 1806  
 Bakema, Peter 1048  
 Baker, Mark C. 199, 203f., 315,  
 321–326, 329–332, 368, 407,  
 542, 916, 922, 924f., 1135,  
 1137, 1140f.  
 Baker, William 1823  
 Bakker, Peter 1646f., 1659  
 Balbi, Adriano 91, 93  
 Baldi, Philip 1214, 1640  
 Baldinger, Kurt 1706f., 1911  
 Ballweg, Joachim 1273f.  
 Bally, Charles 211, 216, 248,  
 514, 520, 945  
 Bammesberger, Alfred 823  
 Banerjee, Satya Ranjan 42  
 Bani, Ephraim 1034  
 Barac-Cikoja, Dragana 1922  
 Bar-Adon, Aaron 1343  
 Baratin, Marc 694  
 Barbaud, Philipp 837  
 Barbe, Katharina 452  
 Barber, Charles C. 759  
 Bareš, Karel 967  
 Barkali, Sha'ul 1349  
 Barnes, Janet 1022  
 Barnetová, Vilma 1305f.,  
 1308f.  
 Barnhart, Clarence L. 836f.  
 Barnhart, Robert K. 930, 1674  
 Barré, Louis 16  
 Barret, Anthony 1507  
 Barron, Roger 1019, 1024  
 Bartke, Susanne 1269, 1820f.  
 Barton, John 88, 696  
 Bartsch, Werner 1273  
 Barz, Irmhild 356, 887–889,  
 891–894, 898f., 902, 931f.,  
 950, 958, 1267, 1276  
 Basbøll, Hans 341, 384f., 433,  
 584  
 Bassarak, Armin 1359f., 1363  
 Basset, Louis 694  
 Bastiaanse, Roelien 1818  
 Bat-El, Outi 556  
 Bates, Dawn 560  
 Bates, Elizabeth 289, 795, 1800,  
 1807  
 Bátori, István 984  
 Battistella, Edwin 273, 276,  
 278f.  
 Baudouin de Courtenay, Jan  
 19f., 91, 98, 125f., 134, 205,  
 344f., 405, 417, 420, 436f.,  
 451f., 454, 460, 490, 492,  
 508, 613, 1601f., 1616, 1620  
 Bauer, Heinrich 109  
 Bauer, Laurie 5, 155, 220, 235,  
 237, 248, 251f., 319, 350,  
 354, 356, 405, 463, 474, 526,  
 532f., 542f., 546, 556, 582,  
 588, 617, 622, 638, 832–835,  
 837f., 874, 886, 888–895,  
 899, 901, 931, 933, 938, 945,  
 955, 957, 968, 979, 993, 1047,  
 1248, 1255, 1259f., 1264,  
 1626–1629, 1631f., 1636–  
 1638, 1640f., 1643, 1781,  
 1888, 1890, 1922  
 Bauer, Winifred 766  
 Baum, Richard 689  
 Bausch, Karl-Heinz 1274  
 Bautista, Maria L. S. 1481  
 Bayer, Josef 364, 829, 1817f.  
 Bazell, Charles E. 1225f.  
 Beard, Robert 153, 200, 205,  
 247, 250, 266, 320, 332, 353,  
 367, 417, 454, 511, 622, 627,  
 657, 842, 849, 866  
 Beauzée, Nicolas de 698  
 Bébian, Roch-Amboise Auguste  
 1555  
 Bech, Gunnar 1267f.  
 Béchade, Hervé D. 1287, 1291,  
 1294  
 Becker, Alton 999  
 Becker, Claudia 1560  
 Becker, Karl Ferdinand 19,  
 104f., 107, 109f., 113f., 157,  
 162, 669f., 672  
 Becker, Thomas 161, 231, 298,  
 456, 866, 883, 1275, 1278  
 Becker, Ulrich 1563  
 Béclard, Pierre Augustin 17  
 Beekman, John 1831  
 Behaghel, Otto 112, 159, 753,  
 1233, 1680, 1690, 1692  
 Behnstedt, Peter 662, 1742f.  
 Behrens, Leila 1067f.  
 Beifuss, Karin 984  
 Beito, Olav Toreson 147  
 Bellmann, Günther 894  
 Bellugi, Ursula 1822f.  
 Belvalkar, Shripad Krishna 41  
 Ben Amotz, Dahn 1356  
 Ben Asher 74  
 Ben Asher, M. 1345  
 Ben Labrat, Dunash 703  
 Bendix, Edward 178, 1091  
 Bendor-Samuel, David 1194  
 Bendor-Samuel, John T. 587  
 Benedict, Paul K. 1733f., 1737  
 Beneš, Brigit 97  
 Bennis, Hans 829  
 Benson, Morton 1890  
 Benveniste, Émile 122, 142, 277,  
 286, 614, 751, 785, 813, 984,  
 998, 1214, 1218, 1295, 1592,  
 1595  
 Berendsen, Egon 389  
 Berezin, Fedor M. 135, 1308  
 Berg, Helma E. van den 1367,  
 1370–1373

- Berg, Thomas 1773, 1791  
 Bergenholtz, Henning 405, 408,  
 439, 445, 452f., 489, 526,  
 542, 596, 600, 676f., 683,  
 699f., 867, 891, 893f., 933,  
 1275, 1615, 1888, 1890f.  
 Bergman, M. W. 1786  
 Bergmann, Rolf 122, 904, 1684  
 Bergsland, Knut 146f., 1389  
 Bergström, Gustav A. 1678  
 Berko, Jean 300, 1615, 1619,  
 1623, 1766, 1769, 1795,  
 1797f.  
 Berman, Ruth Aronson 1343,  
 1345, 1347–1349, 1795–  
 1799, 1802f.  
 Bernard, John 832  
 Berndt, Rolf 1673  
 Bernhardi, A. F. 108, 698  
 Bernstein Ratner, Nan 1798  
 Berthelsen, Christian 1389  
 Bertinetto, Pier Marco 364,  
 660f., 1176  
 Bertocchi, Alessandra 1213  
 Besch, Werner 1267, 1681f.,  
 1684  
 Besnier, Niko 577f., 580f.  
 Best, Günter 1507  
 Best, Karl-Heinz 121  
 Bettelhäuser, Hans-Jörg 1269  
 Beurden, Lisan van 165  
 Bever, Thomas G. 1799f.  
 Bhartṛhari 42  
 Bhat, Darbhe N. Shankara 720,  
 725f., 759, 761, 764, 1112,  
 1209f., 1214, 1216  
 Bhattacharya, Sudhibhusan  
 1004  
 Bhayani, Harivallabh Chunnial  
 1722  
 Bickel, Balthasar 1178, 1835f.,  
 1853  
 Bickerton, Derek 1183, 1764  
 Biedermann, Johann 125, 135,  
 697  
 Bierbach, Mechtild 1704  
 Biermann, Anna 1054, 1060,  
 1064, 1097  
 Bierwisch, Manfred 160, 268–  
 270, 667, 908, 910, 1099,  
 1788  
 Biesold, Horst 1555  
 Biggs, Bruce G. 578  
 Bills, Garland 1157  
 Binnick, Robert I. 1168, 1172,  
 1175, 1180  
 Bird, Charles S. 569, 589  
 Bisang, Walter 1657  
 Bishop, Brian R. 1563  
 Bishop, Dorothy V. M. 1820f.  
 Bittner, Andreas 294, 297, 299,  
 519, 636, 1272  
 Bittner, Dagmar 1269  
 Black, James R. 389  
 Black, Jeremy 35  
 Blainville, Henri Marie  
 Ducrotay de 16f.  
 Blair, Hugh E. 1563, 1567  
 Blake, Barry J. 640, 642f.,  
 809f., 815, 825, 1083, 1085–  
 1088, 1135f., 1482, 1491,  
 1493f.  
 Blake, Frank R. 1474f., 1479f.,  
 1486  
 Blanc, Haim 1065, 1742, 1749  
 Blanche-Benveniste, Claire 1285  
 Blank, Andreas 1616  
 Blanke, Detlev 1563–1565, 1571  
 Blanken, Gerhard 1817  
 Blatz, Friedrich 106  
 Blau, Joshua 1741  
 Bleek, W. H. I. 96  
 Bleser, Ria de 364, 1817, 1819  
 Bloch, Bernard 178f., 180, 185,  
 218f., 241, 379, 437, 440f.,  
 453  
 Bloch, Jules 1712, 1724  
 Blois, Kornelis F. de 589  
 Bloom, Lois 1807  
 Bloomfield, Leonard 20, 26, 55,  
 71, 100, 109, 121, 165, 170–  
 189, 191, 196–198, 202f.,  
 205, 218f., 236f., 252, 304,  
 319, 336f., 345, 350–352,  
 356–359, 379f., 404–406,  
 418, 420, 427, 437f., 440,  
 444f., 453, 463, 490, 494,  
 508, 542, 561, 582, 599, 677,  
 698, 721, 735, 872f., 1005,  
 1012, 1233, 1416, 1474,  
 1476–1479, 1481, 1483,  
 1486, 1488, 1579, 1789, 1836  
 Blust, Robert A. 423f.  
 Boas, Franz 173, 176–178, 186,  
 547, 760, 824  
 Boase-Beier, Jan 908  
 Bochner, Harry 407  
 Bock, Kay 1780  
 Boeder, Winfried 1091, 1136  
 Boersma, Paul 1613  
 Boettcher, Wolfgang 1936  
 Bogaards, Paul 1891  
 Bogdanov, Sergej I. 420, 422  
 Bogorodickij, Vasilij A. 126  
 Bohas, Georges 70, 702f.  
 Bohnhoff, Lee E. 985, 1006  
 Böhtingk, Otto 437  
 Bois, John W. Du 983  
 Boisson, Claude 693  
 Bokamba, Eyamba G. 811  
 Bokarev, Evgenij A. 1367  
 Bol, Gerard 1821  
 Bolinger, Dwight L. 247, 249,  
 252, 262, 418, 423, 766, 953,  
 1136, 1233f., 1591, 1619f.,  
 1623, 1884, 1889, 1922  
 Bolozky, Shmuel 1345  
 Bondárko, Aleksandr V. 132,  
 609, 611, 613–615, 886f.  
 Bonifatius 76, 79, 81f.  
 Bonvillain, Nancy 920, 993  
 Boogaart, Ronny 1178  
 Booij, Geert E. 32, 163, 165f.,  
 195, 198, 203, 230f., 318,  
 327, 336f., 339–342, 351f.,  
 361, 363, 365–367, 380,  
 384f., 395, 414, 433, 476,  
 541, 845, 854, 861f., 864,  
 879, 883, 889, 891f., 902,  
 926, 931, 933, 976, 979, 1214,  
 1288, 1806, 1809  
 Boost, Karl 668  
 Bopp, Franz 6–8, 19, 59, 93,  
 95–98, 100, 537, 1590, 1592  
 Borer, Hagit 314, 331  
 Boretzky, Norbert 290, 1644f.,  
 1647, 1649, 1651  
 Borg, Albert 760  
 Borg, Alexander 1748f.  
 Borgman, Donald M. 1216  
 Borgström, Carl Hjalmar 147,  
 439  
 Börjars, Kersti 656f., 662  
 Bork, Hans Dieter 890, 1698,  
 1704  
 Bornemann, Eduard 1320  
 Borsche, Tilman 696  
 Bortolini, Umberta 1820–1822  
 Bosák, Ján 424  
 Bosch, A. van den 1898f.  
 Botha, Rudolf P. 165f., 195,  
 318, 351f., 357, 845, 878f.,  
 1636, 1816  
 Bouchard, Denis 1265  
 Bouma, Lowell 1809  
 Bouquiaux, Luc 1832  
 Bowe, Heather 715  
 Bowerman, Melissa 1798  
 Bowern, Claire 1491  
 Boyes Braem, Penny 1556  
 Bradley, Dianne C. 1818  
 Brady, Ronald H. 16  
 Braine, Jean Critchfield 803  
 Branch, Michael 1079f.  
 Brandt Corstius, Hugo 772, 775  
 Braune, Wilhelm 100, 1683,  
 1686f., 1689  
 Brauner, Siegmund 715  
 Bréal, Michel 19  
 Brecht, Richard 1157  
 Breedveld, Anneke 1510  
 Breeze, Mary 1007, 1011  
 Brekle, Herbert E. 19, 54, 91f.,  
 161, 834f., 837, 906f.  
 Brend, Ruth M. 192–194  
 Brenninkmeyer, Leo 1760  
 Bresnan, Joan 199, 346, 784,  
 1135  
 Brewster, E. Thomas 1832

Brewster, Elizabeth S. 1832  
 Bright, William 553, 803  
 Brinker, Klaus 114, 1273  
 Brinkmann, Hennig 116, 1693  
 Brinton, Laurel 815  
 Brisard, Frank 983  
 Brittain, Julie 1419  
 Broadwell, George Aaron 584f.  
 Brockelmann, Carl 1096  
 Brockhaus, Wiebke 431  
 Broeder, Peter J. F. J. 1813  
 Bromley, Myron H. 1142, 1203,  
 1205  
 Bronckart, Jean-Paul 1807  
 Brøndal, Viggo 146, 273f.,  
 679f., 699f.  
 Broschart, Jürgen 683, 717, 721,  
 725, 938  
 Broselow, Ellen 206, 548, 555,  
 560, 564  
 Brouwer, Dédé 1050  
 Brown, Dunstan 155  
 Brown, Roger 540, 1001, 1011,  
 1795, 1797  
 Browne, Wayles 571  
 Browning, Robert 809  
 Bruce, Les 1036, 1041  
 Bruche-Schulz, Gisela 135  
 Brugman, Claudia 1596  
 Brugmann, Karl 10, 100f., 112,  
 158, 782, 1626  
 Brunner, Karl 1670  
 Bryan, Margaret A. 1077, 1079  
 Buba, Malami 990f., 993  
 Bubenik, Vit 1712  
 Buchholz, Oda 1213, 1218  
 Buck, Carl Darling 1126  
 Buckingham, Henry W. Jr. 1789  
 Buckingham, Hugh W. 1819  
 Bugge, Aage 1389  
 Bübler, Karl 110, 163, 249, 667,  
 682, 685, 983  
 Bühler, W. 58  
 Bulachóv, Michail G. 135  
 Bulygina, Tat'jana V. 26f., 133,  
 526  
 Bumann, Waltraud 98  
 Bunt, Harry C. 1067f.  
 Burani, Cristina 1772, 1778f.  
 Burdach, Karl Friedrich 16–18  
 Burgschmidt, Ernst 834  
 Burling, Robbins 1196  
 Burrow, Thomas 1150, 1159  
 Bursill-Hall, Geoffrey L. 696f.  
 Burt, Marina K. 1806  
 Burzio, Luigi 813  
 Buscha, Joachim 677, 683  
 Busláev, Fëodor I. 125  
 Bußmann, Hadumod 436, 463  
 Butterworth, Brian 1768, 1771f.,  
 1778f., 1783, 1790, 1818  
 Butzkamm, Wolfgang 1931  
 Buyssens, Eric 211, 214, 216

Buzássyova, Klára 424  
 Bwenge, C. 1886, 1888  
 Bybee, Joan L. 27, 32f., 122,  
 203f., 220, 231, 261, 266f.,  
 286f., 294, 346, 348, 361,  
 366, 371f., 374–376, 407,  
 442–444, 536–543, 548, 550,  
 600, 604, 606, 626, 633,  
 656f., 659–662, 708, 712,  
 753f., 799, 804, 806, 814,  
 869, 871f., 929f., 938, 976,  
 1137, 1140, 1167, 1173, 1175,  
 1177f., 1180, 1184f., 1191,  
 1194–1196, 1199, 1234,  
 1247, 1249–1252, 1407,  
 1423, 1591, 1595, 1603, 1773,  
 1778, 1781, 1795–1797  
 Bynon, Theodora 97, 121, 1223,  
 1225, 1228f.  
 Byrd, Roy J. 206

C

Cabré, M. Teresa 1927  
 Cadot, Pierre 699  
 Caesar 1077  
 Caferoğlu, Ahmet 1579  
 Calder, George 228  
 Caldwell, Robert 96, 813  
 Caluwe, Johan de 166  
 Camaj, Martin 978  
 Campbell, Lyle 442, 760, 979,  
 1666  
 Canello, Ugo Angelo 452  
 Canger, Una 1450  
 Cann, R. 681  
 Cannon, Garland 835, 952f.,  
 955f.  
 Cantwell, Dennis P. 1823  
 Capell, Arthur 373, 1479  
 Capidan, Theodor 1651  
 Caplan, David 1818  
 Caramazza, Alfonso 347, 364,  
 1769, 1772, 1778, 1780f.,  
 1783, 1817–1819  
 Cardona, George 41, 43–46, 48,  
 50f., 1712  
 Carlson, Greg 862, 1261  
 Carlson, Robert 757, 761, 764,  
 1018, 1022f., 1025  
 Carochi, Horacio 1433  
 Carpenter, Kathie 1018, 1028f.  
 Carrier-Duncan, Jill 559, 561,  
 1477f.  
 Carroll, Janet F. 1096  
 Carroll, John M. 837  
 Carroll, Lewis 953, 955  
 Carroll, Mary 1810  
 Carroll, Susanne 1812  
 Carstairs, Andrew 27, 154, 200,  
 231, 338, 346, 389, 460, 463,  
 469, 478f., 514, 516f., 597,  
 599, 603, 617, 626f., 630,  
 632, 634–636, 638f., 645f.,  
 1269, 1514, 1766, 1772  
 Carstairs-McCarthy, Andrew  
 155, 289, 453, 618, 623f.,  
 635, 637, 639, 1255  
 Carter, Michael 74  
 Carvalho, Paulo de 1076  
 Casacchia, Giorgio 703f.  
 Casad, Eugene H. 1121, 1129  
 Cassirer, Ernst 212, 1091  
 Castrén, M. Alexander 96  
 Castren, Mathias A. 1376  
 Catach, Nina 1921  
 Caubet, Dominique 1742  
 Ćavar, Damir 398  
 Cavigneaux, Antoine 37  
 Caxton, William 1668  
 Cerrón-Palomino, Rodolfo 1453,  
 1455  
 Chafe, Wallace L. 1098, 1197  
 Chalilov, Majid Sh. 1367  
 Chanet, Anne-Marie 1319, 1323  
 Channon, Rachel 1157  
 Chantraine, Pierre 121, 1325  
 Chao, Yuen Ren 316, 1730,  
 1732, 1736–1738  
 Chapman, Shirley 987, 993,  
 1040  
 Chappell, Hillary 1093  
 Charax 58  
 Charette, Monik 506  
 Charitonova, Irina I 677, 685,  
 699  
 Charles, Elena 922f.  
 Charles, George 922  
 Charmasson, Henri 966  
 Charney, Jean O. 917  
 Charpin, François 696  
 Chaucer, Geoffrey 1642  
 Chaumeil, Jean-Pierre 1464  
 Chen, Matthew Y. 1732  
 Cheng, Chin-chuan 1732  
 Cherpillod, André 1563  
 Cherubim, Dieter 1858  
 Chervel, André 698  
 Chevalier, Jean-Claude 698  
 Chevillard, Jean-Luc 703  
 Chi, Tedee 950  
 Chialant, Dorian 1780f., 1783  
 Chierchia, Gennaro 1070, 1213  
 Childs, G. Tucker 719  
 Chiu, Bonnie Hui-Chun 549  
 Chmura-Klekotowa, Maria 1802  
 Cholewa, Jürgen 1817, 1819  
 Cholodovič, Aleksandr A. 132,  
 1137, 1140, 1146f.  
 Chomsky, Noam A. 26, 28, 31,  
 121, 143, 155, 160, 165, 170,  
 176, 179, 188f., 194–200,  
 202, 204f., 252, 267, 276,  
 321, 329, 331, 341, 344–347,  
 378, 385, 405, 408, 471, 478,

- 497, 500, 507, 541, 555, 667, 671, 818, 835, 841, 948, 1131, 1250, 1262, 1576, 1602, 1612 f., 1640, 1778, 1789 f., 1822, 1894, 1917, 1921
- Chou, Fa-kao 1731 f., 1735
- Chouinard, Michelle M. 1798
- Chrakovskij, Viktor S. 132, 1134
- Christaller, Johann G. 1518
- Christie, Agatha 1641
- Christie, William M. 422
- Christmann, Hans Helmut 109
- Christoph, Ernst-Michael 1901
- Chrysoloras, Manuel 59
- Chung, Sandra 814, 1134, 1197
- Churchward, Clerk Maxwell 577, 1136
- Cicero 76, 1929
- Cinque, Guglielmo 899, 904
- Civil, Miguel 35
- Clahsen, Harald 1767, 1799, 1813 f., 1820–1822
- Clairis, Christos 241, 488
- Clajus 697
- Clark, Douglas 836
- Clark, Eve V. 445, 835, 853, 947–950, 1093, 1796, 1798, 1800–1803
- Clark, Herbert H. 445, 835, 853, 947–950, 1115, 1796, 1801
- Clarke, Sandra 1412
- Claudi, Ulrike 1094, 1099
- Clements, George N. 506, 563, 1005
- Cloarec-Heiss, France 529
- Cloquet, Hippolyte 17
- Coates, Jennifer 1194 f.
- Cobb, Howard 1783, 1819
- Coenen, Josée 1810
- Cohen, Antonie 164
- Cohen, David 486 f., 1741
- Cohen, Marcel 142
- Cohen, Sophia R. 1802
- Cohn, Abigail C. 341, 414
- Cole, Jennifer 554
- Cole, Peter 714, 1091 f., 1106, 1137–1139, 1194, 1199, 1454, 1459, 1863
- Cole, Ronald A. 540
- Coleman, Robert 639, 641, 1087
- Colombat, Bernard 693, 697, 699
- Coltheart, Max 1770, 1773, 1817
- Combrink, Johan G. H. 166, 858
- Comenius, Johann Amos 1930, 1933
- Comrie, Bernard 14, 245, 252, 282, 374, 401, 407, 481, 538, 543, 601, 603, 609–611, 613, 640, 712, 715, 734, 753, 797, 809, 812–815, 817, 846, 848, 939, 1076 f., 1080 f., 1083,
- 1085, 1096, 1108 f., 1120 f., 1137 f., 1152, 1165, 1174, 1180 f., 1186, 1223 f., 1247, 1432, 1556, 1780, 1863, 1868, 1871 f., 1922
- Condillac, Etienne 698
- Conklin, Harold 178
- Conklin, Nancy F. 1016
- Connell, Phil J. 1821
- Conrad, Rudi 436, 536
- Conrady, August 1730
- Consentius 85
- Constantino, Ernesto 1473 f.
- Contini-Morava, Ellen 1189
- Contreras, Heles 1236
- Cooke, Joseph 716, 1013
- Cooper, Jerrold S. 1916
- Cooper, William E. 540
- Corbett, Greville G. 29, 155, 282 f., 318, 362, 770, 787, 1033 f., 1042, 1053 f., 1057–1060, 1062–1065, 1067–1069, 1071, 1080
- Corbin, Danielle 139, 143 f., 583, 834, 837 f., 841, 852, 869, 871, 873 f., 1285–1292, 1294–1296
- Corbin, Pierre 143, 1286, 1290–1292, 1294
- Corblin, Francis 983
- Cornish, Francis 983
- Correa, Julio 1430
- Corriente, Federico 1741, 1748
- Coseriu, Eugenio 7, 115, 222, 682 f., 686, 688, 834, 1221 f.
- Coulmas, Florian 988, 992
- Couper-Kuhlen, Elisabeth 1136
- Coupez, André 1527
- Court de Gébelin, Antoine 698
- Courtés, Joseph 210
- Couté, Bernard 1285
- Couturat, Louis 1563, 1565
- Cowan, William 588, 1415
- Coward, Harold 42
- Cowell, Mark W. 547 f., 765
- Cowgill, Warren 1237, 1240
- Cowie, Anthony P. 1884, 1888–1890
- Cowper, Elizabeth 569 f., 572
- Crago, Martha B. 1820 f.
- Craig, Colette G. (s. Grinevald, Colette)
- Crates 60
- Creider, Chet A. 1135, 1507
- Creider, Jane Tabsubei 1135, 1507
- Creissels, Denis 1092
- Crocco Galéas, Grazia 458, 460, 585, 875, 1706
- Croft, William 258, 261 f., 275, 277 f., 709, 714, 728, 758, 785, 788, 796 f., 986, 1094, 1112–1115, 1139, 1141, 1169, 1171, 1248, 1250, 1252, 1422 f., 1426
- Cromer, Richard 1822
- Crookston, Ian 1212
- Crössmann, Helga 693
- Crowhurst, Megan J. 1474
- Crowley, Ellen T. 953
- Crowley, Terry 1078
- Crowther, Jonathan 1884
- Cruse, David A. 596, 871
- Crystal, David 437, 677, 693
- Cubar, Ernesto H. 1473
- Curganova, Valerija G. 419, 424
- Curtiss, Susan 1822
- Curtius 1167
- Cusihuamán Gutiérrez, Antonio 1455
- Cutler, Anne 432, 539 f., 550, 867, 941, 1080, 1251, 1779 f., 1783, 1785, 1788, 1790 f.
- Cuzzolin, Pierluigi 1214
- Cyr, Danielle 1412–1414, 1416
- Czapkiewicz, Andrzej 1746
- Czech, Franz Herrmann 1555
- Çabej, Ekrem 1647
- D**
- Daalder, Saskia 984
- Dabir-Moghaddam, Mohammad 1161
- Dacia, Martin von 83–85
- Daelemans, Walter 1896–1899
- Dahl, Östen 32, 374–376, 539, 609–611, 613, 656, 661, 797, 806, 809, 1175 f., 1180, 1184 f., 1208, 1527, 1595
- Dahlstrom, Amy 1151, 1413, 1419
- Dái Xuân Ninh 1550 f.
- Dakin, Karen 1447
- Dal, Georgette 143, 1292
- Dal, Ingerid 1268
- Dalalakis, Jenny 1821
- Daly, Margarita H. 193
- Daniels, Peter T. 1921
- Danon-Boileau, Laurent 983
- Danso, Mary 1519
- Darmesteter, Arsène 138–143, 1287
- Darski, Józef 1271 f.
- Darwin, Charles 8, 17, 98
- Daum, Edmund 1309
- Dauzat, Albert 140, 142
- Davane, G. V. 1722
- Davidson, Donald 727
- Davies, John 713, 767, 800, 1202 f.
- Davis, Carol 1781
- Davis, John F. 1203 f.
- Davis, Phillip W. 923
- Davis, Stuart 430, 560

- Day, A. Colin 1833  
 Day, Christopher 979  
 Dayley, Jon P. 764, 766, 1131,  
 1215  
 Debrunner, Albert 121, 452,  
 1316, 1320, 1325  
 DeCamp, David 1653  
 Déchaine, Rose-Marie 1413,  
 1419  
 Declerck, Renaat 1166  
 Dede, Müserref 1360  
 Dedenbach, Beate 1271  
 Dederding, Hans-Martin 1278  
 Deegener, Günter 300  
 Deely, John 211  
 Dees, Anthonij 1701  
 DeGraff, Michel 1654  
 Deighton, Len 836  
 Delacroix, Eugène 163  
 DeLancey, Scott 1010, 1027,  
 1082, 1592, 1595f.  
 Delbridge, Arthur 832  
 Delbrück, Berthold 100f., 104f.,  
 112, 642  
 Delfitto, Denis 1176  
 Delgaty, Alfa 1098  
 Dell, François 143, 338, 1288f.  
 Deloria, Ella 547  
 Delsing, Lars-Olof 1069  
 DeMatteo, Asa 1560  
 Demers, Richard 721, 723  
 Demuth, Katherine A. 1019,  
 1043, 1799  
 Dench, Alan 1009, 1079, 1086f.,  
 1204, 1491, 1493  
 Denes, Gianfranco 1818f.  
 Denny, J. Peter 922f., 993, 1016,  
 1018, 1121, 1417  
 Denz, Adolf 1746f.  
 Depraetere, Ilse 1166  
 Derbyshire, Desmond C. 742f.,  
 919f., 979f., 987, 993, 1024,  
 1040, 1078, 1098, 1202f.,  
 1208  
 Derwing, Bruce L. 420, 854,  
 866, 1766  
 Dešeriëva, Tamara I. 604  
 Desloges, Pierre 1554  
 Deuchar, Margaret 1556  
 Deutschbein, Max 814  
 DeWolf, Charles M. 1475, 1479,  
 1481  
 Dez, Jacques 702  
 Díaz Cahuachi, Pedro 1471  
 Díaz, Manungo 1471  
 Dibble, Charles 1452  
 Diderichsen, Paul 146  
 Didymus 62  
 Diem, Werner 703, 1096, 1741f.,  
 1748, 1750  
 Diessel, Holger 983  
 Diesterweg, Adolph 109  
 Dietrich, Rainer 1808  
 Diewald, Gabriele 1273  
 Diez, Friedrich 139f., 158  
 Diffloth, Gérard 1128  
 Dijkhoff, Marta 1658  
 Dik, Simon C. 165f., 196, 973f.,  
 1105, 1109, 1130f.  
 Dillon, Myles 569, 571  
 Dimmendaal, Gerrit J. 637, 751,  
 758, 766, 822, 1032f., 1507,  
 1509–1511, 1514f.  
 Dineen, Francis P. 54  
 Diogenes Laertius 54  
 Dionysios Thrax 55–58, 77, 87,  
 674, 694, 708, 732, 745, 1149  
 Dirven, René 1049  
 Dittmann, Jürgen 1817  
 Dixon, Robert M. W. 341f.,  
 373, 400, 432, 443f., 621,  
 637, 685f., 714–716, 719,  
 728, 742, 757, 767, 796f.,  
 817, 985, 1004, 1018, 1021,  
 1026, 1032, 1035, 1038,  
 1040–1042, 1080–1083,  
 1108, 1111–1116, 1135, 1137,  
 1158, 1197, 1209, 1491,  
 1493–1495  
 Dobson, William Arthur Charles  
 Harvey 1053  
 Doerfer, Gerhard 1579, 1649,  
 1651  
 Dogil, Grzegorz 1819  
 Doke, Clement Martyn 719,  
 854, 1538  
 Dokulil, Miloš 446, 888  
 Doleschal, Ursula 1275  
 Dolphyne, Florence 1518  
 Dominicy, Marc 92, 96, 211  
 Domokos, Péter 330  
 Don, Jan 165f., 446, 950  
 Donaldson, Bruce 658  
 Donaldson, Tamsin 822, 1004,  
 1196, 1198f.  
 Donatus 63–65, 77f., 82–84,  
 674, 696, 1929, 1931f.  
 Donegan, Patricia Jane 401,  
 1594  
 Donhauser, Karin 1274  
 Donnellan, Keith S. 983, 985  
 Donner, Kai 1376  
 Donnet, Daniel 59, 65  
 Dordian, Monique 1818  
 Dorian, Nancy 1043, 1655f.  
 Doron, Edit 327, 328  
 Downer, Gordon B. 588, 1731f.  
 Downing, Pamela 352, 837, 852,  
 907, 909, 1258, 1760  
 Dowty, David 201, 907, 909,  
 1168  
 Drabbe, Peter 1501  
 Drach, Erich 108  
 Dragunov, A. A. 677  
 Drapeau, Lynn 1412, 1414, 1418  
 Dressler, Wolfgang Ullrich 26,  
 31, 33, 122, 161, 219f., 270,  
 289f., 292–294, 296, 336,  
 338f., 361, 424, 431, 433,  
 455–457, 460, 485, 490, 492,  
 507, 514f., 518, 530–533,  
 550, 555f., 579, 583–585,  
 737, 753, 755, 781, 838, 868,  
 874, 879, 891, 934, 936, 940,  
 979, 1045f., 1049f., 1276,  
 1602f., 1623, 1789, 1817–  
 1820, 1822, 1902  
 Driem, George van 767, 1214  
 Dringel-Techt, Eva 290  
 Drivaud, Marie-Hélène 1883  
 Droixhe, Daniel 93  
 Dromi, Esther 1795, 1821  
 Drossard, Werner 1475, 1479,  
 1483, 1486  
 Drozd, Lubomír 1924  
 Drude, Sebastian 1836f., 1851  
 Dryer, Matthew 539, 1053, 1062  
 Dubnova, Yelena Zinov'evna  
 1216  
 Dubois, Claude 1886f., 1889f.  
 Dubois, Jean 139, 142f., 536,  
 1285, 1291, 1885–1887,  
 1889f.  
 Duchan, Judith F. 983  
 Dudas, Karen 568  
 Duhoux, Yves 122  
 Dul'zon, Andrei P. 1376, 1382,  
 1384, 1386  
 Dulay, Heidi 1806  
 Duličenko, Aleksandr D. 1563  
 Dunnigan, Timothy 785, 788  
 Durand, Jacques 500  
 Durie, Mark 329, 521, 713, 717,  
 781, 978, 1058, 1082  
 Durkheim, Emile 4, 142  
 Durnovó, N. N. 127  
 Durovic, Lubomir 697, 1309  
 Durrell, Martin 1685  
 Dyk, Janet 362
- E
- Eades, Diana 992, 1095, 1097  
 Ebbinghaus, Horst 1560  
 Ebeling, Carl 164, 1081, 1084  
 Ebert, Karen H. 985  
 Ebert, Robert Peter 1685,  
 1688f., 1693  
 Eckert, Gabriele 1707  
 Eckman, Fred 276  
 Eco, Umberto 211f., 882  
 Efremova, Tatjana F. 131, 889  
 Egede, Paul 1389  
 Egerod, Søren 546, 548  
 Ehrlich, Susan 1174  
 Eichinger, Ludwig M. 931, 1278,  
 1683  
 Eichler, Ernst 1905  
 Eijk, Jan P. van 701f.

