Head-Driven Phrase Structure Grammar

Edited by

Stefan Müller

Anne Abeillé

Robert D. Borsley

Jean-Pierre Koenig

Empirically Oriented Theoretical Morphology and Syntax

language science press

Empirically Oriented Theoretical Morphology and Syntax

In this series:

1. Lichte, Timm. Syntax und Valenz: Zur Modellierung kohärenter und elliptischer Strukturen mit Baumadjunktionsgrammatiken

2. Bîlbîie, Gabriela. Grammaire des constructions elliptiques: Une étude comparative des phrases sans verbe en roumain et en français

3. Bowern, Claire, Laurence Horn & Raffaella Zanuttini (eds.). On looking into words (and beyond): Structures, Relations, Analyses

4. Bonami, Olivier, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer. The lexeme in descriptive and theoretical morphology.

5. Guzmán Naranjo, Matías. Analogical classification in formal grammar.

6. Julia Zinova. Russian Verbal Prefixation: A Frame Semantic Analysis

7. Müller, Stefan, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig. Head-Driven Phrase Structure Grammar.

ISSN: 2366-3529

Head-Driven Phrase Structure Grammar

Edited by

Stefan Müller

Anne Abeillé

Robert D. Borsley

Jean-Pierre Koenig

language science press

Müller , Stefan, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (ed.). 2019. *Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/000

ISBN: no digital ISBN
no print ISBNs!
ISSN: 2366-3529
no DOI
ID not assigned!

Cover and concept of design: Ulrike Harbort
Typesetting: Stefan Müller, Elizabeth Pankratz
Proofreading: Elizabeth Pankratz
Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono
Typesetting software: XƎLaTeX

Language Science Press
Unter den Linden 6
10099 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

Freie Universität Berlin

Contents

Preface v

Acknowledgments vii

I Introduction

1 Basic properties and elements
Bob Borsley & Anne Abeillé 3

2 The evolution of HPSG
Dan Flickinger, Carl Pollard & Tom Wasow 7

3 Formal Background
Frank Richter 45

4 The nature and role of the lexicon in HPSG
Anthony Davis & Jean-Pierre Koenig 47

5 HPSG in understudied languages
Douglas L. Ball 93

II Syntactic phenomena

6 Agreement
Stephen Wechsler 131

7 Case
Adam Przepiórkowski 155

8 Nominal structures
Frank Van Eynde 185

9 **Argument structure and linking**
Stephen Wechsler, Jean-Pierre Koenig & Anthony Davis 225

10 **Constituent order**
Stefan Müller 267

11 **Clitics**
Berthold Crysmann & Gerald Penn 313

12 **Complex predicates**
Danièle Godard & Pollet Samvelian 315

13 **Control and raising**
Anne Abeillé 351

14 **Unbounded dependencies**
Bob Borsley & Berthold Crysmann 353

15 **Relative Clauses in HPSG**
Doug Arnold & Danièle Godard 355

16 **Island phenomena and related matters**
Rui Chaves 415

17 **Coordination**
Anne Abeillé & Rui P. Chaves 469

18 **Idioms**
Manfred Sailer 517

19 **Negation**
Jong-Bok Kim 551

20 **Ellipsis**
Joanna Nykiel & Jong-Bok Kim 585

21 **Anaphoric Binding**
António Branco 613

III Other levels of description

A Phonology
Jesse Tseng 669

B Morphology
Berthold Crysmann 673

C Semantics
Jean-Pierre Koenig & Frank Richter 675

D Information structure
Kordula De Kuthy 709

IV Other areas of linguistics

E Diachronic syntax
Ulrike Demske 747

F Acquisition
Jonathan Ginzburg 773

G Processing
Tom Wasow 775

H Computational linguistics and grammar engineering
Emily M. Bender & Guy Emerson 797

I Grammar in dialogue
Andy Lücking, Jonathan Ginzburg & Robin Cooper 841

J Sign languages
Markus Steinbach & Anke Holler 885

K Gesture
Andy Lücking 887

V The broader picture

L HPSG and Minimalism
Robert D. Borsley & Stefan Müller — **937**

M HPSG and Categorial Grammar
Yusuke Kubota — **1013**

N HPSG and Lexical Functional Grammar
Doug Arnold — **1071**

O HPSG and Dependency Grammar
Dick Hudson — **1073**

P HPSG and Construction Grammar
Stefan Müller — **1077**

Indexes — **1129**

Preface

Acknowledgments

Part I

Introduction

Chapter 1

Basic properties and elements

Bob Borsley
University of Essex

Anne Abeillé
Université Paris Diderot

Change epigram in chapters/01.tex
or remove it there

Change the abstract in chapters/01.tex Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean nec fermentum risus, vehicula gravida magna. Sed dapibus venenatis scelerisque. In elementum dui bibendum, ultricies sem ac, placerat odio. Vivamus rutrum lacus eros, interdum scelerisque est euismod eget. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos.

1 Introduction

Phasellus maximus erat ligula, accumsan rutrum augue facilisis in. Proin sit amet pharetra nunc, sed maximus erat. Duis egestas mi eget purus venenatis vulputate vel quis nunc. Nullam volutpat facilisis tortor, vitae semper ligula dapibus sit amet. Suspendisse fringilla, quam sed laoreet maximus, ex ex placerat ipsum, porta ultrices mi risus et lectus. Maecenas vitae mauris condimentum justo fringilla sollicitudin. Fusce nec interdum ante. Curabitur tempus dui et orci convallis molestie (Chomsky 1957).

Meier (2017)

Bob Borsley & Anne Abeillé. 2019. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 3–5. Berlin: Language Science Press. DOI:??

(1) Latin (personal knowledge)
cogit-o ergo sum
think-1SG.PRES.IND hence exist.1SG.PRES.IND
'I think therefore I am'

Sed nisi urna, dignissim sit amet posuere ut, luctus ac lectus. Fusce vel ornare nibh. Nullam non sapien in tortor hendrerit suscipit. Etiam sollicitudin nibh ligula. Praesent dictum gravida est eget maximus. Integer in felis id diam sodales accumsan at at turpis. Maecenas dignissim purus non libero scelerisque porttitor. Integer porttitor mauris ac nisi iaculis molestie. Sed nec imperdiet orci. Suspendisse sed fringilla elit, non varius elit. Sed varius nisi magna, at efficitur orci consectetur a. Cras consequat mi dui, et cursus lacus vehicula vitae. Pellentesque sit amet justo sed lectus luctus vehicula. Suspendisse placerat augue eget felis sagittis placerat.

Table 1: Frequencies of word classes

	nouns	verbs	adjectives	adverbs
absolute	12	34	23	13
relative	3.1	8.9	5.7	3.2

Sed cursus [1] sapien pulvinar. Sed consequat, magna[2]. Nunc dignissim tristique massa ut gravida. Nullam auctor orci gravida tellus egestas, vitae pharetra nisl porttitor. Pellentesque turpis nulla, venenatis id porttitor non, volutpat ut leo. Etiam hendrerit scelerisque luctus. Nam sed egestas est. Suspendisse potenti. Nunc vestibulum nec odio non laoreet. Proin lacinia nulla lectus, eu vehicula erat vehicula sed.

Abbreviations

COP	copula	NEG	negation
FV	final vowel	SM	subject marker

[1] eros condimentum mi consectetur, ac consectetur

[2] eu scelerisque laoreet, ante erat tristique justo, nec cursus eros diam eu nisl. Vestibulum non arcu tellus

Acknowledgements

References

Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). The Hague/Paris: Mouton.

Meier, Jane. 2017. *Language universals and linguistic typology*. Oxford: Basil Blackwell.

Chapter 2

The evolution of HPSG

Dan Flickinger
Stanford University

Carl Pollard
Ohio State Universitiy

Tom Wasow
Stanford University

HPSG was developed to express insights from theoretical linguistics in a precise formalism that was computationally tractable. It drew ideas from a wide variety of traditions in linguistics, logic, and computer science. Its chief architects were Carl Pollard and Ivan Sag, and its most direct precursors were Generalized Phrase Structure Grammar and Head Grammar. The theory has been applied in the construction of computational systems for the analysis of a variety of languages; a few of these systems have been used in practical applications. This chapter sketches the history of the development and application of the theory.

1 Introduction

From its inception in 1983, HPSG was intended to serve as a framework for the formulation and implementation of natural language grammars which are (i) linguistically motivated, (ii) formally explicit, and (iii) computationally tractable. These desiderata are reflective of HPSG's dual origins as an academic linguistic theory and as part of an industrial grammar implementation project with an eye toward potential practical applications. Here (i) means that the grammars are intended as scientific theories about the languages in question, and that the analyses the grammars give rise to are transparently relatable to the predictions

Dan Flickinger, Carl Pollard & Tom Wasow. 2019. The evolution of HPSG. in Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 7–43. Berlin: Language Science Press. DOI:??

(empirical consequences) of those theories. Thus HPSG shares the general concerns of the theoretical linguistics literature, including distinguishing between well-formed and ill-formed expressions and capturing linguistically significant generalizations. (ii) means that the notation for the grammars and its interpretation have a precise grounding in logic, mathematics, and theoretical computer science, so that there is never any ambiguity about the intended meaning of a rule or principle of grammar, and so that grammars have determinate empirical consequences. (iii) means that the grammars can be translated into computer programs that can handle linguistic expressions embodying the full range of complex interacting phenomena that naturally occur in the target languages, and can do so with a tolerable cost in space and time resources.

The two principal architects of HPSG were Carl Pollard and Ivan Sag, but a great many other people made important contributions to its development. Many, but by no means all, are cited in the chronology presented in the following sections. There are today a number of groups of HPSG researchers around the world, in many cases involved in building HPSG-based computational systems. While the number of practitioners is relatively small, it is a very active community that holds annual meetings and publishes quite extensively. Hence, although Pollard no longer works on HPSG and Sag died in 2013, the theory is very much alive, and still evolving.

2 Precursors

HPSG arose between 1983 and 1985 from the complex interaction between two lines of research in theoretical linguistics: (i) work on context-free Generative Grammar (CFG) intitiated in the late 1970s by Gerald Gazdar and Geoffrey Pullum, soon joined by Ivan Sag, Ewan Klein, Tom Wasow, and others, resulting in the framework referred to as Generalized Phrase Structure Grammar (GPSG: Gazdar, Klein, Pullum & Sag 1985); and (ii) Carl Pollard's Stanford dissertation research, under Sag and Wasow's supervision, on Generalized Context-Free Grammar, and more specifically Head Grammar (HG: Pollard 1984).

2.1 Generalized Phrase Structure Grammar

In the earliest versions of Generative Grammar (Chomsky 1957), the focus was on motivating transformations to express generalizations about classes of sentences. In the 1960s, as generative linguists began to attend more explicitly to meaning, a division arose between those advocating using the machinery of transformations to capture semantic generalizations and those advocating the use of other types of formal devices. This division became quite heated, and was subsequently dubbed "the linguistic wars" (see Newmeyer 1980: Chapter 5). Much of the work in theoretical syntax and semantics during the 1970s explored ways to constrain the power of transformations (see especially Chomsky 1973 and Chomsky & Lasnik 1977), and non-transformational approaches to the analysis of meaning (see especially Montague 1974 and Dowty 1979).

These developments led a few linguists to begin questioning the central role transformations had played in syntactic research of the preceding two decades (notably, Bresnan 1978). This questioning of Transformational Grammar (TG) culminated in a series of papers by Gerald Gazdar, which (in those pre-internet days) were widely distributed as paper manuscripts. The project that they laid out was succinctly summarized in one of Gazdar's later publications as follows:

> Consider eliminating the transformational component of a generative grammar. (Gazdar 1981: 155)

The framework that emerged became known as Generalized Phrase Structure Grammar; a good account of its development is Ted Briscoe's interview of Gazdar in November 2000.[1]

GPSG developed in response to several criticisms leveled against transformational grammar. First, TG was highly underformalized, to the extent that it was unclear what its claims—and the empirical consequences of those claims—amounted to; CFG, by comparison, was a simple and explicit mathematical formalism. Second, given the TG architecture of a context-free base together with a set of transformations, the claimed necessity of transformations was standardly justified on the basis of arguments that CFGs were insufficiently expressive to serve as a general foundation for natural language (NL) grammar; but Pullum & Gazdar (1982) showed all such arguments presented up to that time to be logically flawed or else based on false empirical claims. And third, closely related to the previous point, they showed that transformational grammarians had been insufficiently resourceful in exploiting what expressive power CFGs *did* possess, especially through the use of complex categories bearing features whose values might themselves bear features of their own. For example, coordinate constructions and unbounded dependency constructions had long served as prime exemplars of the need for transformations, but Gazdar (1981) was able to show that both kinds of constructions, as well as interactions between them, did in fact yield straightforward analyses within the framework of a CFG.

[1] https://nlp.fi.muni.cz/~xjakub/briscoe-gazdar/, 2018-08-21.

Gazdar and Pullum's early work in this vein was quickly embraced by Sag and Wasow at Stanford University, both formally inclined former students of Chomsky's, who saw it as the logical conclusion of a trend in Chomskyan syntax toward constraining the tranformational component. That trend, in turn, was a response, at least in part, to (i) the demonstration by Peters & Ritchie (1973) that Chomsky's (1965) Standard Theory, when precisely formalized, was totally unconstrained, in the sense of generating all recursively enumerable languages; and (ii) the insight of Emonds (1976) that most of the transformations proposed up to that time were "structure-preserving" in the sense that the trees they produced were isomorphic to ones that were base-generated. Besides directly addressing these issues of excess power and structure preservation, the hypothesis that NLs were context-free also had the advantage that CFGs were well-known by computer scientists to have decidable recognition problems and efficient parsing algorithms, facts which seemed to have some promise of bearing on questions of the psychological plausibility and computational tractability of the grammars in question.

Aside from serving as a framework for theoretical linguistic research, GPSG also provided the theoretical underpinnings for a natural language processing (NLP) project established in 1981 by Egon Loebner at Hewlett-Packard Laboratories in Palo Alto. This project, which led in due course to the first computer implementation of HPSG, is described below.

2.2 Head Grammar

Pollard, with a background in pure mathematics, Chinese historical phonology, and 1930s–1950s-style American structural linguistics, arrived at Stanford in 1979 with the intention of getting a Ph.D. in Chinese linguistics, but was soon won over to theoretical syntax by Wasow and Sag. He had no exposure to Chomskyan linguistics, but was immediately attracted to the emerging nontransformational approaches, especially the early GPSG papers and the contemporaneous forms of CG in Bach (1979; 1980) and Dowty (1982a,b), in part because of their formal simplicity and rigor, but also because the formalism of CFG was (and is) easy to read as a more technically precise rendering of structualist ideas about syntax (as presented, e.g., in Bloomfield 1933 and Hockett 1958).

Although Pullum & Gazdar (1982) successfully refuted all published arguments to date that CFGs were inadequate for analyzing NLs, by the following year, Stuart Shieber had developed an argument (published in Shieber 1985), which was (and remains) generally accepted as correct, that there could not be a CFG that accounted for the cross-serial dependencies in Swiss German; and Chris Culy showed, in his Stanford M.A. thesis (cf. Culy 1985), that the presence of reduplicative compounding in Bambara precluded a CF analysis of that language. At the same time, Bach and Dowty (independently) had been experimenting with generalizations of traditional A-B (Ajdukiewicz-Bar Hillel) CG which allowed for modes of combining strings (such as reduplication, wrapping, insertion, cliticization, and the like) in addition to the usual concatenation. This latter development was closely related to a wider interest among nontransformational linguists of the time in the notion of discontinuous constituency, and also had an obvious affinity to Hockett's (1954) item-and-process conception of linguistic structure, albeit at the level of words and phrases rather than morphemes. One of the principal aims of Pollard's dissertation work was to provide a general framework for syntactic (and semantic) analysis that went beyond—but not too far beyond—the limits of CFG in a way that took such developments into account.

Among the generalizations of CFG that Pollard studied, special attention was given to HGs, which differ from CFGs in two respects: (i) the role of strings was taken over by headed strings, essentially strings with a designation of one of its words as its head; and (ii) besides concatenation, headed strings can also be combined by inserting one string directly to the left or right of another string's head. An appendix of his dissertation (Pollard 1984) provided an analysis of discontinuous constituency in Dutch, and that analysis also works for Swiss German. In another appendix, Pollard used a generalization of the CKY algorithm to prove that the head languages (HLs, the languages analyzed by HGs) shared with CFLs the property of deterministic polynomial time recognition complexity, but of order n^7, subsequently reduced by Kasami, Seki & Fujii (1989) to n^6, as compared with order n^3 for CFLs. For additional formal properties of HGs, see Roach (1987). Vijay-shanker & Weir (1994) proved that HGs had the same weak generative capacities as three other grammar formalisms – Combinatory Categorial Grammar (Steedman 1987; 1990), Lexicalized Tree-Adjoining Grammar (Shabes 1990), and Linear Indexed Grammar (Gazdar 1988) – and the corresponding class of languages became known as 'mildly context sensitive'.

Although the handling of linearization in HG seems not to have been pursued further within the HPSG framework, the ideas that (i) linearization had to involve data structures richer than strings of phoneme strings, and (ii) the way these structures were linearized had to involve operations other than mere concatenation, were implicit in subsequent HPSG work, starting with Pollard & Sag's (1987) Constituent Order Principle (which was really more of a promissory note than an actual principle). These and related ideas would become more fully fleshed out a decade later within the linearization grammar avatar of HPSG developed by Reape (1996), Reape (1992), and Kathol (1995; 2000). On the other hand, two other innovations of HG, both related to the system of syntactic features, were incorporated into HPSG, and indeed should probably be considered the defining characteristics of that framework, namely the list-valued SUBCAT and SLASH features, discussed below.

3 The HP NLP project

Work on GPSG culminated in the 1985 book *Generalized Phrase Structure Grammar* by Gazdar, Klein, Pullum, and Sag. During the writing of that book, Sag taught a course on the theory, with participation of his co-authors. The course was attended not only by Stanford students and faculty, but also by linguists from throughout the area around Stanford, including the Berkeley and Santa Cruz campuses of the University of California, as well as people from nearby industrial labs. One of the attendees at this course was Anne Paulson, a programmer from Hewlett-Packard (HP) Laboratories in nearby Palo Alto, who had some background in linguistics from her undergraduate education at Brown University. Paulson told her supervisor at HP Labs, Egon Loebner, that she thought the theory could be implemented and might be turned into something useful. Loebner, a multi-lingual polymathic engineer, had no background in linguistics, but he was intrigued, and invited Sag to meet and discuss setting up a natural language processing project at HP. Sag brought along Gazdar, Pullum, and Wasow. This led to the creation of the project that eventually gave rise to HPSG. Gazdar, who would be returning to England relatively soon, declined the invitation to be part of the new project, but Pullum, who had taken a position at the University of California at Santa Cruz (about an hour's drive from Palo Alto), accepted. So the project began with Sag, Pullum, and Wasow hired on a part-time basis to work with Paulson and two other HP programmers, John Lamping and Jonathan King, to implement a GPSG of English at HP Labs. J. Mark Gawron, a linguistics graduate student from Berkeley who had attended Sag's course, was very soon added to the team.

The initial stages consisted of the linguists and programmers coming up with a notation that would serve the purposes of both. Once this was accomplished, the linguists set to work writing a grammar of English in Lisp to run on the DEC-20 mainframe computer that they all worked on. The first publication coming out of this project was a 1982 Association for Computational Linguistics paper. The paper's conclusion begins:

> What we have outlined is a natural language system that is a direct im-

> plementation of a linguistic theory. We have argued that in this case the linguistic theory has the special appeal of computational tractability (promoted by its context-freeness), and that the system as a whole offers the hope of a happy marriage of linguistic theory, mathematical logic, and advanced computer applications. (Gawron et al. 1982: 80)

This goal was carried over into HPSG.

It should be mentioned that the HP group was by no means alone in these concerns. The early 1980s was a period of rapid growth in computational linguistics (due at least in part to the rapid growth in the power and accessibility of computers). In the immediate vicinity of Stanford and HP Labs, there were at least two other groups working on developing natural language systems that were both computationally tractable and linguistically motivated. One such group was at the Xerox Palo Alto Research Center, where Ron Kaplan and Joan Bresnan (in collaboration with a number of other researchers, notably Martin Kay) were developing Lexical Functional Grammar; the other was at SRI International, where a large subset of SRI's artificial intelligence researchers (including Barbara Grosz, Jerry Hobbs, Bob Moore, Hans Uszkoreit, Fernando Pereira, and Stuart Shieber) worked on natural language. Thanks to the founding of the Center for the Study of Language and Information (CSLI) at Stanford in the early 1980s, there was a great deal of interaction among these three research groups. Although some aspects of the work being done at the three non-Stanford sites were proprietary, most of the research was basic enough that there was a fairly free flow of ideas among the three groups about building linguistically motivated natural language systems.

Other projects seeking to develop theories combining computational tractability with linguistic motivation were also underway outside of the immediate vicinity of Stanford, notably at the Universities of Pennsylvania and Edinburgh. Aravind Joshi and his students were working on Tree Adjoining Grammars (Joshi et al. 1975; Joshi 1987), while Mark Steedman and others were developing Combinatory Categorial Grammar (Steedman 1987; 1990).

During the first few years of the HP NLP project, several Stanford students were hired as part-time help. One was Pollard, who was writing his doctoral dissertation under Sag's supervision. Ideas from his thesis work played a major role in the transition from GPSG to HPSG. Two other students who became very important to the project were Dan Flickinger, a doctoral student in linguistics, and Derek Proudian, who was working on an individually-designed undergraduate major when he first began at HP and later became a master's student in computer science. Both Flickinger and Proudian became full-time HP employees after finishing their degrees. Over the years, a number of other HP employees also worked on the project and made substantial contributions. They included Susan Brennan, Lewis Creary, Marilyn Friedman (now Walker), Dave Goddeau, Brett Kessler, Joachim Laubsch, and John Nerbonne. Brennan, Walker, Kessler, and Nerbonne all later went on to academic careers at major universities, doing research dealing with natural language processing.

The HP NLP project lasted until the early 1990s. By then, a fairly large and robust grammar of English had been implemented. The period around 1990 combined an economic recession with what has sometimes been termed an "AI winter" – that is, a period in which enthusiasm and hence funding for artificial intelligence research was at a particularly low ebb. Since NLP was considered a branch of AI, support for it waned. Hence, it was not surprising that the leadership of HP Labs decided to terminate the project. Flickinger and Proudian came to an agreement with HP that allowed them to use the NLP technology developed by the project to launch a new start-up company, which they named Eloquent Software. They were, however, unable to secure the capital necessary to turn the existing system into a product, so the company never got off the ground.

4 The emergence of HPSG

A few important features of GPSG that were later carried over into HPSG are worth mentioning here. First, GPSG borrowed from Montague the idea that each phrase structure rule was to be paired with a semantic rule providing a recipe for computing the meaning of the mother from the meanings of its daughters (Gazdar 1981: 156); this design feature was shared with contemporaneous forms of Categorial Grammar (CG) being studied by such linguists as Emmon Bach (Bach 1979; 1980) and David Dowty (Dowty 1982a,b). Second, the specific inventory of features employed in GPSG for making fine-grained categorial distinctions (such as case, agreement, verb inflectional form, and the like), was largely preserved, though the technical implementation of morphosyntactic features in HPSG was somewhat different. And third, the SLASH feature, which originated in Gazdar's (1981) derived categories (e.g. S/NP), and which was used to keep track of unbounded dependencies, was generalized in HPSG to allow for multiple unbounded dependencies (as in the notorious violins-and-sonatas example in (1) below). As will be discussed, this SLASH feature bears a superficial—and misleading—resemblance to the Categorial Grammar connectives written as '/' and '\'. On the other hand, a centrally important architectural feature of GPSG absent from HPSG (and from HG) was the device of metarules, higher-order rules used to generate the full set of context-free phrase structure rules (PSRs) from an initial inventory of basic PSRs. Among the metarules were ones used to introduce non-null SLASH values and propagate them upward through trees to a position where they were discharged by combination with a matching constituent called a filler (analogous to a *wh*-moved expression in TG).

A note is in order about the sometimes confusing use of the names *Head Grammar (HG)* and *HPSG*. Strictly speaking, HG was a specific subtype of generalized CFG developed in Pollard's dissertation work, but the term *HG* did not appear in academic linguistic publications with the exception of the Pollard & Sag (1983) WCCFL paper, which introduced the distinction between head features and binding features (the latter were incorporated into GPSG under the name *foot features*). In the summer of 1982, Pollard had started working part time on the HP NL project; and the term *HPSG* was first employed (by Pullum) in reference to an extensive reworking by Pollard and Paulson of the then-current HP GPSG implementation, incorporating some of the main features of Pollard's dissertation work in progress, carried out over the summer of 1983, while much of the HP NLP team (including Pullum and Sag) was away at the LSA Institute in Los Angeles. The implication of the name change was that whatever this new system was, it was no longer GPSG.

Once this first HPSG implementation was in place, the NLP work at HP was considered to be within the framework of HPSG, rather than GPSG. After Pollard completed his dissertation, he continued to refer to *HG* in invited talks as late as autumn 1984; but his talk at the (December 1984) LSA Binding Theory Symposium used *HPSG* instead, and after that, the term *HG* was supplanted by *HPSG* (except in publications by non-linguists about formal language theory). One additional complication is that until the Gazdar, Klein, Pullum & Sag (1985) volume appeared, GPSG and HPSG were developing side by side, with considerable interaction. Pollard, together with Flickinger, Wasow, Nerbonne, and others, did HPSG; Gazdar and Klein did GPSG; and Sag and Pullum worked both sides of the street.

HPSG papers, about both theory and implementation, began to appear in 1985, starting with Pollard's WCCFL paper *Phrase structure grammar without metarules* (Pollard 1985), and his paper at the Categorial Grammar conference in Tucson (Pollard 1988), comparing and contrasting HPSG with then-current versions of Categorial Grammar due to Bach, Dowty, and Steedman. These were followed by a trio of ACL papers documenting the current state of the HPSG implementation at HP Labs: Creary & Pollard (1985), Flickinger, Pollard & Wasow (1985), and Proudian & Pollard (1985). Of those three, the most significant in terms of its influence on the subsequent development of the HPSG framework was the second, which showed how the lexicon could be (and in fact was) organized using multiple-inheritance knowledge representation; Flickinger's Stanford dissertation (Flickinger 1987) was an in-depth exploration of that idea.

5 Early HPSG

Setting aside implementation details, early HPSG can be characterized by the following architectural features:

Elimination of metarules Although metarules were a central feature of GPSG, they were also problematic: Uszkoreit & Peters (1982) had shown that if metarules were allowed to apply to their own outputs, then the resulting grammars were no longer guaranteed to generate CFLs; indeed, such grammars could generate all recursively enumerable languages. And so, in GPSG, the closure of a set of base phrase structure rules (PSRs) under a set of metarules was defined in such a way that no metarule could apply to a PSR whose own derivation involved an application of that metarule. This definition was intended to ensure that the closure of a finite set of PSRs remained finite, and therefore still constituted a CFG.

So, for example, the metarule STM1 was used in GPSG to convert a PSR into another PSR, one of whose daughters is [+NULL] (informally speaking, a 'trace'), and feature cooccurrence restrictions (FCRs) guaranteed that such daughters would bear a SLASH value, and that this SLASH value would also appear on the mother. Unfortunately, the finite closure definition described above does not preclude the possibility of derived PSRs whose mother carries multiple, in fact unboundedly many SLASH values (e.g. NP/NP, (NP/NP)/NP, etc.). And this in turn leads to an infinite set of PSRs, outside the realm of CF-ness (see Ristad 1986). Of course, one could rein in this excess power by imposing another FCR that disallows categories of the form (X/Y)/Z; but then there is no way to analyze sentences containing a constituent with two undischarged unbounded dependencies, such as the VP complement of *easy* in the following example:

(1) Violins this finely crafted, even the most challenging sonatas are easy to [play _ on _].

GPSG avoided this problem by not analyzing such examples. In HPSG (Pollard 1985), by contrast, such examples were analyzed straightforwardly by replacing GPSG's category-valued SLASH feature with one whose values were lists (or sets) of categories. This approach still gave rise to an infinite set of rules, but since maintaining context-freeness was no longer at stake, this was not seen as problematic. The infinitude of rules in HPSG arose not through a violation of finite closure (since there were no longer any metarules at all), but because each of the handful of schematic PSRs (see below) could be directly instantiated in an infinite number of ways, given that the presence of list-valued features gave rise to an infinite set of categories.

Lexical rules GPSG, generalizing a suggestion of Flickinger (1983), constrained metarules to apply only to PSRs that introduced a lexical head. Pollard (1985) took this idea a step further, noting that many proposed metarules could be reformulated as lexical rules that (among other effects) operated on the subcategorization frames (encoded by the SUBCAT feature discussed below) of lexical entries. The idea of capturing some linguistic generalizations by means of rules internal to the lexicon had been explored by generative grammarians since Jackendoff (1975); and lexical rules of essentially the kind Pollard proposed were employed by Bach (1983), Dowty (1978), and others working in Categorial Grammar. Examples of constructions handled by metarules in GPSG but in HPSG by lexical rules included sentential extraposition, subject extraction, and the passive. Flickinger, Pollard & Wasow (1985) argued for an architecture for the lexicon that combined lexical rules with multiple inheritance using a frame-based knowledge representation system (Minsky 1975), on the basis of both overall grammar simplicity and efficient, easily modifiable implementation.

CG-like treatment of subcategorization GPSG treated subcategorization using an integer-valued feature called SUBCAT that in effect indexed each lexical item with the rule that introduced and provided its subcategorization frame; e.g. *weep* was listed in the lexicon with SUBCAT value 1 while *devour* was listed with SUBCAT value 2, and then PSRs of roughly the form in (2)

(2) VP → V[SUBCAT 1]
VP → V[SUBCAT 2] NP

guaranteed that lexical heads would have the right kinds of complements. In HPSG, by contrast, the SUBCAT feature directly characterized the grammatical arguments selected by a head (not just the complements, but the subject too) as a list of categories, so that e.g. *weep* was listed as V[SUBCAT ⟨NP⟩] but *devour* as V[SUBCAT ⟨NP, NP⟩] (where the first occurrence of NP refers to the object and the second to the subject). This treatment of argument selection was inspired by

Categorial Grammar, where the same verbs would have been categorized as NP\S and (NP\S)/NP respectively;[2] the main differences are that (i) the CG treatment also encodes the directionality of the argument relative to the head, and (ii) in HPSG, all the arguments appear on one list, while in CG they are 'picked up' one at a time, with as many connectives (/ or \) as there are arguments. In particular, as in the CG of Dowty (1982b), the subject was defined as the last argument, except that in HPSG, 'last' now referred to the rightmost position on the SUBCAT list, not to the most deeply embedded connective. In HPSG, this ordering of the categories on the SUBCAT list was related not just to CG, but also to the traditional grammatical notion of obliqueness, and also to the accessibility hierarchy of Keenan & Comrie (1977). See Müller & Wechsler (2014: Section 4) for a more detailed discussion of these developments from GPSG to HPSG.

Schematic rules Unlike CFG but like CG, HPSG had only a handful of schematic rules. For example, in Pollard (1985), a substantial chunk of English 'local' grammar (i.e. leaving aside unbounded dependencies) was handled by three rules: (i) a rule (used for subject-auxiliary inversion) that forms a sentence from an inverted (+INV) lexical head and all its arguments; (ii) a rule that forms a phrase from a head with SUBCAT list of length > 1 together with all its non-subject arguments; and (iii) a rule that forms a sentence from a head with a SUBCAT value of length one together with its single (subject) argument.

List- (or set-) valued SLASH feature The list-valued SLASH was introduced in Pollard (1985) to handle multiple unbounded dependencies, instead of the GPSG category-valued SLASH (which in turn originated as the *derived categories* of Gazdar (1981), e.g. S/NP). In spite of the notational similarity, though, the PSG SLASH is not an analog of the CG slashes / and \ (though HPSG's SUBCAT is, as explained above). In fact, HPSG's SLASH has no analog in the kinds of CGs being developed by Montague semanticists such as Bach (1979; 1980) and Dowty (1982a) in the late 1970s and early 1980s, which followed the CGs of Bar-Hillel (1954) in having only rules for eliminating (or canceling) slashes as in (3):

(3) $\dfrac{\text{A}\ \ \text{A}\backslash\text{B}}{\text{B}} \qquad \dfrac{\text{B/A}\ \ \text{A}}{\text{B}}$

To find an analog to HPSG's SLASH in CG, we have to turn to the kinds of CGs invented by Lambek (1958), which unfortunately were not yet well-known to

[2] We adhere to the Lambek convention for functor categories, so that expressions seeking to combine with an A on the left to form a B are written 'A\B' (not 'B\A').

linguists (though that would soon change, starting with Lambek's appearance at the 1985 Categorial Grammar conference in Tucson). What sets apart grammars of this kind (and their elaborations by Moortgat (1989), Oehrle et al. (1988), Morrill (1994), and many others), is the existence of rules for hypothetical proof (not given here), which allow a hypothesized category occurrence introduced into a tree (thought of as a proof) to be discharged.

In the Gentzen style of natural deduction (see Pollard 2013), hypothesized categories are written to the left of the symbol ⊢ (turnstile), so that the two slash elimination rules above take the following form, where Γ and Δ are lists of categories, and comma represents list concatenation as in (4):

(4) $$\frac{\Gamma \vdash \mathrm{A} \quad \Delta \vdash \mathrm{A}\backslash\mathrm{B}}{\Gamma,\Delta \vdash \mathrm{B}} \qquad \frac{\Gamma \vdash \mathrm{B}/\mathrm{A} \quad \Delta \vdash \mathrm{A}}{\Gamma,\Delta \vdash \mathrm{B}}$$

These rules serve to propagate hypotheses (analogous to linguists' traces) downward through the proof tree (downward because logicians' trees are upside down with the conclusion, or 'root', at the bottom). In HPSG notation, these same rules can be written as one rule (since SUBCAT is nondirectional) in (5):

(5) $$\frac{\mathrm{B}[\textsc{subcat}\ \langle \ldots, \mathrm{A} \rangle, \textsc{slash}\ \Gamma] \quad \mathrm{A}[\textsc{slash}\ \Delta]}{\mathrm{B}[\textsc{subcat}\ \langle \ldots \rangle][\textsc{slash}\ \Gamma,\Delta]}$$

This in turn is a special case of an HPSG principle first known as the Binding Inheritance Principle (BIP) and later as the Nonlocal Feature Principle (binding features included SLASH as well as the features QUE and REL used for tracking undischarged interrogative and relative pronouns). The original statement of the BIP (Pollard 1986) treated SLASH as set- rather than list-valued:

> The value of a binding feature on the mother is the union of the values of that feature on the daughters.

For example, the doubly-gapped VP in the violins-and-sonatas example in (1) is analyzed in HPSG roughly as is shown in Figure 1 and essentially the same way in Lambek-style CG:

(6) $$\dfrac{\dfrac{\overset{\text{play}}{\vdash ((\mathrm{NP}\backslash\mathrm{S})/\mathrm{PP})/\mathrm{NP}} \quad \overset{t}{\mathrm{NP} \vdash \mathrm{NP}}}{\mathrm{NP} \vdash (\mathrm{NP}\backslash\mathrm{S})/\mathrm{PP}} \quad \dfrac{\overset{\textit{on}}{\vdash \mathrm{PP}/\mathrm{NP}} \quad \overset{t}{\mathrm{NP} \vdash \mathrm{NP}}}{\mathrm{NP} \vdash \mathrm{PP}}}{\mathrm{NP},\mathrm{NP} \vdash (\mathrm{NP}\backslash\mathrm{S})}$$

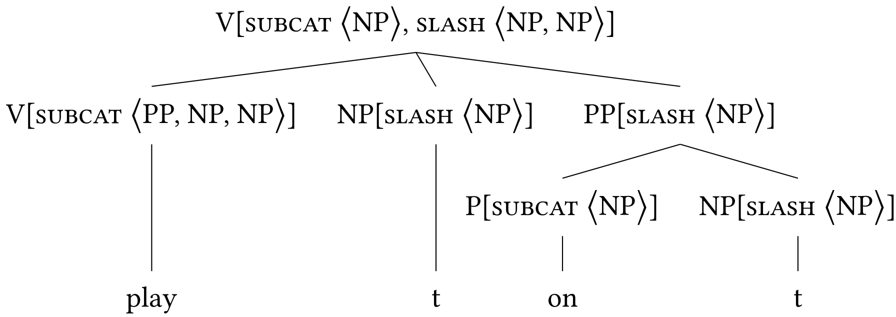


Figure 1: *play on* as part of *Violins this finely crafted, even the most challenging sonatas are easy to play on.*

Aside from the binary branching of the Lambek analysis, the main difference is that HPSG traces of the form A[SLASH ⟨A⟩] correspond to Lambek axioms of the form A ⊢ A, which is the standard mechanism for introducing hypotheses in Gentzen-style natural deduction.

An overview and elaboration of early HPSG is provided by the two books Pollard & Sag (1987) and Pollard & Sag (1994). Confusingly, the former is called *Information-Based Syntax and Semantics, Volume 1: Fundamentals*, and the second simply *Head-Driven Phrase Structure Grammar* (not *Information-Based Syntax and Semantics, Volume 2*). The reason for the title change had to do with a change in the underlying mathematical theory of feature structures. In the first book, following work in theoretical computer science by Rounds & Kasper (1986) and Moshier & Rounds (1987), feature structures were treated as data structures that supplied partial information about the linguistic objects being theorized about; this perspective in turn was based on Scott's (1982) mathematical theory of computation in terms of what he called information systems. Subsequently, Paul King persuaded Pollard and Sag that it was more straightforward to distinguish between feature structures, thought of as formal models of the linguistic objects, and feature descriptions or formulas of feature logic, which provided partial information about them, as described in his Manchester dissertation (King 1989). Although the formal issues involved in distinguishing between the two approaches are of interest in their own right, they seem not to have had a lasting effect on how theoretical linguists used HPSG, nor on how computational linguists implemented it. As for subject matter, Pollard & Sag (1987) was limited to the most basic notions, including syntactic features and categories (including the distinction between head features and binding features); subcategorization and the distinction between arguments and adjuncts (the latter of which necessitated one more rule schema beyond the three proposed by Pollard 1985); basic principles of grammar (especially the Head Feature Principle and the Subcategorization Principle); the obliqueness order and constituent ordering; and the organization of the lexicon by means of a multiple inheritance hierarchy and lexical rules. Pollard & Sag (1994) used HPSG to analyze a wide range of phenomena, primarily in English, that had figured prominently in the syntactic literature of the 1960s–1980s, including agreement, expletive pronoun constructions, raising, control, filler-gap constructions (including island constraints and parasitic gaps); so-called binding theory (the distribution of reflexive pronouns, nonreflexive pronouns, and non-pronominal NPs), and scope of quantificational NPs.

6 Theoretical Developments

Three decades of vigorous work since Pollard & Sag (1987) developing the theoretical framework of HPSG receive detailed discussion throughout the present volume, but we highlight here two significant stages in that development. The first is in Chapter 9 of Pollard & Sag (1994), where a pair of major revisions to the framework presented in the first eight chapters are adopted, changing the analysis of valence and of unbounded dependencies. Following Borsley (1987; 1988; 1989; 1990), Pollard and Sag moved to distinguish subjects from complements, and further to distinguish subjects from specifiers, thus replacing the single SUBCAT attribute with SUBJ, SPR, and COMPS. This formal distinction between subjects and complements enabled an improved analysis of unbounded dependencies, eliminating traces altogether by introducing three lexical rules for the extraction of subjects, complements, and adjuncts respectively. It is this revised analysis of valence constraints that came to be viewed as part of the standard HPSG framework, though issues of valence representation cross-linguistically remain a matter of robust debate.

The second notable stage of development was the introduction of a type hierarchy of *constructions* as descriptions of phrasal feature structures, employed first by Sag (1997) in a richly detailed analysis of a wide variety of relative clause phenomena in English. This extension from the lexicon of the use of descriptions of typed feature structures organized in hierarchies to syntactic rules preserved the ability to express general principles holding for rule schemata while also enabling expression of idiosyncratic properties of phrases. In Borsley & Abeillé (2019), Chapter 1 of this volume, the version of the framework with this extended use of types is termed *Construction-based HPSG*, including further elaboration by

Ginzburg & Sag (2000) to a comprehensive analysis of interrogatives in English.

7 The LinGO Project

In the early 1990s, a consortium of research centers in Germany secured funding from the German government for a large project in spoken language machine translation, called Verb*mobil* (Wahlster 2000), which aimed to combine a variety of methods and frameworks in a single implemented state-of-the-art demonstrator system. Grammars of German and English were to be implemented in HPSG, to be used both for parsing and for generation in the translation of human-human dialogues, with a German grammar initially implemented by Pollard and Tibor Kiss at IBM in Heidelberg, later replaced by one developed at the German AI Research Center (DFKI), coordinator for the Verb*mobil* project. The DFKI contracted in 1993 with Sag at CSLI to design and implement the English grammar, with Flickinger brought over from HP Labs to help lead the effort, forming a new research group at CSLI initially called ERGO (for English Resource Grammar Online), later generalized to the name LinGO (Linguistic Grammars Online). Early LinGO members included Wasow and linguistics graduate student Rob Malouf, who authored the initial implementation of the English Resource Grammar (ERG), along with four other linguistics graduate students, Emily Bender, Kathryn Campbell-Kibler, Tony Davis, and Susanne Riehemann.