- Einhauser, Eveline 100  
 Eisele, John C. 1746  
 Eisenberg, Peter 108, 1269,  
     1916, 1920–1922  
 Ekdahl, Muriel 1210  
 Elbert, Samuel 717  
 Elffers-van-Ketel, Els 105  
 Elimelech, Baruch 571, 573  
 Elson, Benjamin F. 442, 489,  
     541, 1831  
 Emenanjo, E. Nolue 934, 1128  
 Emmorey, Karen 541, 1560  
 Emonds, Joseph E. 681, 811,  
     1257  
 Encrev  , Pierre 1700  
 Endzelin, Jan 1650  
 Engberg-Pedersen, Elisabeth  
     1559  
 Engelen, Bernhard 671, 688  
 Engelhardt, Wolf von 16  
 Engelien, August 113  
 England, Nora C. 1012, 1033,  
     1100  
 Englebretson, Robert 1116  
 Engler, Rudolf 216  
 Eoin, Gear  id Mac 627  
 Epstein, Richard 985  
 Erasmus 1930  
 Erbaugh, Mary S. 1027  
 Erben, Johannes 160, 677, 891–  
     895, 898, 902, 1273, 1278,  
     1621, 1686, 1694f.  
 Erdman, Paul 836  
 Erd  di, J  zsef 823  
 Erickson, Barbara 193  
 Erlebach, Peter 1677  
 Erlenkamp, Sonja 1560  
 Erlinger, Hans Dieter 104f., 110,  
     115  
 Ernout, Alfred 120  
 Eroms, Hans-Werner 1278  
 Ersen-Rasch, Margarete I. 1106  
 Ertel, Suitbert 1811f.  
 Erwin, Wallace 1742, 1747  
 Escalante, Fernando 1092  
 Eschenlohr, Stefanie 1275  
 Espinosa P  rez, L. 1464  
 Ettinger, Stefan 1045, 1047,  
     1051  
 Euler, Wolfram 122  
 Eutyches 77, 80  
 Evans, Nicholas D. 709, 711,  
     743f., 753, 918, 926, 1033f.,  
     1042, 1057, 1079, 1086f.,  
     1108, 1491, 1854, 1880  
 Evans, Roger 1896, 1899  
 Everett, Daniel L. 1214, 1216,  
     1465  
 Evett, Lindsay J. 1771  
 Ewen, Colin J. 507  
 Eynde, Karel van den 222, 1285
- F
- Fabb, Nigel 412, 882, 1255f.  
 Fabricius, Anne H. 1056, 1061  
 Fabricius-Hansen, Cathrine 904,  
     1272–1274  
 Fagan, Sarah 1149  
 Faib  , Klaus 1627, 1668f., 1671,  
     1673–1675  
 Fanselow, Gisbert 161, 904–  
     906, 908–910, 1278  
 Faraclas, Nicholas 759  
 Farrar, Frederic W. 16, 19  
 Faust, Manfred 157  
 Fay, Donald 1784, 1790f.  
 Fearn, John 92  
 Fedson, Vijayarani Jotimuttu  
     1138  
 Fejos, Paul 1464  
 Feldman, Laurie B. 206, 1783,  
     1922  
 F  l  ice, Th  odore de 1285  
 Fellbaum, Christiane 1821  
 Fellman, Jack 1343  
 Fenk-Oczlon, Gertraud 289  
 Ferguson, Charles A. 1741  
 Ferguson, Charles J. 179, 1654  
 Fettes, Mark 1564  
 Feuillet, Jack 983, 993, 1274  
 Fife, James 482  
 Filin, Fedot P. 135  
 Fill, Alwin 1619f., 1623, 1633f.  
 Fillmore, Charles J. 260, 667,  
     983, 1099  
 Finck, Franz Nikolaus 1835,  
     1837  
 Firestone, Homer L. 711  
 Firth, John Rupert 149, 150f.,  
     153, 455, 494, 554, 589, 616  
 Fischer, Olga 1590  
 Fischer, Renate 1555  
 Fischer, Susan 1555, 1559f.  
 Fischer, Wolfdieterich 821, 1746–  
     1751  
 Fischer-J  rgensen, Eli 489, 506,  
     508  
 Fisher, Herbert A. L. 52  
 Flacourt, Etienne de 702  
 Fl  mig, Walter 678, 699  
 Fleisch, Henri 70  
 Fleischer, Wolfgang 160, 250,  
     355f., 744, 886f., 889, 891–  
     894, 898f., 902, 931f., 950,  
     958, 970, 1267, 1276, 1635,  
     1902  
 Fleischman, Susanne 1184, 1708  
 Fliedl, G  nther 453  
 Flood, Walter Edgar 1926  
 Fludernik, Monika 983, 985, 988  
 Flydal, Leif 146f.  
 Fodor, Jerry A. 286, 1822  
 Fogel, Michael 92  
 Foley, James 1285  
 Foley, William A. 760, 795, 859,  
     1012, 1032, 1036, 1038, 1041,  
     1110, 1120, 1126–1128, 1142,  
     1192f., 1211, 1377, 1475,  
     1479, 1501  
 F  nagy, Ivan 214  
 Forzheimer, Paul 716, 786,  
     1000, 1003f., 1008  
 Ford, Alan 871, 1415  
 Ford, Kevin 1041  
 Forrest, R. A. D. 1731  
 Forsgren, Kjell-  ke 104f., 109,  
     113, 115, 665f., 668–670,  
     672  
 F  rstemann, Ernst 1615, 1620–  
     1623  
 Forster, Keith I. 1790  
 Forster, Kenneth 1766, 1768f.,  
     1771f., 1779, 1781, 1784  
 Fortescue, Michael 254f., 768,  
     922, 979, 981, 1005, 1009,  
     1107, 1216, 1389f., 1394,  
     1400, 1407, 1840  
 Fortunatov, Filipp F. 126–130,  
     1308  
 Foster, George 1004  
 Foster, Mary 1004  
 Fought, John 173, 175f.  
 Fourquet, Jean 1267f., 1270,  
     1272f.  
 Fowler, Anne E. 1822  
 Fowler, C. A. 206  
 Fowler, Carol F. 1783  
 Fowler, George 958, 961  
 Fox, Anthony 250, 1667  
 Fox, Barbara 1099  
 Fradin, Bernard 143  
 Fradkin, Robert A. 1060  
 Fraiberg, Selma 999  
 Frajzyngier, Zygmunt 546, 548,  
     1042, 1058  
 Francis, W. Nelson 253  
 Frank, Paul 1135  
 Franks, Steven 789  
 Frantz, Donald G. 924f.  
 Fraser, Norman 155  
 Frauenfelder, Uli 347, 1772–  
     1774, 1799  
 Frede, Michael 54  
 Fredkin, E. 1894  
 Freeland, L. S. 804  
 Frei, Henri 386, 435, 984, 1099  
 French, Koleen Matsuda 547f.  
 Freud, Sigmund 1782  
 Frey, Eberhard 1270  
 Fried, Mirjam 397  
 Friederici, Angela D. 1818f.  
 Friedrich, Paul 1596  
 Friend, Jewell A. 106  
 Fries, Charles C. 668, 672  
 Fritz, Helmut 438  
 Fromkin, Victoria A. 507, 541,  
     558, 589, 1788–1791

Fruyt, Michèle 1049  
 Fuhrhop, Nanna 458, 905, 1276,  
 1908  
 Fukuda, Shinji 1821  
 Fukuda, Suzy 1821  
 Fulmer, S. Lee 543  
 Funaioli, Hyginus 56  
 Funk, Wolf-Peter 1319  
 Funke, Otto 693  
 Funnell, E. 1790

## G

Gabelentz, Georg von der 9,  
 105, 1590, 1858, 1870f.  
 Gabelentz, Hans Georg Conon  
 von der 95, 97, 99f., 1835  
 Gallmann, Peter 1270  
 Galloway, Brent D. 1217  
 Gamillscheg, Ernst 1616  
 Gandour, Jack 1029  
 Ganschow, Gerhard 122  
 Garde, Paul 383, 708  
 Gardiner, Alan H. 216  
 Gardiner, Donna B. 924f.  
 Garey, Howard B. 1166  
 Garlandia, Johannes de 83  
 Garnham, Alan 983  
 Garrett, Merrill F. 1788f.,  
 1791f.  
 Gasparov, Boris M. 127  
 Gathercole, Susan E. 1822  
 Gawełko, Marek 1708  
 Gaynor, Frank 536  
 Gazdar, Gerald 331, 1896, 1899  
 Geckeler, Horst 1702  
 Geeraerts, Dirk 1050  
 Geerts, Guido 166  
 Geiger, Richard A. 983  
 Geiger, Wilhelm 1712  
 Gelb, Ignaz 1915  
 Genetti, Carol 1205f.  
 Geniusienė, Emma S. 1131f.,  
 1146f., 1157, 1159  
 Gentner, Dedre 374, 755, 795f.  
 George, Kenneth E. M. 958  
 Georgius Choeroboscus 58  
 Gerd, Aleksandr S. 420  
 Gerhardt, Julie 1001  
 Gersbach, Bernhard 933, 1267  
 Gesenius, William 1345, 1347–  
 1349  
 Gessinger, Joachim 97  
 Gessmann, Albert M. 453  
 Ghomeshi, Jila 1061  
 Giacalone Ramat, Anna 1590,  
 1808  
 Giannakis, Giorgios K. 1320,  
 1323  
 Gibbon, Dafydd 1899  
 Gibbon, Edward 52  
 Gibson, Eleanor 1769, 1783

Giegerich, Heinz J. 1780  
 Gil, David 721, 737, 1059, 1061,  
 1067, 1070–1072, 1475  
 Gildea, Spike 987  
 Gilliéron, Jules 1615  
 Gilligan, Gary 538–540, 750,  
 752, 867, 941  
 Gilman, Albert 1001, 1011  
 Ginneken, Jacobus van 813  
 Gipper, Helmut 95  
 Girard, Abbé Gabriel 92, 672,  
 698  
 Girault, Louis 1461  
 Giurescu, Anca 1677  
 Givón, Talmy 286, 371, 401,  
 407, 537–540, 727, 733, 735,  
 784, 786, 795, 799, 810, 814,  
 823, 941, 983, 986, 1081,  
 1086, 1091f., 1115, 1148,  
 1161f., 1202, 1204f., 1210,  
 1229, 1232, 1248, 1423,  
 1590–1593, 1595f., 1871  
 Gládkij, Aleksandr V. 132  
 Glazov, Jurij Ja 1237  
 Gleason, Henry A. 219, 345  
 Glebova, Jvetta 1550  
 Gledhill, Christopher 1564  
 Gleitman, Henry 837  
 Gleitman, Lila R. 837  
 Glinert, Lewis 1096, 1343, 1345,  
 1347–1349  
 Glinz, Hans 104, 107–110,  
 113–116, 668f., 672, 677,  
 1273, 1935  
 Glück, Manfred 64  
 Goad, Heather 1821  
 Goddard, Cliff 266, 481, 603,  
 1076  
 Goddard, Pliny E. 764  
 Goddard, R. H. Ives 1004, 1413,  
 1416, 1418  
 Gode, Alexander 1563f., 1567  
 Godel, Robert 217, 946  
 Goethe, Johann Wolfgang von  
 4, 6–10, 16–20  
 Goffman, Erving 984  
 Goldsmith, John A. 150, 189,  
 259, 339, 496, 500, 506, 880,  
 1898  
 Golla, Victor 1012, 1399  
 Golston, Chris 433, 472, 476,  
 1791  
 Gómez Lopez, Paula 243  
 Gonçales, Cristina Helena R. C.  
 919  
 Gonzales, Andrew B. 1475  
 Goodenough, Ward H. 178, 248  
 Goodglass, H. 1769  
 Goossens, Jan 1582f.  
 Gopnik, Myrna 1820f.  
 Gopsill, Peter 1564  
 Goral, Donald R. 1021  
 Gorbačevskij, Anton A. 520  
 Gordon, Lynn 1009f.  
 Görlach, Manfred 1620  
 Goswami, Usha 1922  
 Gottsched, Christoph 1682, 1690  
 Götz, Dieter 1619, 1627, 1885  
 Gove, Philip B. 1883, 1888  
 Goyaerts, Didier L. 1829  
 Graf, Rainer 933, 1267, 1274  
 Gragg, Gene 1001  
 Graham, A. C. 1734f.  
 Grammont, Maurice 577, 1552  
 Granberry, Julian 702  
 Grasserie, Raoul de la 548f.  
 Graßhoff, Ludwig 1555  
 Gray, Louis H. 249, 254  
 Green, John N. 788, 813  
 Greenbaum, Sidney 316, 790,  
 947  
 Greenberg, Joseph H. 121, 202,  
 273, 275, 351, 366, 373, 405,  
 418, 443, 537, 539, 540, 542,  
 548, 550, 715, 738, 741,  
 745f., 752f., 772–776, 781,  
 784f., 791f., 805, 854, 940f.,  
 1000, 1003, 1008, 1028,  
 1041f., 1053, 1055, 1057f.,  
 1062, 1223, 1225, 1229,  
 1235–1244, 1247–1249,  
 1252, 1400, 1598, 1609, 1796  
 Greene, Elton 917–919, 925  
 Greenough, James Bradstreet  
 1619, 1621f.  
 Gregor, Bernd 1269  
 Gregores, Emma 1107, 1423,  
 1427–1429  
 Greimas, Algirdas J. 210  
 Grenoble, Lenore A. 985  
 Grésillon, Almuth 901, 1286  
 Greule, Albrecht 1903  
 Grevisse, Maurice 571, 932  
 Grewendorf, Günther 672  
 Grice, H. Paul 985  
 Grimes, Joseph E. 541, 1202,  
 1210, 1831  
 Grimm, Jacob 6–8, 16, 19, 96,  
 109, 157–159, 162, 898,  
 905f., 1050, 1267, 1269,  
 1683, 1693  
 Grimm, Ursula 1625, 1628  
 Grimm, Wilhelm 16, 162, 1050  
 Grimshaw, Jane 203, 327, 329,  
 331, 407, 864, 1130  
 Grinevald, Colette (Craig,  
 Colette G.) 1018, 1020f.,  
 1024f., 1028f., 1119f., 1123,  
 1135f., 1206  
 Grodzinsky, Yosef 1769, 1818  
 Grønvik, Ottar 1273  
 Groot, A. Willem de 163f.  
 Groot, Casper de 166, 823f.,  
 1109f.  
 Gross, Gaston 699, 1287  
 Grosseteste, Robert 79, 83

- Grotjahn, Rüdiger 1240  
 Grotfeld, Heinz 1742  
 Grube, Henner 447  
 Gruber, Jeffrey S. 165, 1637  
 Guasch, Antonio 711  
 Gudschninsky, Sarah C. 253, 1832  
 Guilbert, Louis 142, 1287, 1883  
 Guillaume, Jean-Patrick 70, 142, 702f.  
 Guillaume, Paul 1798  
 Guinet, Lynn 1769, 1783  
 Guiraud, Pierre 142, 1288  
 Gulya, János 92  
 Gulyga, Elena 666  
 Gundel, Jeannette 791f.  
 Günther, Hartmut 122, 907, 909, 1278, 1768f., 1773f., 1919  
 Gussmann, Edmund 503, 505–507, 880  
 Guzman, Videa P. de 1477, 1479f.  
 Gvozdanović, Jadranka 273, 638, 640, 772  
 Gvózdev, Aleksandr N. 126, 1799f.  
 Gwynn, Aubrey 1929  
 Gwynn, J. P. L. 1033  
 Gyarmathi, Samuel 92f.
- H**
- Haafsten, Ton van 862  
 Haarmann, Harald 91, 1241  
 Haarmann, Henk J. 1818  
 Haas, Mary R. 538, 814, 1592  
 Haas, Walter 582, 585  
 Haas, William 150–152, 218, 440f., 534, 626, 946  
 Haas, Wim de 166  
 Habermann, Mechthild 1684  
 Häckel, Susanne 1693  
 Hacken, Pius ten 166, 350f., 353f., 357–359  
 Häcki Buhofer, Annelies 656  
 Hackl, Martin 1818  
 Hadj-Sadok, Mahammed 1750  
 Haeckel, Ernst 17–19  
 Haegeman, Liliane 321, 829  
 Haenisch, Erich 554, 568  
 Haensch, Günther 890f.  
 Haeringen, Coenraad Bernardus van 163, 1585  
 Haftka, Brigitta 668  
 Hagège, Claude 916, 923f., 1005f., 1009  
 Hahn, Michael 1732  
 Haider, Hubert 681, 1822  
 Haiman, John 259, 282, 286, 375, 397, 548f., 568, 573, 712, 1012, 1091, 1097, 1202–1205, 1232f., 1252  
 Hajdu, Péter 330  
 Haji-Abdolhosseini, Mohammad 1059, 1061  
 Häkkinen, Kaisa 1328  
 Hakulinen, Auli 1330  
 Hale, Kenneth 328, 521, 582, 584, 821, 1009, 1119f., 1123, 1135f.  
 Hale, William Gardner 1126  
 Hall, Christopher J. 453, 537–540, 550, 1251  
 Hall, Robert A. 237, 381, 925  
 Hall, Robert A. Jr. 249, 251, 1657f., 1756  
 Hall, T. Alan 386  
 Halle, Morris 26, 31, 33, 189, 197f., 200, 202–205, 252, 267f., 341, 346f., 378, 385, 408, 427f., 471–479, 497, 500, 507, 541, 841, 948, 1256, 1262, 1272, 1602, 1612f., 1640, 1732, 1766, 1771, 1778, 1790f., 1894, 1917, 1921  
 Halliday, Michal A. K. 222, 977, 1193, 1195, 1198, 1202  
 Halpern, Aaron L. 389, 395–397, 400  
 Halpern, Abraham M. 711  
 Hamanaka, Toshihiko 1817  
 Hammarström, Göran 457  
 Hammerich, Louis L. 380, 701  
 Hammond, Michael 20, 32, 220, 548, 564, 627  
 Hamp, Eric Pratt 453, 1235  
 Hamzé, Hassan 702  
 Hancock, Ian 1647  
 Hankamer, John A. 1785  
 Hankamer, Jorge 347, 364  
 Hanks, Patrick 1883  
 Hanks, William F. 983–986, 988, 993, 1121  
 Hansen, Björn 132  
 Hansen, Detlef 1821  
 Hansen, Klaus 893  
 Hansson, Kristina 1821  
 Happ, Daniela 1560  
 Harada, Sin I. 993  
 Hardman, Martha 1003, 1009  
 Hardy, Heather K. 547, 582, 584  
 Hare, Mary 1767  
 Harell, Richards 765  
 Harford, Carolyn 802  
 Hargreaves, David 809, 813  
 Hargas, Sharon 433, 546  
 Harley, Heidi 1060  
 Harman, John 809  
 Harning, Kerstin E. 1748  
 Harnisch, Rüdiger 289, 1272, 1908  
 Harré, Rom 1006  
 Harrell, Richard S. 1742  
 Harries, Lyndon 804  
 Harris, Alice C. 1082, 1085, 1137  
 Harris, James W. 200  
 Harris, John 430  
 Harris, Maria De Boe 193  
 Harris, Martin 599  
 Harris, Zellig S. 6, 121, 150, 164, 174, 176–182, 184–187, 189, 194f., 200, 205, 345, 379, 405, 418, 436, 440f., 453, 478, 490–492, 494, 496, 554, 573, 582, 584, 589, 620, 622, 666, 672  
 Harrison, Sean P. 559, 1135  
 Hartman, Lawton 1732  
 Hartman, Reinhard R. K. 568, 597  
 Hasan, Ruqaiya 1202  
 Haselbach, Gerhard 669  
 Hashimoto, Mantaro J. 1147  
 Hashimoto, O. K. Y. 781  
 Haspelmath, Martin 361f., 549, 658, 660f., 663, 709, 719, 799, 1041, 1132, 1137, 1140, 1142, 1158, 1213, 1218, 1580, 1590, 1871, 1877  
 Hasselrot, Bengt 1045, 1048f., 1051  
 Hatcher, Anna Granville 1705  
 Haudricourt, André 1545, 1731, 1734  
 Haudry, Jean 122  
 Haupenthal, Reinhard 1563  
 Hauri, Hans Walter 1324  
 Häusler, Frank 508  
 Hausmann, Franz Josef 694  
 Haverkate, Henk 1192  
 Haverkort, Marco 1819  
 Havers, Wilhelm 253  
 Haviland, John B. 993, 1092, 1158  
 Hawkins, John A. 538–540, 550, 750, 752, 867, 941, 983–985, 1251, 1267f., 1780  
 Hay, Jennifer 1786  
 Hayes, Bruce 431, 1262, 1613  
 Hays, David G. 667  
 Head, Brian 788, 1001  
 Healey, Alan 1829  
 Healey, Phyllis M. 558, 561  
 Heath, Jeffrey 1010, 1023, 1041, 1493, 1495, 1745, 1750  
 Hecht, Barbara F. 1798, 1802  
 Hedberg, Nancy 791f.  
 Hedin, Eva 1187  
 Heeschen, Claus 1818  
 Heeschen, Volker 983, 993  
 Hegel, Georg Wilhelm Friedrich 2  
 Heger, Klaus 211, 222, 685  
 Heidolph, K. E. 678, 684  
 Heine, Bernd 32, 743, 1019, 1032, 1035–1037, 1039f., 1042f., 1092, 1094, 1161, 1232, 1507, 1590–1593, 1595f., 1654, 1752

- Heinicke, Samuel 1554  
 Heinemann, S. 87  
 Heinz, Sieglinde 1700  
 Helbig, Gerhard 109, 114, 665,  
   667, 671, 676–679, 683, 690,  
   693  
 Helias, Petrus 84  
 Heller, Louis G. 960  
 Helmbrecht, Johannes 1053,  
   1060, 1249, 1253  
 Helwig, Christoph 697f.  
 Hempel, H. 677  
 Hempen, Ute 1272, 1688  
 Henderson, Leslie 1769, 1771,  
   1773, 1775, 1779, 1783–1785  
 Hendrick, Randall 784  
 Hengeveld, Kees 728f., 938,  
   974f., 1105, 1107, 1192, 1199  
 Henry, David 1399  
 Henry, Patrick Leo 1649  
 Henzen, Walter 19, 160, 891,  
   898, 901, 1275, 1684  
 Herbermann, Clemens-Peter  
   250, 598, 1278  
 Herbers, Birgit 1694  
 Herbert, Robert K. 1043  
 Hercus, Luise 1491  
 Herder, Johann Gottfried von 7  
 Herling, Simon Heinrich Adolf  
   105, 669  
 Hermann, Eduard 104, 106, 677  
 Hermon, Shula 1822  
 Herms, Irmtraud 1045, 1048  
 Hertog, Cornelis H. den 162  
 Hervás y Panduro, Lorenzo 91  
 Hervey, Sándor G. 213f.  
 Hess, Thom 701f.  
 Heßmann, Jens 1560  
 Hetzron, Robert 750, 1077  
 Hewitt, Brian George 800, 802,  
   824, 980f., 1004, 1038, 1096,  
   1203, 1205, 1215  
 Hewitt, Helen-Jo Jakusz 122  
 Heynatz, Johann Friedrich 108  
 Heyse, Karl W.L. 104f.  
 Higginbotham, Jim 862  
 Hilgard, Alfred 54, 56–59  
 Hill, Jane H. 926  
 Hill, Kenneth C. 926  
 Hillis, Argye E. 1818  
 Himmelmann, Nikolaus P. 717,  
   721f., 983, 1475f., 1479f.,  
   1482, 1486  
 Hinch, H. E. 373  
 Hinds, John 1094, 1215  
 Hinnebusch, Thomas 791, 1099  
 Hinrichs, Erhard 1170  
 Hinüber, Oscar von 1712  
 Hirshfeld, H. 74  
 Hirt, Hermann 1686, 1691  
 Hjelmslev, Louis 3, 24, 59, 95,  
   146f., 210f., 220f., 235, 273,  
   275, 378, 380, 405f., 452,

## Verzeichnisse und Sprachenkarten

- 608, 688, 834, 1076, 1084,  
   1086, 1127  
 Hò Lê 1545, 1550  
 Hocart, A. M. 577  
 Hock, Hans Henrich 121, 402,  
   475, 576f., 1667  
 Hockett, Charles Francis 26, 53,  
   150–152, 155, 178–181,  
   183–189, 204, 211, 214, 219,  
   235–237, 241, 248f., 350,  
   379, 383, 405–407, 409f.,  
   418, 427, 439–441, 447,  
   452f., 458, 463f., 475, 546,  
   573, 582, 588, 622, 627,  
   656f., 660, 694, 743, 866,  
   1225, 1732  
 Hoddingott, William G. 1080  
 Hoek, Karen van 1823  
 Hoeksema, Jacob 153, 165f.,  
   201, 206, 851f., 855f., 861,  
   863, 900, 1255  
 Hoekstra, Teun 166, 346, 864,  
   1250  
 Hoenigswald, Henry M. 100,  
   118, 345, 1887  
 Hoff, Berend J. 164, 166  
 Hoffman, Lothar 1924  
 Hoffmann, J. 720, 725f.  
 Hoffmann, Walter 1688  
 Höfler, Manfred 1698, 1705  
 Hofmann, Johann Baptist 1217  
 Hofstede, Ben T. M. 1818  
 Hogben, Lancelot 1926  
 Hohenberger, Annette 1560  
 Höhle, Tilman 161, 351–354,  
   892  
 Holdcroft, David 214  
 Holes, Clive 750, 1741f.  
 Holm, John 1654, 1657  
 Holmberg, Märta A. 905  
 Holmquist, Jonathan Carl 1036  
 Holton, David 1217  
 Holtus, Günter 1703  
 Holtz, Louis 63, 696  
 Homer 55, 58, 122, 1064, 1149,  
   1582, 1611  
 Honda, Isao 1208, 1210f.  
 Honey, Patric J. 1552  
 Hooper, Joan Bybee 189, 336,  
   348, 496f., 1582, 1602  
 Hoppe, Gabriele 1275, 1277  
 Hopper, Paul J. 259, 287, 401,  
   727f., 747, 764, 795, 799f.,  
   809, 1084, 1170, 1174, 1232,  
   1590–1592, 1595f.  
 Horn, Laurence 285  
 Hornby, A. S. 1884  
 Horne Tooke, John 1228  
 Hotzenköcherle, Rudolf 1684  
 Houghton, Donald E. 1643  
 Housie, Maurice 1202  
 Householder, Fred W. 406  
 Housen, Alex 1808  
 Hovdhaugen, Even 242f., 582,  
   1107  
 Howard, David 1783, 1818  
 Howe, Stephen 1270  
 Hristea, Theodor 584  
 Hualde, José Ignacio 1069  
 Huber, Walter 811  
 Hübschmann, J. H. 100  
 Huck, Geoffrey J. 549  
 Huddleston, Rodney D. 248,  
   602, 656, 810, 816  
 Hudson, Grover 337, 463f., 467,  
   474, 479, 1216  
 Hudson, Richard A. 313, 555,  
   1891  
 Huges, Richard 1619  
 Hugutio von Pisa 83  
 Hulst, Harry van der 165, 842  
 Humboldt, Alexander von 7–  
   10, 12  
 Humboldt, Wilhelm von 19, 32,  
   92, 94f., 98f., 108, 116, 127,  
   157, 742, 840, 998, 1221–  
   1224, 1228, 1247, 1474,  
   1575f., 1590, 1835, 1873  
 Humphreys, Glyn 1771  
 Hunnius, Klaus 1285, 1701f.  
 Hurch, Bernhard 1651  
 Hurford, James R. 771f., 774,  
   1064, 1791  
 Hutchison, John P. 374, 1135,  
   1205  
 Hyman, Larry M. 330, 383,  
   1099, 1232  
 Hymes, Dell 173, 175f.

## I

- Iacobini, Claudio 1706  
 Iatridou, Sabine 818  
 Ibn ʕUsfur (al-ʔAshibili) 68f.,  
   73f.  
 Ibn al-ʔAnbari, Abu l-Barakat  
   69, 73  
 Ibn al-Sarraj, Abu Bakr 68f.,  
   71–73, 703  
 Ibn Faris, Ahmad 71  
 Ibn Jinni, Abu l-Fath 67–69,  
   70f., 73  
 Ibn Qureish 74  
 Ibn Sina, Abdallah 69  
 Ibn Yaʕish, Ibn ʕAli 71, 73  
 Ickelsamer, Valentinus 675  
 Ifrah, Georges 437, 773  
 Igla, Birgit 1644, 1650  
 Ilson, Robert 1889f.  
 Ineichen, Gustav 1235  
 Ingham, Bruce 1742, 1746, 1748  
 Ingo, Rune 647f.  
 Ingram, David 716, 784, 1003  
 Inhoff, Albrecht 1770  
 Inkelas, S. 150

- Innes, Gordon 1135  
 Irigoin, Jean 1322  
 Irmer, Roland 446  
 Irvine, Judith T. 984  
 Isaac, Luc 1285  
 Isachenko, Alexander V. 129, 134, 609, 1091, 1099, 1303, 1305 f., 1309  
 Isakov, Isak A. 1367  
 Ishihara, Shinichiro 1060  
 Isserlin, M. 1818  
 Iturrioz-Leza, José Luis 238 f., 242, 1054, 1056, 1058–1061, 1064  
 Ivanova, Irina P. 420  
 Iverson, Gregory K. 878  
 Ivo, Hubert 676, 696
- J**
- Jaberg, Karl 521  
 Jachnow, Helmut 1309  
 Jachontov, Sergej E. 1237, 1239 f., 1552  
 Jackendoff, Ray 161, 188, 197 f., 346, 667, 671, 681, 841 f., 845, 847, 849, 900, 948, 1067–1069, 1250, 1790  
 Jackson, Howard 1884, 1886–1889  
 Jacobsen, Thorkild 37  
 Jacobsen, William H. 721–723  
 Jacobsen, William H. Jr. 985, 1012, 1088, 1202–1204  
 Jacobson, Steven 922  
 Jaeger, Jeri 1778  
 Jaeggli, Osvaldo A. 398  
 Jaggar, Philip J. 990 f., 993  
 Jahr, Ernst Håkon 147  
 Jain, Dhanesh K. 1712  
 Jakobson, Roman 26, 33, 124, 127, 129, 134, 164, 204, 217 f., 238 f., 267–270, 272–277, 279 f., 287 f., 305, 336, 338, 340 f., 380, 436, 438, 442, 444, 493, 536, 608–612, 871, 1076 f., 1192, 1197, 1349, 1377, 1732, 1817 f., 1862  
 James, Deborah 1414  
 Janda, Richard D. 153, 401 f., 437, 580, 585, 877, 936, 1666  
 Jänicke, Otto 1698  
 Janis, Wynne 1559  
 Jankowsky, Kurt R. 1612  
 Janssen, Theo A. J. M. 985–988, 993, 1274  
 Janssen, Theo M. V. 851, 856  
 Janssens, Gerard 1742  
 Jansson, Tore 396  
 Janton, Pierre 1564  
 Janus 65
- Jarceva, Viktoria N. 122, 135  
 Jarema, Gonia 1820  
 Jarvela, Robert J. 983, 1770, 1783 f.  
 Jastrow, Otto 1742, 1746–1751  
 Javed, Naseem 966  
 Jaxontov, Sergej Je. 1177  
 Jedlička, Alois 135  
 Jeffers, Robert J. 401  
 Jelinek, Eloise 721, 723 f., 1092  
 Jellinek, Max Hermann 113, 669, 672, 675, 697 f.  
 Jendraschek, Gerd 1056  
 Jensen, Hans 1212, 1214  
 Jensen, John T. 20, 217, 346, 405, 573, 622, 628, 641, 835  
 Jerome 76  
 Jespersen, Otto 18, 110 f., 145 f., 201, 344, 352, 381 f., 446, 680, 688, 699, 758 f., 945, 948 f., 958, 1068, 1112, 1114, 1584, 1637, 1858, 1871  
 Jetté, Jules 1399–1401  
 Jisa, Harriert 1803  
 Joannes Balbus (John of Genoa) 84 f.  
 Joeres, Rolf 1684  
 Johanson, Lars 753  
 Johnson, Janet H. 40  
 Johnson, Kyle 323, 329  
 Johnson, Mark 286  
 Johnson, Robert E. 423  
 Johnson, Samuel 1638  
 Johnson, Trevor 1557, 1560  
 Johnson-Laird, Philip N. 1091  
 Johnston, Raymond Leslie 560  
 Jolley, Catherine 1151  
 Jolly, Julius 106  
 Joly, André 694, 697, 699  
 Jonasson, Kerstin 986  
 Jones, Daniel 149, 251, 383  
 Jones, Eliza 1399–1401  
 Jones, Peter 985  
 Jones, Sir William 18, 93 f.  
 Jong, Arie de 1563  
 Jönsjö, Jan 1677  
 Joos, Martin 179  
 Jordens, Peter 1810–1812  
 Jørgensen, Mogens Wied 1269  
 Joseph, Brian D. 330, 397 f., 401 f., 815, 1611, 1666  
 Josephs, Lewis 990  
 Joshi, Shivaram D. 508  
 Jouison, Paul 1559  
 Joyce, James 1621  
 Juillard, Alphonse 1637  
 Jung, Walter 677  
 Junger, Judith 1345, 1347–1349  
 Jungraithmayr, Herrmann 555  
 Jurafsky, Daniel 1049, 1893  
 Jurjani, ŠAbd al-Qahir 67, 73  
 Justinian 63
- K**
- Kaburaki, E. 734  
 Kageyama, Taro 882  
 Kahane, Henry 1889  
 Kahane, Renée 1889  
 Kahn, Daniel 336  
 Kainz, Friedrich 1621  
 Kaisse, Ellen M. 393, 411, 1264 f.  
 Kalocsay, Kálmán 1565, 1567  
 Kaltz, Barbara 693 f., 696–700  
 Kamenskikh, Uljana K. 1388  
 Kamhi, Alan G. 1820  
 Kamp, Hans 911, 1170, 1179  
 Kamptz, Hans von 122  
 Kanakin, Igor A. 1376  
 Kandler, Günther 418 f.  
 Kanngießer, Siegfried 887  
 Kant, H. 888  
 Kapesh, An Antane 1419  
 Kaplan, Lawrence 922  
 Kaplan, Ron 1894  
 Karaúlov, Jurij N. 135  
 Karcevski, Serge 719  
 Karger, Nikolaj K. 1376  
 Kari, James 412, 1400  
 Karius, Ilse 446  
 Karlgren, Bernhard 1730–1732, 1735  
 Karlsson, Fred 648, 652, 980, 1114, 1127, 1330 f., 1338  
 Karlsson, Göran 651  
 Karlsson, Keith E. 1709  
 Karmiloff-Smith, Annette 983, 1800  
 Károly, Sándor 1132  
 Karpf, Annemarie 290, 294, 1615, 1619, 1623  
 Karstien, Hans 547–550  
 Karttunen, Frances 400  
 Karttunen, L. 1896  
 Kasevič, Vadim B. 1237, 1239 f.  
 Kastenholz, Raimund 1092  
 Kaster, Robert A. 58  
 Kastovsky, Dieter 19, 195, 306, 441, 446, 465, 834, 847, 882, 906, 1275, 1626, 1628, 1631 f., 1674–1676  
 Katamba, Francis 155, 405, 436, 541, 622, 641  
 Katharina II 91  
 Katre, Sumitra Mangesh 45  
 Katwijk, Albert van 165  
 Kātyāyana 42 f.  
 Katz, Leonard 1922  
 Kaufman, Terence 1645 f.  
 Kay, Martin 1894  
 Kay, Paul 1093  
 Kaye, Alan S. 546, 1080  
 Kaye, Jonathan D. 430, 582  
 Kayne, Richard S. 323  
 Kazakov, Dimitar 1898

- Kaznelson, Solomon D. 677, 679 f., 682 f., 687–689
- Kean, Marie-Louise 341, 1769, 1818
- Keen, Sandra L. 1079
- Keenan, Edward L. 245, 716, 853, 983, 986, 988, 991–993, 1153, 1161
- Keil, Heinrich 53, 60, 62–65, 1930
- Keizer, M. Eveline 1105
- Kelemen, Józef 1235
- Keller, Jörg 1560
- Keller, Rudi 1250
- Kemmer, Suzanne 285, 1134, 1150, 1158
- Kemp, Alan 694, 1149
- Kempen, Willem 166
- Kempgen, Sebastian 697, 1303, 1308 f.
- Kendall, Martha B. 569
- Kennedy, Benjamin Hall 16
- Kennedy, George A. 1731, 1736
- Kenstowicz, Michael 429, 431 f., 469, 499 f., 502, 508, 1613
- Kent, Roland G. 120
- Kerleroux, Françoise 143, 1286
- Kern, Barbara 1214, 1216
- Kern, Franz 105, 107, 110 f., 162
- Kern, Peter Chr. 1683, 1692
- Kessler-Mesguich, Sophie 703
- Kettemann, Bernhard 458
- Key, Mary Ritchie 488
- Keyser, Samuel J. 506, 570
- Khaidakov, S. M. 1038
- Khrakovskiy, Viktor 1146 f.
- Kibbee, Douglas 696
- Kibrik, Aleksandr E. 482, 715, 1033, 1070 f., 1367 f., 1829
- Kibrik, Andrej A. 1136
- Kiefer, Ferenc 195, 1049
- Kienle, Richard von 1683
- Kihm, Alain 1658 f.
- Kilani-Schoch, Marianne 220, 289, 290, 533, 555 f., 579, 585, 962
- Kilbury, James 337, 489, 494
- Kiliaan, H. N. 951
- Killingley, K. S.-Y. 781
- Kim, Kong-On 568 f.
- Kimball, Geoffrey 1056, 1060
- Kimenyi, Alexandre 1135, 1139, 1527–1531, 1533, 1535, 1538 f., 1541
- Kimhi, Rabbi David 1345, 1348
- King, Gareth 482, 660, 792
- King, Robert D. 1612
- King, Ross 1919
- Kinkade, M. Dale 702, 720, 723
- Kiparsky, Paul 26, 31, 121, 165, 189, 203, 331, 339, 346, 367, 412, 433, 471, 505, 508, 570, 653, 841, 854, 873, 879, 880, 946, 948, 1256, 1263, 1612 f., 1790
- Kirsner, Robert S. 983, 993, 1099
- Kirwan, Brian Edmond Renshaw 1187
- Kisseberth, Charles 500, 502
- Kitagawa, Yoshihisa 560
- Kittredge, George Lyman 1619, 1621 f.
- Kiyomi, Setsuko 1061
- Klaiman, Miriam 1078, 1085, 1149 f.
- Klammer, Marian 362
- Klaus, Georg 3
- Klausenburger, Jürgen 1602 f.
- Klavans, Judith L. 327, 393 f., 397 f., 400
- Kleiber, Georges 990, 1290, 1296
- Klein, Hamet 1024
- Klein, Harriet E. Manelis 1421
- Klein, Henny 1217
- Klein, Thomas 1691
- Klein, Wolfgang 983, 1180, 1273, 1810
- Kleinhenz, Ursula 386
- Kleinschmidt, Samuel 1389
- Klimaszewska, Zofia 1045 f., 1049 f.
- Klimov, Georgij A. 1082
- Klingebiel, Kathryn 1705
- Klingenschmitt, Gert 122
- Kloeke, Gesinus G. 163
- Kloeke, Wus van Lessen 419, 1268
- Klokeid, Terry J. 400
- Kluge, Friedrich 121, 1616, 1684, 1911
- Knobloch, Clemens 104 f., 669, 672, 680, 685, 687
- Knobloch, Johann 1622
- Knox, Robert 16 f.
- Kobler-Trill, Dorothea 1275
- Koch, Harold 1491, 1579
- Koch, Max 1905
- Koch, Peter 1702, 1916
- Koch, Walter A. 211, 214
- Kocourek, Rostislav 1292, 1924
- Kodzasov, Sandro V. 1367
- Koefoed, Geert A. T. 233, 1574, 1580
- Koehn, Edward 1214
- Koehn, Sally 1214
- Koenraads, Willy Henri August 1685
- Kofod, Frances M. 1080
- Kohrt, Manfred 1680
- Koivulehto, Jorma 1328
- Kolk, Herman 1818
- Kollien, Simon 1560
- Kølln, Herman 1312
- Kolmer, Agnes 1067
- Kölver, Ulrike 737, 1042
- König, Ekkehard 362, 1205
- Konrad, Reiner 1560
- Kooij, Jan G. 1250
- Köpcke, Klaus Michael 289, 479 f., 1035, 1036, 1268 f., 1272, 1808 f.
- Koptjevska-Tamm, Maria 658, 939, 976, 980 f., 1072, 1093
- Korhonen, Mikko 701
- Kornfilt, Jaklin 1079
- Koschmieder, Erwin 613, 685 f., 688
- Koskenniemi, Kimmo 207, 1833, 1894
- Köstling, Pamela 1323
- Koul, Omkar N. 1062
- Kouloughli, Djamel Eddine 703
- Koutsoudas, Andreas 628
- Kouwenberg, Silvia 236, 1658 f.
- Kozinceva, Natalija A. 1134
- Kozinsky, Ivan S. 1162
- Koziol, Herbert 945, 1615 f., 1619, 1621, 1694
- Krahe, Hans 1674
- Krámsky, Jiří 249, 252–254, 717
- Krashen, Stephen D. 1806
- Kraus, Alfred 1817
- Krause, Scott Russell 560
- Krauss, Michael E. 1022 f., 1028, 1399
- Krebernik, Manfred 1916
- Kreidler, Charles W. 957, 960
- Krejnović, Eruchim A. 1376 f., 1380
- Krifka, Manfred 1070
- Krishnamurti, Bhadriraju 1032 f., 1043
- Kristoffersen, Lars 1393
- Kröber, Günter 2
- Kroeber, Alfred 916, 1237
- Kroeger, Paul R. 1475, 1481
- Krohn, Robert 1640
- Kronecker, Leopold 782
- Kruisinga, Ettsko 163, 427, 944 f.
- Krupa, Viktor 582, 1238–1244
- Kruszewski, Nikolaj 126, 1602 f., 1606
- Kryk-Kastovsky, Barbara 983
- Kubrjáková, Elena S. 133, 406, 419, 421 f., 424, 886, 888, 892–894
- Kučera, Antonin 1888
- Kučera, Henry 253
- Kuczaj, Stan A. 1798, 1801
- Kudo, Takayaki 1819
- Kuen, Heinrich 1703
- Kuhn, Dorothea 16
- Kuhn, Wilfried 1904
- Kühner, Raphael 655
- Kühnhold, Ingeburg 355 f., 890, 933
- Kuiken, Folkert 1821