During the first of the two four-year phases of the Verb*mobil* project, the focus was on designing and implementing core syntactic and semantic analyses, initially using the DISCO/PAGE platform (Uszkoreit et al. 1994) developed at the DFKI, and largely informed by the framework presented in Pollard & Sag (1994). However, a more computationally useful semantic formalism emerged, called Minimal Recursion Semantics (MRS: Copestake, Flickinger, Pollard & Sag 2005), which Ann Copestake, formerly of the European ACQUILEX project, helped to design. Copestake also expanded the LKB system (Copestake 2002) which had been used in ACQUILEX, to serve as the grammar development environment for the LinGO project, including both a parser and a generator for typed feature structure grammars.

The second four years of the Verb*mobil* project emphasized development of the generation capabilities of the ERG, along with steady expansion of linguistic coverage, and elaboration of the MRS framework. LinGO contributors in this phase included Sag, Wasow, Flickinger, Malouf, Copestake, Riehemann, and Bender, along with a regular visitor and steady contributor from the DFKI, Stephan Oepen. Verb*mobil* had meanwhile added Japanese alongside German (Müller &

Kasper 2000) and English (Flickinger, Copestake & Sag 2000) for more translation pairs, giving rise to another relatively broad-coverage HPSG grammar, Jacy, authored by Melanie Siegel at the DFKI (Siegel 2000). Work continued at the DFKI, of course, on the German HPSG grammar, written by Stefan Müller, adapted from his earlier Babel grammars (Müller 1999), and with semantics contributed by Walter Kasper.

Before the end of Verb*mobil* funding in 2000, the LinGO project had already begun to diversify into other application and research areas using the ERG, including over the next several years work on augmented/adaptive communication, multiword expressions, and hybrid processing with statistical methods, variously funded by the National Science Foundation, the Scottish government, and industrial partners including IBM and NTT. At the turn of the millenium, Flickinger joined the software start-up boom, co-founding YY Software funded through substantial venture capital to use the ERG for automated response to customer emails for e-commerce companies. YY produced the first commercially viable software system using an HPSG implementation, processing email content in English with the ERG and the PET parser (Callmeier 2000) which had been developed by Ulrich Callmeier at the DFKI, as well as in Japanese with Jacy, further developed by Siegel and by Bender. While technically capable, the product was not commercially successful enough to enable YY to survive the bursting of the dot-com bubble, and it closed down in 2003. Flickinger returned to the LinGO project with a considerably more robust ERG, and soon picked up the translation application thread again, this time using the ERG for generation in the LOGON Norwegian-English machine translation project (Lønning et al. 2004) based in Oslo.

8 Research and Teaching Networks

The first international conference on HPSG was held in 1993 in Columbus, Ohio, in conjunction with the Linguistic Society of America's Summer Institute. The conference has been convened every year since then, with locations in Europe, Asia, and North America. Two of these annual meetings have been held jointly with the annual Lexical Functional Grammar conference, in 2000 in Berkeley and in 2016 in Warsaw. Proceedings of these conferences since 2000 are available on-line from CSLI Publications.[3] Since 2003, HPSG researchers in Europe have frequently held a regional workshop in Bremen, Berlin, Frankfurt, or Paris,

[3] http://csli-publications.stanford.edu/HPSG/, 2018-08-21.

annually since 2012, to foster informal discussion of current work in HPSG. These follow in the footsteps of European HPSG workshops starting with one on German grammar, held in Saarbrücken in 1991, and including others in Edinburgh and Copenhagen in 1994, and in Tübingen in 1995.

In 1994, the HPSG mailing list was initiated,[4] and from 1996 to 1998, the electronic newsletter, the HPSG Gazette,[5] was distributed through the list, with its function then taken over by the HPSG mailing list.

Courses introducing HPSG to students became part of the curriculum during the late 1980s and early 1990s at universities in Osaka, Paris, Saarbrücken, Seoul, and Tübingen, along with Stanford and OSU. Additional courses came to be offered in Bochum, Bremen, Carnegie-Mellon, Göttingen, Heidelberg, Jena, Leuven, Potsdam, and the University of Washington. Summer courses and workshops on HPSG have also been offered since the early 1990s at the LSA Summer Institute in the U.S., including a course by Sag and Pollard on binding and control in 1991 in Santa Cruz, and at the European Summer School in Logic, Language and Information (ESSLLI), including a course by Pollard in Saarbrücken in 1991 on HPSG, a workshop in Colchester in 1992 on HPSG, a workshop in Prague in 1996 on Romance (along with two HPSG-related student papers at the first-ever ESSLLI student session), and courses in 1998 in Saarbrücken on Germanic syntax, grammar engineering, and unification-based formalisms, in 2001 on HPSG syntax, in 2003 on linearization grammars, and more since. Also in 2001, a Scandinavian summer school on constraint-based grammar was held in Trondheim.

Several HPSG textbooks have been published, including at least Borsley (1991; 1996), Sag & Wasow (1999), Sag, Wasow & Bender (2003), Müller (2007a; 2013a), Müller (2016), Kim (2016), and Levine (2017).

9 Implementations and Applications of HPSG

The first implementation of a grammar in the HPSG framework emerged in the Hewlett-Packard Labs natural language project, for English, with a lexical type hierarchy (Flickinger, Pollard & Wasow 1985), a set of grammar rules that provided coverage of core syntactic phenomena including unbounded dependencies and coordination, and a semantic component called Natural Language Logic (Laubsch & Nerbonne 1991). The corresponding parser for this grammar was implemented in Lisp (Proudian & Pollard 1985), as part of a system called HP-NL (Nerbonne

[4] Its archives can be found at https://hpsg.hu-berlin.de/HPSG/MailingList.

[5] http://www.sfs.uni-tuebingen.de/~gazette, 2018-08-21.

& Proudian 1987) which provided a natural language interface for querying relational databases. The grammar and parser were shelved when HP Labs terminated their natural language project in 1991, leading Sag and Flickinger to begin the LinGO project and development of the English Resource Grammar at Stanford.

By this time, grammars in HPSG were being implemented in university research groups for several other languages, using a variety of parsers and engineering platforms for processing typed feature structure grammars. Early platforms included the DFKI's DISCO system (Uszkoreit et al. 1994) with a parser and graphical development tools, which evolved to the PAGE system; the ALE system (Franz 1990; Carpenter & Penn 1996), which evolved in Tübingen to TRALE (Meurers et al. 2002; Penn 2004); and Ann Copestake's LKB (Copestake 2002) which grew out of the ACQUILEX project. Other early systems included ALEP within the Eurotra project (Simpkins & Groenendijk 1994), ConTroll at Tübingen (Götz & Meurers 1997), CUF at IMS in Stuttgart (Dörre & Dorna 1993), CL-ONE at Edinburgh (Manandhar 1994), TFS also at IMS (Emele 1994), ProFIT at the University of Saarland (Erbach 1995), Babel at Humboldt University in Berlin (Müller 1996), and HDrug at Groningen (van Noord & Bouma 1997).

Relatively early broad-coverage grammar implementations in HPSG, in addition to the English Resource Grammar at Stanford (Flickinger 2000), included one for German at the DFKI (Müller & Kasper 2000) and one for Japanese (Jacy: Siegel 2000), all used in the Verb*mobil* machine translation project; a separate German grammar (Müller 1996; 1999); a Dutch grammar in Groningen (Bouma, van Noord & Malouf 2001); and a separate Japanese grammar in Tokyo (Miyao et al. 2005). Moderately large HPSG grammars were also developed during this period for Korean (Kim & Yang 2003) and for Polish (Mykowiecka, Marciniak, Przepiórkowski & Kupść 2003).

In 1999, research groups at the DFKI, Stanford, and Tokyo set up a consortium called DELPH-IN (Initiative for Deep Linguistic Processing in HPSG), to foster broader development of both grammars and platform components, described in Oepen, Flickinger, Tsujii & Uszkoreit (2002). Over the next two decades, substantial DELPH-IN grammars were developed for Norwegian (Hellan & Haugereid 2003), Portuguese (Branco & Costa 2010), and Spanish (Marimon 2010), along with moderate-coverage grammars for Bulgarian (Osenova 2011), Greek (Kordoni & Neu 2005), Hausa (Crysmann 2012), Hebrew (Arad Greshler, Herzig Sheinfux, Melnik & Wintner 2015), Indonesian (Moeljadi et al. 2015), Mandarin Chinese (Fan et al. 2015), Thai, and Wambaya (Bender 2008), all described at http://delph-in.net. Several of these grammars are based on the Grammar Matrix (Bender, Flickinger & Oepen 2002), a starter kit generalized from the ERG and Jacy for rapid prototyping of HPSG grammars, along with a much larger set of course-work grammars.[6] Out of this work has grown the linguistically rich Grammar Matrix customization system (Bender, Drellishak, Fokkens, Poulson & Saleem 2010), a set of libraries of phenomena enabling a grammar developer to complete a questionnaire about characteristics of a language to obtain a more effectively customized starting grammar.

Broad-coverage grammars developed in the TRALE system (Meurers et al. 2002; Penn 2004) include German (Müller 2007a), Danish (Müller & Ørsnes 2013), and Persian (Müller 2010). Other TRALE grammars include Mandarin Chinese (Müller & Lipenkova 2013), Georgian (Abzianidze 2011), Maltese (Müller 2009), Spanish (Machicao y Priemer 2015), and Yiddish (Müller & Ørsnes 2011). Development of grammars in TRALE is supported by the Grammix system (Müller 2007b); Müller (2015) provides a summary of this family of grammar implementations.

These grammars and systems have been used in a wide variety of applications, primarily as vehicles for research in computational linguistics, but also for some commercial software products. Research applications already mentioned include database query (HP Labs) and machine translation (Verb*mobil* and LOGON), with additional applications developed for use in anthology search (Schäfer, Kiefer, Spurk, Steffen & Wang 2011), grammar tutoring in Norwegian (Hellan, Bruland, Aamot & Sandøy 2013), ontology acquisition (Herbelot & Copestake 2006), virtual robot control (Packard 2014), visual question answering (Kuhnle & Copestake 2017), and logic instruction (Flickinger 2017), among many others. Commercial applications include e-commerce customer email response (for YY Software), and grammar correction in education (for Redbird Advanced Learning, now part of McGraw-Hill Education: Suppes, Flickinger, Macken, Cook & Liang 2012). See Bender & Emerson (2019), Chapter H of this volume for further discussion.

For most practical applications, some approximate solution to the challenge of parse selection (disambiguation) must be provided, so developers of several of the DELPH-IN grammars, including the ERG, follow the approach of Oepen, Flickinger, Toutanova & Manning (2004), which uses a manually-annotated tree-bank of sentences parsed by a grammar to train a statistical model which is applied at run-time to identify the most likely analysis for each parsed sentence. These treebanks can also serve as repositories of the analyses intended by the grammarian for the sentences of a corpus, and some resources, notably the Alpino Treebank (Bouma, van Noord & Malouf 2001), include analyses which the gram-

[6] http://moin.delph-in.net/MatrixTop, 2018-08-21.

mar may not yet be able to produce automatically.

10 Prospects

As we noted early in this chapter, HPSG's origins are rooted in the desire simultaneously to address the theoretical concerns of linguists and the practical issues involved in building a useful natural language processing system. In the decades since the birth of HPSG, the mainstream of work in both theoretical linguistics and NLP developed in ways that could not have been anticipated at the time. NLP is now dominated by statistical methods, with almost all practical applications making use of machine learning technologies. It is hard to see any influence of research by linguists in most NLP systems, though periodic workshops have helped to keep the conversation going.[7] Mainstream grammatical theory, on the other hand, is now dominated by the Minimalist Program (MP), which is too vaguely formulated for a rigorous comparison with HPSG.[8] Concern with computational implementation plays virtually no role in MP research; see Müller (2016) for a discussion.

It might seem, therefore, that HPSG is further from the mainstream of both fields than it was at its inception, raising questions about how realistic the objectives of HPSG are. We believe, however, that there are grounds for optimism.

With regard to implementations, there is no incompatibility between the use of HPSG and the machine learning methods of mainstream NLP. Indeed, as noted above, HPSG-based systems that have been put to practical use have necessarily included components induced via statistical methods from annotated corpora. Without such components, the systems cannot deal with the full variety of forms encountered in usage data. On the other hand, existing NLP systems that rely solely on machine learning from corpora do not exhibit anything that can reasonably be called understanding of natural language. Current technologies for machine translation, automatic summarization, and various other linguistic tasks fall far short of what humans do on these tasks, and are useful primarily as tools to speed up the tasks for the humans carrying them out. Many NLP researchers are beginning to recognize that developing software that can plausibly be said

[7]For example, one on "Building Linguistically Generalizable NLP Systems" at the 2017 EMNLP conference in Copenhagen, and one on "Relevance of Linguistic Structure in Neural NLP" at the 2018 ACL conference in Melbourne.

[8]Most work in MP is presented without precise definitions of the technical apparatus, but Edward Stabler and his collaborators have written a number of papers aimed at formalizing MP. See in particular Collins & Stabler (2016).

to understand language will require representations of linguistic structure and meaning like those that are the stock in trade of linguists. See Bender, Flickinger, Oepen, Packard & Copestake (2015) for more discussion on sentence meaning.

Evidence for a renewed interest in linguistics among NLP researchers is the fact that major technology companies with natural language groups have recently begun (or in some cases, resumed) hiring linguists, and increasing numbers of new linguistics PhDs have taken jobs in the software industry.

In the domain of theoretical linguistics, it is arguable that the distance between HPSG and the mainstream of grammatical research (that is, MP) has narrowed, given that both crucially incorporate ideas from Categorial Grammar (see Retoré & Stabler 2004, Berwick & Epstein 1995, and Müller 2013b for comparisons between MP and CG). Rather than trying to make that argument, however, we will point to connections that HPSG has made with other work in theoretical linguistics. Perhaps the most obvious of these is the work of Peter Culicover and Ray Jackendoff on what they call *Simpler Syntax*. Their influential 2005 book with that title (Culicover & Jackendoff 2005) argues for a theory of grammar that differs little in its architecture and motivations from HPSG.

More interesting are the connections that have been forged between research in HPSG and work in Construction Grammar (CxG). Fillmore (1988: 36) characterizes the notion of *construction* as "any syntactic pattern which is assigned one or more conventional functions in a language, together with whatever is linguistically conventionalized about its contribution to the meaning or use of structures containing it." Among the examples that construction grammarians have described at length are *the Xer, the Yer* (as in *the older I get, the longer I sleep*), *X let alone Y* (as in *I barely got up in time to eat lunch, let alone cook breakfast*), and *What's X doing Y?* (as in *What's this scratch doing in the table?*). As noted above and in Müller (2019), Chapter P of this volume, HPSG has incorporated the notion of construction since at least the late 1990s.

Nevertheless, work that labels itself CxG tends to look very different from HPSG. This is in part because of the difference in their origins: many proponents of CxG come from the tradition of Cognitive Grammar or typological studies, whereas HPSG's roots are in computational concerns. Hence, most of the CxG literature is not precise enough to allow a straightforward comparison with HPSG, though the variants called Embodied Construction Grammar and Fluid Construction Grammar have more in common with HPSG; see Müller (2017) for a comparison. In the last years of his life, Ivan Sag sought to unify CxG and HPSG through collaboration with construction grammarians from the University of California at Berkeley, particularly Charles Fillmore, Paul Kay, and Laura Michaelis.

They developed a theory called *Sign-Based Construction Grammar* (SBCG), which would combine the insights of CxG with the explicitness of HPSG. Sag (2012: 70) wrote, "To readers steeped in HPSG theory, SBCG will no doubt seem like a minor variant of constructional HPSG." Indeed, despite the name change, the main feature of SBCG that differs from HPSG is that it posits an inheritance hierarchy of constructs, which includes feature structure descriptions for such partially lexicalized multi-word expressions as *Ved X's way PP*, instantiated in such VPs as *ad-libbed his way through a largely secret meeting*. While this is a non-trivial extension to HPSG, there is no fundamental change to the technical machinery. In fact, it has been a part of the LinGO implementation for many years.

That said, there is one important theoretical issue that divides HPSG and SBCG from much other work in CxG. That issue is locality. To constrain the formal power of the theory, and to facilitate computational tractability, SBCG adopts what Sag (2012: 150) calls "Constructional Localism" and describes it as follows: "Constructions license mother-daughter configurations without reference to embedding or embedded contexts." That is, like phrase structure rules, constructions must be characterized in terms of a mother node and its immediate daughters. At first glance, this seems to rule out analyses of many of the examples of constructions provided in the CxG literature. But Sag (2012: 150) goes on to say, "Constructional Localism does not preclude an account of nonlocal dependencies in grammar, it simply requires that all such dependencies be locally encoded in signs in such a way that information about a distal element can be accessed locally at a higher level of structure."

Fillmore (1988: 35) wrote:

> Construction grammars differ from phrase-structure grammars which use *complex symbols* and allow the *transmission of information* between lower and higher structural units, in that we allow the direct representation of the required properties of subordinate constituents. (Should it turn out that there are completely general principles for predicting the kinds of information that get transmitted upwards or downwards, this may not be a real difference.) (Fillmore 1988: 35)

SBCG is committed to the position alluded to in the parenthetical sentence in this quote, namely, that general principles of information transmission within sentences makes it possible to insist on Constructional Localism. See Müller (2019), Chapter P of this volume for a much more detailed discussion, and Van Eynde (2015) for a review of the 2012 SBCG book.

Finally, another point of convergence between work in HPSG and other work in both theoretical linguistics and NLP is the increasing importance of corpus data. In the early years of the HP NLP project, the methodology was the same as that employed in almost all work in theoretical syntax and semantics: the grammar was based entirely on examples invented by the researchers. At one point during the decade of the HP NLP project, Flickinger, Pullum, and Wasow compiled a list of sentences intended to exemplify many of the sentence types that they hoped the system would eventually be able to analyze. That list, 1328 sentences long, continues to be useful as a test suite for the ERG, and is also used by various other NLP groups. But it does not come close to covering the variety of sentence forms that are found in corpora of speech and various written genres. As the goals of the HPSG implementations have broadened from database query to dealing with "language in the wild", the use of corpora to test such systems and motivate extensions to them has increased. This parallels a development in other areas of linguistics, which have also increasingly made use of large on-line corpora as sources of data and tests of their theories. This is a trend that we expect will continue.

In short, there are signs of convergence between work on HPSG and work in other areas, and it seems plausible to think that the market for HPSG research will grow in the future.

Acknowledgements

The work on this chapter by Flickinger was generously supported by a fellowship at the Oslo Center for Advanced Study at the Norwegian Academy of Science and Letters. The authors also thank several reviewers for their insightful and detailed comments on drafts of the chapter, including Emily Bender, Robert Borsley, Danielle Godard, and Stefan Müller.

References

Abzianidze, Lasha. 2011. *An HPSG-based formal grammar of a core fragment of Georgian implemented in TRALE*. Charles University in Prague MA thesis.

Arad Greshler, Tali, Livnat Herzig Sheinfux, Nurit Melnik & Shuly Wintner. 2015. Development of maximally reusable grammars: Parallel development of Hebrew and Arabic grammars. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar, Nanyang Technological University (NTU), Singapore*, 27–40. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2015/ahmw.pdf, accessed 2018-2-25.

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 4(10). 515–531.

Bach, Emmon. 1980. In defense of passive. *Linguistics and Philosophy* (3). 297–342.

Bach, Emmon. 1983. On the relationship between Word-Grammar and Phrase-Grammar. *Natural Language and Linguistic Theory* 1(1). 65–89.

Bar-Hillel, Yehoshua. 1954. Logical syntax and semantics. *Language* 30(2). 230–237. DOI:10.2307/410265

Bender, Emily M. 2008. Evaluating a crosslinguistic grammar resource: A case study of Wambaya. In Johanna D. Moore, Simone Teufel, James Allan & Sadaoki Furui (eds.), *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 977–985. Columbus, Ohio: Association for Computational Linguistics. http://aclweb.org/anthology/P08-1111, accessed 2018-2-25.

Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson & Safiyyah Saleem. 2010. Grammar customization. *Research on Language & Computation* 8(1). 23–72. http://dx.doi.org/10.1007/s11168-010-9070-1. 10.1007/s11168-010-9070-1.

Bender, Emily M. & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 797–840. Berlin: Language Science Press. DOI:??

Bender, Emily M., Dan Flickinger, Stephan Oepen, Woodley Packard & Ann A. Copestake. 2015. Layers of interpretation: on grammar and compositionality. In *Iwcs*.

Bender, Emily M., Daniel P. Flickinger & Stephan Oepen. 2002. The Grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, 8–14. Taipei, Taiwan.

Berwick, Robert C. & Samuel David Epstein. 1995. On the convergence of 'Minimalist' Syntax and Categorial Grammar. In Anton Nijholt, Giuseppe Scollo & Rene Steetskamp (eds.), *Algebraic methods in language processing*, 143–148. Enschede: University of Twente. http://eprints.eemcs.utwente.nl/9555/01/twlt10.pdf, accessed 2018-9-25.

Bloomfield, Leonard. 1933. *Language*. London: George Allen & Unwin.

Borsley, Bob & Anne Abeillé. 2019. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 3–5. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 1987. *Subjects and complements in HPSG*. Report No. CSLI-87-107. Stanford, CA: Center for the Study of Language & Information.

Borsley, Robert D. 1988. Subjects, complements, and specifiers in HPSG. unpublished paper, University of Wales, Bangor.

Borsley, Robert D. 1989. Phrase-Structure Grammar and the Barriers conception of clause structure. *Linguistics* 27(5). 843–863.

Borsley, Robert D. 1990. Welsh passives. In Martin J. Ball, James Fife, Erich Poppe & Jenny Rowland (eds.), *Celtic linguistics: readings in the bretonic languages, a festschrift for t. arwynwatkins* (Current Issues in Linguistic Theory 68), 89–107. Amsterdam: John Benjamins Publishing Co.

Borsley, Robert D. 1991. *Syntactic theory: A unified approach*. London: Edward Arnold.

Borsley, Robert D. 1996. *Modern Phrase Structure Grammar* (Blackwell Textbooks in Linguistics 11). Oxford: Blackwell Publishers Ltd.

Bouma, Gosse, Gertjan van Noord & Robert Malouf. 2001. Alpino: wide-coverage computational analysis of Dutch. In Walter Daelemans, Khalil Sima'an, Jorn Veenstra & Jakub Zavrel (eds.), *Computational linguistics in the Netherlands 2000: Selected papers from the Eleventh CLIN Meeting* (Language and Computers 37). Amsterdam/New York, NY: Rodopi.

Branco, António & Francisco Costa. 2010. A deep linguistic processing grammar for Portuguese. In *Computational processing of the portuguese language*, vol. LNAI6001 (Lecture Notes in Artificial Intelligence), 86
bibrangessep–
*bibrangessep*89. Berlin: Springer.

Bresnan, Joan. 1978. A realistic Transformational Grammar. In Morris Halle, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, 1–59. Cambridge, MA: MIT Press.

Callmeier, Ulrich. 2000. PET: A platform for experimentation with efficient HPSG processing techniques. *Natural Language Engineering* 6(1). 99–108. Special Issue on Efficient Processing with HPSG: Methods, Systems, Evaluation.

Carpenter, Bob & Gerald Penn. 1996. Efficient parsing of compiled typed attribute value logic grammars. In Harry Bunt & Masaru Tomita (eds.), *Recent advances in parsing technology* (Text, Speech and Language Technology 1), 145–168. Dordrecht: Kluwer Academic Publishers.

Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). The Hague/Paris: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart & Winston.

Chomsky, Noam & Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8. 425–504.

Collins, Chris & Edward Stabler. 2016. A Formalization of Minimalist Syntax. *Syntax* 19(1). 43–78.

Copestake, Ann. 2002. *Implementing typed feature structure grammars*. Stanford, CA: CSLI.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Creary, Lewis G. & Carl J. Pollard. 1985. A computational semantics for natural language. In *Proceedings of the 23rd annual meeting of the association for computational linguistics*, 172–179. Chicago, Illinois, USA: Association for Computational Linguistics. http://www.aclweb.org/anthology/P85-1022. DOI:10.3115/981210.981232

Crysmann, Berthold. 2012. HaG: A computational grammar of Hausa. In Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison & Tristan M. Purvis (eds.), *Selected proceedings of the 42nd Annual Conference on African Linguistics (ACAL 42)*, 321–337. Somerville, MA: Cascadilla Press. http://www.lingref.com/cpp/acal/42/paper2780.pdf, accessed 2018-2-25.

Culicover, Peter W. & Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Culy, Christopher. 1985. The complexity of the vocabulary of Bambara. *Linguistics and Philosophy* 8(3). 345–351. DOI:10.1007/BF00630918

Dörre, Jochen & Michael Dorna. 1993. *CUF: A formalism for linguistic knowledge representation*. DYANA 2 deliverable R.1.2A. Stuttgart, Germany: IMS.

Dowty, David R. 1978. Governed transformations as lexical rules in a Montague Grammar. *Linguistic Inquiry* 9(3). 393–426.

Dowty, David R. 1979. *Word meaning and Montague Grammar* (Synthese Language Library 7). Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1982a. Grammatical relations and Montague Grammar. In Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic represen-*

tation (Synthese Language Library 15). Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1982b. More on the categorial analysis of grammatical relations. In Annie Zaenen (ed.), *Subjects and other subjects: Proceedings of the Harvard Conference on the Representation of Grammatical Relations*, 115–153. Bloomington: Indiana University Linguistics Club.

Emele, Martin C. 1994. The typed feature structure representation formalism. In *Proceedings of the international workshop on sharable natural language resources, August 10–11, 1994. Ikoma, Nara, Japan.*

Emonds, Joseph E. 1976. *A transformational approach to English syntax*. New York: Academic Press.

Erbach, Gregor. 1995. ProFIT: prolog with features, inheritance and templates. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, 180–187. Dublin: Association for Computational Linguistics.

Fan, Zhenzhen, Sanghoun Song & Francis Bond. 2015. An HPSG-based shared-grammar for the Chinese languages: ZHONG [|]. In Emily M. Bender, Lori Levin, Stefan Müller, Yannick Parmentier & Aarne Ranta (eds.), *Proceedings of the Grammar Engineering Across Frameworks (GEAF) Workshop*, 17–24. The Association for Computational Linguistics.

Fillmore, Charles J. 1988. The mechanisms of "Construction Grammar". In Shelley Axmaker, Annie Jaisser & Helen Singmaster (eds.), *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, 35–55. Berkeley, CA: Berkeley Linguistics Society.

Flickinger, Daniel. 1983. Lexical heads and phrasal gaps. In *Proceedings of the west coast conference on formal linguistics*, 89–101. Stanford.

Flickinger, Daniel. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6(01). 15–28.

Flickinger, Daniel. 2017. Generating English paraphrases from logic. In Martijn Wieling, Gosse Bouma & Gertjan van Noord (eds.), *From semantics to dialectometry: festschrift in honour of john nerbonne*, 99–108. Springer.

Flickinger, Daniel P. 1987. *Lexical rules in the hierarchical lexicon*. Stanford University dissertation.

Flickinger, Daniel P., Ann Copestake & Ivan A. Sag. 2000. HPSG analysis of English. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 254–263. Berlin: Springer Verlag.

Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-*

Third Annual Meeting of the Association for Computational Linguistics, 262–267. Chicago, IL.

Franz, Alex. 1990. *A parser for HPSG*. Laboratory for Computational Linguistics Report CMU-LCL-90-3. Carnegie Mellon University: Laboratory for Computational Linguistics.

Gawron, Jean Mark, Jonathan King, John Lamping, Egon Loebner, E. Anne Paulson, Geoffrey K. Pullum, Ivan A. Sag & Thomas Wasow. 1982. Processing English with a Generalized Phrase Structure Grammar. In *Proceedings of the 20th annual meeting of the Association for Computational Linguistics*, 74–81.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.

Gazdar, Gerald. 1988. Applicability of Indexed Grammars to natural language. In Uwe Reyle & Christian Rohrer (eds.), *Natural language parsing and linguistic theories* (Studies in Linguistics and Philosophy 35), 69–94. Dordrecht/Boston/Lancaster/Tokyo: D. Reidel Publishing Company.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Götz, Thilo & Detmar Meurers. 1997. The ConTroll system as large grammar development platform. In *Proceedings of the Workshop on Computational Environments for Grammar Development and Linguistic Engineering*, 88–97. Madrid: COLIPS Publications.

Hellan, Lars, Tore Bruland, Elias Aamot & Mads H. Sandøy. 2013. A Grammar Sparrer for Norwegian. In *Proceedings of the Nordic Conference on Computational Linguistcs (NoDaLiDa)*, 435–441. Oslo, Norway.

Hellan, Lars & Petter Haugereid. 2003. Norsource – An excercise in the Matrix Grammar building design. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the ESSLLI 2003 Workshop "Ideas and Strategies for Multilingual Grammar Development"*. Vienna, Austria.

Herbelot, Aurélie & Ann Copestake. 2006. Acquiring Ontological Relationships from Wikipedia using RMRS. In *Proceedings of the ISWC 2006 workshop on web content*. Athens, Georgia.

Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10. 210–234.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.

Jackendoff, Ray S. 1975. Morphological and semantic regularities in the lexikon. *Language* 51(3). 639–671.

Joshi, Aravind K. 1987. Introduction to Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *The mathematics of language*, 87–114. Amsterdam: John Benjamins Publishing Co.

Joshi, Aravind K., Leon S. Levy & Masako Takahashi. 1975. Tree Adjunct Grammar. *Journal of Computer and System Science* 10(2). 136–163.

Kasami, Tadao, Hiroyuki Seki & Mamoru Fujii. 1989. Generalized context-free grammars and multiple context-free grammars. *Systems and Computers in Japan* 20(7). 43–52.

Kathol, Andreas. 1995. *Linearization-based German syntax*. Ohio State University dissertation.

Kathol, Andreas. 2000. *Linear syntax*. New York, Oxford: Oxford University Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

Kim, Jong-Bok. 2016. *The syntactic structures of Korean: A Construction Grammar perspective*. Cambridge, UK: Cambridge University Press.

Kim, Jong-Bok & Jaehyung Yang. 2003. Korean phrase structure grammar and its implementations into the LKB system. In Dong Hong Ji & Kim Teng Lua (eds.), *Proceedings of the 17th Pacific Asia Conference on Language, Information and Computation*, 88–97. National University of Singapore: COLIPS Publications.

King, Paul. 1989. *A logical formalism for Head-Driven Phrase Structure Grammar*. University of Manchester dissertation.

Kordoni, Valia & Julia Neu. 2005. Deep analysis of Modern Greek. In Keh-Yih Su, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.), *Natural language processing IJCNLP 2004* (Lecture Notes in Artificial Intelligence 3248), 674–683. Berlin: Springer Verlag.

Kuhnle, Alexander & Ann A. Copestake. 2017. Shapeworld - A new test methodology for multimodal language understanding. *CoRR* abs/1704.04517. http://arxiv.org/abs/1704.04517.

Lambek, Joachim. 1958. The Mathematics of Sentence Structure. *The American Mathematical Monthly* 65(3). 154–170.

Laubsch, Joachim & John Nerbonne. 1991. *An Overview of NLL*. Tech. rep. Hewlett-Packard Laboratories.

Levine, Robert D. 2017. *Syntactic analysis: An HPSG-based approach*. Cambridge, UK: Cambridge University Press.

Lønning, Jan Tore, Stephan Oepen, Dorothee Beermann, Lars Hellan, John Carroll, Helge Dyvik, Dan Flickinger, Janne Bondi Johannessen, Paul Meurer, Torbjørn Nordgård, Victoria Rosén & Erik Velldal. 2004. LOGON. A Norwegian MT effort. In *Proceedings of the Workshop in Recent Advances in Scandinavian Machine Translation*. Uppsala, Sweden.

Machicao y Priemer, Antonio. 2015. *SpaGram: An implemented grammar fragment of Spanish*. Ms. Humboldt Universität zu Berlin. In Preparation.

Manandhar, Suresh. 1994. *User's Guide for CL-ONE*. Technical Report. University of Edinburgh.

Marimon, Montserrat. 2010. The spanish resource grammar. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the seventh international conference on language resources and evaluation (lrec'10)*. Valletta, Malta: European Language Resources Association (ELRA).

Meurers, Walt Detmar, Gerald Penn & Frank Richter. 2002. A web-based instructional platform for constraint-based grammar formalisms and parsing. In Dragomir Radev & Chris Brew (eds.), *Effective tools and methodologies for teaching NLP and CL*, 18–25. Association for Computational Linguistics. Proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, PA.

Minsky, Marvin. 1975. A framework for representing knowledge. In Patrick Winston (ed.), *The psychology of computer vision*, 211–277. New York: McGraw-Hill.

Miyao, Yusuke, Takashi Ninomiya & Jun'ichi Tsujii. 2005. Corpus-oriented grammar development for acquiring a Head-Driven Phrase Structure Grammar from the Penn Treebank. In Keh-Yih Su, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.), *Natural language processing IJCNLP 2004* (Lecture Notes in Artificial Intelligence 3248), 684–693. Berlin: Springer Verlag.

Moeljadi, David, Francis Bond & Sanghoun Song. 2015. Building an HPSG-based Indonesian resource grammar. In Emily M. Bender, Lori Levin, Stefan Müller, Yannick Parmentier & Aarne Ranta (eds.), *Proceedings of the Grammar Engineering Across Frameworks (GEAF) Workshop*, 9–16. The Association for Computational Linguistics.

Montague, Richard. 1974. *Formal philosophy: Selected papers of Richard Montague*. Richmond H. Thomason (ed.). Yale University Press.

Moortgat, Michael. 1989. *Categorical investigations: Logical and linguistic aspects of the Lambek Calculus* (Groningen Amsterdam Studies in Semantics 9). Dordrecht/Cinnaminson, U.S.A.: Foris Publications.

Morrill, Glyn V. 1994. *Type Logical Grammars: Categorial logic of signs*. Dordrecht: Kluwer Academic Publishers.

Moshier, M. Drew & William C. Rounds. 1987. A logic for partially specified data structures. In *Proceedings of the 14th ACM Symposium on Principles of Programming Languages*, 156–167. Munich, Germany.

Müller, Stefan. 1996. The Babel-System: An HPSG fragment for German, a parser, and a dialogue component. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, 263–277. London.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2007a. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 1st edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2007b. The Grammix CD Rom: A software collection for developing typed feature structure grammars. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar Engineering across Frameworks 2007* (Studies in Computational Linguistics ONLINE). Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/, accessed 2018-2-25.

Müller, Stefan. 2009. A Head-Driven Phrase Structure Grammar for Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics: Papers from the 1st International Conference on Maltese Linguistics (Bremen/Germany, 18–20 October, 2007)* (Studies in Language Companion Series 113), 83–112. Amsterdam: John Benjamins Publishing Co.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. DOI:10.1017/S0022226709990284

Müller, Stefan. 2013a. *Grammatiktheorie*. 2nd edn. (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag. https://hpsg.hu-berlin.de/~stefan/Pub/grammatiktheorie.html, accessed 2019-6-30.

Müller, Stefan. 2013b. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950. DOI:10.1353/lan.2013.0061

Müller, Stefan. 2015. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.91

Müller, Stefan. 2016. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2017. Head-Driven Phrase Structure Grammar, Sign-Based Construction Grammar, and Fluid Construction Grammar: Commonalities and differences. *Constructions and Frames* 9(1). 139–174. DOI:10.1075/cf.9.1.05mul

Müller, Stefan. 2019. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

Müller, Stefan & Walter Kasper. 2000. HPSG analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: foundations of speech-to-speech translation* (Artificial Intelligence), 238–253. Berlin: Springer Verlag.

Müller, Stefan & Janna Lipenkova. 2013. ChinGram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In Huei-ling Lai & Kawai Chui (eds.), *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, 240–249. Taipei, Taiwan: Department of English, National Chengchi University.

Müller, Stefan & Bjarne Ørsnes. 2011. Positional expletives in Danish, German, and Yiddish. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington, U.S.A.* 167–187. Stanford, CA: CSLI Publications.

Müller, Stefan & Bjarne Ørsnes. 2013. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In preparation.

Müller, Stefan & Stephen Mark Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Mykowiecka, Agnieszka, Malgorzata Marciniak, Adam Przepiórkowski & Anna Kupść. 2003. An implementation of a generative grammar of Polish. In Peter Kosta, Joanna Blaszczak, Jens Frasek, Ljudmila Geist & Marzena Zygis (eds.), *Investigations into formal Slavic linguistics: contributions of the fourth European conference on formal description of Slavic languages – FDSL IV held at Potsdam University, November 28–30, 2001*, 271–285. Frankfurt am Main: Peter Lang.

Nerbonne, John & Derek Proudian. 1987. *The HP-NL System*. Tech. rep. Hewlett-Packard Laboratories. URL: http://urd.let.rug.nl/nerbonne/papers/Old-Scans/HP-NL-System-Nerbonne-Proudian-1987.pdf.

Newmeyer, Frederick J. 1980. *Linguistic theory in America: The first quarter-century of transformational generative grammar*. New York: Academic Press.

Oehrle, Richard, Emmon Bach & Deirdre Wheeler (eds.). 1988. *Categorial Grammars and natural language structures*. Dordrecht: D. Reidel Publishing Company.

Oepen, Stephan, Daniel P. Flickinger, Kristina Toutanova & Christopher D. Manning. 2004. LinGO Redwoods: A rich and dynamic treebank for HPSG. *Research on Language and Computation* 2(4). 575–596.

Oepen, Stephan, Daniel P. Flickinger, Jun-ichi Tsujii & Hans Uszkoreit. 2002. *Collaborative language engineering: A case study in efficient grammar-based processing* (CSLI Lecture Notes 118). Stanford, CA: CSLI Publications.

Osenova, Petya. 2011. Localizing a core HPSG-based grammar for Bulgarian. In Hanna Hedeland, Thomas Schmidt & Kai Worner (eds.), *Multilingual resources and multilingual applications*, vol. GSCL 2011, 175
bibrangessep–
*bibrangessep*180. Hamburg.

Packard, Woodley. 2014. UW-MRS: leveraging a deep grammar for robotic spatial commands. *SemEval 2014*. 812.

Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In Donia Scott (ed.), *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), main volume*, 239–246. Barcelona, Spain.

Peters, Stanley & R. W. Ritchie. 1973. On the generative power of Transformational Grammar. *Information Sciences* 6(C). 49–83. DOI:10.1016/0020-0255(73)90027-3

Pollard, Carl J. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and natural language*. Stanford University dissertation.

Pollard, Carl J. 1985. Phrase Structure Grammar without metarules. In Jeffrey Goldberg, Susannah MacKaye & Michael Wescoat (eds.), *WCCFL 4: The proceedings of the Fourth West Coast Conference on Formal Linguistics*, 246–. Stanford Linguistics Association.

Pollard, Carl J. 1986. HPSG tutorial. Handout for tutorial presented at ICOT in Tokyo, in July 1986.

Pollard, Carl J. 1988. Categorial Grammar and Phrase Structure Grammar: An excursion on the syntax-semantics frontier. In Richard Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial Grammars and natural language structures*, 391–415. Dordrecht: D. Reidel Publishing Company.

Pollard, Carl J. 2013. Traces exist (hypothetically)! https://web.stanford.edu/dept/linguistics/structures_evidence/website/docs/slides/pollard.pdf, accessed

2018-9-10. Presented at the Workshop on Structure and Evidence in Linguistics, Stanford.

Pollard, Carl J. & Ivan A. Sag. 1983. Reflexives and reciprocals in english: an alternative to the binding theory. In *WCCFL83*, 189–203.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Proudian, Derek & Carl J. Pollard. 1985. Parsing Head-Driven Phrase Structure Grammar. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 167–171. Chicago, IL.

Pullum, Geoffrey K. & Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy* 4(4). 471–504. DOI:10.1007/BF00360802

Reape, Mike. 1992. *A formal theory of word order: A case study in West Germanic*. University of Edinburgh dissertation.

Reape, Mike. 1996. Getting things in order. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 209–253. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

Retoré, Christian & Edward Stabler. 2004. Generative Grammars in Resource Logics. *Research on Language and Computation*. 3–25.

Ristad, Eric Sven. 1986. Computational complexity of current GPSG theory. In Alan W. Biermann (ed.), *Proceedings of the Twenty-Fourth Annual Meeting of the Association for Computational Linguistics*, 30–39. Columbia University, New York.

Roach, Kelly. 1987. Formal properties of head grammars. In Alexis Manaster-Ramer (ed.), *Mathematics of language*, 293–348. Amsterdam: John Benjamins.

Rounds, William C. & Robert T. Kasper. 1986. A complete logical calculus for record structures representing linguistic information. In *Proceedings of the 15th Annual IEEE Symposium on Logic in Computer Science*, 38–43. Cambridge, MA.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.

Sag, Ivan A. & Thomas Wasow. 1999. *Syntactic theory: A formal introduction*. Stanford, CA: CSLI.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Schäfer, Ulrich, Bernd Kiefer, Christian Spurk, Jörg Steffen & Rui Wang. 2011. The ACL Anthology Searchbench. In *Proceedings of the ACL-HLT 2011 system demonstrations*, 7–13. Portland, Oregon, USA.