- Kuipers, Aert H. 1020, 1024  
 Kujore, Obafemi 1312  
 Kukich, Karen 1897  
 Kukuczka, Elena 1097  
 Kummer, Werner 811  
 Kuno, Susumu 734  
 Kupfer, Peter 676, 1736 f.  
 Kuroda, Sige-Yuki 988  
 Kürschner, Wilfried 160, 901,  
     907, 1278  
 Kuryłowicz, Jerzy 32, 121 f.,  
     365, 376, 456, 610, 612, 680,  
     682, 758, 813, 931, 946, 984,  
     1080, 1590, 1593, 1613  
 Kusters, Wouter 1655  
 Kutscher, P. 1760  
 Kuznecov, Petr S. 128  
 Kuznecova, Ariadna I. 131, 889
- L
- La Brosse, Jean-Baptiste 1412  
 Labov, William 241, 243 f., 1665,  
     1761, 1764  
 Laca, Brenda 873  
 Ladd, Robert 899, 1232  
 Lafford, Barbara A. 272  
 Lafitte, Pierre 1199  
 Lafon, René 1002  
 Lagarde, Jean-Pierre 693, 698–  
     701  
 Laine, Matti 1341, 1819  
 Lakoff, George 283, 286, 297 f.,  
     611, 1018  
 Lakoff, Robin 993  
 Lalleman, Josine A. 1807  
 Lallemand-Rietkötter, Annette  
     890 f.  
 Lallot, Jean 694  
 Lambert, Frédéric 694, 696  
 Lamberterie, Charles de 1325  
 Lambrecht, Knud 984  
 Lancelot, Claude 697 f.  
 Landaburu, Jon 740  
 Landau, Sidney I. 1887–1889  
 Landeweerd, Rita 1174  
 Landsberger, Benno 35, 37  
 Langacker, Ronald W. 257,  
     259–262, 289, 297, 524, 727,  
     926, 983, 1068 f., 1097, 1121,  
     1129, 1203, 1205, 1436 f.,  
     1597  
 Langdon, Margaret 569 f., 1023,  
     1063 f., 1088, 1204  
 Langendoen, Donald Terence  
     150  
 Lapenna, Ivo 1564  
 Lapointe, Steven G. 197, 199,  
     321, 327 f., 346, 395, 900  
 Lardiere, Donna 1813 f.  
 Large, Andrew 1563  
 Larochette, Joe 221 f.
- Larrimore, Bonni 711  
 Larsen-Freeman, Diane 1806  
 Laskowski, Roman 289, 433,  
     596  
 Lass, Roger 430, 465, 507, 637  
 Laudanna, Alessandro 1772,  
     1778 f.  
 Launey, Michel 1433, 1437,  
     1439, 1445, 1447–1449, 1451  
 Laurendeau, Paul 699  
 Laury, Ritva 983, 988, 993  
 Lausberg, Heinrich 655, 1703  
 Law, Vivien A. 60, 63, 86, 696,  
     703  
 Laycock, Donald C. 1756  
 Lazarid, Gilbert 990  
 Lê Quang Trinh 1552  
 Lê Văn Ly 1549, 1552  
 Lê Xuân Thi 1551  
 Le, Tam Duy 814  
 Leau, Léopold 1563  
 Leben, William R. 588  
 Leclère, Christian 699  
 Lecours, André Roch 1817  
 Lee, Jennifer R. 926  
 Leech, Geoffrey 596, 607, 609 f.,  
     747, 837, 1627 f., 1632  
 Leer, Jeff 1404  
 Lees, Robert B. 188, 195 f., 204,  
     352, 357, 906, 907, 1255  
 Lefebvre, Claire 828  
 Lehfeldt, Werner 597, 633,  
     1239–1245, 1308 f.  
 Lehmann, Christian 12 f., 32,  
     242, 244, 262, 660 f., 683,  
     685, 714, 737, 754, 809, 859,  
     864, 1065, 1086, 1202, 1229,  
     1248, 1250, 1481, 1590,  
     1593 f., 1611, 1836, 1844,  
     1860, 1870 f.  
 Lehmann, Thomas 761  
 Lehmann, Winfred P. 121 f.,  
     1594  
 Lehnert, Martin 1637, 1639  
 Leibniz, Gottfried Wilhelm 1555  
 Leinonen, Marja 614, 1183  
 Leirbukt, Oddleif 1273  
 Leisi, Ernst 1619  
 Leiss, Elisabeth 641, 1274  
 Lejeune, Michel 585  
 Lely, Heather K. J van der  
     1820 f.  
 Lemaréchal, Alain 689, 702, 722,  
     1475  
 Lemmens, Marcel 1891  
 Lemoine, Georges 1412  
 Lené, Gustav 1709  
 León, María Lourdes de 1021,  
     1025 f., 1028  
 Leonard, Laurence 1820–1822  
 Leontiev, V. V. 315  
 Leopold, Werner F. 1615, 1621,  
     1623
- Lepsius, Richard 96  
 Leser, Martin 891  
 Lessen, Jacoba H. van 163  
 Leumann, Manu 120 f.  
 Leuninger, Helen 1769  
 Levelt, Willem 1769, 1780  
 Leven, Regina 1560  
 Levi, Judith 352 f., 358, 909,  
     1255, 1262  
 Levin, Aryeh 71  
 Levin, Beth 845, 847  
 Levin, Juliette 559  
 Levin, Saul 441, 444  
 Levinsohn, Stephen H. 1197,  
     1454  
 Levinson, Stephen C. 983 f.  
 Levy, Yonata 1797, 1799, 1822  
 Lévy-Bruhl, Lucien 1091  
 Lewandowski, Theodor 436  
 Lewis, Geoffrey L. 232, 314,  
     351, 825, 969, 1073, 1097,  
     1106, 1191, 1193 f., 1197 f.,  
     1215, 1583  
 Lewy, Ernst 1229, 1311  
 Leys, Odo 1269, 1905  
 Li, Charles N. 240, 749, 810,  
     940, 950, 1020, 1113, 1191 f.  
 Li, Fang-kuei 1730, 1733  
 Li, Jinxi 704  
 Liang, Franklin Mark 1897  
 Libera, Alain De 696  
 Liberman, Mark 899, 1232  
 Lichtenberk, Frantisek 999, 1097  
 Liddell, Scott 423, 1557, 1560  
 Lieb, Hans-Heinrich 1836 f.,  
     1851  
 Lieber, Rochelle 31, 199 f., 202,  
     318–321, 337, 350, 354, 357,  
     408 f., 411, 414, 433, 445 f.,  
     479, 568 f., 572–574, 606,  
     860 f., 871, 873, 880, 900,  
     904, 908, 946, 948, 1255  
 Lightfoot, David W. 1265  
 Lightner, Theodore M. 421, 515  
 Lima, Susan D. 1770  
 Limburg, Machiel J. 823  
 Lindemann, Margarete 1707  
 Lindenfeld, Jacqueline 1078  
 Lindner, Thomas 1704  
 Linebarger, Marcia C. 1818  
 Linell, Per 338, 507, 1792  
 Lipka, Leonhard 122, 446,  
     464 f., 834 f., 837, 887, 894,  
     901, 906, 1626 f., 1632, 1634  
 Littell, Robert 836  
 Littré, Émile 16, 18  
 Liu, Jian 703  
 Ljungerud, Ivar 1270  
 Llamzon, Teodoro A. 1479  
 Lombardi, Linda 548  
 Lomonósov, Michail V. 125,  
     1300  
 Lomtadze, Elizbar A. 1367

- Long, Antony A. 54  
 Longacre, Robert E. 179, 192 f.,  
 1202, 1501, 1831 f.  
 Lönnrot, Elias 1338 f.  
 Loos, Eugene 1141  
 Lopatin, Vladimir V. 132, 419,  
 424, 892, 1309  
 Lopez, Cecilio 1474, 1486  
 Loporcaro, Michele 452, 1702  
 Lorentz, James P. 1271  
 Lorentz, Ove 147  
 Lorimer, David L. R. 766  
 Lotz, John 436  
 Lötzscher, Ronald 1569  
 Lounsbury, Floyd 176, 178 f.,  
 181–183, 186 f.  
 Loveland, Katherine 999  
 Lü, Shuxiang 1736  
 Lüdtke, Helmut 1250, 1685  
 Ludwig, Otto 1917  
 Lukatela, Georgije 206, 1783  
 Luraghi, Silvia 639, 641–645  
 Luria, Aleksander R. 1816 f.  
 Luschützky, Hans-Christian 289,  
 454, 458  
 Luther, Martin 1692  
 Lyall, Gavin 836  
 Lynch, John 1204  
 Lyngé, Kristoffer 1397  
 Lyon, Otto 105  
 Lyons, Christopher 327 f., 983 f.  
 Lyons, John 28, 32, 149, 212,  
 228, 248–250, 252–254, 277,  
 496, 596, 671, 686–688, 762,  
 809, 811, 815, 834, 837, 844,  
 973 f., 983, 986, 1071, 1104,  
 1111, 1115, 1192, 1194 f.,  
 1213, 1625, 1627, 1630,  
 1632 f., 1735, 1887, 1889–  
 1891

## M

- Ma, Jianzhong 703  
 Maas, Utz 1917  
 Macaulay, Monica 546, 766,  
 1062  
 MacDonald, John D. 836  
 MacDonald, Lorna 1191, 1195  
 MacKay, Donald G. 576, 1791  
 Mackenzie, J. Lachlan 974 f.,  
 979 f.  
 MacKenzie, Marguerite E. 1412  
 Mackinnon, Roderick 571  
 Mackridge, Peter 814 f.  
 Macris, James 960  
 Macrobius 64  
 MacWhinney, Brian 206, 289,  
 345 f., 364, 795, 879, 1798 f.  
 Madhavan, P. 546  
 Maes, Hubert 701

- Mager, Karl Wilhelm Eduard  
 698  
 Mahapatra, Bijaya P. 1041  
 Mahmud, Ushari Ahmad 1752  
 Maiden, Martin 302, 456, 1069  
 Mailhot, José 1412  
 Majewicz, Alfred F. 526  
 Malchukov, Andrej L. 1151  
 Malicka-Kleparska, Anna 433,  
 834, 878, 1255, 1260  
 Malkiel, Yakov 536, 542 f., 890–  
 892, 1708 f.  
 Mallinson, Graham 825, 1083,  
 1086 f.  
 Malmberg, Bertil 147 f.  
 Malotki, Ekkehart 987, 993 f.  
 Manandhar, Suresh 1898  
 Manandise, Esméralda 437, 817  
 Manaster-Ramer, Alexis 1475  
 Manchester, Martin 1221  
 Mańczak, Witold 121, 1252,  
 1613  
 Manelis, Leon 1785  
 Mann, Thomas 1621  
 Mannheim, Bruce 1004  
 Manning, Christopher D. 1893,  
 1898  
 Manzelli, Gianguidi 739  
 Marantz, Alec 204 f., 321 f., 411,  
 413, 467, 561 f., 565 f., 935,  
 1131 f., 1141, 1477  
 Maratsos, Michael P. 1799  
 Marchand, Hans 160, 195, 198,  
 350, 355 f., 359, 444–446,  
 834 f., 867, 870–872, 874,  
 887, 889–894, 898 f., 901 f.,  
 904–908, 931 f., 935, 944–  
 949, 958, 965, 971 f., 1255 f.,  
 1259 f., 1263 f., 1276, 1631–  
 1633, 1674 f., 1678, 1789  
 Marchese, Lynell 984, 987, 994  
 Marcos, Habte M. 569, 573  
 Marcos, Imelda 284  
 Marcus, Gary F. 479 f., 1798  
 Mardirussian, Galust 916, 924  
 Markey, Thomas L. 511, 519  
 Marle, Jaap van 32, 164 f.,  
 225 f., 230–233, 265, 305,  
 307 f., 310, 317, 342, 347,  
 365, 836, 841, 848, 877 f.,  
 880, 933, 1578, 1580 f.,  
 1585 f., 1641, 1643  
 Marouzeau, Jules 140, 142  
 Marr, N. Ja. 127, 135  
 Marre, Aristide 1474, 1479  
 Marshall, J. H. 696  
 Marslen-Wilson, William D.  
 412, 539 f., 1770, 1773 f.,  
 1778, 1780 f.  
 Martin, Jack 583–585  
 Martin, James H. 1893  
 Martin, Maisa 1341  
 Martin, Pierre 488

## Verzeichnisse und Sprachenkarten

- Martin, Samuel E. 1114, 1732  
 Martinet, André 24, 142, 214,  
 217, 220, 247–249, 251 f.,  
 338, 406, 454, 485–488, 490,  
 492–494, 693, 1285, 1571,  
 1645, 1870  
 Martínez de Sousa, José 958  
 Martínez Moreno, Annette 1699  
 Martini, Francois 1552  
 Marx, Karl 2, 142  
 Mascaró, Joan 340  
 Masica, Colin 1092, 1712  
 Máslav, Jurij S. 129, 612, 648,  
 650  
 Maspero, Henri 1545  
 Mater, Erich 1272, 1637  
 Mates, Benson 54  
 Mathesius, Vilém 668  
 Matisoff, James A. 1593  
 Matonis, Ann 88  
 Matsuda French, Koleen 1474,  
 1476–1478  
 Matsumoto, Yo 1018 f.  
 Mattens, Willy H. M. 165  
 Matthaei, Rupprecht 16 f.  
 Mattheier, Klaus 1682  
 Matthew 284  
 Matthews, George Hubert  
 1197 f.  
 Matthews, John 1820  
 Matthews, Peter Hugo 5, 27, 52,  
 72, 122, 152–155, 172–177,  
 179, 187–189, 195, 200 f.,  
 205, 217, 248 f., 254, 266–  
 268, 320, 346, 350, 356, 363,  
 373, 405, 409, 437, 453, 460,  
 463, 474, 491, 526, 529, 546,  
 548, 588, 595 f., 605, 616,  
 618 f., 622–624, 626 f., 638,  
 641, 647, 655, 657, 816, 837,  
 841, 1074, 1226, 1475  
 Matthiessen, Christian 1202  
 Mauro, Tullio de 699, 1638  
 Maximus Planudes 54, 59, 1076  
 Maxwell, Mike 1899  
 Mayer, Erwin 1615, 1617,  
 1620–1622  
 Mayer, Franz Joseph Carl 17  
 Mayer, Karl 1623, 1768  
 Mayerthaler, Willi 122, 161, 220,  
 270, 274, 278 f., 288–292,  
 294, 443 f., 453, 1650  
 McArthur, Harry S. 981  
 McArthur, Lucille E. 981  
 McBain, Ed 836  
 McCarthy, John J. 151, 189,  
 205 f., 341 f., 385, 410, 431,  
 458, 467, 501, 546–548,  
 553–556, 560–566, 569 f.,  
 572 f., 970, 1045, 1255, 1348,  
 1613  
 McCaughey, Terence P. 1095  
 McCawley, James D. 385 f., 508,  
 965

- McClelland, James L. 1766,  
1780, 1807
- McCloskey, James 328
- McCoard, Robert W. 1186
- McConnell-Ginet, Sally 1213
- McElhanon, Karl 1501
- McFarland, Curtis D. 1473 f.,  
1479, 1481, 1485 f.
- McGregor, Karla 1821
- McGregor, William B. 1093,  
1203, 1863 f., 1871
- Mchombo, Sam 330, 477, 784
- McIntosh, Mary 624 f.
- McIntyre, Linda Lee 563
- Mcire, Ephrem 87
- McKay, Graham 1008
- McMillan, James B. 546
- McNeill, David 300, 540
- McQuown, Norman A. 1869
- Meder, Gregor 447, 650
- Meeussen, Achiel Emiel 166,  
253, 589, 1527, 1532, 1540
- Mei, Tsu-lin 1731
- Meid, Wolfgang 1674
- Meier, Georg Friedrich 4, 91 f.,  
218, 435, 437, 683, 685, 687,  
693, 698 f., 701
- Meier, Michael 1325
- Meier, Richard P. 1823
- Meigret, Louis 697
- Meillet, André 19, 122
- Meillet, Antoine 140, 142, 214,  
245, 249, 537, 813, 1042,  
1590 f., 1611, 1645, 1664 –  
1666
- Meiner, Johann Werner 698
- Meinhof, Carl 1094
- Meisel, Jürgen 1799
- Meisenburg, Trudel 1917–1919,  
1921
- Meijers, Guust 1783–1785
- Mejlach, Michail 1237, 1241
- Mel'čuk, Igor A. 23, 26, 28,  
32 f., 129 f., 135, 235, 237 –  
239, 266 f., 419 f., 424, 436,  
438, 440, 442, 444, 453,  
456 f., 481 f., 511 f., 514 f.,  
517, 523, 525 f., 528, 530,  
532, 536, 542, 552, 582 f.,  
585, 588, 601, 604, 656, 660,  
709, 712, 841, 867, 935 f.,  
945, 1079 f., 1084, 1130,  
1146 f., 1160, 1309
- Melanchthon, Philipp 697, 1930
- Melchert, Craig 641
- Mélis-Puchulu, Agnès 1292
- Méndez Dosuna, Julián 291
- Meneses, Porfirio 1462
- Menn, Lise 879, 1796, 1798,  
1818 f.
- Menninger, Karl 774
- Menovščikov, Georgij  
Alekseevič 1651
- Menyuk, Paula 1820
- Meringer, Rudolf 1623, 1768
- Merlan, Francesca 237, 744, 751,  
835, 916, 924, 926, 978,  
1034, 1082
- Merlini Barbaresi, Lavinia 220,  
290, 457, 460, 583 f., 934,  
1045 f., 1050, 1817
- Merrifield, William R. 405, 622
- Mervis, Carolyn B. 264
- Merx, Adalbert 59
- Metcalf, G. J. 92 f.
- Mey, Jacob 1091
- Meyer, Gustav 1649
- Meyer-Klabunde, Ralf 910 f.
- Meyer-Lübke, Wilhelm 1698,  
1705
- Meys, Willem Johannes 165 f.,  
763, 834
- Miceli, Gabriele 1818 f.
- Michael, Ian 697
- Michels, Victor 1683
- Micläu, Paul 214, 220 f.
- Miettinen, Erkki 1615, 1622 f.
- Mifsud, Manwel 1056, 1745,  
1750
- Mignot, Xavier 1325
- Mikes, Melanie 1797
- Milewski, Tadeusz 134 f.
- Miller, Carol 1821
- Miller, George A. 1091
- Miller, Philip Harold 392, 395,  
402
- Miller, Robert L. 249
- Miller, Roy Andrew 51, 703
- Miller, Ruth 1807
- Miller, Wick R. 1211
- Mills, Anne 1800, 1820
- Milner, George B. 578
- Milner, Jean-Claude 835
- Miloslavskij, Igor' G. 422
- Mimouni, Zohra 1820
- Miner, Kenneth L. 920, 924 f.
- Minnen, Guido 1899
- Mitchell, Terence F. 555, 620
- Mitchell, Tom M. 1898
- Mithun, Marianne 204, 315,  
322, 442, 521, 711, 736 f.,  
785, 801, 805, 889, 916,  
919 f., 924–926, 975, 1022 f.,  
1026, 1028, 1058 f., 1065,  
1135, 1202, 1800
- Miyaoka, Osahito 1147, 1162 f.
- Moder, Carol Lynn 633, 1778
- Moens, Marc 1168 f., 1179
- Mogford-Bevan, Kay 1823
- Mohanam, Karuvannur P. 203,  
479, 507, 1790
- Mohanam, Tara 713
- Mok, Q. I. M. 1285
- Molina, Alonso de 1433
- Molinier, Guilhem 85, 88
- Monnot, M. 960
- Monod-Becquelin, Aurore 1855
- Monsell, Stephen 1770, 1781
- Montague, Richard 201, 851,  
909, 1182
- Montler, Timothy R. 547–549,  
582, 584
- Moortgat, Michael 165 f., 201,  
842, 860, 863
- Moravcsik, Edith A. 205, 280,  
287, 546–549, 558, 561, 710,  
759, 935, 941, 1085, 1645
- Morel, Mary-Annick 983
- Moreno, Martino Mario 577
- Morev, Lev Nikolaevič 1216
- Morgan, William 316 f., 1098,  
1136
- Morgenstern, Christian 456
- Morin, Yves-Charles 363, 582 f.,  
877, 1603
- Moritz, Karl Philipp 108
- Morpurgo-Davies, Anna 91–93,  
95, 98
- Morris, Charles W. 211, 213,  
680
- Morris, Henry Francis 1187
- Morrow, Daniel G. 258, 261
- Morton, John 1790
- Morvan, Danièle 1291
- Mosel, Ulrike 242 f., 582, 993,  
1070, 1107, 1764, 1859,  
1869 f., 1872
- Moser, Hugo 1271
- Moser, Virgil 1680, 1683
- Moshi, Lioba 1135
- Moskal'skaja, Olga I. 678 f.,  
1273
- Moskalev, Aleksej Alekseevič  
1216
- Motapanyane, Virginia 389
- Motley, Michael T. 576
- Motsch, Wolfgang 160 f., 832 f.,  
886, 906, 1278, 1773
- Mous, Maarten 1646
- Mueller-Vollmer, Kurt 94
- Mufwene, Salikoko S. 811, 1070
- Mugdan, Joachim 20, 236 f.,  
251, 405, 408, 417, 439 f.,  
445, 447, 451–453, 460, 474,  
477, 485, 489, 508, 526, 536,  
542, 596, 600, 631, 650, 867,  
891–894, 933, 1268–1270,  
1275, 1601 f., 1616, 1620,  
1638, 1773, 1882, 1888–1891
- Mühlhäuser, Peter 1006, 1756,  
1758 f., 1761 f., 1764 f.
- Mulder, Jan W. F. 213 f.
- Mulisch, Herbert 1309
- Müller, Beat Louis 665
- Müller, Friedrich 1474, 1479
- Müller, Max 91, 98, 101, 162
- Müller, Peter O. 1684
- Müller, Walter M. 1741
- Müller, Wolfgang 1890

- Müller-Schotte, H. 901  
 Mundt, Christoph 1817  
 Munro, Pamela 814, 1109, 1142,  
 1203f.  
 Muratori, Carlo 772  
 Muravyova, Irina A. 477, 538  
 Murethach 79 f.  
 Murjasov, Rachim Z. 1546  
 Murray, Ben 1491  
 Murru, Furio 59  
 Musaev, Kenesbaj M. 439  
 Mussies, Gerard 164  
 Muysken, Pieter C. 165, 412,  
 541, 1140, 1454, 1459, 1461,  
 1647, 1656 f.  
 Myhill, John 921

## N

- Naes, O. 679  
 Nāgeśabhaṭṭa 41  
 Napps Shirely E. 1783  
 Naro, Anthony J. 1654  
 Nash, David 391, 396, 400, 558  
 Nattiez, Jean-Jacques 211f.  
 Naumann, Bernd 92, 115, 698,  
 886, 892, 894, 898, 932,  
 1275f.  
 Nauta, A. H. 228  
 Naville, Adrien 215  
 Naylor, Paz B. 1475  
 Nedjalkov, Vladimir P. 315, 916,  
 918, 1132f., 1142, 1177  
 Nedyalkov, Igor 1151, 1160  
 Neef, Martin 468f., 472, 479f.,  
 1269, 1275, 1922  
 Neelman, Ad 165f.  
 Nekrásov, Nikolaj P. 125  
 Nemčenko, Vasilij N. 894f.  
 Nespor, Marina 341, 385, 413  
 Nespolous, Jean-Luc 1818  
 Neumann, Werner 104  
 Neuß, Elmar 908  
 Nevis, Joel Ashmore 389, 397,  
 400f., 546  
 Newfield, Madeleine 279  
 Newman, Paul 553, 786, 788  
 Newman, Stanley 1119  
 Newmeyer, Frederick 321  
 Newport, Elissa 1557, 1823  
 Ngai, G. 1898  
 Nguyễn, Huy Cần 1551  
 Nguyễn, Kim Thành 1547  
 Nguyễn, Phú Phong 984, 988,  
 992–994, 1550  
 Nguyễn, Đinh-Hoà 720, 1216  
 Nichols, Johanna 458, 553, 975,  
 1045, 1081, 1093, 1097, 1119,  
 1122, 1197, 1203, 1227,  
 1229f., 1248, 1400, 1491,  
 1515  
 Nicodemus, Lawrence 720

- Nida, Eugene A. 6, 20, 179–  
 182, 184f., 193, 217, 219,  
 236f., 336, 350, 405, 407,  
 418, 438, 440–442, 452f.,  
 463–465, 467, 474, 526, 538,  
 541–543, 573, 582, 588, 880,  
 1249, 1789, 1831  
 Niemi, Jussi 1341  
 Niemikorpi, Antero 1329  
 Nikolaeva, Irina 1215  
 Nikoláeva, Tat'jana M. 133  
 Ningomba, M. S. 759, 1209  
 Nissen, Hans J. 1916  
 Nissim, Gabriel M. 590  
 Nitti-Dolci, Luigia 42  
 Noaily, Michèle 1296  
 Nöldke, Theodor 1741  
 Noll, Craig 628  
 Nolte, Gabriele 613  
 Noonan, Michael 20, 32, 220,  
 799, 1200, 1203, 1859, 1863  
 Nootboom, Sieb G. 540, 1791  
 Nordlinger, Rachel 718  
 Noreen, Adolf 452  
 Noreen, Alfred 110, 147f.  
 Norman, Jerry 724f., 937, 1734  
 Noske, Manuela 1512f.  
 Nöth, Winfried 210f.  
 Noyer, Robert Rolf 603  
 Nübling, Damaris 1271f., 1903  
 Nuessel, Frank 1570  
 Nunberg, Geoffrey 983  
 Nybakken, Oscar E. 1926  
 Nyrop, Krystoffer 139, 1285,  
 1288
- O
- ò Cróinin, Donncha 569, 571  
 Ó Cuív, Brian 88  
 Ó Siadhail, Mícheál 503  
 O'Dochartaigh, Cathair 1215  
 O'Leary, Lacy de 1096  
 Ober, Dana 1041  
 Oberlies, Thomas 1712  
 Obler, Loraine 1818f.  
 Odden, David 559, 565, 587  
 Odden, Mary 559, 565  
 Oehrle, Richard 818  
 Oesterreicher, Wulf 1916  
 Oetting, Janna 1821  
 Oettinger, Norbert 122  
 Oflazer, Kemal 1899  
 Ogulnick, Karen A. 446  
 Ohnheiser, Ingeborg 888f.  
 Oinas, Felix J. 401  
 Ojeda, Almerindo E. 549  
 Oka, I. Gusti Ngurah 999  
 Okrand, Marc 577–579  
 Oksaar, Els 1645  
 Olbertz, Hella G. 1195  
 Ölinger, Albert 697

## Verzeichnisse und Sprachenkarten

- Oliverius, Zdeněk F. 132  
 Olmos, Andrei de 1433  
 Olmsted, David 439  
 Olschansky, Heike 1635, 1905  
 Olsen, Gloria 1025  
 Olsen, Susan 161, 444–446, 892,  
 894, 898, 902–905, 933,  
 1278, 1767  
 Olson, David R. 1916  
 Olson, Michael 1110  
 Omar, Margaret 1797  
 Onions, Charles Talbut 1616  
 Onischii, Masayuki 1032, 1043  
 Orlandini, Anna 1213  
 Ornan, Ozi 1345  
 Ortmann, Albert 466  
 Ortmann, Wolf Dieter 1272  
 Ortner, Hanspeter 355, 898,  
 1278  
 Ortner, Lorelies 355, 898, 1278  
 Osada, Toshiki 720, 726, 1209  
 Osbernd von Gloucester 83  
 Osborne, Charles R. 373, 801  
 Oshima-Takane, Yuriko 999  
 Ossner, Jakob 699f.  
 Osthoff, Hermann 10, 100, 510,  
 514, 519  
 Osumi, Midori 764  
 Oswalt, Robert 1202f.  
 Otanes, Fe T. 390, 399, 543, 569,  
 721f., 743, 749, 760, 786,  
 1203, 1474–1478, 1481–  
 1483, 1486  
 Ouhalla, Jamal 330, 1818  
 Owens, Jonathan 67, 69f., 72,  
 702f., 1742, 1752
- P
- Paccia-Cooper, Jeanne 540  
 Padden, Carol 1557  
 Padley, G. Arthur 697f.  
 Pagliuca, William 443, 537f.,  
 550, 814  
 Pajunen, Anneli 1333  
 Pallas, Peter Simon 91  
 Palmer, Abraham Smythe 1621  
 Palmer, Frank R. 151, 667, 809,  
 811, 817, 1197  
 Palomäki, Ulla 1333  
 Pāṇini 41–51, 93, 179, 186, 437,  
 508, 703, 762, 854, 877–879,  
 1712f., 1715f., 1718f.  
 Panagl, Oswald 122, 289, 1213  
 Panfilov, Valerij 716, 1552  
 Panhuis, Dirk 1324  
 Panov, M. V. 132  
 Panzeri, Maria 1819  
 Papias 83  
 Papp, Ferenc 135  
 Paradis, Michael 1820  
 Paris, Marie-Claude 992

- Park, Jeong-Woon 1120  
 Parker, Elizabeth 988  
 Parker, Gary J. 1453, 1455, 1458  
 Parks, Douglas R. 804, 1196  
 Pasch, Helma 1040, 1094  
 Passerieu, Jean-Claude 703  
 Patañjali 42 f.  
 Paul, Hermann 10, 19, 100, 110,  
   112, 157–160, 162, 670, 675,  
   878, 898, 901, 945 f., 1579,  
   1594, 1597, 1601,  
   1612 f., 1623, 1626, 1683 f.,  
   1686–1689, 1691, 1766  
 Paul, Jean 109  
 Paulissen, Det 165  
 Paunonen, Heikki 1341  
 Pawley, Andrew 713, 1142  
 Payne, David L. 987  
 Payne, Doris 978 f., 919 f., 1018,  
   1022–1025, 1141, 1421,  
   1428, 1465, 1467–1469, 1472  
 Payne, John R. 1208  
 Payne, Thomas E. 978 f., 1154,  
   1464 f., 1467, 1861  
 Peano, Giuseppe 771  
 Pearsall, Judy 1883  
 Pecci, Antonio 1430  
 Pedersen, Annegrethe 1559  
 Pedersen, Holger 145, 177  
 Pée, Willem 163  
 Pei, Mario 536  
 Peirce, Charles Sanders 212 f.,  
   220, 282, 288, 290, 300 f.  
 Pellegrini, Ines Angela 1047  
 Pelletier, Francis J. 1067  
 Peltzer-Karpf, Annemarie 1823  
 Pelzter, Louise 990  
 Peña Cahuachi, Hilario 1471  
 Pennanen, Esko V. 444–446,  
   872  
 Pensado, Carmen 291  
 Penzl, Herbert 381, 1681  
 Percival, W. K. 97  
 Perel'muter, Ilja Aronovič 1318  
 Pérennec, Marcel 699  
 Perkins, Revere D. 443, 550  
 Perlmutter, David M. 321, 361,  
   366, 399 f., 1131, 1140, 1148,  
   1448  
 Perner, Josef 983  
 Perpillou, Jean-Louis 1325  
 Perrot, Jean 142  
 Pesetsky, David 203, 320, 414,  
   863, 1261, 1790  
 Peškóvskij, Aleksandr M. 127,  
   612  
 Peters, Ann M. 1796  
 Peters, Elizabeth 836  
 Peters, Toine 1237, 1239  
 Petersen, Walter 905  
 Petersón, M. N. 127  
 Petraeus 701  
 Peuser, Günther 1818  
 Peytard, Jean 142  
 Pfeffer, J. Alan 1271  
 Pfeifer, Wolfgang 16  
 Philippe de Thaun 1620  
 Philippe, Benoît 1564  
 Phillips, Betty S. 1809  
 Phocas 63, 77  
 Picabia, Lélia 994  
 Picard, H. B. 18  
 Pichon, Edouard 139, 142, 1288  
 Pickett, Velma B. 440, 442, 489,  
   541, 1831  
 Piel, J. M. 1698  
 Pierce, Joe E. 1237 f.  
 Pierer, Heinrich August 15  
 Pike, Evelyn G. 192–194, 1831  
 Pike, Kenneth Lee 179, 191–  
   194, 245, 337, 341, 418, 431,  
   452, 484, 587 f., 1270, 1831  
 Pillon, Agnesa 1785  
 Pinault, Georges-Jean 703, 1214  
 Pinchon, Jacqueline 141, 1285  
 Pinker, Steven 880, 882, 1778,  
   1807  
 Pinkster, Harm 813 f.  
 Piron, Claude 1570  
 Pischel, Richard 1712, 1719  
 Pitman, Isaac 149  
 Pittman, Richard 1486  
 Plag, Ingo 1255, 1259, 1780  
 Plank, Frans 94, 255, 289, 298,  
   361, 420, 465, 478 f., 550,  
   585, 618, 620, 624, 626, 639,  
   645, 647 f., 653, 684 f., 687,  
   710, 746, 754, 833, 866, 878,  
   887 f., 890, 893, 895, 930,  
   999, 1057 f., 1061, 1079,  
   1081, 1085, 1491, 1615, 1620,  
   1623, 1868  
 Plante, Elena 1820  
 Platon 53 f., 56 f., 264, 417,  
   694 f., 953  
 Platt, Carol B. 1798  
 Platz, Bärbel 698  
 Pléh, Csaba 1797  
 Pleines, Jochen 667  
 Plénat, Marc 143 f., 583, 1285 f.,  
   1289  
 Plotkin, Vul'f Ja. 1376  
 Plungjan, Vladimir A. 442, 445  
 Pogonowski, Jerzy 526  
 Poitou, Jacques 684 f.  
 Polenz, Peter von 1275, 1278  
 Polinskaja, Maria S. 916, 918  
 Pollak, Hans W. 452  
 Pollok, Jean-Yves 330, 332, 818  
 Pomorska, Krystyna 272, 275,  
   277, 280  
 Poppe, Nicholas N. 442  
 Porízka, Vincenc 1094  
 Porotova, Tel'mina 1382  
 Port, Robert F. 1135  
 Porter, M. F. 1896, 1899  
 Porzig, Walter 116, 1325, 1623  
 Pospelov, Nikolaj S. 127, 677  
 Postal, Paul M. 199, 428, 1131,  
   1140, 1148, 1265  
 Posti, Lauri 1337, 1340  
 Potebnjá, Aleksandr A. 125,  
   610, 612  
 Pott, A. F. 9  
 Pottier, Bernard 141  
 Poulinne, Nanda 1812  
 Powlison, Esther 1465, 1467  
 Powlison, Paul 1464 f., 1467  
 Pratchett, Terry 249  
 Prcic, Tvrto 1889  
 Prell, Heinz-Peter 1684, 1694  
 Premper, Waldfried 737  
 Press, Ian 760, 823, 979, 1063,  
   1065  
 Press, Margaret L. 767  
 Preuss, Fritz 446  
 Pride, Leslie 588  
 Priestly, Tom M. S. 1043  
 Prince, Alan S. 206, 341 f., 385,  
   431, 467, 471, 555 f., 562–  
   566, 569 f., 1045, 1613, 1807  
 Prior, Arthur 1180, 1182  
 Priscian 53, 58–60, 62–65, 77–  
   79, 82–86, 674, 695 f., 732,  
   1929  
 Probus 63  
 Prochazka, Theodore 1742, 1744  
 Proctor, Paul 1885, 1889 f.  
 Progovac, Ljiljana 789  
 Prokosch, Erich 1752  
 Protagoras 1031  
 Przyluski, Jean 1545  
 Puhvel, Jaan 1214, 1216  
 Pulleyblank, Douglas 541, 558,  
   561, 563  
 Pulleyblank, Edwin 1730–1735,  
   1737  
 Pullum, Geoffrey K. 297, 389 f.,  
   617, 818, 965, 967, 969, 972,  
   1078, 1255, 1265, 1666  
 Pustet, Regina 768, 1100  
 Putschke, Wolfgang 1680  
 Putten, Frans van der 166, 864  
 Putzer, Oskar 933  
 Pye, Clifton 1795  
 Pylyshyn, Zeon Walter 1766  
 Pym, Noreen 711
- Q
- Qafishesh, Hamdi A. 980  
 Quine, Willard V. O. 1067 f.  
 Quintilian 64, 1076, 1929  
 Quirk, Randolph 316, 648, 760,  
   765, 790, 947, 1105, 1215,  
   1255, 1625 f.

## R

- Rabain-Jamin, Jacqueline 1001  
 Rabin, Chaim 1741, 1743  
 Radtke, Petra 1273 f.  
 Raecke, Jochen 1309  
 Rainer, Franz 308, 347 f., 367,  
     584 f., 873, 878, 912, 1049,  
     1291, 1704 f., 1710  
 Raja, K. Kunjunni 42  
 Rajaona, Simeon 702  
 Ramat, Paolo 300, 817, 1222  
 Ramers, Akrl Heinz 1276  
 Ramos, Teresita V. 1479, 1481  
 Ramus, Petrus 697 f.  
 Randall, Janet 864  
 Ranjeva, Henri 702  
 Rappaport, Malka 845, 847  
 Rask, Rasmus Kristian 7, 95,  
     145, 1127  
 Raster, Peter 1244  
 Ratcliffe, Robert R. 1055, 1057,  
     1062  
 Rathay, Wilfried 446  
 Rauh, Gisa 983  
 Ravid, Dorit 1801  
 Rayner, Keith 1770  
 Raynouard, François-Juste-  
     Marie 139  
 Reaney, P. H. 1677  
 Redder, Angelika 1274  
 Reed, Carroll E. 1271  
 Reed, Irene 1082  
 Reed, Judy 987  
 Reesink, Ger P. 1003, 1011 f.,  
     1056, 1062 f., 1202–1205  
 Reeves, Jennifer Elaine 435 f.  
 Reformátskij, Aleksandr A. 128,  
     130, 508, 513, 526  
 Reh, Mechthild 802, 805, 1091,  
     1095 f., 1107, 1142, 1590,  
     1592, 1595  
 Rehg, Kenneth 1022  
 Reich, Carl Gottlob 1555  
 Reichenbach 1180, 1182 f.,  
     1186 f.  
 Reichling, Anton 163 f., 852  
 Reichmann, Oskar 1681  
 Reid, Lawrence A. 1474  
 Reid, Nicholas 1042  
 Reifferscheid, Augustus 58  
 Reiner, Erica 35, 40  
 Reinhart, Tanya 789  
 Reis, Marga 1273  
 Remigius von Auxerre 79, 82 f.,  
     85  
 Remmius Palaemon 62, 696  
 Rennison, John R. 759 f., 766,  
     768  
 Renou, Louis 437, 703  
 Restrepo, Maria Adelaide 1821  
 Retsö, Jan 1746  
 Rettig, Wolfgang 1268 f.