Scott, Dana S. 1982. Domains for denotational semantics. In M. Nielson & E. M. Schmidt (eds.), *International colloquium on automata, languages and programs, volume 140 of lecture notes in computer science*, 577–613. Springer Verlag.

Shabes, Yves. 1990. *Mathematical and computational aspects of lexicalized grammars*. University of Pennsylvania dissertation.

Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8(3). 333–343.

Siegel, Melanie. 2000. HPSG analysis of Japanese. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 264–279. Berlin: Springer Verlag.

Simpkins, Neil & Marius Groenendijk. 1994. *The ALEP project*. Technical Report. CEC, Luxembourg: Cray Systems.

Steedman, Mark J. 1987. Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory* 5(3). 403–439. DOI:10.1007/BF00134555

Steedman, Mark J. 1990. Gapping as constituent coordination. *Linguistics and Philosophy* 13(2). 207–265.

Suppes, Patrick, Dan Flickinger, Elizabeth Macken, Jeanette Cook & L. Liang. 2012. Description of the EPGY Stanford University online courses for Mathematics and the Language Arts. In *Proceedings of the international society for technology in education*. San Diego, California.

Uszkoreit, Hans, Rolf Backofen, Stephan Busemann, Abdel Kader Diagne, Elizabeth A. Hinkelman, Walter Kasper, Bernd Kiefer, Hans-Ulrich Krieger, Klaus Netter, Günter Neumann, Stephan Oepen & Stephen P. Spackman. 1994. DISCO—An HPSG-based NLP system and its application for appointment scheduling. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 436–440. Kyoto, Japan: Association for Computational Linguistics.

Uszkoreit, Hans & Stanley Peters. 1982. Essential variables in metarules. Paper presented at the annual meeting of the Linguistic Society of America.

Van Eynde, Frank. 2015. Sign-Based Construction Grammar: A guided tour. *Journal of Linguistics*. DOI:10.1017/S0022226715000341

van Noord, Gertjan & Gosse Bouma. 1997. HDRUG: A flexible and extendible development environment for natural language processing. In *Proceedings of the* *workshop "Computational Environments for Grammar Developmentand Linguistic Engineering (ENVGRAM)" held in conjunction with the 35th Annua lMeeting of the ACL and 8th Conference of the EACL*, 91–98. Madrid, Spain: Universidad Nacional de Educación a Distancia.

Vijay-shanker, K. & David J. Weir. 1994. The equivalence of four extensions of context-free grammars. *Mathematical Systems Theory* 27. 511–546.

Wahlster, Wolfgang (ed.). 2000. *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence). Berlin: Springer Verlag.

Chapter 3

Formal Background

Frank Richter
Goethe-Universität Frankfurt

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse

Frank Richter. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 45–46. Berlin: Language Science Press. DOI:??

rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 4

The nature and role of the lexicon in HPSG

Anthony Davis
Southern Oregon University

Jean-Pierre Koenig
University at Buffalo

This chapter discusses the critical role the lexicon plays in HPSG and the approach to lexical knowledge that is specific to HPSG. We describe the tenets of lexicalism in general, and discuss the nature and content of lexical entries in HPSG. As a lexicalist theory, HPSG treats lexical entries as informationally rich, representing the combinatorial properties of words as well as their part of speech, phonology, and semantics. Thus many phenomena receive a lexically-based account, including some that go beyond what is typically regarded as lexical. We turn next to the global structure of the HPSG lexicon, the hierarchical lexicon and inheritance. We show how the extensive type hierarchy employed in HPSG accounts for lexical generalizations at various levels and discuss some of the advantages of default (nonmonotonic) inheritance over simple monotonic inheritance. We then describe lexical rules and their various proposed uses in HPSG, comparing them to alternative approaches to relate lexemes and words based on the same root or stem.

1 Introduction

The nature, structure, and role of the lexicon in the grammar of natural languages has been a subject of debate for at least the last 50 years. For some, the lexicon is a prison that "contains only the lawless", to borrow a memorable phrase from Di Sciullo & Williams (1987: p.3), and not much of interest resides there. In some recent views, the lexicon records merely phonological information and some world

Anthony Davis & Jean-Pierre Koenig. 2019. The nature and role of the lexicon in HPSG. in Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 47–92. Berlin: Language Science Press. DOI:??

knowledge about each lexical entry (see Marantz 1997). All of the action is in the syntax, save the expression of complex syntactic objects as inflected words. In contrast, lexicalist theories of grammar, and HPSG in particular, posit a rich and complex lexicon embodying much of grammatical knowledge.

This chapter has two principal goals. One is to review the arguments for and against a lexicalist view of grammar within the generative tradition. The other is to survey the HPSG implementation of lexicalism. In regard to the first goal, we begin with the reaction to Generative Semantics, and note developments that led to lexicalist theories of grammar such as Lexical Functional Grammar (LFG|see also Lexical Functional Grammar) and then HPSG. Central to these developments was the argument that lexical processes, rather than transformational ones, provided more perspicuous accounts of derivational morphological processes. The same kinds of arguments then naturally extended to phenomena like passivization, which had previously been treated as syntactic. Once on this path, lexical treatments of other prototypically syntactic phenomena — long distance extraction, *wh*-movement, word order, and anaphoric binding — were advanced as well, with HPSG playing a leading role.

But this does not mean that opposition to lexicalism melted away. Both Minimalism, and in particular Distributed Morphology, (Bruening 2018; Marantz 1997), and Construction Grammar (Goldberg 1995; Tomasello 2003; van Trijp 2011) claim that lexicalist accounts fail in various ways. We discuss some of these current issues, including the apparent occurrence of syntactically complex structures in the lexicon, word-internal ellipsis, and endoclitics, each of which poses challenges for those who advocate a strict separation between lexical and syntactic processes. While we maintain that the anti-lexicalist arguments are not especially strong, and the phenomena they are based somewhat marginal, we acknowledge that these questions are not yet settled. We then turn to the specifics of the lexicon as modeled within HPSG. Lexicalism demands, of course, that lexical entries be informationally rich, encoding not merely idiosyncratic properties of a single lexical item like its phonology and semantics, but also more general characteristics like its combinatorial possibilities. We outline what HPSG lexical entries must contain, and how that information is represented. This leads naturally to the next topic: with so much information in a lexical entry, and so much of that repeated in similar ones, how is massive redundancy avoided? The hierarchical lexicon, in which individual lexical entries are the leaves of a multiple inheritance hierarchy, is a core component of HPSG. Types throughout the hierarchy capture information common to classes of lexical entries, thereby allowing researchers to express generalizations at various levels. Just as all verbs share certain properties, all transitive verbs, all verbs of caused motion, and all transitive verbs of caused motion share additional properties, represented as constraints on types within the hierarchy. We draw on examples from linking, gerunds, and passive constructions as illustrations, but many others could be added.

Constraints specified on types in the hierarchy are deemed to be inherited by their subtypes, but monotonic inheritance of this kind runs into vexing issues. Most obviously, there are irregular morphological forms; any attempt to represent, say, the phonology of English plurals, as a constraint on a plural noun class in the hierarchical lexicon must then explain why the plural of *child* is *children* and not **childs*. Beyond this simple example, there are ubiquitous cases of lexical generalizations that are true by default, but not always. Various mechanisms for modeling default inheritance have therefore been one focus within HPSG, and we furnish an example of their use in modeling the properties of gerunds in English and other languages.

Finally, we discuss lexical rules and their alternatives. Along with the "vertical" relationships between classes of lexical entries modeled by types and their subtypes in the hierarchical lexicon, there is a need for "horizontal" relationships between lexical entries that are based on a single root or stem, such as forms of inflectional paradigms. Yet formalizing lexical rules adequately within HPSG has proven surprisingly difficult; specifying just what information is preserved and what is changed by a lexical rule is one prominent issue. We conclude this chapter by describing alternatives to lexical rules. One is to appropriately underspecify properties of lexical entries so that they cover all relevant variants of a single lexeme or word. The second augments the type hierarchy via online type construction, extending the predefined lexical types specified in the hierarchy to include "virtual types" that combine the information from multiple predefined types.

2 Lexicalism

2.1 Lexicalism and the origins of HPSG

Lexicalism began as a reaction to Generative Semantics, which treated any regularity in the structure of words (derivational patterns, broadly speaking) as only ephiphenomenally a matter of word structure and underlyingly as a matter of syntactic structure (see Lakoff (1970), among others). In the Generative Semantics view, all grammatical regularities are a matter of syntax (much of it, in fact, logical syntax). Chomsky (1970) presented many arguments that lexical knowledge differs qualitatively from syntactic knowledge and should be modeled differently. Jackendoff (1975) is an explicit model of lexical knowledge that follows Chomsky's insights, although it focuses exclusively on derivational morphological processes. The main insight that Jackendoff formalizes is that relations between stems and words (say, between *destruct* and *destruction*) are to be modeled not via a generative device but through a redundancy mechanism that measures the relative complexity of a lexicon where these relations are present or not present (the idea is that a lexicon where *construct* and *construction* are related is simpler than one where they are not). Bochner (1993) is the most formalized and detailed version of this approach to lexical relations. Lexicalist approaches, including LFG|seeLexical Functional Grammar and HPSG, took their lead from Jackendoff's work. LFG has relied heavily on treating relations between stems and between words as lexical rules, rather than the kind of generative devices that one finds in syntax. But, as accounts of linguistic phenomena in LFG focused increasingly on the lexicon, the question of whether lexical rules retain the character of redundancy rules or turn into yet another kind of generative device arose. Consequently, the necessity of lexical rules has been questioned as well (see Koenig & Jurafsky (1994) and Koenig (1999: p29-49) for potential issues that arise once lexical rules are assumed to be involved in the creation of new lexical items).

Lexicalism, at least within HPSG, embodies two distinct ideas. First is the idea that parts of words are invisible to syntactic operations (*lexical integrity*, see Bresnan & Mchombo 1995), so that relations between stems and between word forms cannot be the result of or follow syntactic operations, as in Distributed Morphology (Halle & Marantz 1993), or other linguistic models that assign no special status to the notion of word. Relations between words are therefore not modeled via syntactic operations (hence the appeal to Jackendoff's lexical rules). Second is the idea that the occurrence of a lexical head in distinct syntactic contexts arises from distinct variants of words. For instance, the fact that the verb *expect* can occur both with a finite clause and an NP+VP sequence (see (2) vs. (1)) means that there are two variants of the verb *expect*, one that subcategorizes for a finite clause and one where it subcategorizes for an NP+VP sequence.[1] Not all lexicalist theories, though, cash out these two distinct ideas the same way. The net effect of lexicalism within HPSG is that words and phrases are put together via distinct

[1]As this chapter is an overview of the approach to lexical knowledge HPSG embodies rather than a description of particular HPSG analyses of phenomena, we will sample liberally from various illustrative examples and simplify whenever possible the analyses so that readers can see the forest and not get lost in the trees.

sets of constructions and that words are syntactic atoms. These two assumptions justify positing two kinds of signs, *phrasal-sign* and *lexical-sign*, and go hand in hand with the surface-oriented character of HPSG and what one might call a principle of surface combinatorics: If expression A consists of B⊕C, then all grammatical constraints that make reference to B and C are circumscribed to A.

(1) I expected that he would leave yesterday.

(2) I expected him to leave yesterday.

An evident concern regarding this view of the lexicon is the potential proliferation of lexical items, replete with redundant information. Will it be necessary to specify all the information in these two items for *expect* without regard for the large amount of duplication between them? Will the same duplication be needed for the verb *hope*, which patterns similarly (but not quite identically)? How will somewhat similar verbs, such as *imagine*, *suppose*, and *anticipate* which allow finite complements but not infinitive ones, be represented? We will describe HPSG's solutions to these questions below, in our discussion of the hierarchical lexicon. First, however, we turn to recent arguments against lexicalism, and then discuss in more detail just what kinds of information should be in HPSG lexical items.

2.2 Recent challenges to lexicalism

As there have been several challenges to lexicalism (see Bruening (2018) and Haspelmath (2011) among others for some recent challenges), we now explore lexicalism and lexical integrity in HPSG in more detail. We first note that lexicalism does not imply that word and phrase formation are necessarily different "components" as is often claimed (see Marantz (1997), Bruening 2018). Some lexicalist approaches *do* assume that word formation and phrase building belong to two different components of a language's grammar (this is certainly true of Jackendoff 1975), but they need not. Within HPSG, there are approaches that treat every sign-formation (be it word-internal or word-external) as resulting from typed mother-daughter configurations (this is the hypothesis pursued in Koenig (1999), and is also the approach frequently taken in implementations of large-scale grammars where lexical rules are modeled as unary-branching trees, see the English Resource Grammar, Copestake (2002) and the grammars developed in the CoreGram project Müller (2015); see **Mueller2018** for a similar point in his response to Bruening's paper). Furthermore, recent approaches to inflectional morphology within HPSG model realizational rules through the very same tools the rest of a language's grammar uses (see Crysmann & Bonami (2016) and **chapters/chap-morphology** Chapter **??** of this volume). There are also approaches to phrases where the same analytical tools developed to model lexical knowledge (see Section 4) are employed to model phrase-structural constructions (see Sag's (1997) analysis of relative clauses, for example). So, both in terms of the formal devices and in terms of analytical tools used to model datasets, words and phrases can be treated the same way in HPSG (alhtough they need not be). Somewhat ironically, and despite claims to the contrary, word formation in the syntactocentric approach Marantz or Bruening advocates *does* make use of distinct formal machinery to model word formation, namely realizational rules as well as various readjustement rules, fusion and fission rules, to model inflectional morphology (see Halle & Marantz 1993; Embick 2015).

With this red herring out of the way, we concentrate on the two most important challenges Bruening (2018) and Haspelmath (2011) present to lexicalist views. The first challenge are cases of phrasal syntax feeding the lexicon, purportedly exemplified by sentences such as (3).

(3) I gave her a don't-you-dare! look. (example (1a) in Bruening 2018)

We can provisionally accept for the sake of argument Bruening's contention that *don't-you-dare!* is a word in (3), despite its reliance on the (unjustified) assumption that the secondary object in (3) involves N-N compounding rather than an AP N structure (we refer readers to Bresnan & Mchombo (1995) or **Mueller2018** for counter-arguments to Bruening's claim). Crucially, though, examples such as (3) have no bearing on HPSG's model of lexical knowledge, as HPSG-style lexicalism does not preclude constructions that form words from phrases. Nothing, as far as we know, rules out constructions of the form *stem/word* → *phrase* in HPSG. The two assumptions underlying the HPSG brand of lexicalism we mentioned above do not preclude a *lexical-sign* having a *phrasal-sign* as sole daughter (although we do not know of any HPSG work that exploits this possibility) and examples such as (3) are simply irrelevant to whether HPSG's lexicalist stance is empirically correct.

The second challenge to lexicalism presented in Bruening (2018) bears more directly on HPSG's assumption that words are syntactic atoms. Word-internal conjunction/ellipsis examples, illustrated in (4) (adapted from Bruening's (31a), p. 14), seem to violate the assumption that syntactic constraints cannot "see" the internal structure of words|seemorphology, lexical integrity, as ellipsis in these kinds of examples seems to have access to the internal part of the word *over-application*. In fact, though, such examples do not violate lexical integrity if one enriches the representation of composite words (to borrow a term from Anderson 1992) to include a representation of their internal phonological parts as proposed in Chaves (2008; 2014).

(4) Over- and under-application of stress rules plagues Jim's analysis.

Chaves' analysis assumes that the phonology of compound words and words that contain affixoids (to borrom a term from Booij 2005) is structured. The MorphoPhonology or MP attribute of words (and phrases) is a list of phonological forms and morphs information. The MP of compound words and words that contain affixoids includes a separate member for each member of the compound, or for the affixoid and stem. Thus in (4), the MPs of *overapplication* and *underapplication* each contain two elements: one for *over*/*under*, and one for *application*. Given this enriched representation of the morphophonology of words like *under/overapplication*, a single ellipsis rule can apply both to phrases and to composite words, eliding the second member of the word *overapplication*'s MP. As Chaves (p. 304) makes clear, such an analysis is fully compatible with lexical integrity, as there is no access to the internal structure of composite words, only to the (enriched) morphophonology of the entire word.

Haspelmath (2011) similarly challenges the view that syntactic processes may not access the internal structure of words, although Haspelmath's point is merely that what is a word is cross-linguistically unclear. So-called suspended affixation in Turkish (see (5)) also shows that word parts can be elided. We cannot discuss here whether Chaves' analysis can be extended to cases like (5) where suffixes are seemingly elided or whether lexical sharing (where a single word can be the daughter of two c-structure nodes à la McCawley 1982), as proposed in Broadwell (2008), is needed.

(5) kedi ve köpek-ler-im-e
cat and dog-PL-1SG-DAT
'to my cat(s) and dogs'

What is important for current purposes is that these putative challenges to lexical integrity such as (4) or (5) do not necessarily render a substantive version of it implausible. The same is true of another potential challenge to lexical integrity which neither Bruening nor Haspelmath discuss, endoclitics, which we turns to next.

Endoclitics are clitics that at least appear to be situated within a word, rather than immediately preceding or following it, as clitics often do (**chapters/chap-clitics**

Chapter ?? of this volume). In many cases, endoclitics appear at morphological boundaries, as in the well-studied pronominal clitics of European Portuguese (Crysmann 2000a). An approach similar to what we have referenced above for composite words and elided morphology may extend to these as well. But some trickier cases have also come to light, in which the clitic appears within a morpheme, not at a boundary. Two of the best documented cases come from the Northeast Caucasian language Udi (Harris 2000) (see (6)) and from Pashto (Tegey 1977; Roberts 2000; Dost 2007) (see (7)).

(6) q'ačaɣ-ɣ-on bez tänginax baš=**q'un**-q'-e
thief-PL-ERG my money.DAT steal$_1$-3PL-steal$_2$-AORII
'Thieves stole my money.' (root *bašq'*, 'steal') (Harris, op.cit., (18))

(7) a. ʈəlwɑhə́=**me**
push.IMPF.PST.3SG-cl.1SG
'I was pushing it.' (from Tegey 1977; Dost 2007)

b. ʈə́l=**me**-wɑhə
push$_1$-cl.1SG-push$_2$.PF.PST.3SG
'I pushed it.' (from Tegey 1977; Dost 2007)

In these cases, as with clitics in general, there is a clash between the phonological criteria for wordhood, under which the clitics would be regarded as incorporated within words, and the syntactic constituency and semantic compositionality. But what makes these particularly odd is that these clitics are situated word-internally, even morpheme-internally. Udi subject agreement clitics such as *q'un* in (6) typically attach to a focused constituent, which can be a noun, a questioned constituent, or a negation particle as well as a verb (Harris 2000). Under certain conditions, as in (6), none of these options is available or permitted, and the clitic is inserted before the final consonant of the verb root, dividing it in two pieces, neither of which has any independent morphological status. Its position in this instance is apparently phonologically determined; it cannot appear word-finally or word initially, and as there is no morphological boundary within the word it must therefore appear within the monomorphemic root. Pashto clitics seek "second position", whether at the phrasal, morphological, or phonological level; *me* in (7) appears to be situated after the first stressed syllable (or metrical foot), which, in the case of (7b), also divides the verb into two parts that lack any independent morphological status.

If clitics are viewed as a syntactic phenomenon ("phrasal affixes", as Anderson (2005) puts it), these endoclitics must "see"' into the internal structure of words (be it morphological, prosodic, or something else), thereby seemingly violating lexical integrity. Anderson's brief account invokes a reranking of optimality theoretic constraints from their typical ordering, whereby the clitic's positional requirements outrank lexical integrity requirements. Crysmann (2000b) proposes an analysis, paralleling in many respects his account of European Portuguese clitics in Crysmann (2000a), using Reape's constituent order domains (Reape 1994) and, in particular, Kathol's topological fields (Kathol 1999) (see **chapters/chap-order** Chapter **??** of this volume). The "morphosyntactic paradox" in Udi, to borrow a phrase from Crysmann (2002), is effectively "resolved on the basis of discontinuous lexical items"; this account then "parallels HPSG's representation of syntactic discontinuity" (Crysmann 2000b), (**chapters/chap-order** Chapter **??** of this volume).

For Pashto, researchers generally agree that the notion of second position is crucial, but that it can be defined at various levels — phrasal, lexical, and phonological. In this last case clitics can appear within a word following the first metrical unit, as illustrated above (**chapters/chap-phonology** Chapter **??** of this volume). Dost (2007) invokes the mechanisms of word order domains (Reape 1994) and topological fields (Kathol 1999) at these various levels to account for this distribution of clitics. In this analysis, some words contain more than one order domain at the prosodic level. Lexical integrity is preserved to the extent that, while domains at the prosodic level are "visible" to clitics in Pashto, syntatic processes do not reference the internal makeup of words.

Still, these accounts of endoclitics in Udi and Pashto appear to breach the wall of the strictest kind of lexical integrity, as they require access to some of the internal structure of lexical entries through a partial decomposition of their morphophonology into distinct order domains. Yet we would not wish to advocate models that permit unconstrained violations of lexical integrity, either. The troublesome cases we have noted here are relatively marginal or cross-linguistically rare; they seem to be limited in scope to prosodic or morphophonological information (e.g., ellipsis, insertion). As Broadwell (2008) points out when comparing possible analyses of Turkish suspended affixation, rejecting lexicalism altogether may lead to an unconstrained theory of the interaction between words/stems and phrases and incorrect predictions (e.g., that all affixes in Turkish can be suspended). Likewise, we would not expect to find a language in which endoclitic positioning is utterly unconstrained or where syntactic operations are sensitive to the fact that *anticonstitutional* is based on the nominal root *constitution*, or where coordination of affixes is always possible. Rejecting lexicalism begs the question of why such languages do not seem to exist, why what is visible to syntactic operations of the internal structure of words (morphophonological structure) is so restricted or why even that kind of morphophonological visibility is so rare (particular affixoids and endoclitics, say).

3 Lexical entries in HPSG

3.1 What are lexical entries?

Because lexical entries or items (the terms are used interchangeably, it seems, within the HPSG literature) play a central role in accounting for the syntax of natural languages, lexical entries are informationally rich in HPSG. An additional consequence of HPSG's lexicalist stance is that there will be many lexical entries where one might at first glance expect a single entry. We will see below how HPSG handles multiple entries and classes of entries while avoiding redundancy, but it is important at the outset to clarify what a lexical entry is in HPSG. One misunderstanding about lexical entries conflates descriptions and the entities they describe, or, in other words, fails to distinguish between constructions in the abstract and a particular word or phrase (i.e., a lexical entry vs. a fully instantiated lexeme). As **chapters/chap-formalbackground** Chapter **??** of this volume makes explicit, grammars in HPSG consist of *descriptions* of structures, and the lexicon thus consists of descriptions of what a fully specified lexeme or word can be. To see the importance of the distinction between descriptions (stored entries) and the fully instantiated objects that are being described, consider HPSG's model of subcategorization with reference to the relevant portion of the tree for sentence (1). HPSG's model of the dependency between heads and complements stipulates identity between the syntactic and semantic information of each complement (the value of the SYNSEM attribute) and a member of the list of complements the head subcategorizes for. Since there are indefinitely many SYNSEM values, on the assumption that there are indefinitely many clausal meanings (a point Jackendoff (1990: p. 8–9) emphasizes), there are, in principle, indefinitely many fully instantiated entries for the verb *see* subcategorizing for a clausal complement (as in (1)). But each of these fully instantiated entries for *expect* – one for each clausal sentence that corresponds to the tree in Figure 1 – corresponds to a single abstract description, and it is this description that the lexicon contains.

The formal status of lexical entries has engendered a fair amount of theoretical work and some debate. We will touch on some aspects of this further below, in connection with online type construction. For further discussion of these kinds

$$\left[\text{COMPS } \langle \boxed{1} \rangle\right] \qquad \left[\text{SYNSEM } \boxed{1}\right]$$

Figure 1: A head-complement phrase

of issues, see **chapters/chap-formalbackground** Chapter ?? of this volume and **chapters/chap-properties** Chapter ?? of this volume.

3.2 What information is in lexical entries?

Aside from the expected phonological and semantic information, specific to each lexeme, lexical entries include morphological information and information about their combinatorial potential. Morphosyntactic features can be part of the input to inflectional rules, but are also used to select the appropriate types of phrases (via their projections through the Head-Feature Principle), as shown in (8). Some verbs, for instance, select for a PP headed by a particular preposition; others select for VPs whose verb is a gerund, or a bare infinitive, and so forth. Lexical entries thus include as much morphological information as both (inflectional) morphology and syntactic selection require.

(8) a. John conceived *of/*about* the world's tastiest potato chip.
    b. John regretted *going/*(to) go* to the party.

We illustrate the second leading idea behind HPSG or LFG's lexicalism – that there are different variants of lexical heads for different contexts in which heads occur – with the French examples in (9). The verb *aller* 'go' in (9a) combines with a PP headed by *à* that expresses its goal argument and a subject that expresses its theme argument. The same verb in (9b) combines with the so-called non-subject clitic *y* that expresses its goal argument. We follow Miller & Sag (1997) and assume here that French non-subject clitics are prefixes. Since the context of occurrence of the head of the sentence, *aller*, differs across these two sentences (NP____PP[*à*] and NP *y*____ , respectively and informally), there will be two distinct entries for *aller* for both sentences, shown in (10) and (11) (we simplify the entries' feature geometry for expository purposes). Information in the entry in (11) that differs from the information in the entry in (10) appears in red.

(9) a. Muriel va à Lourdes.
    Muriel go.PRES.3RD.SG at Lourdes.

b. Muriel y va.
Muriel there go.PRES.3RD.SG

(10)
$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{FORM} & \boxed{5} \\ \text{I-FORM} & \boxed{5}\ \textit{va} \\ \text{STEM} & \textit{v-} \end{bmatrix} \\ \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \begin{bmatrix} \text{MOOD} & \textit{indic} \\ \text{TNS} & \textit{pres} \\ \text{AGR} & \textit{3rdsing} \end{bmatrix} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\text{NP}_{\boxed{3}, \textit{3rdsg}}, \boxed{2}\text{PP}[\text{à}]_{\boxed{4}} \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \textit{go-rel} & \\ \text{THEME} & \boxed{3} \\ \text{GOAL} & \boxed{4} \end{bmatrix} \end{bmatrix}$$

(11)
$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{FORM} & \textit{y-va} \\ \text{I-FORM} & \textit{va} \\ \text{STEM} & \textit{v-} \end{bmatrix} \\ \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \begin{bmatrix} \text{MOOD} & \textit{indic} \\ \text{TNS} & \textit{pres} \\ \text{AGR} & \textit{3rdsing} \end{bmatrix} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\text{NP}_{\boxed{3}, \textit{3rdsg}}, \text{PP}[\textit{p-aff,loc}]_{\boxed{4}} \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \textit{go-rel} & \\ \text{THEME} & \boxed{3} \\ \text{GOAL} & \boxed{4} \end{bmatrix} \end{bmatrix}$$

CATegory information in both entries contains part of speech information (including morphologically relevant features of verb forms), ARGument-STructure information and VALence information. MORPH information includes both stem form information, inflected form information (I-FORM) and, in case so-called clitics are present, the combination of the clitic and inflected form information. Both entries illustrate how informationally rich lexical entries are in HPSG. But, postulating informationally rich entries does not mean stipulating all of the information within every entry. In fact, only the stem form and the relation denoted by the semantic content of the verb *aller* need to be stipulated within either entry. All the other information can be inferred once it is known which classes of verbs these entries belong to. In other words, most of the information included in the entries in (10) and (11) is not specific to these individual entries, an issue we take up in Section 4. As mentioned above, the informational difference between the two entries for *va* and *y va* is indicated in red in (11). The first difference between the two variants of *va* 'goes' is in the list of complements: the entry for *y va* does not subcategorize for a locative PP, since the affix *y* satisfies the relevant argument structure requirement. This difference in the realization of syntactic arguments (via phrases and pronominal affixes) is recorded in the types of the PP members of ARG-ST, *p-aff* in (11), but a PP headed by *à* in (10). Finally, the two entries differ in the FORM of the verb, which is the same as the inflected form of the verb in (10) (as indicated by the identically numbered $\boxed{5}$), but not in (11), whose FORM includes the prefix *y*.

One other question arises with regard to the information in lexical entries. Are there attributes or values that occur solely within lexical signs, and not in phrasal ones? If so, they would provide a diagnostic for distinguishing lexical signs from others. Specific phonological information, for instance, is something we would expect to be introduced by lexical entries, and not elsewhere. Such information would thus be specific to lexical signs. Likewise, the ARG-ST list, which we included in the categorial information of signs in (10) and (11), might be regarded as a feature confined to lexical signs (see, among others, Ginzburg & Sag 2000), on the premise that lexical items alone specify combinatorial requirements (but see Przepiórkowski (2001) for a contrary view, and see **chapters/chap-cxg** Chapter **??** of this volume for other views questioning this assumption). But HPSG researchers have generally not explored this question in depth, and we will leave this issue here.

3.3 The role of the lexicon in HPSG

As we hope is evident by now, the lexicon plays a critical role in HPSG's explanatory mechanisms, as lexical entries encode not merely their idiosyncratic phonological and semantic characteristics, but their distributional and combinatorial potential as well. Much of the information contained in lexical entries is geared towards modeling how words interact with one another, as we have already seen. As detailed in **chapters/chap-arg-st** Chapter **??** of this volume, their combinatorial potential is recorded using two kinds of information, a list of syntactic arguments or syntactic requirements to be satisfied, and distinct lists that indicate how these requirements are to be satisfied (as local dependents, as non-local dependents, as clitics/affixes). Not only are syntactic arguments recorded; so is their relative obliqueness (in terms of grammatical function), as per the partial hierarchy in (12) from Pollard & Sag (1992).

(12) SUBJECT < PRIMARY OBJ < SECOND OBJ < OTHER COMPLEMENTS

We illustrate this explanatory role by noting the role of the lexicon in HPSG's approach to binding, as described in Pollard & Sag (1992) (see **chapters/chap-binding** Chapter ?? of this volume for details). As lexical entries of heads record both syntactic and semantic properties of their dependents, constraints between properties of heads and properties of dependents, e.g. subject-verb agreement, or between dependents, e.g. binding constraints, illustrated in (14), can be stated, at least partially, as constraints on classes of lexical entries. The principle in (14) is such a constraint.

(13) a. Mathilda$_i$ saw herself$_i$ in the mirror.
b. * Mathilda$_i$ saw her$_i$ in the mirror.

(14) An anaphor must be coindexed with a less oblique co-argument, if there is one.

Principle (14) is, formally, a constraint on lexical entries that makes use of the required information in an entry's argument structure regarding the syntactic and semantic properties of its dependents. The three argument structures in (15) illustrate permissible and ungrammatical entries. (15a) illustrates exempt anaphors, as there is no less oblique syntactic argument than the anaphoric NP; (15b) illustrates a non-exempt anaphor properly bound by a less oblique, co-indexed non-anaphor; (15c) illustrates an ungrammatical lexical entry that selects for an anaphoric syntactic argument that is not co-indexed by a less oblique syntactic argument, despite not being an exempt anaphor (i.e., not being the least oblique syntactic argument).

(15) a. $\left[\textsc{arg-st}\ \langle \text{NP}_{i,+ana}, \ldots \rangle\right]$
b. $\left[\textsc{arg-st}\ \langle \text{NP}_{i,-ana}, \ldots, \text{NP}_{i,+ana}, \ldots \rangle\right]$
c. * $\left[\textsc{arg-st}\ \langle \text{XP}_{j}, \ldots, \text{NP}_{i,+ana}, \ldots \rangle\right]$

Our purpose here is not to argue in favor of the specific approach to binding just outlined. Rather, we wish to illustrate that in a theory like HPSG where much of syntactic distribution is accounted for by properties of lexical entries, co-occurrence restrictions treated traditionally as constraints on trees (via some notion of command) are modeled as constraints on the argument structure of lexical entries. It is tempting to think of such a lexicalization of binding principles as a notational variant of tree-centric approaches. Interestingly, this is not the case, as argued in Wechsler (1999). Wechsler argues that the difference between argument structure and valence is critical to a proper model of binding in Balinese. Summarizing briefly, voice alternations in Balinese (e.g., objective or agentive voices) do not alter a verb's argument structure but do alter its valence, the subject and object it subcategorizes for. As binding is sensitive to relative obliqueness within ARG-ST, binding possibilities are not affected by voice alternations within the same clause, which are represented with different valence values. In the case of raising, on the other hand, the argument structure of the raising verb and the valence of the complement verb interact, as the subject of the complement verb is part of the argument structure of the raising verb. An HPSG approach to binding therefore predicts that voice alternations within the embedded clause will not affect binding of co-arguments of the embedded verb, but will affect binding of the raised NP and an argument of the embedded verb. This prediction seems to be borne out, as the Balinese examples in (16) show.

(16) a. | Ia$_i$ | nawang | awakne$_i$/Ia$_{*i}$ | lakar | tangkep | polisi.
| --- | --- | --- | --- | --- | --- |
| 3rd | AV.know | self/3rd | FUT | OV.arrest | police |

‘He$_i$ knew that the police would arrest self$_i$/him$_{*i}$.’

b. | Cang | ngaden | ia$_i$ | suba | ningalin | awakne$_i$/ia$_{*i}$
| --- | --- | --- | --- | --- | --- |
| 1sg | AV.think | 3rd | already | AV.see | self/3rd |

‘I believe him$_i$ to have seen himself$_i$/ him$_{*i}$.’

c. | Cang | ngaden | awakne$_i$ | suba | tingalin=a$_i$.
| --- | --- | --- | --- | --- |
| 1sg | AV.think | self$_i$ | already | OV.see=3 |

‘I believe him to have seen himself.’

Sentence (16a) shows that the proto-agent (the first element of ARG-ST) of the subject-to-object raising verb *nawang* ‘know’ can bind the raised subject (which in this case corresponds to the proto-patient of the complement verb *tangkep* ‘arrest’ since that verb is in the objective voice). Sentence (16b) shows that the raised (proto-agent) subject of the complement verb can bind its proto-patient argument. Critically, sentence (16c) shows that the raised proto-patient (second) argument of the complement verb can be bound by the complement verb's proto-agent. The contrast between sentences (16b) and (16c) illustrates that while binding is insensitive to valence alternations (the same proto-agent binds the same proto-patient argument in both sentences), raising is not (the proto-agent argument is raised in (16b) and the proto-patient argument in (16c)). As Wechsler argues, this dissociation between valence subjects and less oblique arguments on the ARG-ST list is hard to model in a configurational approach to binding that equates the two notions in terms of c-command or the like. What is important for our purposes is that a "lexicalization" of argument structure, valence, and binding has explanatory power beyond tree configurations, illustrating some of analytical possibilities informationally rich lexical entries create.

3.4 Lexical vs. constructional explanations

As we have noted above, HPSG posits that much of the combinatorics of natural language syntax is lexically determined; lexical entries contain information about their combinatorial potential and, as a consequence, if a word occurs in two distinct syntactic contexts, it must have two distinct combinatorial potentials. Under this view, phrase-structure rules are boring and few in number. They are just the various ways for words to realize their combinatorial potential. In the version of HPSG presented in Pollard & Sag (1994), for example, there are only a handful of general phrase-structural schemata, one for a head and its complements, one for a head and its specifier, one for a head and a filler in an unbounded dependency and so forth and the structure of clauses is relatively flat in that relations between contexts of occurrence of words is done "at the lexical level" rather through operations on trees.

In a transformational approach, on the other hand, relations between contexts of occurrence of words are seen as relations between *syntactic* trees, and the information included in words can thus be rather meager. In fact, in some recent approaches, lexical entries contain nothing more than some semantic and phonological information, so that even part of speech information is something provided by the syntactic context (see Borer 2003; Marantz 1997). In some constructional approaches (Goldberg (1995), for example), part of the distinct contexts of occurrence of words comes from phrase structural templates that words fit into. So again, there can be a single entry for several contexts of occurrence.

HPSG's approach to lexical knowledge is quite similar to that of Categorial Grammar (to some degree this is due to HPSG's borrowing from Categorial Grammar important aspects of its treatment of subcategorization). As in HPSG, the combinatorial potential of words is recorded in lexical entries so that two distinct contexts of occurrence correspond to two disteinct entries. The difference from HPSG lies in how lexical entries relate to each other. In many forms of Categorial Grammar (be it Combinatorial or Lambek-calculus style), relations between entries are the result of a few general rules (e.g., type raising, function composition, hypothetical reasoning, etc.) (see Dowty (1978) for an approach that countenances lexical rules, though). The assumption is that those rules are universally available; however, those rules may be organized in a type hierarchy and an individual language might avail itself of only a portion of this hierarchy, as in (Baldridge 2002). Relations between entries in HPSG can be much more idiosyncratic and language-specific. We note, however, that nothing prevents lexical rules constituting a part of a Categorial Grammar (see Carpenter 1992a), so that this difference is not necessarily qualitative, but concerns how much of researchers' efforts are typically spent on extracting lexical regularities; HPSG has focused much more, it seems, on such efforts.

4 The hierarchical lexicon

We have now seen that lexicalism demands that lexical entries be information rich, in order to encode what might otherwise be represented as syntactic rules. To avoid massive and redundant stipulation throughout the lexicon, we need mechanisms to represent the regularities within it. Two main mechanisms have been used in HPSG to achieve this. The first mechanism is the organization of information shared by lexical entries or parts of entries into a hierarchy of types, in a way quite similar to semantic networks within knowledge representation systems (see, among others, Brachman & Schmolze 1985). This hierarchy of types (present in HPSG since the beginning, Pollard & Sag (1987) and the seminal work of Flickinger et al. (1985), Flickinger 1987) ensures that individual lexical entries only specify information that is unique to them. The second mechanism is lexical rules, which relate variants of entries, and more generally, members of a lexeme's morphological family (which consists of a root or stem as well as all stems derived from that root or stem) or members of a word's inflectional paradigm.

In this section, we discuss the hierarchical organization of the lexicon into cross-cutting classes of lexical entries at various levels of generality. We examine two distinct techniques for inheritance, which are not mutually exclusive. One is to create subtypes directly, with pertinent additional constraints stated for each subtype. Different classes of words are thus reified as subtypes of *word* (or *lexical-sign*) in the hierarchy, and all lexical items that belong to that subtype inherit its constraints. Another technique, more prevalent in current HPSG work, uses implicational statements. If certain properties hold of a lexical item (for example, if its AUX value is +), then others must hold as well (e.g., it subcategorizes for a VP complement, whose subject is token identical to the auxiliary verb's). These statements need not involve all of the information that's present in the entire *word*, so they need to refer only to substructures within *word* objects, like their SYNSEM values.

4.1 Inheritance

All grammatical frameworks classify lexical entries to some extent, of course. Basic part of speech information is one obvious case. This high-level classification is present in HPSG, too, as part of the hierarchy of types of heads. That information is recorded in the value of the HEAD feature. A simple hierarchy of types of heads is depicted in Figure 2.

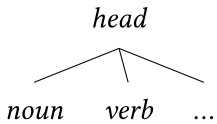


Figure 2: A hierarchy of types of *head*

Each of these types is a partial specification of a lexical entry's head properties. Typing of HEAD information allows the ascription of appropriate properties to different classes of lexical entries. For example, case information is only relevant to nouns and adjectives, and whether a verb is an auxiliary or not is only relevant to verbs. The subtypes of *head* in (2) allow us to define additional specifications of of the properties appropriate for different parts of speech. For example, lexical entries with a HEAD value of *noun* contain an attribute for CASE, while those for *verb* contain the attributes AUX, TENSE, and ASPECT, as shown in (17) (we use implicational statements in (17) to indicate feature appropriateness conditions for the types *noun* and *verb* for perspicuity only; such conditions would be part of the grammar's signature, see **chapters/chap-formalbackground** Chapter **??** of this volume). In other words, the grammar's signature will specify that only for nouns (those lexical entries whose HEAD value is of type *noun*) is the attribute CASE appropriate. Similarly, only for verbs are the attributes *aux*, TENSE, and ASPECT appropriate.

Typing of parts of speech thus lets us specify what it means for a part of speech to be a noun or a verb in a particular language (of course, there will be strong similarities in these properties across languages) and omit for individual noun and verb entries properties they share with all nouns and verbs.

(17) a. If the attrbute CASE is an attribute within a lexical entry's HEAD value, then the value of HEAD is of type *noun*.

b. If the attributes AUX, TENSE, or ASPECT are attributes within a lexical entry's HEAD value, then the value of HEAD is of type *verb*.

The statements in (17) are in some sense merely definitional, as noted. But they allow us to state just once the general information that applies to whole classes of lexical entries. Thus, the pronoun *him* need only include the fact that its HEAD is of type *noun*; the fact that it might bear case can be inferred. Similarly, the entry for the verb *can* need only include information that its head specify [AUX +] for us to be able to infer that it is a *verb* (assuming AUX is not an appropriate attribute for another type).

4.2 Representing lexical generalizations

So far, we have merely shown an HPSG implementation of a part of speech taxonomy, but once we consider subtypes with additional constraints the utility of the hierarchical lexicon within a lexicalist framework becomes apparent. There are interesting generalizations to be made about more specific classes, such as transitive verbs, or past participles, or predicators denoting caused motion (regardless of their part of speech). In the hierarchical lexicon, we can represent these "interesting" classes as types. Which classes are worth positing in the grammar of a given language depends on the properties of its grammar; thus we expect lexical classes to specify a mix of cross-linguistically common (possibly, in some cases, universal) and language-particular constraints.