## Verzeichnisse und Sprachenkarten

- Reuland, Eric J. 789, 818  
 Reuther, J. G. 1491  
 Révizin, Isaak I. 132  
 Rey, Alain 1883, 1927  
 Reyle, Uwe 911, 1170, 1179  
 Rhodes, Richard A. 1049  
 Rhys Jones, Thomas J. 572  
 Rice, Keren D. 407, 412, 569 f.,  
     572  
 Rice, Mabel L. 1820 f.  
 Richards, Maña 1020  
 Richter, Heide 983  
 Richter, Jean Paul 157  
 Riegel, Martin 1286  
 Ries, John 106, 108, 110, 112 f.,  
     665, 672  
 Riis, Hans N. 1518  
 Rijkhoff, Jan 1068  
 Rijksbaron, Albert 1318 f.  
 Rincón, Antonio del 1433  
 Risch, Ernst 121 f., 1320, 1325  
 Rischel, Jørgen 1389  
 Rissanen, Mati 1590  
 Ritchie, Graeme D. 155, 1896  
 Ritner, Robert K. 1917  
 Ritter, Elizabeth 1060  
 Rivet, Paul 1464  
 Rivle, Ancrene 1642  
 Rix, Helmut 122, 1311, 1319  
 Rizzi, Luigi 328  
 Roberts, Ian 323, 329  
 Roberts, John R. 1202–1204  
 Roberts, Lawrence D. 985  
 Robin, Christian 1741  
 Robins, Robert Henry 19, 56,  
     59, 91 f., 100, 151 f., 264,  
     344, 552, 627, 674, 694, 708,  
     770, 1074, 1226, 1235  
 Robinson, Orrin W. 1603  
 Roche, Emmanuel 1896  
 Roché, Michel 1707  
 Rodrigues, Ayrton D. 742  
 Roepert, Thomas 356 f., 862,  
     1259, 1261, 1263  
 Rogers, David E. 175, 437  
 Rohrer, Christian 906  
 Rolecke, Thorsten 1680  
 Romaine, Suzanne 1638–1640,  
     1756, 1761  
 Römer, Paul G. 1659  
 Romero-Figueroa, Andrés 1216  
 Rondal, Jean A. 1822  
 Ronneberger-Sibold, Elke 293,  
     455, 520, 1268, 1270 f., 1275,  
     1684  
 Rooij, Vincent de 245  
 Rosch, Eleanor 264  
 Rose, James H. 834 f.  
 Rosen, Carol 925  
 Rosén, Häim B. 655 f., 1343,  
     1345  
 Rosen, Sarah Thomas 204, 322,  
     919, 924  
 Rosenhouse, Judith 1742  
 Rosetti, A. 249, 252 f.  
 Rosier, Irène 84, 696 f.  
 Ross, John Robert 264, 699, 976  
 Ross, Malcom D. 1135  
 Ross, Sir David 264  
 Rosten, Leo C. 971  
 Roth, Juliana 1650  
 Rothweiler, Monika 1767,  
     1820 f.  
 Rousseau, André 983  
 Rowley, Anthony 1271  
 Royen, Gerlach 163  
 Rozental, Ditmar E. 536  
 Rozwadowska, Bozena 863  
 Rubach, Jerzy 203, 339, 341 f.,  
     384, 395, 414, 541  
 Rubin, Gary 1774  
 Rubino, Carl R. G. 1476  
 Rubio, Lisardo 1076  
 Rude, Noel 1082, 1087  
 Rudes, Blair A. 514, 606  
 Ruegg, David Seyfort 437, 703  
 Rugemalira, Josephat Muhozi  
     1124  
 Ruhlen, Merritt 577  
 Ruipérez, Martín Sánchez 1319  
 Ruiz de Montoya, Antonio 1421  
 Rukeyser, Alison 1056 f., 1060  
 Rumelhart, David E. 1766, 1807  
 Rumsey, Alan 1001  
 Runciman, Sir Steven 63  
 Rundgren, Frithiof 74  
 Rupp, Heinz 1273  
 Russ, Charles 1683, 1687  
 Russell, I. Willis 958  
 Russell, Kevin 471  
 Rutherford, William 276

## S

- Saadiah Gaon 74, 703  
 Sabeau-Jouannet, Emilie 1001  
 Šáchmatov, A. A. 127  
 Sadler, Louisa 327 f.  
 Sadock, Jerrold M. 204, 327,  
     331 f., 394, 922–924, 1191,  
     1395  
 Saeed, John Ibrahim 1135  
 Sager, Juan C. 1924 f., 1927 f.  
 Sagey, Elizabeth 574  
 Saint-Hilaire, Auguste  
     Prouvensal de 16 f.  
 Sakaguchi, Alicja 1563  
 Šákalya 43  
 Sala, Marius 1648, 1651  
 Sallust 76  
 Salminen, Tapani 1328  
 Salmon, P. B. 91  
 Saloni, Zygmunt 1889, 1891  
 Saltarelli, Mario 569, 759, 761,  
     813, 817, 980, 1078, 1085

- Saltveit, Laurits 1274  
 Salzmann, Zdenek 773  
 Samarin, William J. 1759, 1829  
 Sanctius 697  
 Sandberg, Bengt 690, 895, 1275  
 Sanders, Gerald 363, 442 f., 446,  
     628, 750, 871, 946 f., 950  
 Sanders, Lawrence 836  
 Sandfeld, Kristian 1649  
 Sandmann, Manfred 677, 680,  
     683, 690  
 Sandoval, María 877  
 Sandra, Dominiek 421, 1768 f.,  
     1770, 1772, 1774 f., 1779  
 Sands, A. Kristina 1034, 1037,  
     1042 f.  
 Sankoff, David 243, 245  
 Sankoff, Gillian 1755 f., 1761  
 Šanskij, Nikolaj M. 132, 1309  
 Santandrea, Stefano 1094  
 Santen, Ariane J. van 164–166  
 Sanzeev, G. D. 1218  
 Sapir, Edward 163, 170, 172–  
     178, 186, 202, 218, 248, 253,  
     373, 376, 383, 432, 507 f.,  
     513, 524, 526, 538 f., 561,  
     588, 716, 757, 764, 916,  
     923–925, 932, 935, 940,  
     1213, 1221, 1223–1225,  
     1228, 1232 f., 1235 f., 1248,  
     1311, 1400, 1647, 1655  
 Saporta, Sol 440  
 Sardžveladze, Z. A. 87  
 Sarmiento, Ramón 697  
 Sasse, Hans-Jürgen 728, 890,  
     919, 975, 1068, 1646  
 Sassen, Albert 165, 231, 233  
 Sauer, Hans 1627, 1630, 1632–  
     1634, 1676  
 Šaumjan, Sebast'jan K. 1309  
 Saunders, Ross 923  
 Saussure, Ferdinand de 91, 100,  
     126 f., 139, 141 f., 154, 171,  
     210 f., 214–217, 219, 221,  
     225 f., 239, 244, 304, 344,  
     404, 437 f., 441, 609, 698 f.,  
     851, 1349, 1565, 1576, 1612,  
     1615, 1622, 1633, 1915  
 Saussure, René de 1565 f.  
 Sauvageot, Serge 987, 994  
 Savişir, İskender 1364  
 Scalise, Sergio 198 f., 201, 220,  
     320 f., 351 f., 354 f., 361,  
     407 f., 540, 542, 583, 859,  
     869, 875, 879, 882, 1047,  
     1255, 1288  
 Scarborough, D. L. 1770  
 Ščeglov, Jurij K. 1216  
 Ščerba, Lev V. 126, 129  
 Schaaik, Gerjan van 1195  
 Schabes, Yves 1896  
 Schachter, Paul 390, 399 f., 543,  
     558, 569, 589, 710, 714,  
     718 f., 721 f., 743, 746, 749,  
     760, 784, 786, 790 f., 815,  
     817, 1113, 1124, 1203, 1474–  
     1478, 1481–1483, 1486  
 Schaefer, Burkhard 676 f., 683,  
     699 f.  
 Schaff, Adam 212  
 Schane, Sanford A. 1285  
 Schaner-Wolles, Chris 1822  
 Schaub, Willi 978, 1006, 1137,  
     1192, 1199  
 Scheben-Schmidt, Marietheres  
     1684, 1694  
 Scheer, Tobias 1272  
 Scheerer, Eckart 1766  
 Scheerer, Otto 1479  
 Schegloff, Emmanuel A. 984  
 Scheidweiler, Gaston 1050  
 Scheler, Manfred 1619, 1674  
 Schellinger, Wolfgang 1058  
 Schenk, Werner 1309  
 Schenkel, Wolfgang 1917  
 Scherer, Wilhelm 1680  
 Schick, Brenda 1559  
 Schifko, Peter 435 f.  
 Schiller, Friedrich von 16, 1621  
 Schindler, Jochem 1611  
 Schipper, Joleen 165  
 Schirmunski, Victor M. 1688,  
     1691  
 Schlachter, Wolfgang 122  
 Schlegel, August Wilhelm von 7,  
     94, 1221, 1223, 1247, 1698  
 Schlegel, Carl Wilhelm Friedrich  
     von 7, 18, 94, 1221–1223,  
     1247  
 Schleicher, August 8–10, 18 f.,  
     91 f., 97–99, 217, 1221,  
     1223, 1227 f., 1247  
 Schleicher, Charles 925  
 Schleiermacher, Friedrich 108  
 Schleyer, Johann Martin 1563  
 Schlieben-Lange, Brigitte 676  
 Schmid, Günter 16–18  
 Schmid, Stephan 290  
 Schmid, Wolfgang P. 680, 699–  
     701  
 Schmidt, Annette 1029, 1043,  
     1655 f.  
 Schmidt, Gernot 122  
 Schmidt, Günther Dietrich 351,  
     355 f., 892, 931  
 Schmidt, Johann 1563, 1567  
 Schmidt, Karl Horst 1687  
 Schmidt, Reiner 1934  
 Schmidt, Wilhelm 675, 677, 693,  
     699  
 Schmitt, Christian 1704, 1708  
 Schmitter, Peter 91, 95  
 Schmitthenner, F. 157, 698  
 Schneider, Iris 1050  
 Schneider, Klaus P. 1046, 1049 f.  
 Schooneveld, Cornelis H. van  
     125, 127 f., 277, 279  
 Schøsler, Lene 1699  
 Schott, Wilhelm 96  
 Schottelius, Justus Georg 19,  
     157, 1682, 1931  
 Schreiner, Markus 1616  
 Schretter, Manfred K. 35  
 Schreuder, Robert 347, 1922,  
     1770, 1772–1774, 1779,  
     1799,  
 Schreyer, Rüdiger 92  
 Schubert, Klaus 1563, 1565,  
     1570  
 Schubert, Lanhart K. 1067  
 Schuchardt, Hugo 1653, 1659  
 Schuh, Russel G. 1094, 1216  
 Schultink, Hendrik 163–166,  
     195, 201, 305, 342  
 Schultz-Lorentzen, Christian  
     Wilhelm 1389  
 Schuster-Šewc, Heinz 1648  
 Schütz, Albert 1136, 1139  
 Schütze, Carson T. 400  
 Schütze, Hinrich 1893, 1898  
 Schwartz, Catherine 837  
 Schwartz, Linda 275, 785, 788 f.  
 Schwartz, Myrna F. 1817 f.  
 Schwarz, Catherine 1883  
 Schwarz, Hans 1684  
 Schwarze, Christoph 514, 889 f.,  
     892, 895  
 Schwarzwald, Ora 1345, 1349  
 Schwegler, Armin 1698  
 Schweickard, Wolfgang 1708  
 Schwyzer, Eduard 1312 f., 1316,  
     1320  
 Sciumbo, Anna-Maria Di 31, 201,  
     249, 315 f., 319, 322, 325 f.,  
     350, 354, 358, 418, 596,  
     860–862, 900, 904, 911,  
     1255, 1287, 1295  
 Scott, Graham 1078, 1081,  
     1203 f.  
 Scragg, Donald G. 1921  
 Sebeok, Thomas A. 91, 135,  
     210, 213  
 Sedulius Scottus 79 f., 83, 86  
 Seibold, Elmar 1049, 1912–  
     1914  
 Segui, Juan 1780  
 Seibicke, Wilfried 1903, 1924  
 Seidenberg, Mark 1780  
 Seiler, Hansjakob 237, 242, 247,  
     251, 253, 735, 740, 1019,  
     1091, 1097, 1101, 1224  
 Seiler, Walter 1022 f., 1026, 1028  
 Selkirk, Elisabeth O. 31, 33, 189,  
     200, 202 f., 319, 338, 352,  
     354, 357, 408, 540 f., 550,  
     838, 860–862, 880, 899 f.,  
     904 f., 907 f., 911, 1255, 1264,  
     1288 f., 1296  
 Senft, Gunter 1012  
 Serbat, Guy 638, 641, 645, 1076

- Serebrennikov, Boris A. 129  
 Sérégal, Philippe 1272  
 Serzisko, Fritz 1018 f., 1024,  
     1040, 1098 f., 1904  
 Seuren, Pieter 815  
 Sextus Empiricus 56  
 Sgall, Petr 1224  
 Shafer, Robert 1733  
 Shakespeare, William 1002,  
     1674, 1678  
 Sharma, Rama Nath 45  
 Shaumyan, Sebastian 1152  
 Shaw, George Bernard 1921  
 Shaw, J. Howard 901  
 Shaw, Patricia A. 411  
 Sheckley, Robert 836  
 Shepherd, Susan 814  
 Sheppard, Helen E. 953  
 Sheridan, Richard Brinsley 1621  
 Sheridan, Thomas 1638  
 Sherwood, David 407  
 Sherzer, Joel 1077, 1080  
 Shibatani, Masayoshi 1013 f.,  
     1123, 1131 f., 1134, 1137,  
     1140, 1148, 1153–1156,  
     1160–1162, 1225, 1479  
 Shkarban, Lisa I. 1475  
 Shopen, Timothy 1861  
 Showalter, Catherine 987, 994  
 Sibawayhi, Ibn ۹۸۷  
     67–73, 702  
 Sibomana, Leonidas 1538  
 Siegel, Dorothy L. 189, 198 f.,  
     202 f., 320, 341, 350 f., 354,  
     411, 433, 541, 879, 948,  
     1255 f., 1790  
 Siegel, Muffy 356 f.  
 Siemund, Peter 1070  
 Sievers, Eduard 100  
 Siewierska, Anna 809, 813, 1140,  
     1160 f.  
 Sifianou, Maria 1050  
 Sil'nickij, Georgij G. 1132, 1235,  
     1237–1243  
 Silverstein, Michael 788, 1011,  
     1082 f., 1153, 1493  
 Šimečková, Alena 892  
 Simmel, Georg 3  
 Simmler, Franz 1685, 1687  
 Simon, H. F. 677  
 Simon, Walter 725  
 Simons, Gary F. 1836  
 Simons, Linda 1097  
 Simpson, Jane 330, 470  
 Simpson, John A. 1882  
 Sims-Williams, Nicholas 602  
 Sinclair, Hermine 1796, 1807  
 Sinclair, John 1884, 1888  
 Singh, Rajendra 871  
 Sitnikova, Antonina 1550  
 Sitta, Horst 1936  
 Sittig, Ernst 1074  
 Skalička, Vladimír 133 f., 293,  
     630, 913, 1224, 1235  
 Skorik, Pjotr Ja. 314 f.  
 Skousen, Royal 571, 854, 866,  
     1338, 1602 f., 1766  
 Slavíčková, Eleonora 1235  
 Slobin, Dan I. 290, 300, 348,  
     796, 1769, 1778, 1795–1797,  
     1799 f.  
 Slotkin, Alan R. 1641  
 Sluiter, Ivoke 60  
 Smaragdus 80, 83 f.  
 Smedts, Willy A. J. 165  
 Smeets, Ineke 1194 f., 1197, 1205  
 Smirnickij, Aleksandr I. 23, 129,  
     656, 660  
 Smirniotopoulos, Jane C. 330  
 Smirnóv, S. 135  
 Smith, Adam 92, 1221 f.  
 Smith, Carlota S. 1166, 1168 f.,  
     1172 f., 1179  
 Smith, Geoff P. 1756  
 Smith, John Ch. 984  
 Smith, Neil 1822  
 Smith, Norval 555, 1121, 1656–  
     1658, 1863, 1871 f.  
 Smith, Philipp T. 1771  
 Smith, Wayne H. 1557  
 Smith-Lock, Karen M. 1820  
 Smith-Stark, T. Cedric 1059 f.  
 Smith-Stark, Thomas C. 711,  
     736  
 Smits, Caroline 1587  
 Smoczyńska, Małgorzata 1799 f.  
 Smolensky, Paul 471, 566, 1613,  
     1766  
 Smyth, Herbert 1150  
 Sneddon, James Neil 241  
 Snodgrass, John G. 1770  
 Snyman, J. W. 797  
 Soberano, Rosa 1474  
 Soboleva, Polina A. 648, 650,  
     1309  
 Sodmann, Timothy 1681  
 Soeteman, Carl 1272  
 Sohn, Ho-min 577  
 Sokrates 53  
 Solms, Hans-Joachim 1684,  
     1686, 1688–1691, 1693 f.  
 Sommer, Bruce A. 751  
 Sommer, Ferdinand 120  
 Sommerfeldt, Karl-Ernst 693  
 Sommerfelt, Alf 147  
 Son, Jos van 1237, 1239  
 Sonderegger, Stefan 1681 f.,  
     1684 f., 1687 f., 1691–1693  
 Song, Jae Jung 1137, 1139, 1142  
 Sophronius 58  
 Sørensen, Finn 882  
 Sorenson, Arthur P. 374, 803  
 Sorin, Carmen Dobrovie 400  
 Sosenskaja, Tat'jana B. 1215  
 Spang-Hanssen, Henning 214  
 Speas, Margaret 330  
 Spence, N. C. W. 1707  
 Spencer, Andrew 5, 23, 27, 32,  
     155, 188 f., 315 f., 319, 321 f.,  
     325–327, 329–331, 340, 391,  
     405, 409, 413 f., 451, 463,  
     466, 471, 474, 481, 489, 583,  
     627, 639, 647, 841, 854 f.,  
     1256, 1259, 1261, 1263 f.,  
     1780  
 Spencer, Herbert 1258  
 Spitzer, Leo 1290, 1645  
 Spring, Cari Louise 563  
 Sproat, Richard W. 199, 321,  
     326, 332, 413, 899, 904, 1894  
 Spruit, Arie 1109  
 Sridhar, Shikaripur N. 203, 252,  
     414, 541, 880, 993, 1137,  
     1256  
 Staal, J. Frits 165, 703  
 Städtler, Thomas 87, 89, 451,  
     696  
 Stalin, Joseph 127 f.  
 Stalph, Jürgen 1917  
 Stampe, David 288, 401, 770,  
     772, 1594  
 Standop, E. 814  
 Stanley, Richard 385 f., 429  
 Stanners, Robert F. 1770, 1773,  
     1783, 1785  
 Stårcz, John 1688  
 Stark, Jacqueline A. 1820  
 Stark, John A. 1789  
 Starosta, Stanley 1086  
 Stassen, Leon 710, 759, 766,  
     1107, 1212, 1216  
 Steedman, Mark 1168 f., 1179  
 Steele, Susan 396, 400 f., 806,  
     817 f.  
 Steever, Sanford B. 920 f., 1080  
 Stegmann, Carl 655  
 Stehr, A. 92  
 Steinberg, Danny 1640  
 Steinhauer, Anja 1275  
 Steinitz, Renate 1277  
 Steinthal, Heymann 9, 92, 97–  
     99, 105  
 Štejnfel'dt, Évi A. 652  
 Stemberger, Joseph Paul, 206,  
     345 f., 364, 392, 1766, 1788,  
     1817  
 Stepanov, Arthur V. 1238  
 Stepanov, Jurij S. 417  
 Stepanova, Marija D. 665, 667,  
     690, 892 f., 931, 1565  
 Steriade, Donca 430, 436, 490,  
     562–564  
 Stern, Theodore 1735  
 Stevens, Alan M. 400, 560  
 Stewart, John 1517 f.  
 Stiebels, Barbara 380, 902, 905  
 Stocking, George 173  
 Stockwell, Robert P. 1791  
 Stokes, Bronwyn 1078  
 Stokoe, William C. 1555, 1557

- Stoltz, Thomas 693  
 Stong-Jensen, Margaret 346, 628  
 Stonham, John T. 411, 578, 581,  
     584 f.  
 Stopp, Hugo 169  
 Stork, Francis C. 568, 597  
 Strang, Barbara M. H. 835  
 Strauß, Gerhard 694  
 Strauss, Steven L. 203, 880,  
     1256  
 Strauss, Susan 993  
 Strawson, Peter Frederick 974  
 Strong, David R. 815  
 Strunk, Klaus 1318  
 Stubbs, Michael 1921  
 Stump, Gregory T. 153, 321,  
     367, 442, 541, 784, 877, 1047  
 Sturtevant, Edgar H. 121, 866  
 Suárez, Jorge A. 1077, 1107,  
     1216, 1423, 1427–1429  
 Suchotín, A. M. 132  
 Sugamoto, N. 755  
 Summers, Della 1884  
 Suñer, Margarita 327 f.  
 Sunik, Orest P. 129  
 Supalla, Ted 1022, 1557–1560  
 Sütterlin, Ludwig 104, 106–113,  
     115, 670, 672, 675 f., 684, 699  
 Suyuti, Jalal al-Din 73  
 Svartvik, Jan 747  
 Švedova, Natal'ja J. 1303,  
     1305 f., 1308 f.  
 Swadesh, Morris 178, 186, 251,  
     452, 490 f., 508, 721, 781  
 Sweet, Henry 149, 344, 346, 944  
 Sweetser, Eve 1591  
 Swiggers, Pierre 93, 211, 216,  
     218–220, 222, 245, 249,  
     697 f., 1285  
 Swisher, Linda 1821  
 Syncellus 59  
 Szantyr, Anton 1217  
 Szczepaniak, Renata 1904  
 Szemerényi, Oswald 121, 772  
 Szerdahelyi, István 1564  
 Szober, St. 134  
 Szpyra, Jolanta 252  
 Szymanek, Bogdan 266, 267,  
     353, 433, 511, 526, 529, 532,  
     877, 881, 1255
- T
- Tabakowska, Elżbieta 1049  
 Taeldeman, Johan 165  
 Taft, Marcus 1766, 1768–1772,  
     1774, 1778 f., 1784  
 Tagliavini, Carlo 458  
 Talen, Jan G. 162  
 Tallal, Paula 1820  
 Talmy, Leonard 258 f., 261, 374,  
     796, 1125 f., 1402
- Tamba, Irène 988, 992, 1296  
 Tanenhaus, Michael K. 837,  
     1771  
 Tanz, Christine 983  
 Tatwine 78 f., 84  
 Taylor, Ann 398  
 Taylor, Charles 253, 768, 980 f.,  
     1210  
 Taylor, Gerald 1455  
 Taylor, John R. 392, 424, 1049,  
     1091  
 Tedeschi, Philip J. 809  
 Tegey, Habibullah 399, 546  
 Tekinay, Alev 1363  
 Teleman, Ulf 148  
 Telenkova, Margarita A. 536  
 Temple, Martine 143, 1292 f.  
 Terence 76  
 Terrell, Peter 1886  
 Tesák, Jürgen 1817 f.  
 Tesar, Bruce 1613  
 Tesnière, Lucien 107, 109, 667,  
     680, 689, 699 f., 1117–1119,  
     1130  
 Testen, David 556  
 Teuber, Oliver 1271  
 Tharp, David A. 1785  
 Theobald, Elke 1272  
 Theodorides, Chr. 58  
 Theodosius von Gaza 59  
 Theodosius 57–59, 63  
 Thiel, Gisela 836, 906, 1638  
 Thiele, Johannes 888 f., 895,  
     1287, 1291  
 Thieroff, Rolf 1273 f.  
 Thomas von Aquin 54  
 Thomas von Erfurt 7, 86, 697  
 Thomas, Jacqueline M. C. 1832  
 Thomason, Sarah Grey 1645–  
     1647, 1651  
 Thompson, Chad 1400  
 Thompson, E. David 1100  
 Thompson, Laurence 254, 577 f.,  
     580, 936, 1546, 1734  
 Thompson, M. Terry 577 f., 580,  
     936  
 Thompson, Sandra A. 240, 287,  
     374, 712, 727 f., 747, 749,  
     758, 764, 795, 799 f., 809 f.,  
     846, 848, 882, 939 f., 950,  
     1020, 1084, 1112 f., 1116,  
     1191 f., 1202, 1637  
 Thomsen, Vilhelm 145  
 Thomson, David 82, 87  
 Thorndike, Edward L. 1235  
 Thornton, Anna M. 294  
 Thrane, Torben 818  
 Thümmel, Wolf 655  
 Thun, Nils 1679  
 Thurman, Robert 1204  
 Thurneysen, Rudolf 1215, 1219  
 Thwaites, Reuben Gold 1411
- Tichonov, Aleksandr N. V. 677,  
     1307, 1309  
 Tiefenbach, Heinrich 894  
 Tiersma, Pieter Meijes 166, 278,  
     337  
 Timberlake, Alan 814, 1197  
 Ting, Pang-hsin 1732  
 Tobin, Yishai 1343, 1345–1347,  
     1350, 1352 f., 1356  
 Todaeva, B. X. 1218  
 Todd, Loreto 1759  
 Togeby, Knud 379–381, 1704  
 Toivainen, Jorma 1341  
 Tolksaya, Maria 1215  
 Toman, Jindrich 161, 900, 904,  
     908, 1278  
 Tomblin, J. Bruce 1820  
 Tommola, Hannu 1186  
 Tonelli, Livia 289  
 Tonkin, Humphrey 1564  
 Tookey, John Horne 92  
 Toorn, Maarten C. van den 165  
 Topping, Donald M. 560, 568  
 Torero, Alfredo 1453  
 Törnqvist, Nils 905  
 Totanes, Sebastian de 1474 f.  
 Touratier, Christian 1076  
 Tournier, Jean 1903  
 Tovar, Antonio 91  
 Tracy, Rosemarie 1799  
 Trager, George L. 184, 219, 379,  
     453  
 Trask, R. Larry 655, 813, 970,  
     1061, 1667  
 Traugott, Elizabeth Closs 259,  
     263, 401, 814, 1199, 1205,  
     1232, 1590, 1592, 1595 f.  
 Treffry, Diana 1883  
 Trefry, David 1491  
 Trier, Jost 116  
 Trnka, Bohumil 134  
 Troll, Wilhelm 16 f.  
 Trommelen, Mieke 165 f., 201,  
     860  
 Trost, Harald 1896  
 Troupeau, Gérard 703  
 Tru'o'ng Đóng San 1551  
 Truax, Catherine Elaine 584 f.  
 Trubetzkoy, Nikolaj S. 132,  
     235 f., 241, 244, 267, 272,  
     275, 335 f., 383, 427, 431–  
     433, 442, 490, 492–494, 772  
 Tschenkéli, Kita 439, 442  
 Tschirch, Fritz 1694  
 Tsimpli, Ianthi-Maria 1822  
 Tsunoda, Tasaku 1158 f.  
 Tucker, Archibald N. 1077,  
     1079, 1095  
 Tuite, Kenin 1215  
 Twaddell, W. Freeman 178  
 Tyler, Lorraine K. 540, 1775,  
     1778, 1783 f., 1786, 1819

## U

- Uc Faidit 696  
 Uehara, Satoshi 1114  
 Uhlenbeck, C. C. 163  
 Uhlenbeck, Eugenius M. 164,  
 166, 867, 951  
 Uhlig, Gustav 56, 59, 63f.  
 Uhrström, Wilhelm 1678  
 Ułazyn, Henryk 134, 490  
 Uldall, Hans Jørgen 254, 452  
 Ullman, Michael 1821  
 Ullmann, Stephen 251 f., 1615,  
 1621 f., 1633  
 Ultan, Russell 547–550, 576 f.,  
 790, 937, 1093, 1595  
 Uluchanov, Igor S. 888, 1309  
 Ulvestadt, Bjarne 1272  
 Underhill, Robert 711, 981  
 Ungeheuer, Gerold 115  
 Ungerer, Friedrich 1615, 1619  
 Upādhyāya, Baladeva 50  
 Urbaniak, Gertrud 902, 1276  
 Ušákov, D. N. 127

## V

- Vachek, Josef 133  
 Vago, Robert M. 621  
 Vaidya, P. L. 1722  
 Vairel, Hélène 999  
 Valin, Robert van 859, 1192 f.,  
 1377, 1475  
 Vanvolsem, Serge 697  
 Varro, Marcus Terentius 56,  
 59–62, 82, 696 f., 732  
 Vasu, Śrīśa Chandra 437  
 Vater, Heinz 681, 698, 1270,  
 1274  
 Vater, J. S. 91  
 Veith, Werner 1272  
 Velázquez-Castillo, Maura 1233,  
 1421–1428, 1430  
 Vendler, Zeno 1165, 1168–1173,  
 1176  
 Vendryès, Joseph 693, 813, 1218,  
 1611  
 Veneziano, Edy 1796  
 Vennemann, Theo 290, 337, 468,  
 498, 577, 840, 1620, 1905  
 Vergnaud, Jean-Roger 203, 347  
 Verkuyl, Henk J. 1167, 1170 f.  
 Verner, Genrich K. 1377, 1382  
 Verner, Karl 100, 145, 1673  
 Verrac, Monique 693  
 Versaw, Larry 1836  
 Versteegh, Cornelis H. M. 54,  
 57, 60, 702  
 Versteegh, Kees 70, 74, 1741,  
 1752  
 Vesper, Wilhelm 104, 115, 669

## Vet, Co 1174

- Vietor, Wilhelm 1930 f.  
 Vihman, Marily M. 1799  
 Vilkuna, Maria 1183  
 Villa Dei, Alexander von 79  
 Villiers, Jill de 1823  
 Villiers, Reinette De 352  
 Vincent, Nigel 442, 648, 656,  
 661 f., 808, 814  
 Vinogradov, Viktor V. 127 f.,  
 1307–1309  
 Vinokúr, G. O. 127  
 Virgiliius 76  
 Vjalkina, L. V. 1218  
 Voegelin, Charles F. 178, 186,  
 508, 1235  
 Voelitz, Erhard F. K. 1099  
 Vögeding, Joachim 902  
 Vogel, Irene 341, 385, 413  
 Vogel, Petra Maria 932, 936,  
 1275  
 Vogel, Thomas 1808  
 Vogt, Hans 146 f.  
 Voigt, Wolfram 4  
 Volodin, Aleksandr P. 443  
 Voorhoeve, Clemens L. 164, 166  
 Voorhoeve, Jan 589, 950, 1660  
 Vooyis, Cornelis G. N. de 163  
 Vorlat, Emma 697  
 Vries, Jan W. de 164  
 Vries, Lourens J. de 1205,  
 1501 f., 1506  
 Vries, Wobbe de 163
- W**
- Waag, Albert 110  
 Wackernagel, Jacob 121, 396 f.,  
 401, 644  
 Waddell, Helen 65  
 Waddy, Julie 1033  
 Wagner, Robert-Léon 142, 1291  
 Wahrig, Gerhard 1884  
 Walch, Maria 1693  
 Waldron, Ronald A. 1620 f.  
 Wali, Kashi 1062  
 Wall, Margarita N. 1376, 1388  
 Wallace, Anthony 178  
 Wallis, Ethel 569, 572  
 Wallis, John 19, 149, 660  
 Walter, Henriette 485–488  
 Walter, Heribert 687, 709  
 Wang, Li 1731, 1735–1738  
 Wanner, Dieter 1701  
 Ward, Dennis 958  
 Wares, Alan C. 193  
 Waringhien, Gaston 1565, 1567  
 Warnke, Camilla 2  
 Warren, Beatrice 1630  
 Wasow, Thomas 810, 818  
 Waterhouse, Viola G. 192  
 Waterson, Natalie 150
- Watkins, Calvert 122, 286, 998,  
 1663, 1665  
 Watkins, Laurel J. 917  
 Watkins, Ruth 1820  
 Watson, Janet C. E. 1742  
 Waugh, Linda R. 272, 275, 279,  
 493  
 Webb, Charlotte 569  
 Weber, David 1011, 1454, 1833  
 Weekley, Ernest 1618, 1620–  
 1622  
 Weerman, Fred 166  
 Wegener, Heide 1268 f., 1808 f.,  
 1814  
 Wegera, Klaus-Peter 1684–  
 1686, 1691–1694  
 Weggelaar, Cas 926  
 Weiers, Michael 1830  
 Weigand, Georg 669  
 Weigand, Hans 976  
 Weiner, Edmund S. C. 1882  
 Weinhold, Karl 1688  
 Weinreich, Uriel 178, 243, 519,  
 1647, 1650 f.  
 Weinrich, Harald 1274  
 Weir, E. M. Helen 921, 1136  
 Weisgerber, Leo 109, 115, 160,  
 1278, 1620–1622  
 Weiß, Joseph 74  
 Weißenborn, Jürgen 983  
 Weist, Richard M. 1807  
 Wekker, Herman 1891  
 Welke, Klaus 668  
 Wellens, Ineke 1752  
 Wellmann, Hans 890, 902, 933  
 Wells, Christopher J. 1681  
 Wells, John C. 484 f., 507  
 Wells, Rulon S. 173, 184, 186–  
 189, 226, 381, 406, 452, 464,  
 475, 672  
 Welmers, William E. 588 f.,  
 1038, 1112, 1538  
 Wentworth, Harold 954  
 Werner, Otmar 91, 289, 301,  
 519, 738, 1267–1270, 1272,  
 1274, 1680, 1683, 1902  
 Westney, Paul 662  
 Westrum, Peter N. 987, 1100  
 Wetzer, Harrie 766 f., 1113  
 Wexler, Ken 1821  
 Wheatley, Julian K. 747  
 Wheeler, Deirdre W. 878  
 Wheeler, Max W. 294  
 Whitaker, Katherine P. K. 589  
 Whitney, William Dwight 98,  
 358 f., 568  
 Whorf, Benjamin Lee 883, 1450,  
 1563  
 Wicentowski, R. 1898  
 Wiederspiel, Brigitte 986  
 Wiegand, Herbert Ernst 1888  
 Wiersma, Robinia 1501  
 Wierzbicka, Anna 32, 259, 262,  
 286, 457, 708, 719, 758, 1046,

- 1048–1050, 1068 f., 1077,  
1083, 1113, 1423 f.
- Wiese, Bernd 1270
- Wiese, Richard 476, 558, 900,  
904, 1275
- Wiesemann, Ursula 785, 983,  
987, 1013, 1100, 1832
- Wiesinger, Peter 1267
- Wiik, Kalevi 1338
- Wijk, Nicolaas van 162 f.
- Wilbur, Ronnie Bring 566
- Wilder, Chris 398
- Wildgen, Wolfgang 4, 7
- Wilkins, David 1203
- Willett, Thomas 1197 f.
- Williams, Edna R. 86
- Williams, Edwin S. 31, 200–203,  
249, 268 f., 271, 315 f., 319 f.,  
322, 325 f., 350, 352, 354,  
358, 366, 408, 418, 540, 596,  
603, 855, 860–862, 900, 904,  
911, 1141, 1255, 1287, 1295
- Williams, Jennifer 1161
- Williams, Stephen J. 823
- Williamson, Kay 1128
- Willis, Penny 478, 482
- Wills, Dorothy D. 987, 991, 994
- Wilmanns, Wilhelm 19, 104,  
106, 158 f., 898, 901, 1683 f.
- Wils, J. 254
- Wilske, Ludwig 887–889, 892 f.,  
950
- Wilson, William H. 1094 f.
- Wimbish, John S. 1833
- Wimmer, Hans 1922
- Windfuhr, Gernot L. 746, 1086
- Winfield, W. W. 766
- Winford, Donald 1665
- Winkler, Christian 453
- Winkler, Gertraud 1684
- Winter, Werner 584 f., 1202,  
1204
- Winters, Margaret E. 289
- Wirth, Jessica 280
- Wisch, Fritz-Helmut 1823
- Wise, Mary Ruth 1135 f., 1138 f.
- Wissing, Daniel P. 165
- Withgott, Margaret 330
- Wittgenstein, Ludwig 1233
- Wittwer, Michael 1315
- Woidich, Manfred 1742
- Woisetschlaeger, Eric 259
- Wölcken, Fritz 958
- Wolf, Heinz Jürgen 1708
- Wolf, Lothar 16 f.
- Wolf, Norbert Richard 1681,  
1684, 1687, 1691 f.
- Wolfart, H. Christoph 1096,  
1233, 1413, 1419
- Wolfenden, Elmer 1478 f.
- Wolff, John U. 1474, 1476, 1478
- Wolff, Susanne 875
- Wolfram, Walt 241
- Wonderly, William L. 546
- Woodard, Roger D. 122
- Woodbury, Anthony 1152
- Woodbury, Hanni 919, 924,  
1844
- Woodworth, Nancy L. 374, 376,  
800 f.
- Worth, Dean S. 87
- Wouters, Alphonse 56
- Wrede, Ferdinand 163
- Wright, Elizabeth Mary 1668,  
1670
- Wright, Fiona 1761
- Wright, Joseph 16, 1668, 1670
- Wright, William 555
- Wuest, Jakob 1702
- Wüllner, Franz 59, 608
- Wunderlich, Dieter 380, 902, 905
- Wundt, Wilhelm 105, 110, 122,  
162
- Wurm, Stephen A. 999, 1010,  
1501, 1756
- Wurst, Raimund Jakob 405
- Wurzel, Wolfgang Ullrich 6,  
160 f., 195, 200, 231, 266 f.,  
279, 288–294, 338 f., 433,  
479, 514, 516, 519 f., 556,  
570, 573, 597, 604 f., 622,  
632–637, 648 f., 653, 878,  
894, 1267–1272, 1277, 1603,  
1607
- X
- Xenophon 1326
- Xerberman, K. R. 1212, 1214
- Xhuvani, Aleksandër 1647
- Y
- Yallop, Colin 1079
- Yamada, Jeni E. 1822
- Yang, L. 1162
- Yang, Paul Fu-mien 1734
- Yarowsky, D. 1898
- Yaska 703
- Yip, Moira 587, 1045
- Young, Philip D. 810
- Young, Robert 316 f., 1098,  
1136
- Youssi, Abderrahim 487
- Yu, Alan C. 1058, 1062
- Z
- Zacharski, Ron 791 f.
- Zaenen, Annie 809
- Zager, David 456
- Zajjaji, Abu l-Quasim 67 f., 71,  
73
- Zajjaji, Ibn ?Ishāq 73
- Zaliznjak, Andrej A. 129–132,  
1308 f., 1868, 1876
- Zamaxshari, Abu l-Qasim 71, 73
- Zamenhof, Ludwig L. 1564 f.
- Zandvoort, Reinhard W. 835
- Zavadovskij, Jurij 1746
- Zavalà, Roberto M. 1020 f.,  
1024, 1026
- Zawadowski, Leon 655
- Zeegers Vander Vorst, Nicole  
1323
- Zemb, Jean Marie 700
- Zemskaja, Elena A. 132, 419,  
421, 424, 1309
- Zgusta, Ladislav 250, 1887–  
1890
- Zhang, Zheng-Sheng 559
- Zhen, Yuan 1737
- Zhou, Xiaolin 1780
- Zide, Norman 1128
- Zimmer, Karl 195 f., 907, 909,  
1093
- Zipf, George Kingsley 286 f.,  
289, 1233, 1252
- Žirmūnskij, Viktor M. 129
- Živova, Galina T. 1377
- Zonneveld, Ronaldus M. van  
165
- Zonneveld, Wim 165, 201, 860,  
970
- Zribi-Hertz, Anne 990
- Zubin, David A. 613, 1035 f.,  
1269, 1811
- Zubizareta, Maria-Luisa 1780
- Žuravlev, Aleksandr P. 423
- Zutt, Herta 1683, 1692, 1694
- Zwanenburg, Wiecher 266, 847,  
1287 f.
- Zwettler, Michael 1741
- Zwický, Arnold M. 200, 266 f.,  
297, 319, 346, 389 f., 392 f.,  
395, 401, 451, 464, 472, 482,  
546, 617, 815, 861, 877, 899,  
965, 967, 969, 984, 1008,  
1191, 1255, 1262, 1264 f.,  
1666