A seemingly straightforward way to "capture generalizations about the elements of the lexicon" is to posit a hierarchy of subtypes of *word*. Thus types such as *verb-word* and *noun-word* specify properties of verbs and nouns, and types such as *subj-control-pred* and *obj-control-pred* specify properties of predicates that exhibit subject and direct object control. Individual lexical items belong to multiple types in the hierarchy; the verbs *try* and *attempt* inherit the information from *verb-word* and *subj-control-pred*, while the nouns *attempt* and *effort* inherit the information from *noun-word* and *subj-control-pred*.

Ackerman & Webelhuth (1998) use this kind of hierarchy of subtypes of *word* in their accounts of German passives and other phenomena, which we will disucss briefly in the following section. In this case, the information involved in their account is both morpohological and syntactic, and they propose a hierarchy of verb types at the *word* level.

However, a hierarchy of subtypes of *word* is, while formally feasible, potentially rather inelegant. Note first that types like *verb* and *noun* are already defined as subtypes of the type *head*. There is an obvious danger of redundancy if we additionally posit parallel subtypes of *word* such as *verb-word* and *noun-word*, serving no other function than as types with the corresponding HEAD values. Furthermore, signs in HPSG are structured objects, with their various kinds of information deliberately arranged in a way that associates pieces of information that "travel together." The information within HEAD, for example, is grouped there because it is all subject to the Head Feature Principle. Both part-of-speech and control information are found within SYNSEM, as phonological information has no bearing on these things. So rather than creating subtypes of *word* to capture regularities in the lexicon, we would prefer to express those regularities as constraints on subtypes that encompass only the information that's pertinent. These are the smallest, "narrowest" portions of *word* objects that include all that information; the remaining portions of a *word* can be ignored in this context. In other words, we take advantage of HPSG'sfeature geometry and of the hierarchies of types appropriate for a particular substructure within signs to express generalizations as "locally" as possible (see **chapters/chap-formalbackground** Chapter **??** of this volume).

Implicational statements can serve well for expressing generalizations as" locally" as possible; they constrain the range of possible values of attributes and can stipulate structure sharing among them. As a simple example, consider the possible complements of prepositions. Unlike verbs, which, at least in some languages, can have multiple elements on their COMPS lists, prepositions are limited to at most one. There are no ditransitive prepositions (as far as we are aware). The following statement expresses this generalization in English as well as more formally.

(18) a. If a *category* object has a HEAD of type *prep*, then the value of its VAL | COMPS is a list that contains at most one element.

b. $\begin{bmatrix} \textit{cat} \\ \text{HEAD} \;\; \textit{prep} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{VAL}|\text{COMPS} \;\; \textit{elist} \vee \langle \textit{synsem} \rangle \end{bmatrix}$

A more extensive example concerns linking of semantic roles to syntactic arguments, and is drawn from the work of Davis & Koenig (2000); Davis (2001); Koenig & Davis (2003). Informally, the linking constraints we wish to capture are:

(19) a. Causal verbs link the causer to the subject.

b. Caused motion verbs link the causer to the subject and the moving entity, if distinct from the causer, to the direct object.

c. However, caused motion verbs in which the causer and moving entity are the same thing can link both to the subject (and needn't have a direct object).

The second and third statements are subcases of the first, so ideally we prefer to state the substance of the first statement just once, rather than repeat it. We could posit subtypes of *word*, along the lines of the approach mentioned above, such as *transitive-verb* and *caused-motion-verb*. But implicational statements provide an arguably simpler way to model the facts of linking. Since the constraints we wish to express concern both ARG-ST and CONT, our implications are stated on *local* objects, which are the minimal type of object containing these attributes. We presuppose here a hierarchy of semantic relation types as values of CONT, including *cause-rel*, *motion-rel* and their subtype *caused-motion-rel*, each of which licenses attributes for the required participant roles.

First, we require that the causer, denoted in (20b) by the value of ACT, be linked to the subject (the first element of ARG-ST):

(20) a. If a *synsem* object's CAT | HEAD value is of type *verb*, and its CONT value is of type *cause-rel*, then its value of CONT | ACT is token identical to the index of the first element of its ARG-ST list.

b. $$\begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{verb} \\ \text{CONT} & \textit{cause-rel} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CONT} & \begin{bmatrix} \text{ACT} & \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$$

Then, we link the moving entity in a caused motion verb, denoted in (21b) by the value of UND, to any NP on ARG-ST:

(21) a. If a *synsem* object's CAT | HEAD value is of type *verb*, and its CONT value is of type *move-rel*, then its value of CONT | UND is token identical to the index of some NP element of its ARG-ST list.

b. $$\begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{verb} \\ \text{CONT} & \textit{move-rel} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CONT} & \begin{bmatrix} \text{UND} & \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \ldots\text{NP}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$$

Both of these implicational statements apply to a verb with a CONT value of type *caused-motion-rel*. Note that if the causer and the moving entity are distinct, they will be realized as separate NPs on the ARG-ST list. This is the linking pattern we find in numerous verbs, such as *throw*, *lift*, *expel*, and so on. In some cases, however, the causer and the moving entity may be one and the same. If the ACT and UND values are identical in CONT, then the second implication allows the moving entity to be realized as the subject also, or as a reflexive direct object, as in:

(22) The kids rolled (themselves) down the hill.

What is ruled out be this pair of statements, though, is a hypothetical verb *quoll*, with a linking pattern like this:

(23) *The rock quolled the kids down the hill.
(intended: The kids rolled the rock) down the hill.)

Additional restrictions may apply to some verbs of motion. For instance, many verbs of locomotion entail that the causer and moving entity are identical, and allow only an intransitive variant:

(24) The kids strolled (*themselves) down the hill.

We could represent this identity using another constraint, solely within CONT, as follows, where the type *self-move-rel* is a subtype of *move-rel*:

(25) a. If a *synsem* object's CAT | HEAD value is of type *verb*, and its CONT value is of type *self-move-rel*, then its values of CONT | ACT and CONT | UND are token identical.

b. $\begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{verb} \\ \text{CONT} & \textit{self-move-rel} \end{bmatrix} \rightarrow \begin{bmatrix} \text{CONT} & \begin{bmatrix} \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{1} \end{bmatrix} \end{bmatrix}$

When we consider the most specific types of the lexical hierarchy, where individual lexical entries reside, the same kinds of constraints, pertaining solely to a given lexical entry's phonological form, inflectional class, specific semantics, register, and so forth, can be employed. This lexeme or word-specific information needs to be spelled out somewhere in any grammatical framework. We can now view this as just the narrowest, most particular case of specifying information about a class of linguistic entities. At the same time, information shared across a broader set of lexical entries need not be stated separately for each one. Thus, the phonology of the word *spray* and the precise manner of motion of the particles or liquid caused to move in a spraying event are unique to this lexical entry. However, much of its syntactic and semantic behavior – it is a regular verb, participating in a locative alternation, involving caused ballistic motion of a liquid or other dispersable material – is shared with other English verbs such as *splash*, *splatter*, *inject*, *squirt*, and *drizzle*. To the extent that these "narrow conflation classes", as Pinker (1989) terms them, are founded on clear semantic criteria, we can readily state syntactic and semantic constraints on the appropriate types in the relevant type hierarchy. Thus much of the semantics of a verb like *spray* need not be specified at the level of that individual lexical entry. Apart from the broad semantics of caused motion, shared by numerous verbs, the verbs in the narrow conflation class containing *spray* share the selectional restriction, noted above, that their objects are set in motion by an initial impulse and that they are liquid or particulate material. We might therefore posit a subtype of the type *caused-motion-rel* to represent this shared semantics triggering the locative alternation, with further subtypes for the semantics of the individual verbs. Note that not all these constraints apply to precisely the same class (there are other verbs with somewhat different semantics, like *load* and *wrap*, exhibiting the locative alternation, for example), so several types might be required.

To sum up the import of these brief examples, the substance of the hierarchical lexicon need not be directly expressed in terms of subtypes of *word*, but rather in implicational statements that express constraints among types in the structures inside lexical entries. Interactions among these statements provide a way for classes of lexical items to inherit and share properties, so that they need not specify the same information over and over again.

4.3 Crosscutting types in the lexicon

Having now illustrated the use of implicational statements to specify constraints on classes of lexical entries at various levels of generality, we present in this section an example of crosscutting types, each expressing some generalization about a class of words. The information in a particular lexical entry is a result of multiple inheritance from several of these types, Drawn from Ackerman & Webelhuth (1998), this sample analysis concerns German passives, which come in several varieties, each with its own constraints. Each passive construction uses a different auxiliary (*werden*, *sein*, or *bekommen*) and two of these constructions require a participial form of the verb, while the *sein* passive requires *zu* followed by an infinitive VP. Additionally, passives appear attributively, as NP modifiers, as well as predicatively. Here are two examples of the *zu* + infinitive passive, the first attributive, the second predicative:

(26) a. die dem Mann von Johann zu schenkenden Blumen
the the man by Johann to give flowers
'the flowers that must be given to the man by Johann'

b. weil die Blumen den Mann von Johann zu schenken sind
because the flowers the man by Johann to give are
'because the flowers must be given to the man by Johann'

Ackerman & Webelhuth's account of German passives posits a multiple inheritance hierarchy of lexical types (note that these are all subtypes of a type *word*, not subtypes of values within it). A portion of their hierarchy of German passive types is shown in Figure 3.

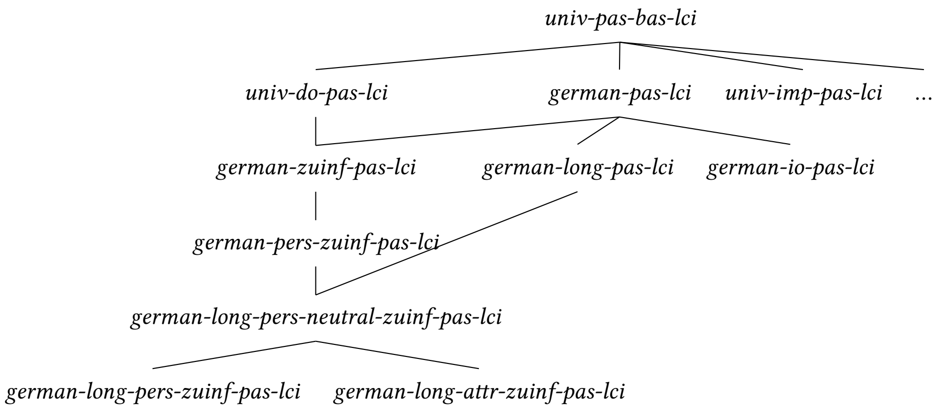


Figure 3: A portion of the hierarchy of passive lexical types according to Ackerman & Webelhuth 1998: 244

While all passives share the constraint that a logical subject is demoted, as stipulated on a general *univ-pas-bas-lci* passive type, the other requirements for each kind of passive are stated on various subtypes. The *zu*+infinitive passive, for instance, requires not only that *sein* is the auxiliary and that the main verb is infinitive, but that the semantics involves necessity or obligation. This differs from the other passives, which simply maintain the semantics of their active counterparts. However, the types of the passive verb *schenken(den)* in (26a) and (26b) both inherit from several passive verb supertypes. As mentioned, at a general level, there is information common to all German passives, or indeed to passives universally, namely that the “logical subject” (first element of the basic verb’s ARG-ST list) is realized as an oblique complement of the passive verb, or not at all. A very common subtype, which Ackerman & Webelhuth also regard as universal, rather than specific to German, specifies that the base verb’s direct object is realized as the subject of its passive counterpart; this defines personal passives. Once in the German-specific realm, an additional subtype specifies that the logical subject, if realized, is the object of a *von*-PP; this holds true of all three types of German personal passives. Among its subtypes is one that requires *zu* and the infinitive form of the verb; moreover, although Ackerman & Webelhuth do not spell this out in detail, this subtype specifies the modal force associated with this passive construction but not of the others. Finally, both the predicative and attributive forms are subtypes of all the preceding, but these inherit also from distinct supertypes for predicative and attributive passives of all kinds. The supertype for predicative passives constrains them to occur with an auxiliary; its subtype for *zu* + infinitive passives further specifies that the auxiliary is *sein*. The attributive passive type, on the other hand, inherits from modifier types generally, which do not allow auxiliaries, but do require agreement in person, number, and case with the modified noun. In summary, the hierarchical lexicon is deployed here to factor out the differing properties of the various German passive constructions, each of which includes its particular combination of properties via multiple inheritance.

4.4 Default inheritance in the lexicon

So far, we have assumed rigid, monotonic inheritance of all information in supertypes to their subtypes; none of the inherited information can be overridden. This runs into difficulties when dealing with lexical entries that appear to be exceptional in some way, the obvious examples being morphological irregularities. How can productive regular forms such as **childs* be blocked, and only *children* allowed as a lexical entry? Under default unification, although the plural of *child* might inherit the information from the pertinent lexical entry and from the *plural-noun* type, which would entail the phonology for **childs*, this regular plural form would be overridden.

Several approaches to exceptions and irregularities have been proposed; we will focus first on *default unification*, and examine an alternative involving type underspecification, in the following section. Various complex issues arise in attempting to formulate a workable system of default unification and inheritance. See, e.g., Briscoe & Copestake (1999) for a brief overview of various ways that default unification might be defined. Lascarides & Copestake (1999) list several desirable criteria, including these:

- Non-default information is always preserved; this implies some means of distinguishing non-default from default (overridable) information.
- Default unification behaves like monotonic unification whenever possible; that is, if monotonic unification is possible, the default unification mechanism should yield the same result.
- Default unification is order-independent; this means that it is commutative and associative, like monotonic unification.

They explore the properties of their system, called YADU, in considerable detail. The intent is to preserve the behavior of non-default unification in cases where no default information is present, and for defeasible information at a more specific level in the type hierarchy to override defeasible information at a more general level.

We now sketch how YADU functions, using the example of English verb forms in Lascarides & Copestake (1999). The pertinent linguistic facts here are as follows: English past and passive participles are always identical in form, (simple) past tense suffixes are usually the same as the corresponding participles', and the past tense suffix of most verbs is *-ed*. The last two statements are defeasible, while the first is not. In YADU, each type is represented with a nondefeasible typed feature structure, plus a set of defeasible feature structures, each with an associated type. The type hierarchy in Figure 4 provides an example (here, the nondefeasible information comes first, and the set of defeasible structures follows the slash.

$$\begin{bmatrix} \textit{verb} & \\ \textsc{past} & \top \\ \textsc{pastp} & \boxed{2} \\ \textsc{passp} & \boxed{2} \end{bmatrix} / \left\{ \left\langle \begin{bmatrix} \textsc{past} & \boxed{1} \\ \textsc{pastp} & \boxed{1} \end{bmatrix}, \textit{verb} \right\rangle, \left\langle \begin{bmatrix} \textsc{past} & \boxed{1} \\ \textsc{pastp} & \boxed{1} \end{bmatrix}, \textit{verb} \right\rangle \right\}$$

$$\begin{bmatrix} \textit{regverb} & \\ \textsc{past} & \top \end{bmatrix} / \left\{ \left\langle \begin{bmatrix} \textsc{past} & +\textit{ed} \end{bmatrix}, \textit{regverb} \right\rangle \right\}$$

$$\begin{bmatrix} \textit{pst-tverb} & \\ \textsc{past} & \top \end{bmatrix} / \left\{ \left\langle \begin{bmatrix} \textsc{past} & +\textit{t} \end{bmatrix}, \textsc{pst-t-verb} \right\rangle \right\}$$

Figure 4: A hierarchy of past-tense formation "rules" from Lascarides & Copestake (1999: 61)

In (4), the most general type *verb* stipulates the identity of the past participle and passive participle forms as nondefault information. The value of the PAST, the simple past tense form, is unspecified (the symbol ⊤ refers to the most general type in the hierarchy). In the default information, the value of PAST is shared with both the values of both participle forms, whatever they may be. For regular verbs (type *regverb*), this value will be, by default, the result of a function that suffixes *-ed* to the verb stem (Lascarides and Copestake gloss over the details of morphology and phonology here). In the more specific type *pst-t-verb*, however, the default *-ed* is overriden by (default) information that the suffix is *-t*.

Thus a *pst-t-verb* like *burn*/*burnt* inherits the nondefault information from *regverb* and *verb*, but overrides the regular past forms. The default information in *pst-t-verb* is associated with a more specific type than that in *regverb*, so it takes precedence in YADU's unification procedure. And as Lascarides & Copestake note (p. 62): "This is the reason for separating *verb* and *regverb* in the example above, since we want the *+t* value to override the *+ed* information on *regverb* while leaving intact the default reentrancy which was specified on *verb*. If we had not done this, there would have been a conflict between the defaults that was not resolved by priority." For morphological irregularities such as *children*, the same devices can be used, with a type for the lexical entry of *child* that overrides the regular plural form.

As an example of the use of default, nonmonotonic inheritance outside of morphology, consider the account of the syntax of gerunds in various languages developed by Malouf (2000). Gerunds exhibit both verbal and nominal characteristics, and furnish a well-known example of seemingly graded category membership, which does not accord well with the categorical assumptions of mainstream syntactic frameworks. Roughly speaking, English gerunds, and their counterparts in other languages, act much like verbs in their "internal" syntax, allowing direct objects and adverbial modifiers, but function distributionally ("externally") as NPs. To take but a couple of pieces of evidence (see Malouf (2000: 27-33) for more details), gerunds can be the complement of prepositions, whereas finite clauses cannot (as in (27)); however, adverbs, not adjectives, can modify gerunds, while adjectives must be used to modify deverbal nouns (as in (28)).

(27) a. Pat is concerned about Sandy('s) getting arrested.

b. * Pat is concerned about (that) Sandy got arrested.

(28) a. Pat disapproved of (me/my) *quiet/quietly departing before anyone noticed.

b. Pat disapproved of my quiet/*quietly departure.

One approach to modeling these distinctions is directly, via syntactic rules that allow an NP to be expanded as a constituent internally headed by a verb. As Malouf notes, this offers no account of the observed behavior of gerund-like forms across languages. Some possible combinations of noun-like and verb-like attributes are frequently attested cross-linguistically in gerunds and their equivalents, while others are rare or unattested. Cross-linguistically, gerunds vary in their subcategorization possibilities: some allow subjects and complements, while some allow only complements and no subjects. But there appear to be no cases of gerund-like lexical items that can take a subject but cannot take complements.

Instead of such unmotivated syntactic rules, Malouf posits a lexical rule, which converts the lexical category of a verb to *noun*, but otherwise preserves its verbal properties, such as subcategorization. With strictly monotonic inheritance, this poses problems, as it would force us to abandon useful generalizations about nouns other than gerunds (e.g., they do not take direct object complements, as many verbs and their gerunds do). Default inheritance provides one way to model the observed phenomena, without weakening the constraints on parts of speech to the point where no meaningful constraints distinguish them.

In Malouf's account, there are both "hard" constraints – a verb lexical entry, for example, must have a HEAD value of type *relational* (encompassing verbs, adjectives, and adpositions) – and "soft," overridable constraints – a verb lexical entry by default has a HEAD value of type *verb*. In addition, following Bouma et al. (2001), he posits the types *ext-subj* and *ext-spr*. The former constrains the HEAD value to *relational* and the first element of the ARG-ST list to be the SUBJ (only adjectives, adpositions, verbs, and predicative NPs have subjects), while the latter constrains the HEAD value to *noun* and the first element of the ARG-ST list to be the SPR (only nouns have specifiers), as shown in (29).

(29) a. $\begin{bmatrix} \textit{ext-subj} & \\ \text{HEAD} & \textit{relational} \\ \text{VAL} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}, \ldots \rangle \end{bmatrix}$ b. $\begin{bmatrix} \textit{ext-spr} & \\ \text{HEAD} & \textit{noun} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \boxed{1} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}, \ldots \rangle \end{bmatrix}$

Malouf then specifies default HEAD values for the lexical classes *n* and *v* (see (30) for the latter's definition). As gerunds have both properties of nominal and relational heads, they are subtypes of both, as shown in the multiple inheritance hierarchy in Figure 5. The *v* type, which concerns us here, has a default HEAD value *verb*, as shown in (30) in addition to the non-default, more general type *relational* it also includes (default information follows /).

(30) $\begin{bmatrix} \textit{v} & \\ \text{HEAD} & \textit{relational} \,/\textit{verb} \\ \text{CONT} & \textit{psoa} \end{bmatrix}$

However, the default value *verb* is overridden in the subtype *vger*, in which the HEAD value is *gerund*, which is a subtype of both *noun* and *relational*, but not of *verb*. The type *vger* is shown in (31); where *f-ing* is a function that produces the *-ing* form of an English verb from its root.