## Sprachenregister / Index of languages

### A

Abchasisch/Abkhaz 739–741, 798, 800, 802, 823 f., 980 f., 1004, 1008, 1038, 1096, 1109, 1227  
– Abaza 1039  
Abipon 541  
Achagua 1040  
Acholi (Acoli) 539, 757  
Acoma 1211  
Adnyamathanha 1491  
Adygeisch/Adigei (Adyge) 330, 1096  
Afar 543  
Afrikaans 13, 166, 232, 307, 352, 365, 658 f., 845, 847, 858, 879, 1186, 1578 f., 1586  
afrikanische Sprachen/African languages 91, 166, 548, 588 f., 867, 1001, 1005–1007, 1013, 1016, 1019, 1031 f., 1040, 1043, 1112, 1128, 1208, 1512, 1515, 1647  
– west-/West 592, 1128, 1183, 1659  
afroasiatische Sprachen/Afroasiatic languages 91, 540 f., 554, 981, 1013, 1032–1034, 1037, 1040, 1096, 1216  
Aghem 590  
Agta 558 f., 561 f.  
Aguatekisch/Aguacatec 980  
Ainu 927, 1123, 1131, 1159 f., 1162  
Akan 558, 560 f., 563, 589, 717, 790, 800, 815, 1517 f.  
– Fante 539  
Akasele 541  
Akhvakh 1035  
Akkadisch/Akkadian 35–40, 780  
– Alt-/Old- 35  
– Altbabylonisch/Old Babylonian 35  
– Standardbabylonisch/Standard Babylonian 35  
Alabama 547, 549, 583 f.  
Alamblak 1033 f., 1036, 1038, 1041, 1128  
Alawa 540, 1010  
Albanisch/Albanian 516, 710, 717, 978, 1310, 1647, 1649, 1651, 1662 f., 1665

Aleutisch/Aleut 926, 1651  
algonkianische Sprachen/Algonquian languages 476, 541, 926, 940, 1004 f., 1010, 1012, 1032, 1151, 1411–1416, 1418  
Altägyptisch/Ancient Egyptian 99, 1917  
altaische Sprachen/Altaic languages 741 f., 1077 f., 1083–1085  
Alur 1142  
Alutor 532  
Alyawarra 539, 1079  
amazonische Sprachen/Amazonian languages 1078, 1214, 1216, 1428  
Amele 1203 f.  
Amerikanische Gebärdensprache/American Sign Language 1022, 1027, 1556 f.  
amerikanische (amerindische) Sprachen/American (Native American, Amerindian) languages 91, 94, 99, 121, 176, 180, 548, 553, 588, 739, 742, 1006, 1010, 1016, 1019, 1021 f., 1024, 1028, 1032, 1045, 1065, 1077, 1082, 1087, 1202–1204, 1216, 1247, 1425, 1433, 1453  
– meso-/Meso- 1018, 1021  
– nord-/North 94, 736, 1071, 1125, 1135  
– süd-/South 94, 1019, 1038, 1041, 1208, 1432, 1469  
Amharisch/Amharic 540, 552, 555, 981, 1216, 1741  
anatolische Sprachen/Anatolian languages 585, 645,  
andische Sprachen/Andean languages 1037, 1206  
Andoke 740  
Angloromani 1647  
Anindilyakwa 1033, 1038  
Annamisch/Annamite (s. Vietnamesisch)  
Apabhraṃśa 1712, 1719–1722, 1725  
Apalai 1214  
Apuriña 1032  
Arabana 748, 1491, 1494  
Arabisch/Arabic 60, 67–72, 74, 151, 206, 238, 431, 486 f., 528, 533 f., 547–549, 554, 556, 558, 587, 620, 702 f., 715, 717, 732, 737, 740 f., 750, 758, 777, 779, 781, 798, 800 f., 821–823, 827, 830, 841, 934, 940, 1007, 1045, 1056 f., 1060, 1062, 1069, 1071 f., 1080, 1096, 1100, 1188, 1215, 1348, 1441, 1740–1742, 1745 f., 1748, 1750–1753, 1797, 1835, 1840, 1853, 1921 f.  
– ägyptisches/Egyptian 620, 1742–1744, 1746–1749, 1751 f.  
– algerisches/Algerian 1742 f., 1750 f.  
– anatolisches (Daragözü)/Anatolian 1741, 1747 f., 1750 f.  
– andalusisches/Andalusian 1748, 1751  
– Bagdad 1746, 1748  
– beduinisches/Bedouin 555, 1742, 1744–1746, 1748  
– Čiblah 1751  
– Golf/Gulf 750, 980  
– Hassāniyya 1742, 1746  
– irakisches/Iraqi 1744, 1747  
– jemenitisches/Yemeni 1742, 1748–1750  
– Juba/Juba 1752 f.  
– Kairener/Cairene 151, 555, 1751  
– Khartoum 1752 f.  
– klassisches/Classical 72, 540, 553, 555, 579, 935, 953, 1055, 1058, 1096, 1213, 1740, 1742–1751  
– libanesisches/Lebanese 1742, 1751  
– libysches/Libyan 1742  
– maghrebinisches/Maghrebine 1746, 1750  
– marokkanisches/Moroccan 765, 953, 1742–1751, 1813  
– mesopotamisches/Mesopotamian 1742  
– Nağdī 1746, 1748  
– Neu-/New 1096, 1741 f., 1744, 1746, 1748–1750  
– Neusüd-/Modern South 1750  
– Ost-/Eastern 1742  
– palästinensisches/Palestinian 1751  
– saudiarabisches/Saudi-Arabian 1744, 1746, 1751

- sudanesisches/Sudanese 1742, 1748, 1751
  - syrisches/Syrian 547, 765, 1057, 1742, 1746–1749, 1751
  - tschadisches/Chad 1742, 1748, 1751
  - tunesisches/Tunisian 579, 1217, 1742, 1751
  - usbekisches/Uzbekistan 1741, 1747–1750
  - zypriotisches/Cypriot 1748 f., 1751 f.
  - Arapesh 633
  - Arawá 1032
  - Arawak 773, 798, 821, 823, 827 f., 830, 927, 1032, 1035, 1039 f., 1039 f., 1135 f., 1139
  - Ignaciano 1039
  - Nord-/North 1039, 1043
  - Archi 482, 715, 1034, 1038
  - Arikara 1082
  - Arinisch/Arin 1376
  - Armenisch/Armenian 59, 94, 1041, 1064, 1134, 1185, 1310
  - Alt-/Old 122
  - Asamiya 1712
  - Asheninca 987
  - asiatische Sprachen/Asian languages 1016, 1018, 1021, 1028, 1032, 1113
  - ost-/East 993, 1004, 1187
  - südost-/Southeast 1053, 1216
  - Ásóka (Aśokan) 1239, 1242, 1712, 1719 f., 1722
  - Assamese 1042
  - Assanisch/Assan 1376
  - Atayal 546–549
  - athapaskische Sprachen/Athabaskan languages 763, 926, 1012, 1022, 1028, 1031, 1136, 1399–1401, 1403, 1407, 1409, 1467
  - Athapaskisch-Eyak-Sprachen/Athabaskan-Eyak languages 1399 (s. auch Eyak)
  - Äthiopisch/Ethiopian 1013
  - Atjeh/Acehnese 713, 934, 1082
  - atlantische Sprachen/Atlantic languages
    - ost-/East 1061
    - west-/West 1038
  - Atsugewi 796, 934, 1125, 1126
  - Auetö 773
  - Australische Gebärdensprache/ Australian Sign Language (Auslan) 1556
  - australische Sprachen/Australian (Aboriginal) languages 432, 540, 548, 603, 643 f., 744, 748, 751, 999, 1003, 1007, 1009–1012, 1016, 1021 f., 1026, 1032–1034, 1037 f., 1041–1043, 1056 f., 1060 f., 1076–1080, 1082–1088, 1092, 1095, 1097, 1112, 1135 f., 1158, 1202–1205, 1208, 1491, 1493–1495
  - austroasiatische Sprachen/
    - Austroasiatic languages 549, 561, 1128, 1731, 1545
  - austronesische Sprachen/Austronesian languages 362, 540, 548 f., 560 f., 577, 760, 1012, 1017, 1032, 1053, 1056, 1082, 1094, 1124, 1214, 1473, 1476, 1501
  - Avarisch/Avar 715, 1081, 1083, 1085, 1142
  - Awyu 1501 f., 1506
  - Aymara 1003, 1009
  - aymarische Sprachen/Aymaran languages 1453, 1461
  - Azerbaijani (Azeri) 1578 f.
- B**
- Babungo 978, 1006, 1137, 1192, 1199
  - Bagavalal 1035
  - Bahwana 1043
  - Balante 785
  - Balbi 93
  - Balochi 541
  - baltische Sprachen/Baltic languages 548, 1069, 1214, 1310
  - baltofinnische Sprachen/Balto-Finnic (Baltic-Finnic) languages 407, 1042, 1078, 1080, 1083, 1328, 1650
  - baltoslawische Sprachen/Balto-Slavic languages 401, 642
  - Bambara 589, 592, 715, 717, 1092
  - Bamileke
    - Dschang 590
    - Ghomálá? 590
  - Banda-Linda 529
  - Bangla (s. Bengali)
  - Baniwa 1039
  - Bantusprachen/Bantu languages 134, 182, 323 f., 329, 336, 530, 538, 750, 777, 791, 854, 1013, 1019, 1033, 1037–1040, 1042, 1070, 1077, 1120, 1123, 1186, 1189, 1092, 1208, 1210, 1216, 1527–1533, 1538, 1542
  - Barai 815, 1110
  - Bare 1043
  - Baskisch/Basque 94, 100, 569 f., 741, 759, 761, 806, 813, 816 f., 970, 980, 1000, 1002, 1007, 1011, 1013, 1045, 1054, 1056, 1061, 1064, 1067, 1078, 1085, 1199, 1214, 1651
  - West-/West 1061
  - Zentral-/Central 1061
  - Bayso 1054 f., 1057 f.
  - Bedauye (Beja) 400, 555, 777
  - Belhare 1178
  - Bella Coola 1101
  - Belorussian (s. Weißrussisch)
  - Beludzhi 1043
  - Bemba 757, 1092, 1210
  - Bembe 1114
  - Bengali (Bangla) 737, 779, 1043, 1048, 1078, 1085, 1092, 1239, 1242, 1712, 1723–1729
  - Benue-Congo-Sprachen/Benue-Congo languages 774, 778, 1043
  - Berber 330
  - Berbice 236, 1658 f.
  - Berik 987, 1100
  - Bété 738
  - Bini 569
  - Blackfoot 1048
  - Bodisch/Bodic 1205
  - Bodo-Garo 1733
  - Bolanci 971
  - Botocudo 773
  - Bowili 1036
  - Bretonisch/Breton 328 f., 367, 442, 516, 760, 823, 979, 1063, 1065
  - Britische Gebärdensprache/British Sign Language 1556
  - Budukh 1071
  - Bulgarisch/Bulgarian 29, 400, 492, 520, 814, 1064, 1176, 1245, 1649 f.
  - Alt-/Old 1649
  - Burjat 1218
  - Burmesisch/Burmese 745, 1013, 1021, 1029, 1732 f., 1736, 1887
  - Buruschaski/Burushaski 766, 1038
  - Bəhzənī 1751
- C**
- Cabecar 1020 f.
  - Caddoan 926
  - Cahuapana 927
  - Callahuaya 1461
  - Cambodian (s. Kambodscha-nisch) 1141
  - Capanahua 1141
  - Cariban languages (s. karibische Sprachen)
  - Caspian languages (s. kaspische Sprachen)
  - Catalan (s. Katalanisch)
  - Caucasian languages (s. kaukasi sche Sprachen)
  - Cayuga 728, 926, 1022

Cayuvava 927  
 Cebuano 1153 f., 1156  
 Celtic languages (s. keltische Sprachen)  
 Cemuhi 589  
 Ch'ol 935  
 Chacobo 539  
 Chadic languages (s. tschadische Sprachen)  
 Chaha 569 f., 572  
 Chama 488  
 Chamalal 1038  
 Chamorro 545, 560 f., 568  
 Chatino 588  
 Chechen (s. Tschetschenisch)  
 Chemakua 1032  
 Chemehuevi 569, 767, 1077  
 Cherokee 1023  
 Cheyenne 541, 805  
 Chibcha-Sprachen/Chibchan languages 1021  
 Chi-Bemba 1039  
 Chicewa 322–324, 477, 536, 784  
 Chickasaw 584, 1110, 1142  
 Chimariko 538  
 Chi-Mwini 324  
 chinantekische Sprachen/Chinanteck languages 539  
 Chinesisch/Chinese 92, 94, 96–99, 134, 164, 227, 239, 241, 251, 258, 316, 454, 525, 536, 559, 561, 591 f., 677, 682, 703 f., 714, 720, 724–726, 728, 758, 807, 913, 934, 936 f., 940, 947, 950, 960, 992, 1004, 1017, 1020, 1027–1029, 1045, 1051, 1070, 1151, 1218, 1222, 1232, 1545, 1730 f., 1733–1738, 1808, 1813 f., 1819, 1843, 1875, 1897, 1903, 1916, 1919  
 – Alt-/Old 937, 1147 f., 1053, 1730–1735, 1737  
 – Amoy 1734  
 – archaisches/Archaic 1730  
 – Changsha 591, 1735  
 – Chungan 1734  
 – Fuzhou 1734  
 – kantonesisches/Cantonese 589, 781, 1736  
 – Kienyang 1734  
 – klassisches/Classical 588, 940, 1735, 1737, 1871  
 – Mandarin 239, 376, 727, 736, 748 f., 773–775, 781, 800, 810, 937, 1113, 1191, 1192, 1730, 1732, 1735–1738  
 – Min 1734  
 – Mittel-/Middle 1730–1735, 1737  
 – Neu-/Modern 937, 1730  
 – Neumandarin/Modern Mandarin 1731 f.

– Proto- 1731  
 – Shanghai 592  
 – Spätmittel-/Late Middle 1730 f.  
 Chinook 1007, 1048  
 ChiNyanja 588  
 Chipewyen 546  
 Chirag 1070  
 Chishona 802  
 Chitimacha 781, 785  
 Choctaw 538, 584, 625, 1000  
 Chontal 547  
 Chrau 934  
 Chuave 1204  
 Chuckchee (s. Tschuktschisch)  
 Chukotko-Kamtchatkan languages (s. Tschukotko-Kamtschatka-Sprachen)  
 Chuvasch (s. Tschuvaschisch)  
 Clallam 578, 936  
 Cocama 798  
 Cœur d'Alène 553 f., 720  
 Comanche 917, 1005, 1077, 1433  
 Comox 923, 926  
 Coos 537  
 Cora 1129  
 Cree 1096, 1233, 1411, 1419, 1647  
 Cross River 1043  
 Cushitic languages (s. kuschitische Sprachen)  
 Czech (s. Tschechisch)

**D**

Dagbani 593  
 daghestanische Sprachen/Dagestanian languages 526, 1070 f.  
 Dakota 442, 517, 547–549, 739, 806, 1082, 1099, 1100, 1835  
 Dangbon 1042  
 Dani 1142, 1203, 1205  
 Dänisch/Danish 95, 145–147, 249, 253, 378, 380, 382, 384, 386, 584, 741, 784, 832, 960, 1058, 1070, 1160, 1328, 1389, 1905, 1914  
 – Alt-/Old 146  
 Dargwa (Dargi) 526, 543, 1203  
 Deutsch/German 1, 3, 11, 13–15, 26, 61, 91–93, 96, 104 f., 108–111, 114, 118, 122, 135, 157, 159 f., 163, 174, 216, 236–239, 241, 249 f., 255, 266–269, 273 f., 291, 298 f., 301, 338, 340 f., 351–353, 355 f., 362, 366, 378–380, 383, 412, 418, 420, 423, 440–442, 445, 454–459, 468, 479, 482, 485, 491 f., 495, 497 f., 517 f., 520, 528, 535 f., 541 f., 546, 568–570, 573 f., 576, 582–584, 595–597, 599 f., 602, 604 f., 607, 612 f., 621 f., 626, 628, 630–637, 646, 648–650, 655 f., 659, 661, 670, 674–677, 682, 684 f., 693, 697–699, 703, 732, 734–739, 744 f., 749, 751 f., 754 f., 774, 777, 779–782, 787–793, 798, 800, 803, 811, 829, 833 f., 836, 858, 881, 887–895, 899 f., 902–905, 907–909, 911, 929–937, 939–941, 943–945, 950, 956, 958, 960–962, 970, 975, 1007, 1013, 1035–1037, 1046–1051, 1054–1056, 1061, 1063 f., 1067 f., 1072, 1076, 1078 f., 1087, 1091, 1093, 1099 f., 1114, 1135 f., 1139 f., 1142, 1147, 1149, 1153, 1160 f., 1167, 1178, 1181, 1212, 1217 f., 1224, 1228 f., 1255, 1267–1278, 1309, 1325, 1328, 1491, 1546, 1564, 1582, 1600, 1602 f., 1608 f., 1612, 1615–1621, 1625–1634, 1636–1638, 1643, 1648, 1650, 1671–1674, 1681–1686, 1690 f., 1694, 1705, 1756–1758, 1767 f., 1770, 1773, 1775, 1783, 1797, 1799 f., 1803, 1808–1814, 1817, 1819–1821, 1835 f., 1838, 1840–1843, 1852–1854, 1860, 1862–1864, 1873, 1878, 1880, 1884–1888, 1890, 1894, 1897 f., 1901, 1903–1906, 1912–1916, 1919–1922, 1924, 1931 f., 1936  
 – Althoch-/Old High 13, 122, 159, 636, 930, 936, 956, 1604–1610, 1618, 1637, 1680–1694, 1910  
 – Bayrisch/Bavarian 1067  
 – Fränkisch/Franconian 584  
 – Hessisch/Hessian 475, 582, 584  
 – Mittel-/Middle 1069, 1682, 1691  
 – Mittelhoch-/Middle High 649, 930, 1217, 1604–1606, 1609 f., 1616 f., 1621, 1637, 1680–1694, 1914, 1917  
 – Neuhoch-/Modern High 675, 1087, 1228, 1602–1604, 1626, 1631, 1680–1686, 1688–1694, 1910, 1917, 1920  
 – Nieder-/Lower 1681  
 – Nord-/Northern 1625  
 – Ober-/Upper 1682, 1685, 1688, 1690, 1692 f.

- Sächsisch/Saxonian 1681, 1692
- Süd-/Southern 1178, 1186, 1625
- Westmittel-/Western Middle 1689
- Westober-/Western Upper 1689
- Deutsche Gebärdensprache/German Sign Language 1554–1561
- Dholuo 802, 1142
- Dhuwal 1003
- Diegueño 569 f., 1023, 1088
- Digul-Wambo 1501
- Dinka 571, 851
- Diyari 987, 993, 1035, 1078, 1158, 1159, 1490–1500
- Djabugay 258, 262
- dravidische Sprachen/Dravidian languages 561, 813, 1013, 1032, 1037, 1039, 1043, 1077 f., 1080, 1085
- Dutch (s. Niederländisch)
- Dyirbal 481, 643, 714, 716, 719, 742, 758, 817, 1004, 1011, 1029, 1034–1038, 1043, 1080, 1112, 1116, 1655 f.
- Giramay 1011
- Dyola 541
- 'Dongo-ko 1040
  
- E
- Efik 778
- Englisch/English 13 f., 16, 19, 38, 52 f., 58, 61 f., 64, 82, 86–88, 95, 98, 118, 120, 133, 135, 149, 152, 154 f., 163, 172, 175, 178, 180 f., 185, 195 f., 198–203, 205, 216 f., 232, 238, 245, 248–253, 255, 258, 263–267, 272–279, 282, 284–286, 288, 296–300, 308, 312–314, 316, 318–320, 322–324, 329, 331, 335 f., 340, 342, 346–348, 351 f., 354 f., 358 f., 362–365, 367, 370–373, 379–381, 383, 393–395, 402, 408, 410, 412–414, 427 f., 430, 433, 440–442, 453, 457, 464–466, 471–474, 481, 484 f., 490–494, 497, 500–502, 506 f., 512, 514 f., 517, 525, 529–531, 533 f., 536 f., 541, 545–547, 568–571, 576–578, 583, 587, 589, 596, 599, 601 f., 607, 610, 617 f., 620 f., 625 f., 628, 633, 648 f., 651, 653, 655–658, 660–663, 668, 670, 675, 682, 684, 708, 713–718, 727 f., 732, 741, 743–747, 749, 751, 755, 757, 759 f., 762–765, 770–773, 775, 777–782, 787–793, 795–801, 803, 806–808, 810–818, 824, 827 f., 830, 832–834, 840–843, 845 f., 849, 853, 860, 864, 870, 874, 881 f., 886, 889–891, 893–895, 899–905, 908 f., 911–913, 921 f., 925 f., 930 f., 935–937, 939–941, 943–951, 953 f., 957–962, 964–981, 988 f., 998–1004, 1006–1009, 1011, 1013, 1020, 1033 f., 1036, 1043, 1053, 1060 f., 1063, 1067–1069, 1072, 1077, 1086 f., 1094, 1105, 1107 f., 1112–1116, 1118 f., 1121 f., 1127, 1130, 1134–1136, 1138, 1147, 1149, 1153 f., 1156 f., 1161 f., 1165–1178, 1180 f., 1183–1189, 1193, 1197–1200, 1202, 1205, 1208–1210, 1213–1215, 1217, 1222–1224, 1226, 1228 f., 1232–1234, 1236, 1255–1266, 1276, 1278, 1329, 1334, 1356, 1400, 1411 f., 1415, 1417, 1422, 1424, 1433 f., 1437, 1439, 1443, 1466, 1473, 1480, 1512, 1529, 1533, 1541, 1546, 1564, 1567, 1584, 1591, 1593 f., 1596, 1612, 1615–1622, 1630–1633, 1636–1640, 1642 f., 1653, 1656 f., 1664–1666, 1668, 1671–1679, 1694, 1704–1706, 1724, 1752, 1755–1758, 1762 f., 1766, 1769 f., 1773–1775, 1778, 1780, 1782–1785, 1789, 1796–1798, 1800–1803, 1806–1809, 1812–1814, 1819–1822, 1831, 1836, 1839 f., 1852, 1862, 1882–1890, 1894, 1896–1899, 1912, 1914, 1920–1922, 1924–1927
- Alt-/Old 86 f., 118, 372, 573, 576, 599, 696, 1043, 1080, 1087, 1228, 1236, 1239, 1242, 1256, 1343, 1616–1618, 1620, 1627–1630, 1632, 1634, 1636, 1641, 1666, 1668–1677, 1679
- amerikanisches/American 180, 772, 970, 1264 f., 953, 1473, 1616, 1620, 1625, 1633
- angelsächsisches/Anglo-Saxon 95, 1236, 1621, 1921
- anglo-irisches/Hibernian English 1649
- britisches/British 1625, 1628, 1633, 1071, 1264 f., 1883
- chinesisches Pidgin/Chinese Pidgin 1657
- Frühneu-/Early Modern 1668, 1671–1675, 1677–1679
- Hindi-Englisch/Hindi English 232
- Inglis (Scottis) 1668–1670
- kentisches/Kentish 1668, 1670
- Mercian 1668, 1669
- Midland 1668, 1670, 1672 f.
- Mittel-/Middle 88, 649, 1579, 1617, 1626, 1630, 1633, 1642, 1668, 1671–1679
- Neu-/Modern (Present-day) 118, 1239, 1242, 1256, 1343, 1626–1631, 1633 f., 1642 f., 1663, 1666, 1668, 1671, 1674 f., 1678 f.
- Neufundland-/Newfoundland 1070
- Neuseeland-/New Zealand 1071
- Nord-/Northern 1668–1670, 1672–1674
- Spätneu-/Late Modern 1668, 1677
- Süd-/Southern 1668, 1672 f.
- tasmanisches/Tasmanian 1070
- West Saxon 1668 f., 1671
- West Somerset 1070
- Ennemor 569 f., 573
- Epun 1822
- eritreische Sprachen/Eritrean languages 1808
- Erzya Mordva/Erza Mordvin 821, 823, 825–830
- Ese'ejja 488
- Eskimo 204, 254, 380, 539, 739, 922 f., 926, 1005, 1009, 1152, 1162, 1223, 1236 f., 1239, 1241 f., 1400, 1407, 1651
- Eskimo-aleutische Sprachen/Eskimo-Aleut languages 1032, 1077
- Esperanto 453, 933, 1563–1571
- Estnisch/Estonian 546, 774, 778, 814, 1042, 1076, 1085, 1218, 1328
- Etsako 571, 573
- europäische Sprachen/European languages 91, 367, 381, 620, 1001, 1016, 1019, 1067, 1071, 1147, 1157, 1183–1186, 1208, 1218, 1228 f., 1319, 1564, 1570 f., 1647–1649, 1682, 1657, 1704, 1706, 1708, 1886 f., 1931
- nord-/North 1189
- west-/West 381, 1189, 1647
- Even 1151

Evenki 1151, 1160  
Ewe 30, 719, 751, 1005 f., 1036,  
1092, 1094  
Eyak 1023, 1028, 1399

## F

Fang 469  
Färöisch/Faroese 1268  
Farsi 779  
Fidschi/Fijian (Fiji) 1108, 1208,  
1135 f., 1139  
– Boumaa 1108  
Filipino (s. Philippinisch)  
Finnisch/Finnish 252, 330, 388,  
389–391, 396, 400, 569 f.,  
622, 647, 649–652, 701, 753,  
758, 809, 811, 817, 825, 941,  
980, 988, 1060, 1072, 1079 f.,  
1084–1086, 1114, 1121, 1127,  
1160, 1185, 1186, 1328–  
1342, 1602 f., 1650, 1797,  
1894, 1897  
finnopermische Sprachen/Finno-  
Permic languages 401  
finnougriische Sprachen/Finno-  
Ugric languages 541, 1056,  
1083, 1218, 1328  
Flämisch/Flemish 829, 1038  
Fore 1078, 1081, 1127, 1203 f.  
Französisch/French 16 f., 86–89,  
93, 109, 118, 135, 138–142,  
144, 159, 175 f., 183 f., 216,  
232, 235–237, 245, 249, 253,  
274 f., 279, 291, 297, 316,  
318, 338, 352, 358, 361, 363,  
378–383, 385, 393, 395,  
398 f., 402, 454, 457 f., 476,  
484–488, 492, 495, 506, 512,  
514–517, 521, 525, 535 f.,  
571, 582–585, 599, 618 f.,  
622, 626, 655, 659 f., 662,  
675, 680, 696–698, 702 f.,  
708, 732, 745, 750, 772, 774,  
776 f., 781, 800, 812, 832,  
833 f., 841, 843 f., 846–849,  
851 f., 859, 869, 872–874,  
881, 886, 888–890, 892, 894,  
908, 911–913, 929, 935, 940,  
950, 957–960, 962, 964, 966,  
979, 1006 f., 1011, 1013,  
1016, 1031, 1037, 1039, 1042,  
1051, 1063, 1067–1069,  
1112, 1160, 1166 f., 1173–  
1175, 1177–1179, 1181, 1186,  
1217, 1228 f., 1258 f., 1264,  
1285–1297, 1411, 1417,  
1564, 1581, 1583, 1590, 1592,  
1594–1596, 1603, 1615–  
1618, 1620–1622, 1628,  
1630, 1634, 1637 f., 1640–  
1642, 1647 f., 1653, 1666,

1671, 1675–1677, 1679,  
1698–1710, 1745, 1750,  
1785, 1799–1801, 1803,  
1808, 1811 f., 1819, 1821,  
1836, 1843, 1852, 1859, 1861,  
1863, 1883–1885, 1887–  
1889, 1891, 1897, 1910–  
1912, 1914, 1920 f., 1924,  
1931  
– Alt-/Old 86 f., 278, 645, 1086,  
1579, 1616 f., 1621, 1698–  
1709  
– Alt nord-/Old Northern 1618  
– elsässisches/Alsatian 1617  
– kanadisches/Canadian 488  
– Mittel-/Middle 1698–1701,  
1704 f., 1707  
– Neu-/Modern 1699–1705,  
1707–1709  
– normannisches/Norman 1257  
– West-/West 444  
Französische Gebärdensprache/  
French Sign Language 1560  
Friesisch/Frisian 95, 278, 1070,  
1255  
– West-/West 444, 754  
Fula 568–571, 573 f., 762, 790,  
1040, 1047, 1049, 1134 f.,  
1150  
Fur 1042

G

Gaagadju 1034  
Gälisch/Gaelic 228, 1656  
– schottisches/Scottish 1655  
– hebridisches/Hebridean 623  
Gallisch/Gallic 1218  
Gāmid 1744  
Garo 537, 1196  
Georgisch/Georgian 87, 325,  
439, 442–444, 521, 541 f.,  
597, 814, 934, 1079, 1082,  
1085, 1136–1138, 1186, 1215  
– Alt-/Old 754, 1079  
– Neu-/Modern 1215  
germanische Sprachen/Germanic  
languages 94–96, 121, 157 f.,  
300, 314, 335, 338, 341, 361,  
363, 382, 384, 386, 444, 520,  
531, 541, 548, 642, 653, 849,  
857 f., 861, 880, 911, 937,  
950, 1045, 1070, 1078, 1080,  
1167, 1172, 1186, 1228,  
1257 f., 1262, 1264, 1268,  
1310, 1563 f., 1596, 1609,  
1643, 1670, 1674, 1675, 1681,  
1685–1687, 1689, 1691–  
1693, 1704, 1707, 1709, 1801,  
1903, 1927  
– nord-/Northern 1914  
– nordost-/North East 380

– süd-/Southern 1681  
– Urgermanisch/Proto-Ger-  
manic 520, 1268  
– west-/West 118, 1255, 1267  
Giha 1527  
Gimi 549  
Gimira 1007, 1011  
Godié 984, 987, 994  
Godoberi 1035, 1037  
Gokana 588, 590, 592  
Gola 741  
Golin 1053  
Gombe Fulani 538, 814  
Gooniyandi 1078, 1203, 1871  
Gotisch/Gothic 93, 96, 104, 159,  
936, 1056, 1132, 1139, 1141,  
1239, 1242, 1621, 1662, 1666,  
1685, 1691  
Grebo 1135  
Greenlandic (s. Inuktitut)  
Griechisch/Greek 18, 40, 52–65,  
74, 77–79, 86 f., 92–96, 139,  
152, 174, 242, 245, 251, 253,  
304, 367, 388, 406, 422, 435,  
451 f., 457, 539, 561, 563 f.,  
642, 644 f., 654, 694, 696,  
701, 709, 748, 777 f., 780,  
809, 812, 814 f., 818, 869,  
874, 901, 945, 957, 968, 971,  
1031, 1074, 1076, 1139 f.,  
1145, 1149 f., 1156, 1181,  
1222, 1224, 1239, 1242, 1247,  
1261, 1264, 1288, 1310,  
1313–1321, 1323, 1325,  
1343, 1617, 1619 f., 1648,  
1662–1666, 1671, 1704 f.,  
1709, 1750, 1835, 1862, 1878,  
1918, 1925, 1927 f., 1930–  
1932  
– aiolisches/Aeolic 1310  
– Alt-/Ancient 119, 121 f., 239,  
291, 354, 516, 624, 732,  
1056 f., 1074, 1076–1078,  
1080, 1086 f., 1214, 1218,  
1224, 1228, 1239, 1310–  
1326, 1616, 1621, 1853, 1914,  
1917  
– anatolisches/Anatolian 585  
– arkadisch-kyprisches/Arcado-  
Cypriot 1310  
– attisches/Attic 956, 1311,  
1313, 1316, 1319, 1321  
– dorisches/Doric 1310  
– homerisches/Homeric 1064,  
1133, 1242, 1310, 1315, 1325  
– ionisch-attisches/Ionic-Attic  
1310, 1325  
– klassisches/Classical 342,  
658, 988, 1149 f., 1155, 1157,  
1159, 1233  
– Koiné 1311  
– makedonisches/Macedonian  
585

- mykenisches/Mycenaean 1310
- Neu-/Modern 58, 235, 359, 662 f., 1055, 1134, 1187, 1217, 1239, 1242
- Nordwest-/Northwest 1310
- Nord-/Northern 585
- West-/West 1310
- Grönländisch (s. Inuktitut) Gta? 553 f., 570
- Guahibo 1032
- Guajajara 1194
- Guajiro 1034
- Guana 773
- Guaraní 711, 927, 1107, 1233, 1421–1432
- Guinea-Bissau Portuguese Creole (s. Portugiesische Kreolsprache von Guinea-Bissau) Gujarati 1712, 1723–1729
- Gumbaynggir 992, 1095, 1097
- Guoyu 1730
- Gur 936
- Gurindji 796, 1003
- Guugu Yimidhirr 797, 993, 1092, 1099
- Gu-Yalanji 539
- Gwari 592
- Gyarong 1082
  
- H
- Haitisch/Haitian 237, 1653, 1657 f.
- Harakmbet 927
- Harari 540
- Haruui 817
- Hatam 1056, 1062 f.
- Hausa 29, 467, 518, 590, 593, 714 f., 718, 741, 752, 758, 777, 786–788, 793, 817, 990 f., 993, 1035, 1216
- Havyaka 1208–1210
- Hawaiisch/Hawaiian 1095, 1862
- Haya 590 f.
- Hebräisch/Hebrew 79, 331, 364, 458, 474 f., 538, 702 f., 711, 737, 781, 799, 957, 1061, 1092, 1094, 1096, 1139, 1228, 1343–1356, 1650, 1796 f., 1799–1803, 1819, 1821
- Alt-/Old 362
- klassisches/Classical (Biblical) 362, 1343, 1345, 1352 f.
- mittelalterliches/Medieval 1343
- Neu-/Modern (Israeli Hebrew, Contemporary Hebrew, Ivrit) 1039, 1343, 1345, 1352 f., 1355
- rabinisches/Rabbinical (Mishnaic) 1343
- Herero 1094
- Hethitisch/Hittite 35, 122, 641–643, 735 f., 738 f., 1042, 1239, 1242, 1310, 1662 f.
- Hidatsa 584, 1197, 1198
- Hiligaynon 1061
- Hindi 232, 245, 546, 713, 954, 1039, 1094, 1138, 1142, 1161, 1712, 1723–1729
- Hixkaryana 739, 741–743, 823, 979 f., 1098, 1203, 1208
- Hoka/Hokan 541, 773, 1125, 1126
- Hopi 746, 987, 994, 1064, 1433
- Hottentot (s. Nama)
- Houailou 1094
- Hua 547, 549, 568–570, 573, 712, 1012, 1204 f., 1232 f.
- Huahtla 926
- Hualapai 1063 f.
- Huichol 242 f., 536, 1054, 1056, 1058–1063, 1433
- Hungarian (s. Ungarisch) Hunzib 1062, 1367–1375
- Hurrisch/Hurrian 1079
  
- I
- Iatmul 1036, 1127
- Igbo 589, 767, 934, 1112, 1128
- Ijo 741, 1658 f.
- Ik 1092, 1098 f.
- Ika 1135
- Ila 777
- Illyrisch/Ilyrian 1310
- Ilokano/Ilocano 971, 1061
- Imonda 1023, 1026 f.
- Indisch/Indic 1042, 1085, 1914, 1919
- Alt-/Old 96, 100, 548, 642, 644 f., 1056, 1316, 1879, 1914
- Mittelindische Umgangssprache/Middle Indic vernacular 42, 50
- indoarische Sprachen/Indo-Aryan languages 737, 1013, 1078, 1209, 1712–1728
- alt-/Old 1712–1723, 1728
- mittel-/Middle 1712, 1719–1722, 1725–1729
- neu-/New 1712, 1720, 1722–1729
- indogermanische (indoeuropäische) Sprachen/Indo-European languages 18, 94–96, 100, 119, 121 f., 126, 145, 147, 149, 157 f., 361, 376, 396, 536, 539, 541, 548 f., 561, 568, 602, 631, 633, 639, 642, 644 f., 700, 710, 725, 735, 737–739, 746, 748, 772, 809, 911, 940, 945, 971, 1013, 1031–1035, 1040–1043, 1056, 1068, 1076–1078, 1080, 1083–1087, 1112, 1141, 1147, 1149 f., 1156, 1182, 1203, 1205, 1215–1218, 1227 f., 1237, 1301, 1307, 1310, 1312 f., 1317, 1328, 1334, 1336, 1343, 1412, 1439, 1560, 1649, 1655, 1663, 1686 f., 1689, 1712, 1717 f., 1724, 1730, 1752, 1853
- Urindogermanisch/Proto-Indo-European 100, 158, 514, 612, 642, 1087, 1142, 1159, 1214, 1216, 1328, 1662–1667
- indoiranische Sprachen/Indo-Iranian languages 122, 1085 f., 1310
- indonesische Sprachen/Indonesian languages 94
- Indonesisch/Indonesian 241, 414, 540 f., 549, 676, 715, 721, 739, 758, 849, 921, 934, 951, 1011, 1017, 1134, 1218, 1501, 1503, 1853
- Inga 1197
- Ingrisch/Ingrian 1328
- Ingusch/Ingush 1039
- Interlingua 1563–1571
- Inuktitut (Grönländisch/Greenlandic) 254 f., 538, 701, 768, 922, 924, 1121, 1152, 1218, 1400, 1800, 1821
- Ostgrönländisch/East Greenlandic 1389
- Polareskimo/Polar Eskimo 1389
- Westgrönländisch/West Greenlandic 979, 981, 1107, 1216, 1389–1398, 1840
- Irakw 1035
- iranische Sprachen/Iranian languages 314, 541, 602, 642, 1042, 1808, 1651
- Irisch/Irish 328, 500, 503–505, 517, 570 f., 574, 627 f., 806, 823, 826 f., 1095, 1218, 1233, 1622, 1649
- Alt-/Old 87, 1215, 1219, 1664 f.
- Neu-/Modern 1215 f., 1218 f.
- irokesische Sprachen/Iroquoian languages 926, 975, 1022, 1263
- süd-/Southern 1023
- Isländisch/Icelandic 444, 500, 502, 584, 620, 1610, 1618, 1085, 1268, 1914
- Alt-/Old 13, 86, 1604, 1606, 1607
- Neu-/Modern 13, 1601, 1607, 1609

Isoko 593  
 Italienisch/Italian 14, 59, 95,  
 198, 199, 201, 249 f., 255,  
 330, 352, 358, 361, 364, 366,  
 378, 384, 398, 444, 447, 452,  
 456–458, 469, 471, 485 f.,  
 542, 585, 587 f., 661, 697,  
 718, 743, 750, 775 f., 812 f.,  
 832, 841, 859, 869, 874 f.,  
 878, 882, 890, 892 f., 934,  
 936, 988, 1031, 1046 f., 1049,  
 1051, 1068, 1216–1218,  
 1264, 1288, 1563 f., 1592,  
 1617 f., 1671, 1709, 1741,  
 1745, 1750, 1784, 1797, 1808,  
 1810, 1813 f., 1819, 1821,  
 1931

Italisch/Italic 1310  
 Itelmenisch/Itelmen 443  
 Itzá 257  
 Iwaidja 711 f.

## J

Jakaltekisch/Jacalteco 516, 979,  
 1017, 1021, 1026–1028, 1094  
 Jakutisch 1236, 1239, 1242  
 Japanisch/Japanese 14, 88, 236,  
 238, 244, 375, 536, 712, 717,  
 745, 755, 805, 807, 832, 882,  
 961, 969, 987–989, 992 f.,  
 998, 1000, 1003 f., 1006,  
 1009, 1012, 1017, 1019, 1028,  
 1045, 1051, 1060, 1063 f.,  
 1077, 1094, 1114, 1128, 1129,  
 1146, 1150 f., 1153, 1156,  
 1161, 1185, 1187, 1208, 1215,  
 1755, 1819, 1821, 1844, 1862,  
 1897, 1916, 1917  
 – Alt-/Old 970  
 – Neu-/Modern 969  
 Javanisch/Javanese 95, 568–570,  
 883, 951, 1012  
 – Alt-/Old 999  
 Jarawara 1034, 1038  
 Jawony 1010, 1034, 1042  
 Jê 1032  
 Jenissei-Sprachen/Yenisei lan-  
 guages 1376  
 Jiddisch/Yiddish 243, 482, 971,  
 1186, 1650  
 Jívaro 539, 1654  
 Jukagirisch/Yukaghír 543

## K

K'iche 1800  
 Kabardian 1071  
 Kabyl 539  
 Kachin 1733

Kaingang 1035, 1832  
 Kalam 713, 795 f.  
 Kaliana 773  
 Kalkatungu 1084–1087, 1136  
 Kalmuk 1218  
 Kambata 467, 474  
 Kambera 362  
 Kambodschanisch/Cambodian  
 1225  
 Kannada 758, 761, 765, 1137,  
 1139, 1208, 1211  
 Kanukuru 798  
 Kanuri 374, 807, 1135, 1205  
 Karaboro 1183  
 Karaimisch/Karaim 439, 441 f.  
 Karelisch/Karelian 1086, 1328  
 karibische Sprachen/Cariban lan-  
 guages 1208  
 – Ge-Pano 541  
 – Kreolsprachen/Creoles 1655  
 Karok 541, 546, 553, 781, 803  
 kartvelische Sprachen/Kartvelian  
 languages (s. kaukasische  
 Sprachen)  
 Kaschmiri 1062  
 Kasem 576  
 Kashaya 1203  
 kaspische Sprachen/Caspian lan-  
 guages 541  
 Kassitisch/Kassite 35–37  
 Katalanisch/Catalan 619 f., 622  
 Katcha 1035  
 kaukasische Sprachen/Caucasian  
 languages 482, 541, 548,  
 1032, 1035, 1041, 1077,  
 1084 f., 1096, 1203, 1205  
 – kartvelische Sprachen/Kart-  
 velian languages 1082, 1215,  
 1218  
 – nordost-/Northeast 1034,  
 1038, 1083, 1085, 1142, 1203  
 – nordwest-/Northwest 549,  
 980 f., 1004, 1008, 1038,  
 1096, 1109, 1136, 1227  
 – süd-/South 1032  
 Kawi 94 f.  
 Kayardild 708 f., 711, 717 f., 744,  
 753 f., 1057, 1087, 1108, 1880  
 Kazakh 529  
 Kegboid 1043  
 Kekchi 536, 541, 543  
 keltische Sprachen/Celtic lan-  
 guages 93 f., 327, 478, 503,  
 569–571, 574, 1071 f., 1218,  
 1310, 1620, 1665  
 Kerasan 1211  
 Ketisch/Ket 536, 1035 f., 1376–  
 1388  
 Kewa 1011, 1128  
 Khinalug 1033  
 Khmer-Sprachen/Khmer lan-  
 guages 539, 720, 805  
 Khmu 934  
 Khoisan-Sprachen/Khoisan lan-  
 guages 541  
 – Nord-/Northern 1032  
 – Zentral-/Central 1037  
 Kihehe 559, 561, 565  
 Kikuyu 1527  
 Kilivila 1012  
 Kimberley 1038  
 Kimbundu 1160  
 Kinande (s. Nandi)  
 Kinubi (s. Nubi)  
 Kinyarwanda (Rwanda) 1058,  
 1062, 1135, 1139, 1527–  
 1544, 1867, 1904  
 Kiowa 917  
 Kiowa-Tanoa-Sprachen/Kiowa-  
 Tanoan languages 926, 1042  
 Kiranti 1082  
 Kiriwinisch/Kiriwinian 757  
 Kirundi 530 f., 1038, 1527, 1536  
 Kiswahili (s. Swahili)  
 Kitimaat 723  
 Kituba 1654  
 Kiwai 536, 541  
 Koasati 584  
 Kobon 713, 736, 767, 1202 f.  
 Kolami 1035, 1039  
 Kombai 1501  
 Komi 330, 1183  
 Komy-Permyak 1085  
 Konkani 767, 1712, 1724  
 kordofanische Sprachen/Kordo-  
 fanian languages 1037, 1208  
 Koreanisch/Korean 275, 568–  
 570, 708, 714, 716, 1012,  
 1064, 1077 f., 1120, 1128,  
 1151, 1153, 1160 f., 1208,  
 1663, 1842, 1897, 1919  
 Koromfe 759, 760, 766–768  
 Koryak 528  
 Kottisch/Kott 786, 1376  
 Koyo 741  
 Koyra Chiini 1006  
 Koyukon 917–919, 1399–1410  
 Kresh 1094  
 Krongo 1091, 1095, 1098, 1107  
 Kru 1043  
 Kui 539, 766  
 Kunjen 751  
 Kunwinyku 1010  
 Kurdisch/Kurdish 1186, 1741  
 Kuria 1037  
 kuschitische Sprachen/Cushitic  
 languages 555, 577, 777, 919,  
 927, 1007, 1013, 1054 f.,  
 1057 f., 1077, 1135  
 – Ost-/East 474  
 – Süd-/Southern 1647  
 Kuyani 1491  
 Kwa 569  
 Kwakw'ala 739, 742 f., 934  
 Kwara'ae 577, 580  
 Kxoe 743, 1036, 1042

## L

- Lahu 587, 589
- Lak 516, 715, 1038
- Lakhota 758, 846
- Lango 802, 1142, 1203
- Lappisch/Lappish 95, 330, 786
- Larike 1057
- Latein/Latin 7, 11, 13 f., 18, 25, 28, 52–65, 76–79, 81–83, 86–89, 93–96, 118–120, 139–141, 153 f., 159, 164, 174, 187, 227 f., 235, 237, 250, 268 f., 271, 277 f., 291, 293 f., 304, 362 f., 367, 376, 393, 396, 406, 418, 422, 536–538, 548, 561, 563 f., 576, 584 f., 588, 597–600, 602–605, 618, 620, 631, 634, 637–639, 642–645, 647 f., 651, 654–657, 662, 674 f., 694, 696 f., 701, 704, 732 f., 735–738, 743 f., 746–748, 750–752, 754, 761, 770, 772, 774, 777 f., 780–782, 799, 808, 812–814, 852, 857, 869, 874, 881, 901, 911, 935, 940, 945 f., 950, 956 f., 966 f., 988, 1000, 1006, 1031, 1039, 1051, 1064, 1074–1077, 1079, 1081, 1084–1087, 1126, 1135 f., 1139, 1140, 1145, 1147, 1149, 1173, 1181, 1186 f., 1202, 1213–1215, 1217 f., 1222, 1224–1226, 1228 f., 1247, 1264, 1273, 1288, 1295, 1310, 1314, 1316, 1325, 1343, 1491, 1545, 1563 f., 1613, 1615–1619, 1621, 1628–1630, 1633 f., 1642, 1648 f., 1651, 1662–1666, 1671, 1675 f., 1681, 1684, 1687, 1698–1710, 1834–1836, 1838–1840, 1842, 1852–1854, 1861 f., 1877–1879, 1911, 1917–1919, 1926–1933, 1935
- Alt/Old 1706
- Früh-/Early 1611 f.
- klassisches/Classical 278, 660, 1612, 1698–1701, 1704
- Spät-/Late 538, 1617, 1649, 1703, 1706, 1708
- Vulgär-/Vulgar 1592, 1698–1703, 1705, 1708
- Laz 1082
- Lendu 588
- Lesgisch/Lezgi (Lezgian) 517, 604 f., 658, 660, 663, 1041, 1080, 1137, 1871
- Lettisch/Latvian 92, 693, 1069 f., 1650
- Limbu 1214

## Lingala

- 810
- Lingua Franca 1653
- Litauisch/Lithuanian 96, 401, 619, 1056, 1069, 1132, 1140, 1142
- Livisch/Livonian 1070, 1328, 1650
- Logo 1095
- Lokono 1034
- Lotuxo 772
- Lude 1328
- Luganda 593, 757, 1039, 1070, 1120, 1123 f., 1527
- Luiseño 396, 400, 818, 1204
- Lule 1835
- Lumbi 767
- Lunda 1042
- Luo 1188
- Luoravetlan 927
- Lwo 802
- Lyélé 987, 994
- M
- Ma 1037
- Ma'a 1647
- Maasai 516, 530, 536, 588
- Maba 1079
- Macushi 1097
- Madagassisch/Malagasy 702, 991 f.
- Madura/Madurese 400
- Maidu 539, 1203
- Maipura-Sprachen/Maipuran languages 927, 1467
- Maithili 1161
- Makedonisch/Macedonian 29, 327, 391, 396, 398, 400, 585, 1176, 1245
- Makro-Gê Sprachen/Macro-Ge languages 773
- Makú-Puinave 921
- Malabang 1762
- Malaiisch/Malay 528, 547, 549, 751, 784, 1045, 1048, 1758
- Bazaar 1501
- Malak-Malak 1039
- Malayalam 537, 546
- malaio-polynesische Sprachen/ Malayopolynesian languages 1061, 1473
- Maltesisch/Maltese 658, 760, 1056, 1071, 1741, 1745, 1748, 1750
- Malto 1035, 1040 f.
- Mam 258, 262, 541, 1012, 1022, 1026, 1033, 1100
- mamische Sprachen/Mamean languages 1028
- Manam 999, 1022
- Manambu 1033, 1036–1038
- Manchu (s. Mandschu)
- Mandan/Mandan 1024
- Mandara 546, 548
- Mandarin (s. Chinesisch)
- Mande-Sprachen/Mande languages 588, 1092
- Mandinka 1659
- Mandobo 1501
- Mandschu/Manchu 495, 554, 568–570, 1738
- Mandschu-tungusische Sprachen/Manchu-Tungus languages 1032
- Mangarayi 744, 750 f., 835, 978, 1082
- Mangbetu 539
- Manipuri 758 f., 1209
- Manjaku 1659
- Maori 29, 766, 935
- Mapuche 1194 f., 1197, 1205
- Mara 998
- Marathi 1712, 1724, 1729
- Margi 540
- Mari 330
- Maricopa 1009, 1203
- Marind 1038
- Marrithiyel 1042
- Marshallese 935
- Mayasprachen/Mayan languages 532, 539, 541, 774, 927, 1012, 1028, 1098, 1100, 1152, 1835, 1842, 1863
- Mayali 529, 711 f., 918, 926, 1042, 1060, 1854
- Mba 1040, 1094
- Mbangi 1099
- Mbugu 1647
- Media Lengua 1647
- Mednyj Aleut 1651
- Meglenoromanisch/Meglenoromanian 1651
- Melanesisch/Melanesian 1045, 1755 f.
- Mende 569 f., 572, 574, 588
- Menomini 735
- Miao 1734
- Mikasuki 584
- Mikir 1733
- mikronesische Sprachen/Micronesian languages 1021 f., 1024
- Miskito 548, 750, 1216
- Miwok 555 f., 578, 803 f.
- Mixtekisch/Mixtec 431, 591, 927, 1021, 1027
- Caotzoquitengo 1021
- Chalcatongo 766
- Mixtepekisch/Mixtepec 541
- Miya 1216
- Mohawk 711 f., 924, 993, 1027, 1029, 1800
- Mokil/Mokilese 559 f., 563
- Mongolsprachen/Mongol languages 1217

Mongolisch/Mongolian 1217 f.  
 Mon-Khmer-Sprachen/Mon  
   Khmer languages 548, 1128,  
   1734  
 Montagnais (Innu-aimun) 588,  
   1411–1419  
 Mordvinisch/Mordvinian 330  
 Motu 935  
 Motuna 1032, 1039, 1041  
 Mparntwe Arrernte 1203  
 Mpi 590  
 Mubi 554  
 Munda-Sprachen/Mundan lan-  
   guages 548, 553, 927, 1004,  
   1013, 1128  
 Mundani 988, 1209  
 Mundari 720, 725 f., 728  
 Mundurukú 919, 1022, 1024  
 Mupun 1042  
 Mura 927  
 Murle 1100  
 Muskogee-Sprachen/Muskogean  
   languages 584, 814, 926,  
   1088, 1110, 1142  
 Mutsum 578 f.  
 Muyu 1501  
 Mwera 804, 1173

## N

Nadēb (Nadeb) 921, 1136  
 Nahuatl 442, 541, 701, 711 f.,  
   739, 797, 1433–1452  
   – Huasteca 1056, 1060  
   – Isthmus 1436  
   – klassisches/Classical 538,  
   621 f., 741, 926, 939, 1433,  
   1436, 1442, 1835  
 Naiki 1039  
 Nakanai 548, 560 f., 798  
 Nakhsprachen/Nakh languages  
   1032, 1039  
 Nama (Hottentot) 539, 763,  
   1045  
 Nanai 1085  
 Nandi (Kinande) 591, 593, 1135,  
   1507  
 Nasioi 1032  
 Naskapi 1411, 1419  
 Nass 524  
 Naukan 778  
 Navajo 316, 330, 539, 739, 796,  
   1005, 1011, 1098, 1100, 1136  
 Ndu 1032  
 Nembe 593  
 Nenets 1101  
 Neo-Melanesian (s. Tok Pisin)  
 Nepali 1043  
 Nera 1100  
 Neuaramäisch/Neo-Aramaic  
   1741 f.

Neuguinea-Pidgin/New Guinea  
   Pidgin (s. Tok Pisin)  
 Newari 1205 f.  
 Nez Perce 1082, 1087  
 Ngäbére 810  
 Ngalakan 1034  
 Ngamini 1491  
 Ngandi 1010, 1023  
 Ngan'gityemerri 1042  
 Ngiyambaa 394, 396 f., 822,  
   1004, 1195–1199  
 Nicobarese 803  
 Niederländisch/Dutch 13, 162–  
   166, 198, 229–232, 250, 300,  
   304, 306–310, 318, 336–342,  
   347, 351 f., 357–359, 361–  
   366, 367, 412, 467 f., 476,  
   482, 516 f., 541, 715, 728,  
   800, 803, 821, 824, 829 f.,  
   833, 841, 848, 856–862, 864,  
   883, 902, 904, 911, 926, 933,  
   943 f., 949 f., 977, 979, 1046,  
   1048–1051, 1063, 1070–  
   1172, 1177 f., 1198, 1214,  
   1218, 1255, 1272, 1288, 1328,  
   1578–1581, 1583–1586,  
   1617, 1641, 1643, 1653, 1659,  
   1783 f., 1807–1813, 1819,  
   1821, 1891, 1894, 1897 f.  
   – amerikanisches/American  
   232, 1579, 1587  
   – Ost-/East 1037  
 Niger-kordofanische Sprachen/  
   Niger-Kordofanian lan-  
   guages 937  
 Niger-Kongo-Sprachen/Niger-  
   Congo languages 541, 561,  
   978, 1006, 1032, 1042 f.,  
   1094, 1112, 1135, 1137, 1187,  
   1192, 1199, 1208, 1517, 1538  
 nilo-saharanische Sprachen/Nilo-  
   Saharan languages 772,  
   1077, 1135, 1208, 1507–  
   1509, 1513  
 nilotische Sprachen/Nilotic lan-  
   guages 772, 1032, 1507, 1513  
 Nimboran 539, 803  
 Niue 1059, 1061  
 Nivkh 475, 716  
 Nkore-Kiga (Runyankore, Ru-  
   nyankole) 253, 768, 980 f.,  
   1187, 1189, 1210, 1527  
 Nomatsiguenga 1136, 1138  
 Noni 590  
 Nootka 539, 721–723, 935, 940  
   – Nitinaht 251, 400, 722  
 Nordische Sprachen/Nordic lan-  
   guages 1604  
   – Alt-/Old 95, 1621, 1914  
   – Urnordisch/Proto-Nordic  
   1601, 1604–1606, 1608–  
   1610, 1914

Norwegisch/Norwegian 146 f.,  
   384, 784, 960, 1100, 1617,  
   1914  
 nostratische Sprachen/Nostratic  
   languages 121  
 Nubi (Kinubi) 1654, 1752  
 Nubische Sprachen/Nubian lan-  
   guages 1752  
 Nuer 569  
 Nungali 1980  
 Nunggubuyu 1003, 1008, 1041  
 Nyigina 1078

## O

O'doham 806  
 Obolo 759  
 Ojibwe (Ojibwa) 785, 788, 1049  
 Ok 1501  
 Okrand 579  
 Okzitanisch/Occitan 85 f., 88,  
   1186  
 Ollari 1035  
 Omotische Sprachen/Omotic lan-  
   guages 1007, 1013  
 Onondaga 1844  
 Oriya 774, 778, 1043  
 Ormuri 541  
 Osage 1835  
 Ossetisch/Ossete 1085  
 Ostyak 774  
 Otomangue-Sprachen/Otoman-  
   guean languages 541, 1021,  
   1032  
 Otomi 569 f., 572  
 ozeanische Sprachen/Oceanic  
   (Oceanian) languages 1016,  
   1021, 1025, 1094, 1097,  
   1135 f., 1142

## P

Paiute 716  
 Palantla 539  
 Palauan 990 f.  
 Pāli 1712, 1719–1722  
 Palikur 1039  
 Pama-Nyunga-Sprachen/Pama-Nyun-  
   gan languages 481, 987, 993,  
   1035, 1077–1079, 1085,  
   1158 f., 1491, 1493 f.  
 Pamir-Sprachen/Pamir lan-  
   guages 541  
 Panara 1022  
 Panjabi 1723  
 Panoan 927, 1467  
 Panyjima 1009, 1079  
 Papago 584, 801, 1204  
 Papiamentu 1654, 1656–1659  
 Papua-Sprachen/Papuan lan-  
   guages 987, 1012, 1022,

- 1032, 1037, 1041, 1078, 1100, 1110, 1120, 1126f., 1142, 1202–1205, 1501, 1506  
**Parji** 1035, 1039  
**Pashto** 399 f., 541, 546  
**Paumari** 987, 993, 1041, 1043, 1071  
**Pawnee** 543, 804, 1196  
**Peba** 1464  
**Peba-Yagua-Sprachen/Peba-Yaguán languages** 1464  
**Penuti-Sprachen/Penutian languages** 541, 770, 1077  
**Persisch/Persian** 93–95, 458, 520 f., 739, 741, 746, 1043, 1061, 1086, 1093, 1161, 1185, 1186, 1236, 1649, 1651  
– Alt-/Old 520, 1239, 1242  
– Avestisch/Avestan 520  
– Frühneu-/Early Modern 541  
– Neu-/Modern 1239, 1242, 1649  
**Philippinisch/Filipino** 1473  
**Philippinische Sprachen/Philippine languages** 1152–1156, 1473, 1479  
**Phrygisch/Phrygian** 1310  
**Pima** 801, 1204  
**Pipil** 760, 979  
**Pirahã** 1053  
**Pitjantjatjara** 715  
**Pitta-Pitta** 1491  
**Plains Cree** 1151  
**Polnisch/Polish** 134, 290 f., 318, 335, 339, 365, 391, 395 f., 414, 427 f., 432 f., 437, 441 f., 469, 490, 508, 520, 529, 548, 583–585, 661, 863, 881, 936 f., 1048–1050, 1068 f., 1085, 1245, 1801, 1809, 1819, 1889, 1904  
**polynesische Sprachen/Polyesian languages** 94, 99, 585, 780, 1045, 1225, 1263  
**Pomo** 1203, 1208  
– Zentral-/central 1058  
**Ponape** 467, 1022, 1024  
**Popoloc** 593  
**Portugiesisch/Portuguese** 271, 380 f., 697, 737, 799, 806, 1032 f., 1039, 1047, 1051, 1217, 1564, 1653, 1656 f., 1659, 1710, 1813 f.  
– brasilianisches/Brazilian 1042  
**Portugiesische Kreolsprache von Guinea-Bissau/Guinea-Bissau Portuguese Creole** 1658  
**Potawatomi** 476 f., 742 f.  
**Potou-Tano** 1517  
**Prakrit** 41, 50, 1712, 1719–1722  
**Puglian** 1069  
**Puluwat** 717  
**Pumpokolisch/Pumpokol** 1376  
**Puquina** 1461  
**Putonghua** 1730
- Q**
- Qafar** 1035  
**Q'anjeb'al (Kanjobal)** 1020 f., 1024–1028  
**Qawasqar** 488  
**Quechua** 412, 701, 714, 739 f., 746, 828, 1004, 1091, 1101, 1106, 1137, 1140, 1194, 1199, 1453–1463, 1647, 1656  
– Amazon 1654 f.  
– Ancash 1219, 1455, 1458, 1460 f.  
– Ayacucho 1455–1462  
– bolivianisches/Bolivian 1139, 1460  
– Cajamarca 1457  
– Cochabamba 1157  
– Cuzco 1457–1461, 1654 f.  
– ekuadorianisches/Ecuador 1458, 1460 f.  
– Huallaga 1011  
– Huancayo 1455, 1459–1461  
– Imbabura 1092, 1137 f., 1457, 1459, 1461  
– Pacaraos 1454, 1456–1460  
– Puno 1458  
– San Martín 1457  
– San Pedro de Cajas 1458  
– südperuanisches/Southern Peruvian 1004  
– Tarma 1454–1462  
**Quiché (s. K'iche)**  
**Quileute** 547, 549, 560 f., 564, 565 f.
- R**
- Rama** 1119, 1136, 1206  
**Rembarrnga** 1042  
**Romani** 1645, 1648–1651  
– lettisches/Latvian 1070  
**romanische Sprachen/Romance languages** 88, 141, 159, 323, 327, 358, 361, 367, 380–382, 384, 398, 422, 444, 521, 538, 645, 658, 660, 814 f., 849, 857, 911 f., 950, 961, 1045, 1051, 1070, 1077, 1157, 1167, 1175 f., 1184, 1186, 1217, 1228, 1233, 1261, 1264, 1310, 1440, 1563–1565, 1567, 1570 f., 1596, 1603, 1616, 1643, 1651, 1674, 1698, 1700–1704, 1706 f., 1710, 1798, 1801, 1810, 1904, 1921, 1927, 1931  
– balkan-/Balkan- 1649  
– ibero-/Ibero- 1654, 1656  
– räto-/Rheto- 444, 521, 1651, 1701  
– west-/West 1087  
**Rotuman** 475, 529, 577, 580, 936  
**Ruc** 1734  
**Rukai** 541  
**Rukiga** 1527  
**Rumänisch/Romanian** 26, 444, 529, 533 f., 584, 625, 645, 657, 741, 1047, 1086 f., 1217, 1648–1650  
– Istro- 1648  
**Runyambo** 1123 f., 1129  
**Runyankole (s. Nkore-Kiga)**  
**Runyoro** 1527  
**Russisch/Russian** 11, 14, 125, 127–133, 135, 155, 237, 239, 251, 254, 261, 269 f., 273, 277, 282, 285, 314, 338, 414, 418–421, 427, 431–433, 444, 466, 477, 481, 489 f., 492 f., 498, 513, 515 f., 518, 520 f., 525, 528, 531, 534, 535 f., 583, 601 f., 605, 607–613, 626, 632 f., 650–652, 655 f., 663, 697, 700, 708–711, 714 f., 717, 719, 735–737, 743, 746, 770–772, 776 f., 779, 782, 806, 814, 832, 886, 889, 892–894, 929, 930, 934–937, 939 f., 950, 957 f., 961, 975, 1033–1035, 1037, 1039, 1067–1072, 1077, 1080, 1085, 1132 f., 1139 f., 1142, 1157, 1159, 1161, 1167, 1172 f., 1175 f., 1181, 1183 f., 1186, 1188, 1210, 1234, 1245, 1256, 1277, 1300–1309, 1328, 1376, 1441, 1546, 1551 f., 1564, 1601, 1651, 1799, 1809, 1867 f., 1876, 1879, 1912, 1921 f.  
– Alt-/Old 937, 1070, 1218  
– Mittelgroß-/Middle Great 1300  
– Neu-/Modern 1218  
– Nordgroß-/North Great 1300  
– Südgroß-/South Great 1300  
**Rwanda (s. Kinyarwanda)**
- S**
- Saami (Saame, Sámi)** 330  
– Enontekio 401 f.  
– South 147  
**Saek** 1216  
**Sait** 1501  
**Sakao** 539  
**Salish** 411, 702, 721–724, 726, 729, 923, 926, 1032, 1119, 1208

- Lummi 400, 578
- Saanich 411, 548 f., 578
- Samoanisch/Samoan 242 f., 541, 549, 560 f., 564 f., 582, 758, 786, 1071, 1107 f., 1214, 1217
- samoanisches Plantagen-Pidgin/  
Samoa Plantation Pidgin  
1755, 1758
- samojedische Sprachen/Samo-  
yedic (Samoyed) languages  
330, 1328
- Sango 757
- Sanskrit 18, 41–43, 45, 49–51,  
93–96, 98 f., 121, 174, 253,  
358 f., 383, 454, 568 f., 623,  
703, 737, 749, 758, 762 f.,  
908, 911, 956, 971, 1057 f.,  
1061, 1085 f., 1142, 1149 f.,  
1155–1157, 1159, 1214,  
1222 f., 1228, 1236, 1239,  
1242, 1247, 1592, 1662–  
1666, 1712, 1717–1722,  
1726, 1728, 1837, 1862, 1914
- Sanuma 1216
- Saramaccan 1657–1659
- Sayula Popoluca 622
- Schwedisch/Swedish 14, 147 f.,  
380, 384, 648, 650, 662, 741,  
902, 960, 1038, 1051, 1068–  
1072, 1189, 1328 f., 1648,  
1793, 1821, 1903 f., 1914
- Sekani 546
- semitische Sprachen/Semitic lan-  
guages 35, 83, 99, 134, 205,  
314, 336, 339, 364, 410,  
548 f., 552, 554, 558, 562,  
741, 750, 869, 1056, 1077 f.,  
1080, 1085, 1095 f., 1215,  
1228, 1343, 1655, 1741, 1750,  
1835, 1917
- alt-/Old 1096
- Seneca 1034, 1040, 1082, 1098
- Sepik 1211, 1032, 1037, 1042
- Serbo-Kroatisch/Serbo-Croatian  
206, 327, 400, 520, 542, 603,  
803, 1080, 1245, 1783, 1797,  
1922
- Kroatisch/Croatian 1648,  
1922
- Serbisch/Serbian 1648, 1922
- Sesotho 1019, 1799
- Seychellen-Kreolsprache/Sey-  
chelles Creole 1218
- Shilluk 1062, 1232
- Shona 1038
- Siane 549
- sibirische Sprachen/Siberian lan-  
guages 475
- ost-/East 475
- paläo-/Paleo- 477
- Sidamo 577
- Sierra Popoluca 1004
- Si-Luyana 846 f., 939

- Sinhala/Sinhalese 1086
- sinotibetische Sprachen/Sino-  
Tibetan languages 549, 805,  
1731, 1737
- Sioux-Sprachen/Siouan lan-  
guages 926, 1032, 1037
- Siriono 711
- Siuslaw 1080
- Sizang (Siyin) Chin 1735
- skandinavische Sprachen/Scandi-  
navian languages 145, 911,  
1621, 1819
- festland-/continental 13
- Skandinavische Gebärdenspra-  
che/Skandinavian Sign Lan-  
guage 1556
- Slave 407, 412, 798
- slawische Sprachen/Slavic (Sla-  
vonic) languages 314, 318,  
340, 375, 422, 444, 520, 548,  
588, 648, 650, 772, 1034,  
1042, 1045 f., 1049, 1068,  
1071, 1142, 1157, 1167, 1172,  
1176, 1228, 1301, 1307, 1310,  
1647 f., 1651, 1800, 1904,  
1922
- (Alt-)kirchenlawisch/Old  
Church Slavonic (Old Slavo-  
nic) 87, 517, 520, 548, 1040,  
1071, 1239, 1242, 1300, 1308
- Ost-/East 1245
- Slowakisch/Slovak 133, 477,  
1245
- Slowenisch/Slovenian 291, 1056,  
1245
- Solomoninsel-Pidgin/Solomon  
Islands Pijin 1755
- Somali 778, 1093, 1096, 1099,  
1135
- Songhay 1006
- Soninke 569 f.
- Sonrai 1160
- Sora 547, 549, 1086
- Sorbisch/Sorbian 362, 1056,  
1648
- Ober-/Upper 318, 1245
- Sotho 1186
- Spanisch/Spanish 29, 95, 183 f.,  
193, 236 f., 243 f., 272 f.,  
275 f., 327, 371, 373, 390 f.,  
398 f., 485, 493, 497, 513,  
517 f., 530, 542, 583 f., 621 f.,  
626, 658, 660, 697, 786–793,  
796, 800, 806, 869, 873, 880,  
883, 911–913, 926, 950, 954,  
957 f., 960 f., 970, 988, 1009,  
1019, 1034, 1036, 1039, 1041,  
1049, 1069, 1071, 1082 f.,  
1086, 1064, 1091, 1106, 1186,  
1194, 1217, 1229, 1264,  
1421 f., 1426, 1433, 1437,  
1461, 1464, 1467, 1471,  
1473 f., 1564, 1582, 1592,
- 1603, 1617, 1646 f., 1653,  
1656 f., 1666, 1671, 1709,  
1807 f., 1813 f., 1919–1922,  
1931
- Alt-/Old 576, 881
- argentinisches/Argentinian  
327
- kantabrisches/Cantabrian  
1036, 1042
- mexikanisches/Mexican 953
- Lena 1069
- Squamish 539, 1020, 1024, 1027,  
1080
- Sranan (Sranan Tongo) 950,  
1657, 1660
- Sumerisch/Sumerian 35–40, 751,  
1217, 1916
- Sundanesisch/Sundanese 151,  
545, 548 f., 846, 935, 939, 951
- Supyire 757, 761, 764
- Svan 1215
- Swahili (Kiswahili) 546, 711,  
743, 762, 786, 789, 791, 793,  
814, 934 f., 939, 1031–1033,  
1037 f., 1042, 1045, 1047 f.,  
1062, 1094, 1135, 1141, 1187,  
1189, 1236, 1239–1242,  
1508, 1527, 1529, 1533, 1542,  
1655, 1842 f., 1880, 1888
- Syrisch/Syriac 59, 74

## T

- Tabassaranisch/Tabassaran 516,  
642, 1083
- Tadzhik 458
- Tagalog 269, 390, 396, 399 f.,  
409, 411, 528, 530, 541, 543,  
547 f., 559, 560 f., 569, 702,  
718, 721 f., 724, 726, 729,  
743, 749, 760, 786, 841, 1070,  
1072, 1123 f., 1203, 1218,  
1473–1488, 1836
- Taglisch/Taglish 1473
- Tahitisch/Tahitian 780, 798, 990
- Tai 1027
- Taino 1617
- Taiwanische Gebärdensprache/  
Taiwan Sign Language 1557
- Tajik 520
- Takelma 757, 764, 814, 926,  
1101
- Talysh/Talysh 541
- Tamil 557, 558–561, 563, 703,  
738, 761, 920 f., 1035, 1037,  
1039, 1060, 1080, 1085, 1097,  
1138, 1142, 1189
- Tangkisch/Tangkic 1087
- Taraft 823
- Tarahumara 1433
- Tarascan 1027, 1077, 1085
- Tariana 1039

- Tauya 1191 f., 1195  
 Tazerwalt Shilh 1041  
 Telugu 1033, 1138  
 Temiar 539  
 Temne 465  
 Tera 553  
 Terena 587, 1210  
 Teso-Turkana-Sprachen/Teso-Turkana languages 1507, 1509  
 Thaisprachen/Thai languages 548, 1545  
 Thai 92, 251, 561, 716, 745, 758, 941, 1012 f., 1017 f., 1020 f., 1028 f., 1186  
 Thargari 747  
 Thirrari 1491  
 Thrakis 1310  
 Tibetisch/Tibetan 96, 1085, 1593, 1731–1734, 1736 f.  
 – klassisches/Classical 1732 f.  
 – Lhasa- 1595, 1597  
 – zentrales/Central 1592  
 tibetobirmanische Sprachen/Tibeto-Burman languages 591, 1010, 1013, 1032, 1209, 1733–1735, 1737  
 Tigre 806  
 Timucua 701  
 Tinrin 764  
 Tiwa 924 f.  
 Tiwi 539, 801, 926, 1034, 1038  
 Tlapanec-Subtiaba 927  
 To'abaita 1097  
 Toba 1024  
 Tocharisch/Tocharian 1310  
 Tojolabal 805  
 Tok Pisin (Neo-Melanesian, New Guinea Pidgin) 1755–1765  
 Tolai 993, 1070, 1758, 1762–1764  
 Tondano 541  
 Tongisch/Tongan 717, 721, 724–726, 729, 934, 938, 1136, 1173  
 Tonkawa 1085  
 Totonakisch/Totonac 442, 541, 927, 1869  
 trans-neuguineische Sprachen/Trans-New Guinean languages 1208  
 Trique 715  
 Trukese 549, 806, 938, 1160  
 Trumai 1855  
 Tsachurisch/Tsaxur 1215  
 Tsakwambo 1501  
 Tsangla 1733  
 tschadische Sprachen/Chadic languages 553 f., 1013, 1094, 1216  
 Tschechisch/Czech 133, 253, 440 f., 471, 491, 584, 956, 1245, 1643, 1905 f.
- Tschetschenisch/Chechen 1039, 1058, 1062, 1203, 1227  
 Tschukotka-Kamtschatka-Sprachen/Chukotko-Kamtschatskan languages 541, 1032, 1263  
 Tschuktschisch/Chuckchee (Chuckchi) 252, 314 f., 477 f., 528, 538, 541, 560, 916–918, 980, 1010 f., 1162  
 Tschuvaschisch/Chuvash 753, 799, 1072  
 Tsimshian 524, 760, 926  
 Tsova-Tush 1034, 1039  
 Tuareg 540  
 Tübatulabal 508  
 Tuburi 1006  
 Tucano 374, 803, 1039, 1040  
 Tümpiso (Tümpisa) 766, 1215  
 tungusische Sprachen/Tungusic languages 1151, 1160, 1208  
 Tupi 701, 927  
 Tupi-Guaraní-Sprachen/Tupi-Guarani languages 1421  
 Tupinamba 742  
 Türksprachen/Turkic languages 314, 531, 1013, 1032  
 Turkana 588, 637, 751, 758, 766, 822, 1032, 1507–1516  
 Türkisch/Turkish 11, 28 f., 98, 150 f., 227–229, 232, 275, 285, 292 f., 314, 317, 330, 347, 351, 364, 370, 429, 465, 492, 495 f., 528, 539, 604, 619, 623–625, 630, 711, 715, 717, 733, 736–742, 753, 778, 786, 805, 825, 846, 939–941, 969 f., 981, 1000, 1056, 1061, 1064, 1073 f., 1077, 1079–1081, 1083–1085, 1096, 1106 f., 1133 f., 1138, 1191, 1193–1195, 1197–1199, 1215, 1222–1224, 1226, 1237, 1239, 1242, 1358–1366, 1578 f., 1583–1585, 1649–1651, 1741, 1785, 1797, 1800, 1809 f., 1813, 1835, 1841 f., 1873, 1877, 1879, 1894, 1897  
 Tuscarora 916–919, 922, 925, 975  
 Tututni 1012  
 Tuyuca 1022  
 Twi 561, 1062, 1517–1526  
 – Akuapem 1517–1519, 1522  
 – Asante 1517 f.  
 Tzeltal 1020, 1028, 1831  
 Tzotzil 1017, 1020, 1025 f., 1098, 1134  
 Tzutuhil (Tzutujil) 532, 760, 764, 766
- U  
 Udihe 1215  
 Udmurtisch/Udmurt 330, 541, 799, 1132  
 Ugaritisch/Ugarit 35  
 Uigurisch/Uigur 539  
 Ukrainisch/Ukrainian 520, 633, 1244 f.  
 Ungarinjin 1034, 1039  
 Ungarisch/Hungarian 13, 166, 252, 285, 298 f., 330, 338, 340, 469, 477, 495, 516 f., 583, 598 f., 618 f., 621, 632, 644, 657 f., 717, 737, 746, 753 f., 771, 778, 821, 823–826, 829, 846, 939 f., 1049, 1064, 1068, 1080, 1085, 1097, 1109 f., 1132 f., 1214, 1218, 1227, 1233, 1263, 1312, 1797, 1801, 1820  
 Uradhi 1078  
 Ural-altaische Sprachen/Ural-Altaic languages 99, 134  
 uralische Sprachen/Uralic languages 330, 1032, 1077, 1083, 1085, 1121, 1208, 1328  
 Urdu 1138, 1142  
 Urhobo 589  
 Usan 1003, 1011 f., 1203–1205  
 Usbekisch (Özbekisch)/Uzbek 1651, 1741  
 uto-aztekische Sprachen/Uto-Aztec languages 541, 917, 926, 1005, 1077, 1088, 1097 f., 1129, 1203, 1205, 1433
- V  
 Vanuatu Bislama 1755  
 Vedisch/Vedic 41–43, 1214, 1712, 1716–1719, 1914  
 – Rigvedisch (Rgvedic) 1239, 1242, 1724  
 Venetian 1217  
 Veps 401, 1328  
 Vietnamesisch (Annamitisch)/Vietnamese 30, 201, 254, 636, 708, 720, 724 f., 745, 813 f., 937, 988, 992 f., 1013, 1021, 1070, 1216, 1223–1226, 1236, 1239, 1241–1243, 1247, 1545–1552, 1731, 1734 f., 1897  
 Vogul 778  
 Volapük 1563–1571  
 voltaische Sprachen/Voltaic languages 541, 1042  
 Vot/Vote 401, 1328

## W

- Wagaya 1034  
 Wakaschisch/Wakashan 1032  
 Walisisch/Welsh 328, 478, 481 f.,  
   532, 572, 574, 619, 660, 717,  
   774, 792, 823, 1071, 1218 f.  
 Walmatjari 796  
 Walpiri (s. Warlpiri)  
 Wambaya 718  
 Wambon 1501–1506  
 Wangkangurru 1491, 1494 f.  
 Wangkumara 1034, 1491  
 Wappo 939  
 Warekna 1043  
 Wargamay 985, 1209  
 Wari' 1214, 1216  
 Warlpiri 391, 396, 400, 469,  
   558 f., 561, 643, 718 f., 796,  
   821, 1003, 1056, 1800  
 Warndarang 1003  
 Warray 1034, 1041  
 Warrungu 1158  
 Warumungu 470  
 Weißenrussisch/Belorussian 1245  
 Wemba-Wemba 643  
 Western-Desert-Sprachen/West-  
   ern Desert languages 603,  
   621  
 Western Torres Strait 1034 f.,  
   1041  
 Wik-Munkan 993  
 Winnebago 1123  
 Woarani 927  
 Wolof 779, 939, 987, 991–994,  
   1108

- Worora 540, 1038  
 Wunambal 809, 810, 1038

## X

- Xhosa 1217

## Y

- Yağ Dii 985, 1006  
 Yagua 978 f., 1017, 1020, 1464–  
   1472  
   – Cahocuma 978  
 Yameo 1464  
 Yami 561  
 Yana 539, 926  
 Yandrruwandha 1491  
 Yankunytjatjara 603  
 Yanyuwa 1038  
 Yao 1734  
 Yap 1062  
 Yaqui 1077 f., 1086, 1092  
 Yawarrawarra 1491  
 Yessian-Mayo 1041  
 Yiddish (s. Jiddisch)  
 Yidiny 341 f., 400, 467, 644, 796,  
   1021, 1135, 1158  
 Yimas 796, 1032, 1036, 1038,  
   1127, 1211  
 Yokuts 622, 1004  
 Yonggom-Wambo 1501  
 Yoruba 151, 539, 558, 560 f.,  
   563, 745, 800, 934, 1094,  
   1182, 1185 f.

## Yukaghir (s. Jukagirisch)

- Yukatekisch/Yucatec 247, 736 f.,  
   739 f., 749, 809, 984–987,  
   989–991, 993, 1022, 1065,  
   1838, 1841–1843, 1852–  
   1854, 1861, 1863, 1871, 1878

## Yukulta 1079, 1087

- Yuma-Sprachen/Yuman lan-  
   guages 569 f., 710, 739, 741,  
   773, 1022 f., 1077, 1088,  
   1202, 1208  
 Yupik (Yup'ik) 922 f., 926, 1081,  
   1389  
   – zentrales/Central 1147, 1162  
   – sibirisches/Siberian 537  
 Yurok 552 f., 770

## Z

- Zande 1037, 1042  
 Zapotekisch/Zapotec 541 f.  
 Zend 96  
 Zhuang 1216  
 Zoque 546, 548 f., 569 f., 927,  
   934, 1077, 1085  
 Zulu 539, 636, 854, 939 f., 1048,  
   1186, 1904  
 Zuni 925 f.

## !

- !Kung 541, 797, 804  
 !Xu 1032

# Sachregister / Index of subjects

## A

abbreviation (s. Abkürzung)  
Abbreviaturbildung 1308  
abessive 1085  
ability, acquired 1194  
ability, intrinsic 1194  
Abkürzung/abbreviation 958, 1875  
Ablaut/ablaut 174, 568, 620, 971, 1271 f., 1275, 1671, 1683, 1685  
– Ablautausgleich 1680  
– Ablautklasse 1272  
– Ablauttyp 1272  
– *guna* ablaut 568  
– *vriddhi* ablaut 568  
Ableitung (s. Derivation)  
absentative 1413  
absolute 748, 1434, 1441  
absolutive 739 f., 1081, 1514, 1727  
abstraction 747  
abstractness of underlying representations 500  
Abundanzplural 1057  
accent (s. Akzent)  
accidence 344  
accommodation 1793  
accountability principle 1665  
accountability, total 183  
accusative 1493 f.  
accusative language 1081  
accusative system 1081  
acquisition process 1741  
acronym 956, 958, 1256, 1619, 1678  
actant 1146  
activa tantum 1150  
active language 1082  
active nominalizer 1458  
actuation 308  
ad hoc formation (s. Einzelprägung)  
adaptation 1581, 1587  
– morphologische Adaptation/ morphological adaptation 1581, 1645  
adessive 1080  
adjacency condition 199, 411  
Adjektiv/adjective 1111, 1114, 1301, 1306, 1369, 1466, 1510, 1537  
– adjectivization 940

– deadjectival adjective 1709  
– denominal adjective 1708  
– deverbal adjective 1708  
– possessive adjective 743  
– relational adjective 1708  
– semantic type for adjectives 1112  
adjustment rule 198  
admonitive 1191, 1440  
adstratal language 1741  
Adverb/adverb 675, 1301, 1315, 1466  
– adverbial suffix 1709  
– mimetic adverb 1128  
affective 1085  
affirmation 1207, 1210  
affixation 527, 750  
– affix substitution 165  
– Affix/affix 257, 406, 524, 1218, 1255, 1853  
– Affixoid/affixoid 892, 931, 1566  
– affix-stripping 1779  
– derivational affix 1801  
– Halbaffix 892  
– phrasal affix 395  
– relatives Affix 892  
agent 979, 1081, 1146  
Agglutination/agglutination 1271, 1615, 1655  
agglutinierende Sprache/agglutinative language 254, 619, 630, 1235, 1222, 1247, 1300, 1358, 1415, 1465  
Agrammatismus/agrammatismus 364, 1769, 1817  
agreement (s. Kongruenz)  
agreement class 1033  
Ähnlichkeit 236  
Aktionsart 1165, 1277, 1509  
– dynamic Aktionsart 1168  
– phasal Aktionsart 1167, 1172  
– stative Aktionsart 1168  
Aktor 1100  
Akzent/accent 251 1319  
– Akzentierungsmuster 1244  
– Akzentverschiebung 1063  
alienability 1025  
– alienable noun 740  
– alienable possession 1748  
– inalienable noun 740  
– inalienable possession 1748  
alignment 341  
allativ/allative 1494, 1521  
alliterative concord 1038  
allocutive 1002  
Allomorphie/Allomorphik/allomorphy 848, 874, 1268, 1289, 1311, 1655, 1876  
– Allomorph, Auswahl 1270  
– Allomorph/allomorph 336, 463, 473  
– stem allomorphy 363  
Alphabet 1564  
– Alphabetschrift 1916  
Alternation/alternation 336, 485, 489, 527, 1685  
– automatische Alternation/automatic alternation 178, 464  
– morphologisches Alternationsverhalten 1244  
– non-neutralizing phonetic alternation 497  
– phonetic alternation 497  
– revised alternation condition 339  
– weak beat alternation 1539  
Ambiguität 1270  
ambiguity, structural 837  
analog-holistisch 888  
Analogie/analogy 363, 1574, 1611, 1637, 1680, 1766  
analogization 1615  
analytisch/analytic/analytical 1217, 1306, 1671, 1698  
– analytic structure 1223  
– analytical form 524  
– analytisches Ausdrucksverfahren 1268  
– Analytismus/analyticity 1238, 1742  
anaphoric island 199  
animacy 1016, 1412  
– animacy hierarchy/empathy hierarchy 734, 1068, 1762  
– animate 747, 1465  
– inanimate 747  
Anschluß 1914  
anterior 1177  
anticausative 1132  
antipassive 1148, 1152  
antonym, gradable 1115  
antonymy 1115  
aorist 1176, 1715  
Aphasie 1817  
– amnestische Aphasie 1819  
– aphasic 1782  
– aphasic speech 1789

Apophonie/apophony 527, 529, 568, 620, 739, 1687  
 – accentual apophony 530  
 – segmental apophony 529  
 – tonal apophony 530  
 – zero apophony 531  
 Appellativ 1901  
 applicative 1134  
 – benefactive applicative 1134  
 – comitative applicative 1135  
 – comprehensive applicative 1135  
 – direct applicative/goal applicative 1135  
 – instrumental applicative 1135  
 – locative applicative 1135  
 appositional 1002  
 appropriateness principle 1928  
 Äquivalenzdual 1057  
 Arbitraritätsprinzip 1767  
 archiphoneme 336, 492  
 argument 1104, 1130, 1434  
 – argument position 1434  
 – argument structure 845  
 argumentation 1863  
 Aristotelian classification 264  
 articulation, double 1232  
 Artikel/article 674, 717, 783, 1752  
 – Artikelflexion 1692  
 – bestimmter Artikel/definite article 791, 1315, 1700  
 – demonstrative article 792  
 – indefinite article 792, 1700, 1748  
 – interrogative article 793  
 – partitive article 1069  
 – possessive article 792  
 Aspekt/aspect 1165, 1302, 1509, 1702, 1746  
 – Aristotelian aspect 1172  
 – aspect marker 1536, 1743  
 – aspectual derivation 1402  
 – aspectual system 1414  
 – Aspektform 1274  
 – Aspektstamm 1319  
 – Aspektualität/aspectuality 1165, 1569  
 – grammatical aspect 1166  
 – lexical aspect 1166  
 – phasal aspect 1172  
 – situation aspect 1166  
 – successive aspect 1737  
 assimilation 1390  
 assistive 1138  
 atelic 1168  
 Attraktion 1913f.  
 attribution 1655  
 – attribute 1112  
 – attributiv/attributive 1113, 1301  
 Aufbauprozeß 1645  
 Augenblicksbildung (s. Einzelprägung)

augment 1529  
 augmentative 743, 1045f., 1049  
 Ausgleichsprozeß 1688  
 austauschbar 235  
 autolexical syntax 204  
 autonomous action/state/event 1346  
 autosegmental analysis 189, 554  
 Autosemantika 1565  
 auxiliary 1002, 1199, 1595  
 auxiliare 809  
 aversive 1086  
**B**  
 back formation/derivation (s. Rückbildung)  
 background 1170, 1174  
 bahuvrihi compound 1631, 1675  
 Basisform/base form/basic form 336, 464, 489, 499, 604, 865, 932  
 – base form inflection 604  
 – morphological structure of the base 1046  
 – postbase 1468  
 beat addition 1539  
 beat deletion 1539  
 beat movement 1539  
 Bedeutung/meaning 404, 1278  
 – basic meaning 607  
 – Bildungsbedeutung 1913  
 – derivational meaning 524  
 – Gebrauchsbedeutung 1913  
 – general meaning 607  
 – Gesamtbedeutung 259, 1076, 1841  
 – grammatical meaning 257, 524  
 – Grundbedeutung 259  
 – lexical meaning 257  
 – Sachbedeutung 677  
 – systematische Bedeutung 1911  
 – verallgemeinerte abstrahierte Bedeutung 677  
 bedingt segmentierbar 419  
 Bedingung 235  
 Beibehaltung der Abteilungsrichtung 1608  
 beneficiary 1122  
 Benennungsmittel 1912  
 Benennungsmotiv 1912  
 Bestimmungswort 1275  
 Bewegung/motion/movement 736, 1469, 1557  
 Bewegungsverb 1557  
 Bezeichnungsmotiv 1912  
 Bildungsbedeutung 1913  
 bilingual dictionary 1885  
 bilinguale Konzept 1931  
 Bilingualismus/bilingualism 1647  
 bioprogrammatic 1764  
 bleaching 813, 1593  
 bleeding 505  
 blending 957, 1256, 1678  
 blocking 308, 653, 848, 933, 1642  
 blocking, morphological 347  
 body-part tally system 1504  
 borrowing (s. Entlehnung)  
 bottom-up progression 1872  
 boundary 377f.  
 boundedness/bounding 1068, 1166, 1168  
 – bound 1864  
 – obligatorily bound 406  
 – bound pronoun 783  
 – unbounded 1469  
 bracket erasure convention 411  
 bracketing paradox 413, 841  
 broad phonetic transcription 1875  
 Broca-Aphasia 1817  
**C**  
 calque (s. Lehnübersetzung)  
 case (s. Kasus)  
 category 261, 264, 1865  
 category, derivational 265  
 causal 1086  
 causation 1427  
 causative 1418, 1449, 1485, 1718, 1728, 1733, 1742, 1762f.  
 – causative suffix 1370  
 – direct causative 1138  
 – immediate causative 1138  
 – indirect causative 1138  
 cause 1122  
 change of state 1168  
 change, analogical 1611  
 change, morphological 1574  
 change, sound 1611  
 child-etymology 1621  
 circumfix (s. Zirkumfix)  
 circumfixation (s. Zirkumfigierung)  
 circumstantial 1117  
 citation form (s. Zitierform)  
 class (s. Klasse)  
 classification of lexemes 1864  
 classifier (s. Klassifikator)  
 clause linkage 1514  
 clause  
 – clause-chaining 1501, 1592, 1595  
 – implicated clause 1497  
 – lest clause 1498  
 – sequential clause 1498  
 cleft-sentence 1872  
 climbing 390  
 clipping 956, 1256, 1678

- clitic 253, 719, 1002, 1264, 1465
  - clitic boundaries 1852
  - clitic doubling 391
  - cliticization 1594
  - simple clitic 393
  - special clitic 393
- coercion 1179
- cognitive domain 1870
- cohesion 389
  - cohesion, internal 252
- coinage 835, 966, 1801
  - analogical coining 165
  - creative coining 1576
  - expressive coinage 963
  - unusual coinage 963
- collection 1071
- collective 737, 747, 1071, 1706, 1736, 1749, 1752
  - collective count noun 1068
  - collective mass noun 1068
  - collective number 1071
  - collective numeral 1071
- combining form 895
- comitative 1085, 1122
- comment 1434
- commercial name 966
- communicative domain 1870
- comparative (s. Komparation)
- comparative method 1661
- comparison (s. Komparation)
- complementary 1115
- completed (form of the verb) 1440
- completive 1172
- complexity 1796
- complexity, level of 1872
- component of a word form 523
- component, lexical 345
- component, long 189
- composition (s. Komposition)
- compositionality 519, 851, 1286, 1288
- compositionality principle 840
- compound (s. Komposition)
- compression 300, 1927
- conceptualisation/construal 259
- concession 1122
- concomitant 1085
- concord, grammatical 1799
- conditional (s. konditional)
- configuration (of entities) 1017
- fixation 528
- conflation, inclusive 1010
- conjugation 631
- connectionist view 1807
- conservative mutation 1595
- consistency (of entities) 1017
- consonant (s. Konsonant)
- constituent, immediate 407
- constituent, prosodic 413
- construct state 362
- construction
  - coverb construction 1393
  - experiential construction 1187
  - mensurative construction 1072
  - periphrastic construction 1199, 1595
  - pseudopartitive construction 1072
  - contamination 953
  - content (s. Significatum)
  - contentive 249
  - continuative 1725, 1734
  - continuous 1725
  - contraintes 1289
    - contraintes de disponibilité 1289
    - contraintes historiques 1290
    - contraintes phonologiques 1289
    - contraintes sélectionnelles 1289
    - contraintes sémantiques 1289
    - contraintes structurelles 1289
  - conventionalization 1577
  - converb 362
  - convergence (diachronic process) 1741
  - conversion (s. Konversion)
  - coordination 1202
  - co-ordinator 1492
  - copula 718, 811, 1737
  - coreference 1000, 1465
  - corepresentable linguistic units 511
  - corpus 1860, 1869
  - cranberry morph 893
  - creativity, morphological 164
  - creativity, rule-changing 1576
  - creativity, rule-governed 1576
  - creole 1653, 1761
  - creolisation 1752, 1755, 1759–1761
  - culture-specific 1050
  - cumulation 812
  - cumulative reference condition 1068
  - cyclic rules 340
  - cyclicity 340
- D
- data 1860
- data source 1461
- Dativ/dative 1092, 1494, 1862
- dative, ethical 1002
- deaccusative 1132
- Deadjektiva 1566
- declarative 1191
- declension (s. Deklination)
- decompositional hypothesis 1778
- de-compression 301
- defective paradigm 1067
- defectivity 602, 745
- defektiv segmentierbar 419
- definite 1434
- definiteness 986, 1413, 1747
- definition 1863
- deglutination 1615
- degree 1115
- deixis 983, 1120, 1724
  - deictic base 993
  - deictic center 999
  - deictic infix 993
  - deictic prefix 993
  - deictic reduplication 994
  - deictic suffix 993
  - far deixis 1751
  - middle deixis 1751
  - near deixis 1751
  - vertical deixis 1011
- Deklination/declension 631, 733, 1668
- Hauptdeklinationsklasse 1305
- Deklinationsklasse/declension class 735 f., 1316
- Indeklinabilitum 1305
- Dekomposition 1771
- Dekompositionshypothese 1768
- demonstrative 1011, 1391, 1433 f., 1701
- demotivation 1615
- denominative 1718
- deobjective 1131
- dependence, conceptual 1422
- dependency 754
- dependency stemmata 1879
- dependent marking 1122, 1491
- depersonalization 1010
- deponent 1139
- deprecatve 971
- deproninalization 1010
- Derivation/Ableitung/derivation 344, 499, 595, 887, 1199, 1275, 1325, 1416, 1467, 1567, 1693, 1867, 1889
  - abgeleitetes Wort 1546
  - Ableitungsmorphem 1236
  - adjectival derivate 1708
  - affixal derivation 1258
  - aspectual derivation 1402
  - derivateme 524
  - derivational affix 1801
  - derivational category 265
  - derivational meaning 524
  - derivational morphology 370, 833, 1248, 1426, 1801, 1867
  - derivational suppletion 515
  - derivational type 266
  - Derivationsformant 1546
  - Derivationstiefe 1307
  - derivative 932
  - derived form 604
  - direction of derivation 866

- explizite Derivation/explicit derivation 890, 932, 1566
- implizite Ableitung/implizite Derivation/implicit derivation 893, 932, 1275
- kombinatorische Derivation 890
- Konfixderivat 895
- L-derivation 266
- lexical derivation 266
- multiple derivation 1762
- Nullableitung/zero derivation 444, 932, 1275, 1531, 1660, 1801
- prefix derivation 1531
- primary derivation 1417
- proper derivation 932
- Pseudoderivat 894
- secondary derivation 1417
- semantic derivative 931
- syntactic derivative 931
- taddhita derivate 1717
- verbal derivational suffix 1459
- description 1863, 1869
  - descriptive concept 1862
  - descriptive device 1875
  - descriptive grammar 1860
- desiderative 1718
- Determination
  - Determinans-Determinatum-Struktur 887
  - determinant 931
  - Determinativkompositum 1275
  - determinatum 931
- diachronic analysis 1591
- diagrammaticity 867
- dialectale Gliederung 1267
- Diathese/diathesis 1140, 1146, 1318
- dictionary 1874
  - dictionary entry 1886
  - learners' dictionary 1884
  - monolingual dictionary 1882
- Diffusion 1647
- diglossia 1741
- diminutive 743, 1045, 1048, 1173, 1413, 1586, 1737
- diminutive tantum 1046
- direct discourse 1001
- directional marker 1438
- directionality, principle of 866
- discontinuity, morphological 546
- discourse role 988
- disguise 1010
- disharmonic 1009
- Distanz (des Sprechers vom Sachverhalt) 1274
- Distanzkompositum 1277
- Distanzstellung 889
- distinctive unit 493
- Distribution 676

- distribution, non-contrastive 185
- distributive reference condition 1068
- ditransitive 1434
- documentation of the language 1869
- domain, abstract 260
- domain, concrete 260
- Dual/dual 1056, 1465, 1743, 1749
  - Äquivalenzdual 1057
  - arbiträrer Dual 1057
  - Dualia Tantum 1061
  - Oppositsdual 1057
  - Paritätsdual 1056
  - pseudo-dual 1749
- durative 1168 f., 1171, 1734
- durativity 1746
- Durchsichtigkeit 1774
- dyadic kin term 1493
- Dyslexie 1769
- E
- eclecticism 1858
- economy (s. Sprachökonomie)
- Eigenname 1901
- Einzelprägung/Augenblicksbildung/ad hoc formation/hapax legomenon 1627
- elative 1213
- ellipsis 962
- empathy hierarchy (s. animacy hierarchy)
- endoclitic 392, 546
- Endungsreihe 1244
- Enklise 1271
  - enclitic 392, 1002, 1396, 1456, 1465
  - enclitic presyllable 1734
  - locative enclitic 1536
- Entfaltung 1912
- entity 733, 973
  - first-order entity 974
  - fourth-order entity 974
  - second-order entity 974
  - third-order entity 974
- Entlehnung/borrowing 1586, 1644, 1646, 1656
- entrapment 549, 937
- Entwicklungs dysphasie 1820
- environment 614
- equative form 1213
- equativity 1218
- ergative 1081
- ergative language 1081, 1493 f.
- ergative morphology 1391
- ergative system 813, 1081
- Erleichterungsrückbildung 894
- error 1789, 1796
- etabliert 1091
- Etymologie/etymology
  - étymologie-histoire du mot 1911
  - étymologie-origine 1911
  - folk etymology 1587, 1615
  - popular etymology 1615
- evaluation, domain of 1193
- evaluation, target of 1192
- evalutive 1048
- evalutive morphology 361
- event 1165, 1168
- event point 1182
- evitative 1086
- example 1880
- Expansion/expansion 301, 887, 931
- expletive 970
- explicative 1048
- exponence
  - cumulative exponence 619, 1074
  - exponent 616
  - extended exponence 622
  - fused exponence 621
  - main exponent 623
  - multiple exponence 590, 622, 812
  - separate exponence 618
- expression type 370
- expressive process 524
  - grammatical expressive process 524
  - morphological expressive process 524
  - non-grammatical expressive process 524
  - non-morphological expressive process 524
- extramorphological characteristic 633
- F
- feeding rules relationship 504
- feminin/feminine 1305, 1748
- field structure 611
- finiteness 1334
- flexional language (s. fusionierende Sprache)
- Flexion 676
  - contextual inflection 365, 1806, 1809
  - edge inflection 395
  - first inflection 1796
  - Flexionsabbau 1270
  - Flexionsklasse/inflection class 200, 363, 597, 630, 1269, 1877
  - Flexionsmorphem/inflectional morpheme 370, 1423, 1236
  - Flexionsmorphologie/inflectional morphology 1248, 1600, 1819

- Flexionsprinzip 1685
- Flexionstyp/inflectional type 647, 1269
- general inflectional rules 363
- inflection 174, 344, 595, 1198, 1467, 1867
- inflection class (s. Flexionsklasse)
- inflectional generality 656
- inflectional meaning 524
- inherent inflection 365, 1806
- productive inflectional rule 363
- pronominale Flexion 1095, 1270
- schwache Flexionsklasse 1269 f.
- starke Flexionsklasse 1269 f.
- stem inflection 604
- Trennbarkeit von Kasus- und Numerusflexion 1268
- verbal inflection 1671
- Flexivhomonymie** 1305
- flexivische Profilierung 1684
- fluctuation 486
- focal point 653
- focus 1434
- foreground 1170, 1174
- form (s. Significans)
- formal segmentierbar 419
- formation
  - ad hoc formation (s. Einzelprägung)
  - dynamic formation 1483
  - established formation 1627
  - nonce formation 835
  - playful formation 838
- formative element 172
- Formenlehre** 1683
- frame (s. semantic domain)
- frame-external 1117
- frame-internal 1117
- free form, minimum 252, 379
- free, potentially 406
- frei 240
- Fremdwort/foreign loan** 1268, 1750
- frequency 298, 1252, 1636, 1807
- Fugenelement** 447, 891, 1276, 1683
- full-listing hypothesis 1778
- function (s. Funktion)
- function word 248, 258, 1474, 1538
- function, metasemantic 1049
- functional constraint 1613
- functional domain 1873
- functor 249
- Funktion/function** 3, 1017, 1122, 1467
- funktionaler head 681
- Fusion** 25
- fusionierende Sprache/fusional (inflecting, flectional) lan-
- guage 254, 619, 630, 1222, 1235, 1247
- Futur/future** 1184, 1274, 1303, 1521, 1703, 1746
- future perfect 1187
- future stem 1725
- futurum exactum 1703
- imperfektives Futur 1303
- perfektives Futur 1303
  
- G**
- gap, accidental 834
- gating 1782
- Gebärde** 1556
- Gebärdenraum** 1556
- Gebrauchsbedeutung** 1913
- Gebrauchsgleichheit** 1914
- gemischte Flexionsklasse 1269
- gender (s. Genus)
- general drift 1741
- generality 259
- generalization 1577
- genitive 744, 1514, 1748
  - genitive classifier 1021
  - genitively marked object 1069
  - objective genitive 1078
  - subjective genitive 1078
- genus verbi 1304
  - aktivisch 1304
  - (s. voice)
- Genus/gender** 735, 1007, 1031, 1269, 1304, 1316, 1406, 1412, 1509, 1568, 1698, 1713, 1743
  - controller gender 1033
  - double gender 1033
  - gender assignment 1070
  - gender system 1019
  - Genusprofilierung 1692
  - multiple gender 1033
  - semantic gender resolution 1039
  - syntactic gender resolution 1039
  - target gender 1033
- gerund 1373, 1482
- Gesamtbedeutung** 259, 1076, 1841
- glossary 1875
- goal 1120, 1121, 1135
- government/rection 365
- gradation 1212, 1424
- grammar (s. Grammatik)
- grammatical category label 1842
- grammatical content 373
- grammatical feature 1877
- grammatical function-dependent
  - 346
- grammatical item 1867
- grammatical meaning 257, 524
- grammatical process 523
  
- grammatical relation 1131
- grammatical rule 297
- grammaticalization (s. Grammatikalisierung)
- grammatically corepresentable 512
- grammaticography 1857
- Grammatik/grammar** 1857, 1863
  - categorial grammar 201
  - comparative grammar 1859
  - contrastive grammar 1859
  - general comparative grammar 1859
  - generative Grammatik 1267
  - grammar and dictionary 1870, 1874
  - grammar checker 1897
  - grammar-lexicon interface 1865
  - historical grammar 1859
  - historical-comparative grammar 1859
  - lexical grammar 199
  - lexical-functional grammar 199
  - onomasiological grammar 1873
  - prescriptive grammar 1860
  - reference grammar (s. Sprachlehre)
  - semasiological grammar 1872
  - structure of grammar 1869
  - subdivision of grammar 1873
  - theory of grammar 1858
- Grammatikalisierung/grammaticalization 12, 257, 262, 300, 930, 1232, 1250, 1590, 1736
- grammaticalization of zero 1178
- Grammatikalisierungsprodukt 1273
- grammatische Wechsel 1685
- grammeme 524, 616
- Graphem 1564
- grounding 261
- Grundbedeutung** 259
- Grundform** 1268
- Grundmorphem** 1565
- Grundwort** 1275
- Gruppenflexion 1692
  
- H**
- habitual past 1458
- habitualis/habitual 1520, 1725
- habituality 1746
- Halbaffix 892
- Handlungsnummerus 1058
- hapax legomenon (s. Einzelprägung)
- Haplologie/haplogy** 583, 893
  - haplogenic word 953

– morphological haplogy 392  
 harmonic 1009  
 Hauptdeklinationsklasse 1305  
 Hauptmarker 1609  
 head 201  
 – head class 1033  
 – head movement 368, 398  
 – head operation 201, 861  
 – righthand head rule 201  
 – morphological head 165  
 – head-marking 1122  
 Herkunftsgleichheit 1913  
 hesternal past 1187  
 heteroclysis 738  
 hierarchical asymmetry 1010  
 hierarchy of grammatical levels 1872  
 Homonymie/homonymy 250, 618, 944, 1841, 1914  
 honorific 993, 1724  
 honorific formative 993  
 hortative 1191  
 host 388  
 human 747  
 Humboldtian principle 1575  
 hybrid 1033  
 Hybridisierung 887  
 hypercharacterization 1581, 1587  
 hyphenation, automatic 1897  
 hypostatization 835, 974, 1628, 1640  
 hypotaxis 1202

## I

iambic reversal 1539  
 iconicity (s. Ikonizität)  
 Identitätsoperation 437  
 ideophone 570, 719, 964, 1128, 1538  
 idiom 249  
 idiom suppletion 515  
 idiomatization 837, 930, 1626  
 Ikonismus, konstruktioneller 443  
 Ikonizität/iconicity 618, 867, 1559, 1757  
 illocution 1190  
 – basic illocution 1190  
 – illocutionary force 1050  
 – illocutionary modification 1192  
 Imperativ/imperative 1191, 1209, 1274, 1372, 1521, 1569, 1703, 1716, 1726  
 imperfect 1176, 1715, 1746  
 – imperfect indicative 1702  
 – imperfect subjunctive 1703  
 imperfektiv/imperfektive 1166, 1302, 1481, 1725  
 – derived imperfective 1176  
 – imperfektives Futur 1303

imperative 1191  
 inactive 1425  
 inanimacy 1412  
 inanimate 747  
 inceptive 1172  
 inchoative 1172, 1371, 1737  
 incorporation (s. Inkorporation)  
 indefinite phrase 1434  
 indefiniteness 1742, 1747  
 Indeklinabilium 1305  
 Index 1237  
 index of grammatical items 1875  
 indexical 998  
 indikativ/indicative 1191, 1274, 1303, 1569, 1702  
 – imperfect indicative 1702  
 – pluperfect indicative 1702  
 – present indicative 1702  
 indirect discourse 1001  
 individual (entity type) 1104  
 individual interpretation 1071  
 Individuum 888  
 infinitive 1372, 1458, 1703, 1727  
 infixation 528, 750, 934, 1045, 1330  
 – deictic infix 993  
 – infix 545, 1853  
 – infixal reduplication 545  
 inflection (s. Flexion)  
 information retrieval 1897  
 ingressive 1172  
 Inhalt (s. Significatum)  
 Inhärenz 1091  
 inheritance 166  
 initialism 957  
 initialization 1615  
 Inkorporation/incorporation 284, 527, 889, 916, 1118, 1428, 1451  
 – modifying incorporation 1451  
 – noun incorporation 916  
 – saturating incorporation 1451  
 innate categories and principles 1250  
 instantiation 261  
 institutionalization 837, 1233, 1626  
 instruction 1879  
 instrument 979, 1121  
 instrument, negative 1122  
 intégrateur 1288  
 intensive 1513, 1718  
 interclausal relation 1202  
 Interferenz/interference 1644  
 Interfix 891, 934  
 interfixation 528  
 Interjektion/interjection 1316, 1492  
 interlinear morphemic gloss 1834  
 Interlinguistik 1563

International Auxiliary Language Association 1563  
 interrogative 1191  
 intersegmental 428  
 intonation 1198  
 intransitive 1413, 1434  
 intraparadigmatischer Ausgleich 1690  
 intrasegmental 428  
 invariance 618  
 Invariante 234, 238  
 inverse form of the verb 1010, 1151, 1414  
 Inversion 1608  
 irrealis 1440  
 irreality 1747  
 isolating language 254, 1222, 1247  
 Isomorphie 1694  
 item and arrangement 26, 409, 627  
 item and process 26, 409, 627  
 item familiar 834  
 item familiarity 1627  
 Iteration 1550  
 iterative 1169

## J

jargon 1653, 1757  
 Jargonaphasie 1819  
 juncture 341  
 junggrammatisch 1680  
 Juxtaposition 1093, 1202, 1214  
 Juxtaposatum 1308

## K

Kardinalzahl 1315  
 Kasus/case 1011, 1073, 1304, 1331, 1568, 1595, 1699, 1713, 1747, 1862  
 – abessive 1085  
 – ablative 1494, 1514  
 – accusative 1493 f.  
 – case form 1074  
 – case language 1074  
 – case marker 1074  
 – case marking 737  
 – case morphology 1812  
 – case relation 1074  
 – Dativ/dative 1092, 1494, 1862  
 – direct case 1076, 1080, 1723  
 – ergative 1081  
 – genitive 744, 1514, 1748  
 – independent case 604  
 – indirect case 1080  
 – instrumental 1085, 1514  
 – Kasusnivellierung 1691  
 – Kasussynkretismus 1766

- Kasussystem 1316
  - Lokalkasus 1092
  - nominative 1493 f., 1514
  - oblique case 738, 1076, 1080, 1086, 1723
  - obviative 1005
  - primary case 604
  - secondary derivative case 604
  - Trennbarkeit von Kasus- und Numerusflexion 1268
  - vocative 1086, 1444
  - Kayne's Generalization 398
  - Klammerform 894
  - Klasse/class
    - class-free suffix 1456
    - closed class 249, 1112
    - complementary class 634
    - default class 635
    - Deklinationsklasse/declension class 735 f., 1316
    - distributional class 733
    - form class 677
    - Klassenwechsel 1690
    - macroclass 632, 634
    - morphological class 1507
    - open class 249, 1112
    - unmarked class 635
  - Klassifikator/classifier 735, 1070, 1094, 1467
    - general classifier 1017
    - genitive classifier 1021
    - incorporated classifier 1022
    - mensural classifier 1020, 1070
    - noun classifier 1021
    - numeral classifier 1019
    - possessive classifier 1862
    - sortal classifier 1020, 1070
    - specific classifier 1017
    - typology of classifiers 1018
    - unique classifier 1017
    - verbal classifier 1022
  - klassifikatorischer Schlüssel 1243
  - Klassifizierung 1271
  - Kognitionswissenschaft 1766
  - koine 1741
  - Kombinationsfähigkeit 677
  - Komparation/comparison 1212, 1306, 1569
    - absolute comparative 1213
    - comparative construction 1212
    - comparee/topic of comparison 1212 f.
    - comparison of equality 1213
    - comparison of majority 1213
    - comparison of minority 1213
    - komparativ/comparative 1213, 1257, 1306 f., 1717
    - marker of comparison (s. comparison, pivot of)
    - pivot of comparison/marker of comparison, 1212
  - standard of comparison 1212 f.
  - Kompatibilität 1565
  - Kompositbildung 1307 f.
  - Komposition/composition/com-  
pounding 344, 370, 524, 527, 833, 1262, 1275, 1325, 1416, 1428, 1511, 1567, 1675 f., 1693, 1759, 1801, 1852
    - bahuvrihi compound 1631, 1675
    - composite word/complex word 405
    - compound lexeme 1627
    - compound phrase 889
    - denominales Kompositum 1566 f.
    - deverbales Kompositum 1566
    - Distanzkompositum 1277
    - eigentliches Kompositum 1683
    - endozentrisches Kompositum 444
    - exozentrisches Kompositum/exocentric compound 444, 1276, 1631
    - fossilized compound 1627
    - imperative compound 1631, 1677
    - Kompositum/compound 888, 1308, 1677, 1704, 1717, 1728, 1813, 1890, 1926
    - Konfixkompositum 895
    - Kopulativkompositum 887, 1276
    - phrasal compound 166, 1657
    - Possessivkompositum 1276
    - präpositionales Kompositum 889
    - Pseudokompositum 890
    - stub compound 1256
    - syntactic compound 1627
    - synthetic compound 166, 891
    - verbales Pseudokompositum 1277
    - Verbalkompositum 1277
    - Verb-Ergänzung-Kompositum 890
  - Kompositionalität, synchrone 1273
  - conditional/conditional 1204, 1458, 1569, 1703, 1715
  - Konditionierung 239
    - distributionelle phonologische Konditionierung 463
    - formale phonologische Konditionierung 463
    - grammatische Konditionierung 474
    - lexikalische Konditionierung 474
    - morphologische Konditionierung 474
  - phonologische Konditionie-  
rung 463
  - primäre phonologische Kon-  
ditionierung 465
  - reziproke Konditionierung 475
  - sekundäre phonologische Konditionierung 465
  - semantische Konditionierung 477
  - syntaktische Konditionierung 480
  - Konfix 893, 895, 1276
  - Konfixderivat 895
  - Konfixkompositum 895
  - Kongruenz/agreement 365, 1002, 1270, 1514, 1725, 1799, 1811
  - Konjugation 1569
  - Konjunktion 675
  - Konjunktiv 1274, 1303, 1569
  - Konnektionismus 1766
  - Konnektor 1093
  - konsekutiv 1521
  - Konsonant/consonant
    - appendix consonant 384
    - consonant gradation 568, 1337
    - consonant mutation (s. kon-  
sonantische Veränderung)
    - konsonantische Veränderung/  
consonant mutation 568, 1272
  - Kontamination 1819
  - Konversion/conversion 27, 165, 527, 932, 944, 1275, 1660, 1709, 1759, 1927
    - categorical conversion 530
    - paradigmatic conversion 530
    - principe d'orientation des conversions 871
    - rectional conversion 530
    - zero conversion 531
  - Kopiesprache (replica language) 1646
  - Kopulativkompositum 887, 1276
  - Körperteil 1557
  - Korrelationskoeffizient 1242
  - Kreolsprache 1646
  - Kunstsprache, internationale 1563
  - Kürzungssuffigierung 894
  - Kurzwort/truncation 529, 583, 958, 1286, 1289, 1295, 1390, 1537
- L**
- language death 1653, 1655
  - language external 1043
  - language internal 1042
  - language island 1741
  - language processing 1251

language system 1870  
 language type 1224  
 language use 253  
 langue analogue 1222  
 langue transpositive 1222  
 lapse (s. word, portmanteau)  
 Lautsymbolik 418  
 Lautverschiebung, althochdeutsche 1681  
 Lautverschiebung, zweite 1681  
 Lautwechsel 1685  
 leaner 391  
 lect-marking 307  
 Lehnübersetzung/loan translation/calque 833, 1645, 1649  
 Lehnwortbildung 895  
 lemmatization 1896  
 Lernen, deduktives 1933  
 Lernen, induktives 1932  
 letter try 1894  
 level 189  
 level ordering 165  
 leveling 1587  
 – paradigm leveling 1581  
 – analogical levelling 1613  
 lexeme 152, 248, 596, 868, 1106  
 – lexeme-class 596  
 – lexemic domain hypothesis 197  
 – phrasal lexeme 249  
 lexical category, major 196  
 lexical content 373  
 lexical coverage 1897  
 lexical decision task 1781  
 lexical enrichment 306  
 lexical entries, fully-specified 346  
 lexical entries, impoverished 346  
 lexical expansion 1656  
 lexical expression 370  
 lexical formative 345  
 lexical generality 1251  
 lexical hypothesis, generalized 199  
 lexical hypothesis, strong 199  
 lexical hypothesis, weak 199  
 lexical insertion 196  
 lexical integrity 197, 397  
 lexical item 250, 596  
 lexical meaning 257  
 lexical representation 346  
 lexical retrieval 1807  
 lexical rule (s. *lexikalische Regel*)  
 lexical split 364  
 lexical unit 250, 596  
 lexicalism, strong 197, 346, 367  
 lexicalism, weak 197, 367  
 lexicalist hypothesis 165, 188, 195  
 lexicalization 300, 346, 524, 834, 837, 1577, 1597, 1625 f., 1641  
 lexicalization, synchronic 933  
 lexicography 1857, 1882

lexicon, mental 364  
 lexicon, permanent 345  
 lexicon, potential 345  
 lexifier language 1757 f.  
 lexikalische Entscheidung 1770  
 lexikalische Regel/lexical rule 201, 1767  
 liaison 393  
 lingua franca 1752, 1758  
 linguistic type 1873  
 linguistic union 511, 524  
 link, empty 183  
 linkage, asyndetic 1202  
 linkage, tail-head 1506  
 list fallacy 297  
 listeme 249, 596  
 loan translation (s. *Lehnübersetzung*)  
 loan word 833, 1528  
 loan, foreign (s. *Fremdwort*)  
 local 1081, 1083  
 localistic 1076  
 location 1413, 1469  
 locative 1433, 1445, 1494, 1514  
 logographisch 1916  
 logophoric 1005  
 logophoric, second-level 1006  
 Lokalkasus 1092

**M**

macroclass 632, 634  
 macroparadigm 632, 634  
 malapropism 1615, 1790  
 manner 1121  
 mark 492, 608  
 – Hauptmarker 1609  
 – marker paradigm 1878  
 – Marker/marker 10, 616  
 – marking, double 1491  
 masdar 1750  
 masking 1010  
 maskulin 1305  
 material (semantic feature) 1016  
 materielle Gleichheit 1914  
 meaning (s. *Bedeutung*)  
 means 1121  
 media tantum 1150  
 megamorph 511  
 megamorph suppletion 515  
 Meillet principle 1664  
 mensurative 747  
 metagrammaticography 1858 f.  
 meta-lexicography 1857  
 metanalysis 1615  
 metaphorony 567, 751  
 metaphor, grammatical 977  
 metaphorical extension 1596  
 Metathese/metathesis 467, 549, 576, 936, 1390, 1465  
 – contact metathesis 577  
 – morph metathesis 397  
 – morphological metathesis 576  
 – morphophonological metathesis 577  
 methodological reflection 1869  
 metonymy 1631  
 Minimalzeichen/minimal sign 406, 417, 511  
 mirror principle 407  
 Mischsprachen/mixed languages 1646  
 mitigation 1192  
 mixed languages (s. *Mischsprachen*)  
 mixed processing 1778  
 modality 1190  
 – deontic modality 1193  
 – doxastic modality 1196  
 – dubitative modality 1196  
 – epistemic modality 1193  
 – event-oriented modality 1193  
 – evidential modality 1193, 1197, 1414  
 – facultative modality 1193  
 – hypothetical modality 1196  
 – inferential modality 814, 1197  
 – modal distinction 1087, 1414, 1743  
 – modal function  
 – modal suffix 1459  
 – reportative modality 1197  
 Modalverb 1274  
 model, process 572  
 Modellsprache/source language/model language 1646  
 Modifikation/modification 527, 888, 938, 1451  
 – internal modification 751, 932, 935  
 – modifikatorische morphologische Regel 1607  
 – modifying incorporation 1451  
 – nominal modifier 1511  
 module 1790  
 Modus/mood 1190, 1198, 1303, 1318, 1509, 1673, 1703  
 – superordinate clause mood 1392  
 – subordinate clause mood 1392  
 Monoflexion, kooperierende 1270  
 monogenetic origin 1741  
 monolingual Konzept 1930  
 monosystemic 494  
 Morph/morph 1225  
 – cranberry morph 893  
 – diskontinuierliches Morph 1270, 1275  
 – empty morph 183, 626  
 – leeres Morph 447

- morph metathesis 397
- Quasimorph 419
- replaceive morph 568, 573
- Submorph 417
- zero morph 1852
- Morphe/morpheme 22, 1225, 1564, 1852, 1864
- autonomous morpheme 258
- bound morpheme 1864
- dependent morpheme 257
- Flexionsmorphem/inflectional morpheme 370, 1423, 1236
- free grammatical morpheme 371
- freies Morphem 1062
- grammaticches Morphem 1302
- morpheme boundary 1852
- morpheme structure constraint 429
- morpheme structure rule 428
- morpheme-based 197
- morpheme-oriented morphology 1864
- Nullmorphem/zero morpheme 438, 1268, 1675
- obligatorily bound morpheme 406
- Quasimorphem 893
- submorphemic 1045
- unbounded morpheme 1469
- morphemically irrelevant material 183
- morphemic 1867, 1874
- morpholexical alternant 176
- morpholexical variation 337
- morphological asymmetry 842
- morphological description 1863
- morphological domain 307
- morphological expression 370
- morphological means 523
- morphological model, dynamic 489, 499
- morphological model, static 490
- morphological network 164
- morphological operation (s. morphologische Operation) 1865
- morphological process, zero 531
- morphological realization 200
- morphological rule, productive 298
- morphological spelling operation 266
- morphological structure 1864
- Morphologie/morphology 4
  - autosegmental morphology 573
  - derivational morphology 370, 833, 1248, 1426, 1801, 1867
  - finite-state morphology 1894
  - inflectional morphology (s. Flexionsmorphologie)
  - level-ordered morphology 367
  - lexeme-based morphology 197, 868
  - lexical morphology 165, 411, 833
  - model of morphology 1886
  - morphology-syntax interface 1594, 1865, 1871
  - nominal morphology 1719
  - paradigmatic morphology 164
  - position-class morphology 412
  - prosodic morphology 342
  - split morphology 364, 366
- morphologisch determiniert 1300
- morphologische Operation/morphological operation 25, 1255
- morphologische Regel 1605
  - morphologische Regel, additive 1607
- morphologisches Prinzip 1686
- Morphologisierung/morphologization 13, 936, 1582, 1584, 1588, 1594, 1600, 1685
- morphologisch initiierte Morphologisierung 1605
- Parameter der Morphologisierung 1604
- phonologisch initiierte Morphologisierung 1605
- morphology (s. Morphologie)
- Morphophonemik/morphophonemics/morphophonology/morphophonology 335, 337, 470, 485
- morphonologische Anpassung 893
- morphonologische Anpassungsregel 891
- morphophoneme 490, 1225
- morphophonemic level 490
- morphophonemic representation 490, 1875
- morphophonemic rule 497
- morphophonology, generative 496
- morphosyntax 204
  - morphosyntactic category 266
  - morphosyntactic feature 266
  - morphosyntactic property 266, 616
  - morphosyntactic representation 346
- motion (s. Bewegung)
- motivation 1122, 1633
- movement, bounded 1469
- Movierung 444
- MS-operation 266
- multifunctionality 1660
- multisegmental 894
- Mutation/mutation 888, 1670
- N
- nahe Zukunft 1521
- naturalness 1253
- Nebenmarker 1609
- negation 1207, 1374
  - constituent negation 1210
  - narrow scope negation 1210
  - sentential negation 1210
  - wide scope negation 1210
- neologism/neology 835, 1861
- Neologismus, morphologischer 1819
- neue Wurzel 1275
- neutral 1305
- neutralisation 485, 493, 497, 1210
  - neutralization, absolute 189
- Nomen/noun 674, 733, 746, 1114, 1257, 1391, 1433, 1466, 1528, 1668
  - abstract noun 747
  - action noun 1708, 1721, 1727
  - agent noun 1707, 1721, 1738
  - collective count noun 1068
  - collective mass noun 1068
  - common noun 746
  - compound noun 1450, 1531
  - concrete noun 747
  - corporate noun 1071
  - count noun 747, 1067
  - deadjectival abstract noun 1708
  - deverbal noun 1536, 1707
  - expressive noun 1128
  - inalienable noun 740
  - instrumental noun 1707
  - locative noun 1707
  - mass noun 747, 1067
  - Nomina-agentis Bildung 1684
  - noun of multitude 737
  - noun stripping 920
  - nouniness squish 976
  - personal noun 1707
  - possessed noun 1434
  - proper noun 746
  - relationales Nomen/relational noun 1098, 1595
  - relator noun 1077
  - short noun 1441
  - unit noun (s. Singulativum)
  - verbal noun 748, 1750
- nominal 1114, 1329, 1412, 1492
  - derived nominal 196, 1716
  - gerundive nominal 196
  - location nominal 1494

- nominal modifier 1511
- nominal system 1713
- nominalization 747, 939, 974, 1260
  - nominalizer 1737
  - future nominalizer 1458
  - stative nominalizer 1458
  - generalized nominalizing markers 978
- Nominalklasse 477
- nominative 1493 f., 1514
- non-continuous 1725
- non-human 747
- non-pro-drop 1002
- non-real 1481
- non-sensory evidence 1197
- non-suppletive allomorphy 336
- norm 834
- noun (s. Nomen)
  - noun class 735, 1007, 1031, 1367
    - noun class system 1019
    - noun classification 1528
  - Nullableitung/zero derivation 444, 932, 1275, 1531, 1660, 1801
  - Nullallomorph/zero allomorph 183, 438
  - Nullausdruck 27
  - Nullmorphem/zero morpheme 438, 1268, 1675
  - Nulloperation 437
  - Nullsuffix 894
  - Nullzeichen/zero sign 438, 531
  - Numerals (s. Zahl)
  - Numerus/number 1304, 1316, 1331, 1371, 1508, 1568, 1673, 1699, 1702, 1713, 1747
    - Trennbarkeit von Kasus- und Numerusflexion 1268
    - distributive number 1071, 1734
    - mass number 1069
    - Numerusablaut 1680
    - Numerusprofilierung 1691
    - Handlungsnummer 1058
    - collective number 1071
  - nunation 1747

## O

- object 1009, 1405
  - genitively marked object 1069
  - indirect pronominal object 1747
  - morphological object 596
  - object marking 1508
  - object pronoun 1533
- objective conceptualisation 1346
- obligatory contour principle 1613

- obligatory elements in the sentence 1118
- oblique form 604
- obscurance 1615, 1627
- obsolescence 1655
- obstruent, initial 1732
- obviation 1413
- obviative 1005
- omission 1010
- omnipredicative language 1433
- one-form-one-meaning principle 840
- onomasiology 1870
  - onomasiologische Basis 887
  - onomasiologisches Merkmal 888
- onomatopoetic formation 965
- opaque 1216
- opaqueness 1615
- opérateur de construction de sens 1290
- Opposition 234
  - bilateral opposition 492
  - constant opposition 492
  - exclusive opposition 493
  - phonological opposition 492
- Oppositionsdual 1057
- optativ/optative 1191, 1440, 1458, 1521, 1716
- optimality theory 1613
- optional elements in the sentence 1118
- order
  - order of morphemes 1249
  - conjunct order 1416
  - independent order 1416
  - linear order 1010
  - scope order 1880
  - sequential order 1880
  - template order 1880
- Ordinalzahl 1315
- ordinary 1724
- orientation 1481
- orthographic representation 1875
- overgeneralization 1807
- overt analogue criterion 363, 443, 871

## P

- Paradigma/paradigm 27, 171, 596, 630, 808, 1867, 1878
  - defective paradigm 1067
  - dominant paradigm structure condition 635
  - implicative paradigm structure condition 635
  - inflectional paradigm 647
  - intégrateur paradigmatischer 1289, 1294

- intraparadigmatischer Ausgleich 1690
- macroparadigm 632, 634
- marker paradigm 1878
- paradigm cell 596
- paradigm economy 635
- paradigm leveling 1581
- paradigm symmetry 656
- paradigmatic conversion 530
- paradigmatic deviation 618
- paradigmatic gap 363
- paradigmatic morphology 164
- paradigmatic relation 494
- paradigmatic uniformity 1613
- potential paradigm 808
- subparadigm/slub/screeve 597, 633
- specimen paradigm 1878
- Paragrammatismus 1817 f.
- parameter 1212 f.
- parameter marker 1213
- Paraphasie 1819
- paraphrase 1001
- parasitic formation 526
- parasitic word 605
- Parasyntthese/parasyntesis 890, 1706
- parataxis 1202
- Paritätsdual 1056
- parole 834
- parser 1893
- part of speech 938, 1233, 1864
- participant 1146
  - participant role 998
  - participant-oriented modality 1192
- participle (s. Partizip)
- particle (s. Partikel)
- particularism 1858, 1862
- Partikel/particle 258, 1198, 1316, 1330, 1412, 1433, 1492
- partitive 1048, 1085
- Partizip/participle 674, 1362, 1373, 1441, 1569, 1703, 1747
  - attributive participle 1728
  - past participle 1671
  - present participle 1672
- passé (s. past)
- Passiv/passive 1414, 1448, 1703, 1727, 1742
  - *bekommen*-Passiv 1273
  - eigentliches Passiv 1304
  - internal passive 1746
  - passivische Verwendung der Verben 1304
- past 1182
  - habitual past 1458
  - hesternal past 1187
  - passé antérieur 1703
  - passé composé 1702
  - passé simple 1702

- passé surcomposé 1703
- recent past/hodiernal past 1186 f.
- remote past 1187
- path 1120 f.
- patient 979, 1081, 1146
- pattern 173
- Paukal 1057
- pejorative 1413
- perception, multiple 1346
- perception, single 1346
- percolation 860
- percolation, feature 200
- perfect (s. Perfekt)
- perfective (s. perfektiv)
- perfectivity 1166, 1168, 1746
- Perfekt/perfect 1177, 1185, 1274, 1414, 1521, 1702, 1715, 1732
- doppelte Perfektform 1273
- future perfect 1187
- intransitive de-resultative perfect 1186
- perfect stem 1715
- perfect subjunctive 1703
- possessive de-resultative perfect 1186
- perfektiv/perfective 1166, 1302, 1481, 1568, 1725, 1734, 1736, 1747
- perfektives Futur 1303
- peripherality constraint 478
- Periphrase/periphrasis 750, 1684
  - periphrastic form 600, 808
  - periphrastische Konstruktion, grammatisierte 1273
- permissive 1138
- permutation 529
- person 998, 1304, 1317, 1508, 1673, 1702
  - active personal prefix 1425
  - consciousness model person 998
  - first plural exclusive person 986, 1457, 1465
  - first plural inclusive person 986, 1457, 1465
  - fourth person 1005
  - Personalpronomen/personal pronoun 716, 784 f., 1270, 1369, 1412, 1701
  - second person pronoun 986
  - third person pronoun 986
- perspective 999, 1174
- phase 1172
- Phonem/phoneme 22, 1225
- phonematic unit 149, 495
- phonemic representation 1875
- phonestheme 953
- phonetic representation 1875
- phonetic rule 497
- Phonologie/phonology
  - automatic phonological rule 336
  - autosegmental phonology 496
  - generative phonology 499
  - lexical phonology 346, 411
  - natural generative phonology 496
  - phonological restriction 1047
  - phonologisch determiniert 1300
  - phonologische Regel/phono-logical rule 428, 499, 1605
  - prosodic phonology 494
- phonosymbolism 963
- phonotactics 427
- phrase marking 1491
- phrase-grammar 201
- phraseme 249
- pidgin 1653, 1757, 1759, 1761, 1764
- pidginization 1752
- place 1121
- Plansprache 1563
- pluperfect (s. Plusquamperfekt)
  - pluperfect indicative 1702
  - pluperfect subjunctive 1703
- Plural/plural 1055, 1234, 1465, 1736, 1749
  - Abundanzplural 1057
  - kollektiver Plural 1056
  - assoziativer Plural 1056
  - broken plural 750, 1749
  - distributiver Plural 1056
  - indefinite plural 1736
  - Plurale Tantum/plurale tan-tum 364, 737, 1061, 1969
  - plurality 1486
  - sortaler Plural/sortal plural 747, 1056
  - sound plural 1749
- Plusquamperfekt/pluperfect/past perfect 1187, 1274, 1702 f.
- Plusquamperfektform, doppelte 1273
- point 1169
- point state 1169
- Polyfunktionalität/polyfunction- ality 238, 841 f.
- Polymorphie 238
- Polysemie/polysemy 238, 250, 259, 526, 638, 842, 1841
- polysynthetic language 254, 916, 1247, 1389, 1465
- polysynthetic structure 1223
- polysystemic 150, 494
- polysystemicity 150
- Poniästhem 418
- Portmanteau/portmanteau
  - portmanteau affix 1455
  - portmanteau noun 1010
  - portmanteau word/lapse 953
  - Portmanteau-Morph (port-manteau morph) 183, 439, 573, 622
- Position 676
- Positiv/positive 1213, 1306
- Possession/possession 743, 1091, 1413
  - alienable possession 1748
  - attributive Possessivkon- struktion 1093
  - inalienable possession 1748
  - inhärente Possession 1091
  - Possessivaffix 1099
  - possessive adjective 743
  - possessive case 1085
  - possessives Prädikat 1091
  - Possessivkompositum 1276
  - Possessivsuffix 1360
  - Possessor/possessor 1091, 1442
  - Possessum 1091
- postbase 1468
- potential 1440, 1458
- potential deobjective 1132
- potentive 1483
- Prädikation/predication 1104 f.
  - possessives Prädikat 1091
  - predicate 1104, 1106, 1112
  - predicate determiner 1492
  - predicate, basic 1400
- prädikativ/predicative 1301, 1433
- Präfigierung/prefixation 528, 934, 1045, 1260, 1308, 1320, 1685
  - cross-referencing prefix 1425
  - deictic prefix 993
  - Halbpräfix 1276
  - Präfix/prefix 535, 1062, 1218, 1236, 1275, 1465, 1674 f., 1677, 1705, 1926
  - Präfixoid 1276
  - Präfixtausch 893
  - prefix-stripping 1779
  - Pseudopräfix 1770
- pragmatic 999
- pragmatic effect 963
- pragmatische Frequenzuntersu- chung 1236
- Präposition/preposition 675, 1511, 1737
- Präsens 1274, 1303
- Präteritopräsentien 1685
- Präteritum 1521
- Präverb/preverb 892, 1135, 1277
- precative 1716
- precision (reliability measure- ment) 1897
- precision principle 1928
- predicable 1433
- predication (s. Prädikation)
- predicative (s. prädikativ)
- prefix (s. Präfix)
- present 1185, 1372
  - present in the past 1175
  - present indicative 1702

- present subjunctive 1703
- presentation of language 1869
- preterite 1372
- preterite participle 1671
- preverb (s. Präverb)
- Primärsuffix 1914
- Priming/priming 1770, 1781
  - cross-modal priming 1781
  - masked priming 1781
  - repetition priming 1781
- privative 492, 1085
- process, diachronic 653
- process, formal 523
- process, morphological 523
- proclitic (s. Proklise)
- pro-drop 1002
- productivity 163, 298, 1577, 1636 f., 1641, 1802
- Profilierung 1684
- pro-form 783
- progressiv/progressive 1521, 1737
- progressive Steuerung 1685
- prohibitive 1191
- Proklise 1271
  - proclitic 392, 1002, 1465, 1532
- Pronomen/pronoun 675, 746, 783, 983, 1099, 1301, 1305, 1313, 1330, 1392, 1465, 1492
  - bound pronoun 783
  - demonstrative pronoun 716, 788, 1370, 1713
  - dummy pronominal 1007
  - free pronoun 783
  - impersonal pronoun 1006
  - incorporated pronoun 922
  - indefinite pronoun 716, 790, 1713
  - independent pronoun 1002
  - interrogative pronoun 790, 1370, 1702
  - intimate pronoun 1724
  - Personalpronomen/personal pronoun 716, 784 f., 1270, 1369, 1412, 1701
  - portmanteau pronoun 1010
  - possessive pronoun 788, 1002, 1701
  - reciprocal pronoun 790
  - reflexive pronoun 789, 1370
  - Relativpronomen/relative pronoun 716, 791, 1270, 1702
- property 1105
  - permanent property 1169
  - property concept 1112
  - temporary property 1169
- propositional content 1104
- proposition-oriented modality 1193
- propositus 1003
- propriete 744, 1085

- prosody 149, 495
  - prosodic analysis 494
  - prosodic contrast 1735
  - prosodic inversion 397
- prototype 611
- prototypical meaning 845
- proximate 1005
- pseudoaffix 931
- Pseudoderivat 894
- pseudo-dual 1749
- punctual 1168, 1169, 1734
- purpose 1122, 1205
- purposive 1861
- Q**
- Quadral 1057
- quadruplication 529
- Quasimorph 419
- Quasimorphem 893
- quasi-representable linguistic unit 511
- R**
- readjustment rule 198
- realis 1481
- Realisation/realization 235, 617
- reanalysis 936
- reason (circumstantial notion) 1122
- recall (reliability measurement) 1897
- recategorialization 1594
- recent past/hodiernal past 1186 f.
- recipient 1122
- reconstruction 1661
- reconstruction, internal 1661
- rection (s. government)
- recursiveness 1389
- reduction 1742
- redundancy rule, lexical 188
- Redundanz/redundancy 271, 346, 427, 1567
- Reduplikation/reduplication 467, 529, 676, 751, 935, 970, 1045, 1061, 1217, 1320, 1415, 1477, 1540, 1549, 1659, 1678, 1735, 1758 f., 1763,
  - deictic reduplication 994
  - expressive reduplication 1440
  - inexact reduplication 558, 1217
  - infixal reduplication 545
  - intensive reduplication 1440
  - partial reduplication 558
  - reduplicative system 1512
  - total reduplication 558
- reference 1122
  - frame of reference 984
  - general reference 1924
  - generalized referent 1000
  - homogeneous reference 1068
  - non-participant referent 998
  - reference point 1182
  - referent conflation 1596
- reference form 631, 636
- reflex (result of a shortening process) 957
- reflexive marker 1437
- Reflexivmedium 1304
- regelmäßig 236
- Regelordnung, extrinsische 471
- regressiv 1685
- regularization 1577 f.
- regulärkompositionell 888
- reification 1426
- reinforcement (s. Verstärkung)
- reinterpretation 1579, 1587, 1615
- relation (type of entity) 1105
- relational 748
- relational form 1413
- relational meaning 1222
- relational rule/via-rule 337, 498
- Relationierung 1098
- Relationsmorphen 1685
- relative clause 1113, 1498
- Relativpronomen/relative pronoun 716, 791, 1270, 1702
- relevance 203, 600, 1140, 1251
- relevant feature 493
- remoteness 1187
- remotivation 1615
- repeater 1038
- replacement 529
- replacive signified 526
- replacive signifier 526
- replica 527
- replica language (s. Kopiersprache)
- replication 527
- representable (linguistic unit) 511
- representation, mental 1788
- resegmentation 1580
- resultative 1133, 1177, 1747
- reversal 1210
- rhematic 1434
- rhotacism 1612
- rhyme 971
- root (s. Wurzel)
- rote-learning 1797
- Rückbildung/back formation/back derivation 532, 583, 932, 937, 957, 1256, 1275, 1579, 1587, 1615, 1638, 1678, 1926
- rule 1893
  - rule change 1575
  - rule creation 1575
  - rule discovery 1575
  - rule fallacy 297
  - rule generalization 1578, 1587

- rule of construal 1067
- rule of referral 617
  
- S**
- Sachbedeutung 677
- salience of the parts 1049
- same subject subordinator 1459
- sandhi 251, 383, 389
- saturation 1434, 1451
- Satzname 894
- schema 633, 1797
- Schriftsystem, Tiefe 1917
- screeve (s. subparadigm)
- secondary motivation 1615
- secretion 1584
- segmental 524
- segmental geometry 574
- Sekundärsuffix 1914
- selected type 261
- selection 182
- semantic domain/frame 260
- semantic field 260
- semantic integrality 1350
- semantic map 1877
- semantic potential 611
- semantic restriction 1047
- semantic shift 1596
- semasiological 1870
- semasiology 1870
- semelfactive 1169
- semi-indirect discourse 1001
- sensory evidence 1197
- sentence type 1191
- sentence, conditional 1725
- sentential marker 1423
- serial verb 1558, 1595
- serial verb construction 815, 1538
- serialization 1592
- serialization, verb 1595
- set interpretation 1071
- setting of the language 1869
- shape (semantic feature) 1016
- sign (s. Zeichen)
- signal 210
- Significans/Form/signifier/significant/form/expression 1, 173, 210, 378, 404 f., 1349
- Significatum/Inhalt/signified/signifié/content 2, 210, 378, 405, 1349
- Silbenschrift 1916
- Silbenwort 894
- similative 1213
- simplicity 1802
- Singular/singular 1055, 1465, 1749
- Singulare Tantum/singulare tantum 737, 1061, 1069
- Singulativum/singulative/unit noun 737, 888, 1055, 1071, 1509
- situated type 261
- situation 611, 1050
- skeletal tier 496
- slab (s. subparadigm)
- slip of tongue 1782
- sonority hierarchy principle 550
- sound symbolism 570
- source 210, 957, 1120 f.
- source, language external 1042
- source, language internal 1042
- source, material 1122
- span 1005
- speaker-and-spouse 1004
- speaker-hearer centric 985
- special language 1924
- special reference 1924
- specified quantity 1171
- speech error 1788
- speech level 1004
- speech point 1182
- speech role 998
- speech situation 984
- speech-act participant 1412
- spelling checking 1897
- spezifisch onymische Morphe 1904
- spezifische Sprachentwicklungsstörung (SSES) 1820
- split-S language 1082
- spoonerism 576
- Sprachbequemlichkeit 1567
- Sprachebene 22
- Sprachentwicklungsauffälligkeit 1816
- Spracherwerbsforschung 1769
- Sprachklassifikation 1241
- Sprachlehre/reference grammar 1860
- Sprachökonomie/economy 1252, 1567, 1928
- Sprachstörung 1769, 1816
- Sprachunterricht 1929
- Sprachverfall/language decay 1646
- Sprachwechsel/language shift 1646
- Sprechfehler 1768
- stabilisation 1757
- Stamm/stem 250, 257, 523, 605, 869, 1687, 1914
  - Aspektstamm 1319
  - discontinuous stem 1853
  - final stem component 1417
  - initial stem component 1416
  - medial stem component 1416
  - oblique stem 1080
  - Stammalternation 1062
  - Stammklasse 1687
  - stem allomorphy 363
  - stem form 363
  - stem-forming element 605, 637
- Standard Average European 1563
- standard marker 1213
- standard theory 196
- standardisation 1927
- Standardsprache 1267
- state 1165, 1168
- state of affairs 1104, 1165
- state, absolute 362
- statistical approach, deductive 1898
- statistical approach, inductive 1898
- stativity 1115
- status constructus 741, 1095
- status suffix 1707
  - honorific status 1003
  - humble status 1003
- stem (s. Stamm)
- stemming 1896
- Stichprobengröße 1237
- stratum 165, 189, 1586
- stress 383, 587
  - phonemic stress 1232
  - stress assignment 1477
  - stress-shift 1658
- strict cyclicity 340
- structural relationship 172
- structure, conceptual 845
- structure/configuration/pattern 173, 494
- Strukturalismus 1267
- subgender 1034
- subjacency 199
- subject 1009, 1405, 1514
  - indirect subject 1085
  - subject index 1875
- subjectification 263
- subjective conceptualisation 1346
- subjunctive 1200, 1862
  - imperfect subjunctive 1703
  - perfect subjunctive 1703
  - pluperfect subjunctive 1703
  - present subjunctive 1703
- sublanguage 1924
- Submorph 417
- submorphemic 1045
- subordination 1202, 1737
- subordinator, different subjects 1459
- subparadigm/slab/screeve 597, 633
- Substantiv/substantive 1114, 1301, 1441
  - Artikelflexion des Substantivs 1270
  - Substantiv, schwaches 1684
- Substitution 676
- Substrat/substratum 1646
- substratal influence 1741, 1763
- substratum 1764
- subsystem 494
- Subtraktion/subtraction 475, 581, 751, 936

- subtractive signified 525
- subtractive signifier 525
- Suffixierung/suffixation 528, 934, 1045, 1307, 1319, 1465
  - class-free suffix 1456
  - cohering suffix 341
  - deictic suffix 993
  - detransitive suffix 1371
  - diminutive suffix 1707
  - directional suffix 1460
  - grammatical suffix 1534
  - Halbsuffix 1276
  - hypocoristic suffix 969
  - independent suffix 1456
  - instrumental suffix 1707
  - interrogative suffix 1374
  - lexical suffix 1534
  - non-cohering suffix 341
  - Nullsuffix 894
  - participial suffix 1439
  - Primärsuffix 1914
  - pronominal suffix 1747
  - Sekundärsuffix 1914
  - sentential suffix 1456
  - suffix stripping 1896
  - Suffix/suffix 535, 1061, 1218, 1236, 1258, 1275, 1477, 1674, 1676 f., 1926
  - Suffixaufnahme 1491
  - Suffixerweiterung 891
  - suffixing preference 941
  - Suffixkombination 891
  - Suffixoid 1276
  - Suffixtausch 893
  - Suffixvariante 891
  - Suffixwechsel 893
  - unconventional suffix 967
  - verbal suffix 1709
  - verbalizing suffix 1371
  - superclassing 1034, 1039, 1041
  - superfix, deictic 993
  - superlativ/superlative 1213, 1257, 1306 f., 1717
    - absolute superlative 1213
    - relative superlative 1213
  - superstratum 1764
  - supervised machine learning algorithm 1898
  - supine 1703
  - Suppletion/suppletion 465, 469, 510, 874, 935, 1062, 1727
    - affixal suppletion 515
    - derivational suppletion 515
    - inflectional suppletion 515
    - morph suppletion 515
    - radical suppletion 515
    - suppletive 512, 605
    - suppletive allomorphy 336
    - suppletive Form 1685
    - suppletive stem 1716
    - Suppletivismus 1302
    - suprafixation 527
      - accentual suprafixation 529

- suprafix 527
- tonal suprafixation 529
- suprasegmental 524, 587
  - suprasegmental apophony 529
  - suprasegmental process 751, 936, 1415
- switch-reference 1008, 1203, 1459, 1501
- syllable 1512, 1731
  - syllable structure constraint 430
  - syllable word 958
- symbolic framework 1806
- symbolization type 370
- Synaffix 891
- synaffixation 934
- synchronic analysis 1591
- Synkretismus/syncretism 363, 602 f., 618, 812, 1300, 1842
- synonym 240
- synonymy 842
- syntactic category 264
- syntactic expression 370
- syntactic reanalysis 1596
- syntactic restriction 1046
- syntagma-interne Beziehung 1565
- syntagmatic deviation 618
- syntagmatic relation 494
- syntaktische Funktion 676
- syntax, lexical 346
- synthesis 176, 1223
  - interrupted synthesis 1400
  - synthetic language 1698
  - synthetic structure 1223
  - synthetisch-analytische Sprache 1301
  - synthetische Bildungsweise 1306
  - synthetische Ersatzkonstruktion 1271
  - synthetische Wortform 1271
  - synthetischer Prozeß 1307
  - synthetischer Zug 1268
  - Synthetismus 1238
- system 173, 494, 614, 834
- system, pronominal 1713, 1720
- systematic gap 834
- systematische Bedeutung 1911

## T

- tactic of sign 379
- tagmemics 179
- takeover 639
- target language (s. Zielsprache)
- telic 1168
- telicity 1166
- temporal relation 1202
- temporality 1501

- Tempus/tense 1181, 1302, 1414, 1673, 1702, 1867
  - absolute tense 1182, 1188
  - perfect tense 1458
  - relative tense 1188
  - sequence-of-tenses 1189
  - sudden discovery tense 1458
  - Tempusform 1274
  - tempusneutrale Form 1274
  - tense logic 1182
  - tense stem 1715
  - tense-aspect-modality 1423, 1533
  - tense-mode 1405
- terminative 1171
- terminativity 1168
- terminologische Kombinatorik 895
- terminology 1861
  - conventional term 1801
  - innovative term 1801
  - term 1924
- textbook 1860
- Textfrequenz 1272
- Themavokal/thematic vowel 447, 751
- theme 1401
- Tilgung 436, 468
- time-stability 733, 1422
- token blocking 347
- Ton/tone 587, 937, 1062, 1232, 1538, 1731
  - floating tone 1539
  - tonal morph 588
  - tonal morphology 588
  - tone-shift 1658
- top-down progression 1872
- tracking 1000
- Transfer/transfer 1644
- transfixation 528, 750
  - transfix 552
  - transfixional process 934
- transitivity 1496
  - transitive 1413, 1434
  - transitivity marker 1752
- transkortikale Aphasie 1819
- Transnumeral 1054
- transparency 814, 932, 1615, 1802
- Transposition/transposition 361, 888, 938, 1278
- transpositional inflection 362
- tree diagram 1879
- triadic kin term 1493
- Trial 1057
- triangular kin term 1003
- trilinear representation 1881
- triplication 529, 559
- truncation (s. Kurzwort)
- tun*-Periphrase 1273
- type 261
- type blocking 347
- type familiar 834
- type-familiarity 1627

## U

- Übergeneralisierung 1766
- Umlaut/umlaut 567, 1272, 1275, 1582, 1685
- unaccusative 813
- uncompleted 1440
- unconventional word-formation process 963
- underlying form 489, 499
- underlying representation 336, 499
- underlying segment 428
- underspecified 270
- unidimensional 492
- unikal 418
- unikales Restelement 893
- unisegmental 894
- Univerbierung 889, 1277, 1694
- universal 1764
- universal grinder 1067
- universal packager 1067
- universal sorter 1067
- universal, implicational 1249
- universalism 1858, 1862
- Universalsprache 1563
- unmarked 608
- unregelmäßig 1272
- Unregelmäßigkeit 1272
- unsupervised machine learning algorithm 1898
- utterance, minimum 379

## V

- Valenz/valence/valency 1118, 1130, 1402, 1565
  - morphological valence 164
  - valence class 1879
  - valency pattern 1130
  - valency shifter 1395
- validator 1461
- vantage point, egocentric 985
- vantage point, sociocentric 985
- Variation 234, 1860
  - asymmetrische Beziehung der Varianten 236
  - automatische Beziehung von Varianten 236
  - conditioned variant 484
  - freie Variation/freie Variante/fakultative Variante/free variation/free variant 180, 240, 484
  - funktionale Beziehung von Varianten 242
  - graduelle Beziehung zwischen Varianten 237
  - kombinatorische Variante 241
  - kontextuelle Beziehung von Varianten 242

- optional variant 484
- selbständige Beziehung von Varianten 236
- symmetrische Beziehung der Varianten 237
- Variante 234
- Varietätendifferenzierung 1682
- variety of language 1861
- Vendler class 1168
- venitiv 1521
- Verb/Verbum/verb 674, 1301, 1370, 1433, 1466, 1492
  - adjectival verb 1113
  - analytische Verbform/zusammengesetzte Verbform 1271 f.
  - aspectual verb 1172
  - auxiliary verb 1257, 1495
  - denominal verb 1531
  - dependent clause verb 1497
  - derived verb 1508
  - direct verb 1010, 1151
  - generic verb linkage 1506
  - impersonal verb 1448
  - lexical verb 1257
  - main clause verb 1497
  - main verb 1495
  - narrative verb 1187
  - ornative verb 744
  - Partikelverb/particle verb 166, 1277
  - phrasal verb 889
  - privative verb 744
  - reciprocal verb 1134
  - reduplicierendes Verb 1689
  - reflexive verb 1133, 1157
  - rückumlautende Verbflexion 1685
  - schwaches Verb 1271, 1685
  - sound verb 1743
  - starkes Verb 1272, 1684
  - stative verb 1113, 1520
  - synthetische Verbform/einfache Verbform 1271
  - trennbares Verb 1277
  - vector verb 1729
  - verb class 1168, 1412
  - verb morphology 1532
  - verb prefix 1436
  - verbal derivational suffix 1459
  - verbal measure 1744
  - verbal morphology 1400
  - verbalization 939, 1107
  - verbum finitum 1317
  - verbum infinitum 1317
  - Verbzusatz 892, 1277
  - weak verb 1744
- Verbreitungsgleichheit 1914
- Verdeutlichung 1913
- Verdichtung 1278
- Verfremdung 1275
- Verknüpfung 1914
- Verlaufsform 1273

- Verschmelzungsform 1271
- Verstärkung/reinforcement 1192, 1217, 1648
- Verteilung, inklusive 241
- Verteilung, komplementäre 241
- Verteilung, überlappende 241
- via-rule (s. relational rule)
- viewpoint 1166
- voice 1146, 1703
  - active voice 1146, 1148
  - conveyance voice 1480
  - inverse voice 1148
  - middle voice 1149, 1508
  - passive voice 1146, 1148
  - patient voice 1480
  - undergoer voice 1480
- Vokal/vowel
  - epenthetic vowel 1512
  - Vokalharmonie/vowel harmony 251, 466, 495, 568, 1359, 1512
  - Vokalwechsel 1685
  - vowel gradation 568
  - vowel lengthening 994
  - vowel mutation 567
  - vowel pattern 1746
- volitional 1746
- volitiv/volitive 1193, 1569
- Vollformenlexikon 1771
- Vorform 1910
- vowel (s. Vokal)

## W

- Wechsel, *e-i-* 1272
- Welthilfssprache 1563
- werden 1274
- Wernicke-Aphasie 1817
- Wort/word 5, 152, 377, 595, 1231 f., 1864, 1924
  - actual word 197, 833
  - content word 249, 258
  - echo word 570
  - Entstehung der Wort 1910
  - full word 1474
  - generalized word class 1867
  - grammatical word 152, 248, 378, 596, 868
  - lexical word 248
  - morphosyntactic word 248
  - orthographic word 251, 377
  - phonological word 251, 378
  - possible word 833
  - potential word 197, 833
  - stump word 958
  - syllable word 958
  - synthetisches Wort 895
  - word and paradigm 26, 409, 627
  - word blend 1790
  - word boundary 377 f., 1852

- word class 394, 596, 938, 1864
- word manufacture 833
- word marking 1491
- word order 391, 1198
- word stress 251
- word substitution 1790
- word-and-paradigm 1887
- word-based morphology 197f., 868
- word-form paradigm 647
- word-grammar 201
- word-oriented morphology 1864
- Wortform/word form 22, 152, 248, 378, 523, 595, 868
- Wortart 1301, 1565
- Wortbildung/word formation 198, 344, 676, 833, 1325, 1801, 1867, 1887
  - analytische Wortbildung 1307
  - identifikatorische Wortbildung 1307
  - innere Wortbildung 1683
  - Lehnwortbildung 895
  - mutatorische Wortbildung 1307
  - word formation rule 198, 932
  - Wortbildungsaktivität 889
  - Wortbildungsart 887, 1274
  - Wortbildungsmittel 1274
  - Wortbildungsmodell 1274
  - Wortbildungspotenz 889
  - Wortbildungssystem 1307
  - Wortbildungstyp 1275
- Wortform/word form 22, 152, 248, 378, 523, 595, 868
- Wortgeschichte 1910
- Wortgleichung 1910
- Wortgleichung, reine 1914
- Wortklasse, grammatische 1551
- Wortkreuzung 1275
- Wortkürzung 1275
- Wortlehre 1682
- Wortprägung 1911
- Wortschöpfung 886, 1275
- Wortstammvariante 1564
- Wortstand 1278
- würde*-Konjunktiv 1274
- Wurzel/root 257, 406, 524, 869, 1236, 1401, 1914
  - empty root 258
  - full root 258
  - root creation 965
  - root noun 752
  - root word 752
  - Wurzeletymologie 1914
  - Wurzelflexion 1683
  - Wurzelnomen 1914
  - Wurzelverb 1914
- X
- X-bar theory 408
- Z
- Zahl/numeral
  - collective numeral 1071
  - Kardinalzahl 1315
- Ordinalzahl 1315
- Zeichen/sign 378
  - elementary linguistic sign 511, 523
  - grammatical sign 524
  - leeres Zeichen 447
  - linguistic sign 404, 523
  - morphological sign 523
  - Nullzeichen/zero sign 438, 531
  - quasi-elementary sign 511
  - quasi-sign 406
  - sign vehicle 210
  - unique sign 406
- zero allomorph (s. Nullallo-morph)
- zero marking 745
- zero sign (s. Nullzeichen)
- zero suffixation 531
- zeroing 1010
- Zielsprache/target language 1646
- Zirkumfigierung/circumfixation 528, 934, 1308
- Zirkumfix/circumfix 891, 535, 1215, 1275, 1747, 1853
- Zitierform/citation form 248, 442, 604, 869, 1887
- zugrundeliegend 234
- Zusammenbildung 891, 1276
- zusammengesetztes Wort 1547
- Zusammenrückung 250, 889, 1277
- Zusammensetzung 250
- Zweisprachigkeit 1647

Die Register wurden unter Mitarbeit von Silke Fließ, Margret Schulze und Isabel Schlotte erstellt.

The indices were produced with the assistance of Silke Fließ, Margret Schulze and Isabel Schlotte.