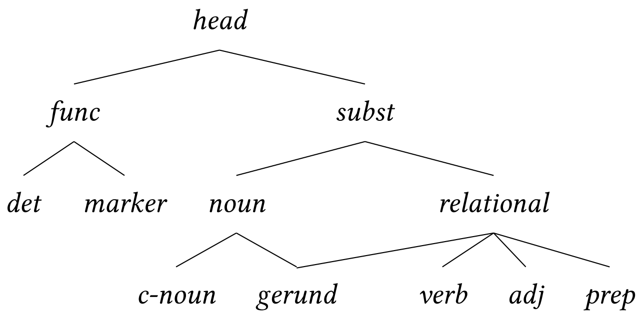


Figure 5: A cross-cutting hierarchy of types of *head* according to Malouf (2000: 65)

(31) $\begin{bmatrix} \textit{vger} & \\ \textsc{morph} & \begin{bmatrix} \textsc{root} & \boxed{1} \\ \textsc{i-form} & \textit{f-ing}(\boxed{1}) \end{bmatrix} \\ \textsc{head} & \textit{gerund} \end{bmatrix}$

The type *vger* is thus compatible with "verb-like" characteristics. But, as its HEAD is also a subtype of *noun*, it lacks a SUBJ attribute and instead has a SPR attribute. Gerunds therefore allow complements (unlike ordinary nouns), but not subjects (unlike ordinary verbs). Malouf's hierarchy of types makes this prediction, in effect, because the *ext-spr* type requires that the "external argument" (the first on the ARG-ST list) is realized as the value of SPR.

While it would be possible to construct type hierarchies of lexical types, HEAD types, and so on that would allow for "anti-gerunds" – those that would act externally as nouns, allow subjects, but not permit complements – this would require reorganizing these type hierarchies to a considerable extent. Given that many nouns besides gerunds – nominalizations, for example – are relational (that is, have a CONTENT value of type *psoa*), it could be difficult to model a hypothetical language that permits only the anti-gerunds rather than the normal ones.

Malouf further notes a key difference between gerunds and exceptions like **childs*/*children*: English gerunds are productive (and completely regular morphologically). If the same mechanisms of default unification are involved in both, what accounts for this difference? His answer is that productive and predictable processes involve on-line type construction (see Section 5.3 for details). The irregular form *children* must of course be learned and stored, not generated online. The default mechanisms described above, however, are employed at higher levels of the lexical hierarchy, and the individual gerunds forms *are* productively generated online. Note that, in contrast to the morphological and syntactic consistency among gerunds, English nominalizations display some idiosycrasies that suggest at least some of them must be stored as distinct lexical items. Thus, as Malouf emphasizes, modeling prototypicality in the lexicon within HPSG can draw on both default inheritance and on-line type construction; together, they make "the connection between prototypicality, and productivity" (p.127).

5 Lexical rules

In this section we describe the role lexical rules play in HPSG as well as their formal nature, i.e., how they model "horizontal" relations among elements of the lexicon. These are relations between variants of a single entry (be they subcategorizational or inflectional variants) or between members of a morphological family, as opposed to the "vertical" relations modeled through inheritance. Thus they provide a means to represent the intuitive notion of "derivation" of one lexeme from another.

5.1 What is the nature of lexical rules in HPSG?

While lexical rules or similar devices have been invoked within HPSG since its inception, formalizing their nature and behavior still continues. The intent, however, has always been, as Lahm (2016) stresses, to treat lexical rules (typically written $A \mapsto B$) to mean that for every lexeme or word described by A there is one described by B that has as much in common with A as possible.

Copestake & Briscoe (1991), Briscoe & Copestake (1999), Meurers (2001), and many others formalize the notion of lexical rule within HPSG by introducing a type, say *lex-rule*, with the attributes IN and OUT, whose values are respectively the rule's input and output lexical descriptions. As Briscoe & Copestake (1999) note, lexical rules of this form also bear a close relationship to default unification. The information in the input is intended to carry over to the output by default, except where the rule specifies otherwise and overrides this information. But, as Lahm (2016) points out, a sound basis for the formal details of how lexical rules work is not easily formulated. Meurers' careful analysis of how to apply lexical rules to map a description A into the description B does not always work as intended, in that what we would expect to be licit inputs are not always actually such, and no output description results as a consequence. Fortunately, it is not clear that this is a severe problem in practice, and Lahm notes that he has not found an example of practical import where Meurers' lexical rule formulation would encounter the problems he raises.

In a slight variant of the representation of lexical rules proposed by Copestake & Briscoe and Meurers, the OUT attribute can be dispensed with; the information in the lexical rule type that is not within the IN value then constitutes the output of the rule. The difference between the two representations with only the attributes SYNSEM and PHON included for expository purposes is shown in (32).

(32) a. $$\begin{bmatrix} \textsc{in} & \begin{bmatrix} \textsc{synsem} & a \\ \textsc{phon} & b \end{bmatrix} \\ \textsc{out} & \begin{bmatrix} \textsc{synsem} & c \\ \textsc{phon} & d \end{bmatrix} \end{bmatrix}$$

b. $$\begin{bmatrix} \textsc{synsem} & c \\ \textsc{phon} & d \\ \textsc{in} & \begin{bmatrix} \textsc{synsem} & a \\ \textsc{phon} & b \end{bmatrix} \end{bmatrix}$$

In the variant in (32b), lexical rules are treated as subtypes of a *derived-word* type, which can combine with other types in the lexical hierarchy, merely adding the derivational source via the IN value. Formulated in either fashion, lexical rules are essentially equivalent to unary syntactic rules, with the IN attribute corresponding to the daughter and the OUT attribute to the mother (or the rest of the information in the rule, if the OUT attribute is done away with). This is the way lexical rules are implemented in the English Resource Grammar (see http://www.delph-in.net/erg/ for demos and details about this large-scale implemented grammar of English) as well in the CoreGram Project and the Grammix grammar development environment (see Müller (2007) and https://hpsg.hu-berlin.de/Software/Grammix/ for details on the Grammix software) (**chapters/chap-cl** Chapter **??** of this volume).

One clear advantage of this kind of representation, i.e. a representation in which the attribute OUT is dispensed with and lexical "rules" are simply subtypes of *derived-word* or *derived-lexeme*, is that they are then positioned in the lexical hierarchy and subject to the same implicational constraints as other classes of words. They can also be organized in complex networks of more or less general rules. As Riehemann (1998) and Koenig (1999) show, if one includes in the lexical hierarchy unary-branching rules to model derivational morphology, a unified account of derivational processes that apply both productively to an open-ended set of lexemes as well as unproductively to another closed set of lexemes becomes possible. Consider the approach to derivational morphology taken by Riehemann (1998). Example (33) (Riehemann's (1)) illustrates *-bar* suffixation in German, a process by which an adjective that includes a modal component can be derived from verb stems (similar to English *-able* suffixation). A lexical rule approach could posit a verb stem input and derive an adjective output. As Riehemann stresses, though, there are many different subtypes of *-bar* suffixation, some productive, some unproductive, all sharing some information. This combination of productive and unproductive variants of a lexical process is exactly what the type hierarchy is meant to capture and what Riehemann's *Type-Based Derivational Morphology* capitalizes on. The structure in (34) presents the relevant information about Riehemann's type for regular *-bar* '-able' adjectives (see Riehemann (1998: 68) for more details). Critically, *-bar* adjectives include a singleton-list base (the value of MORPH-B) that records the information of the adjective's verbal base (corresponding to the would-be lexical rule's input). Because of this extra layer, the local information in the base ($local_1$) and the *-bar* adjective ($local_2$) can differ without being in conflict.

(33) Sie bemerken die Veränderung. Die Veränderung ist bemerkbar.
They notice the change. The change is noticeable.

(34)
$$\begin{bmatrix} \textit{reg-bar-adj} & \\ \textsc{phonology} & \boxed{1} + \langle \textsc{bar} \rangle \\ \textsc{morph-b} & \left\langle \begin{bmatrix} \textit{trans-verb} & \\ \textsc{phon} & \boxed{1} \\ \textsc{local} & \textit{local}_1 \end{bmatrix} \right\rangle \\ \textsc{synsem}|\textsc{local} & \textit{local}_2 \end{bmatrix}$$

5.2 Phenomena accounted for by lexical rules

Lexical rules have been put to many uses, derivational and inflectional morphology (Copestake & Briscoe 1995; Emerson & Copestake 2015), complex predicate formation and conversion (Müller 2010), negation (Kim & Sag 2002; Müller 2010) and diathesis alternations (Davis 2001). Moreover, proposals for lexical rules in HPSG have extended beyond what are traditionally or evidently viewed as lexical phenomena, to include treatments of affixal realization of arguments, extraction, unbounded dependencies, and adjuncts. In this section, we describe the use of lexical rules to model the realization of arguments as extracted dependents or affixes, rather than complements. We concentrate on two of these cases (affixal realization of arguments and complement extraction), which we will contrast with alternative analyses not involving lexical rules presented by the same authors (see the next section). They thus provide a good illustration of some of the analytical choices available to model relations between variant lexical entries based on a single stem.

We begin with the Complement Extraction Lexical Rule (hereafter, CELR) proposed in Pollard & Sag (1994) shown in (35). The input to the rule is any lexeme that selects for a syntactic argument ($\boxed{3}$) that the lexeme requires to be expressed as a complement (as indicated, this syntactic argument is also a member of the COMPS list). The ouput stipulates that this same syntactic argument is no longer a member of the COMPS list; however, the SLASH set now includes a new element, which is the local information of this syntactic argument ($\boxed{1}$). Informally stated, the input entry specifies that a syntactic argument must be realized as a complement, whereas the output entry specifies that the same syntactic argument must be realized by a non-local dependent (see Pollard & Sag (1994: Chapter 4) for the distinction between LOCAL and NON-LOCAL information).

(35)
$$\begin{bmatrix} \text{ARG-ST} & \langle \ldots, \boxed{3}, \ldots \rangle \\ \text{COMPS} & \langle \ldots, \boxed{3}[\text{LOC}\ \boxed{1}], \ldots \rangle \\ \text{SLASH} & \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \text{ARG-ST} & \left\langle \ldots, \boxed{3}\begin{bmatrix}\text{LOC} & \boxed{1} \\ \text{SLASH} & \boxed{1}\end{bmatrix}, \ldots \right\rangle \\ \text{COMPS} & \langle \ldots \rangle \\ \text{SLASH} & \{\boxed{1}\} \cup \boxed{2} \end{bmatrix}$$

A similar use of lexical rules to model alternative realizations of arguments can be found in Monachesi (1993), who analyzes alternations between complements and pronominal object affixes (traditionally called object clitics) in Italian in a way that parallels the French examples in (9). $\bigcirc$ in the rule, shown in (36), a.k.a. the "shuffle" operation, stands for the unordered concatenation of two lists, since any member of the input's COMPS list can be realized as a clitic and therefore not included in the output's COMPS list (see **chapters/chap-order** Chapter **??** of this volume for a more formal explanation of $\bigcirc$). In the output of her lexical rule, in (36), a subset of the list of complements in the input ($\boxed{2}$) corresponds to a list of clitic SYNSEMs, realized as prefixes through inflectional rules not shown here.

(36)
$$\begin{bmatrix} \textit{word} & \\ \text{HEAD} & \textit{verb} \\ \text{VAL}|\text{COMPS} & \boxed{1} \bigcirc \boxed{2} \\ \text{CLTS} & \textit{elist} \end{bmatrix} \mapsto \begin{bmatrix} \textit{word} & \\ \text{VAL}|\text{COMPS} & \boxed{1} \\ \text{CLTS} & \boxed{2}\textit{list(cl-ss)} \end{bmatrix}$$

Here as well, a lexical rule is employed in an analysis of what might well be considered a syntactic phenomenon. The possibility of treating phenomena like extraction and pronominal object affix placement at a lexical level, however, makes sense when they are considered fundamentally as matters of the combinatorial requirements of predicators, rather than effects of movement.

Before turning to the alternatives, we note in passing that lexical rules are inherently "directional", with an input and an output. This seems intuitively correct in the cases we have discussed, but might not always be so. Is there inherent directionality, for example, between the causative and inchoative alternants of verbs such as *melt* or *slide* or between the ditransitive and prepositional object frames of verbs such as *give*, as Goldberg (1995) asks? In constrast, the alternatives to lexical rules described in the following section lack this notion of directionality.

5.3 Alternatives to lexical rules

In this section we briefly examine two alternatives to lexical rules, each involving underspecification. The types of members of the ARG-ST list might be underspecified so that a single lexical description can correspond to more than one subcategorization frame. Or the type of the entry itself may be underspecified, so that it subsumes multiple inflectional or derivational forms. In both cases, the intent is that sufficiently underspecified information covers multiple entries that would otherwise have to be specified and related by lexical rules. We begin with alternatives to the complement extraction and clitic lexical rules in (35) and (36), proposed in Bouma et al. (2001) and Miller & Sag (1997).

In both cases, the idea is to distinguish between "canonical" and "non-canonical" realizations of syntactic arguments, as shown in the hierarchy of *synsem* types in Figure 9. "Canonical" means local realization as a complement or subject/specifier, and "non-canonical" means realization as an affix or filler of an unbounded dependency. Linking constraints between semantic roles (values of argument positions) and syntactic arguments (members of ARG-ST) do not specify whether the realization is canonical or not; thus they retain their original form. Only canonical members of ARG-ST must be structure-shared with members of valence lists. The two constraints that determine the non-canonical realization of fillers are shown in (37). (37a) specifies what it means to be a *gap-ss*, namely that the argument is extracted (its local information is "slashed") whereras (37b) prohibits any *gap-ss* member from being a member of the COMPS list (see Bouma et al. 2001: 23). As these two constraints are compatible with either a canonical or extracted object, there is no need for the lexical rule in (35). (DEPS in (37b) is an attribute Bouma et al. introduce that includes not only syntactic arguments, the value of ARG-ST, but also some syntactic adjuncts.)

(37) a. $\textit{gap-ss} \longrightarrow \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{SLASH} & \boxed{1} \end{bmatrix}$

b. $\textit{word} \longrightarrow \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \ominus \textit{list(gap-ss)} \\ \text{DEPS} & \langle \boxed{1} \rangle \oplus \boxed{2} \end{bmatrix}$

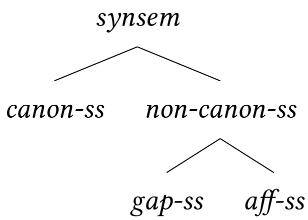


Figure 6: Types of *synsem*

Miller & Sag (1997) make a similar use of non-canonical relations between the ARG-ST list and the valence lists, eschewing lexical rules to model French pronominal object affixes (traditionally called clitics) and proposing instead the constraint in (38), where a subset of ARG-ST members, those that are realized as affixes (of type *aff*) are not also subcategorized for as complements.

(38)
$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix}\text{FORM} & F_{PRAF}(\boxed{1}, \ldots) \\ \text{I-FORM} & \boxed{1}\end{bmatrix} \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC|CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{VAL} & \begin{bmatrix}\text{SUBJ} & \boxed{2} \\ \text{COMPS} & \boxed{3}\textit{list(non-aff)}\end{bmatrix} \\ \text{ARG-ST} & (\boxed{2}\oplus\boxed{3}) \bigcirc \textit{nelist(aff)}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

In both of these analyses, related sets of lexical entries that could be thought of as "generated by lexical rules" are instead regarded as the various possible ways of obeying constraints like those in (37) or (38), This comes at a cost of additional types and constraints for extraction, and a loosening of requirements for the correspondence between the ARG-ST list and the valence lists. However, these approaches, in dispensing with lexical rules, sidestep the conceptual and representational issues that we noted earlier and attempts to restrict lexical rules to cases where they cannot be avoided, e.g., derivational morphology.

The second alternative to lexical rules based on underspecification was presented in Koenig & Jurafsky (1994) and Koenig (1999). Typically in HPSG, all possible combinations of types are reified in the type hierarchy (in fact, they must be present, per the requirement that the hierarchy be sort-resolved, Carpenter 1992b, Pollard & Sag 1994), or, equivalently, that each linguistic entity be assigned exactly one maximally specific type – a.k.a. *species* Richter (2000: 78)). Thus, if one partitions verb lexemes into transitive and intransitive and, orthogonally, into, say, finite verbs and gerunds (limiting ourselves to two dimensions here for simplicity), the type hierarchy must also contain the combinations transitive+finite, transitive+gerund, intransitive+finite, and intransitive+gerund. Naturally, this kind of fully enumerated type system is unsatisfying. For one thing, there is no additional information that the combination subtype *transitive+finite* carries that is not present in its two supertypes *transitive* and *finite*, and similarly for the other combinations. In contrast to the "ordinary" types, posited to represent information shared by classes of lexemes, these combinations seem to have no other purpose than to satisfy a formal requirement on the mathematical structure of a type hierarchy (namely, that it forms a lattice under meet and join). Second, and related to the first point, this completely elaborated type hierarchy is redundant. Once you know that all verbs fall into two valence classes, transitive and intransitive, and simultaneously into two inflectional classes, finite and gerund, and that valence and inflection are two orthogonal dimensions of classification of verbs, you know all you need to know; the type of any verb can be completely predicted from these two orthogonal dimensions of classification and standard propositional calculus inferences.[2]

Figure 7 is a simplified hierarchy of verb lexemes we use for strictly expository purposes, where the boxed labels in small caps VFORM and ARG-ST are mnemonic names of orthogonal dimensions of classification of subcategories of verbs (and are not themselves labels of subcategories). Inheritance links to the predictable subtypes are dashed and their names grayed out; this indicates that these types can be inferred, and need not be declared explicitly as part of the grammar. A grammar of English would include statements to the effect that head information about verbs includes a classification of verbs into finite or base forms (of course, there would be more types of verb forms in a realistic grammar of English) as well a classification into intransitive and transitive verbs (again, a realistic grammar would include many more types).

Crysmann & Bonami (2016) have shown how this *online type construction*, where predictable combinations of types of orthogonal dimensions of classification are not reified in the grammar, is useful when modeling productive inflectional morphology. Consider, for example, exponents of morphosyntactic features whose shape remains constant, but whose position within a word's template (to speak informally here) varies. One case like this is the subject and object markers of Swahili, which can occur in multiple slots in the Swahili verb template

[2]One possible way of making formally explicit the idea behind on-line type construction within the model-theoretic approach to HPSG that is now standard (King 1989; Richter 2000) is to allow maximally specific sorts, or species, to be non-atomic – either sets of species or non-atomic sums of species – just in cases where orthogonal dimensions of classification have been used since Flickinger (1987). For reasons of space, we do not pursue this line of inquiry in this chapter.

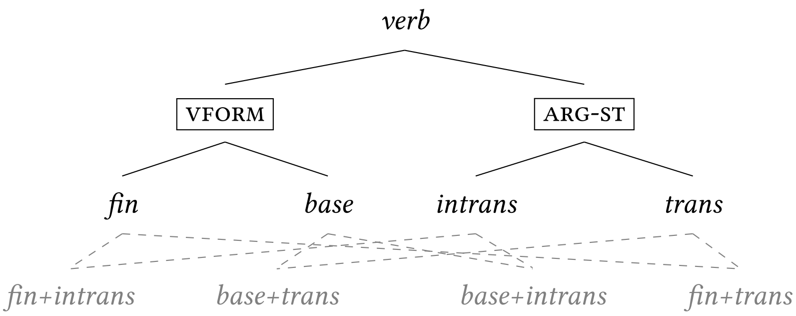


Figure 7: An example of on-line type construction

(Stump 1993; Bonami & Crysmann 2016).

For reasons of space we illustrate the usefulness of this dynamic approach to type creation, the Type Underspecified Hierarchical Lexicon (TUHL) with an example from Koenig (1999), the cross-cutting classification of syntactic/semantic information and stem form in the entry for the French verb *aller* (see Bonami & Boyé (2001) for a much more thorough discussion of French stem allomorphy along similar lines; Crysmann and Bonami's much more developed approach to stem allomorphy would model the same phenomena differently and we use Koenig's simplified presentation for expository purposes only). The forms of *aller* are based on four different suppletive stems: *all-* (1st and 2nd person plural of the indicative and imperative present, infinitive, past participle, and imperfective past), *i-* (future and conditional), *v-* (1st-3rd person singular and 3rd person plural of the indicative present), and *aill-* (subjunctive present). These four suppletive stems are shared by all entries (i.e., senses) of the lexeme *aller*: the one which means 'to fit' as well as the one which means 'to leave', as shown in (39) (see Koenig, p.40–41). The cross-cutting generalizations over lexemes and stems are represented in Figure 8. Any *aller* stem combines one entry and one stem form. In a traditional HPSG type hierarchy, each combination of types (grayed out in Figure 8), would have to be stipulated. In a TUHL, these combinations can be dynamically created when an instance of *aller* needs to be produced or comprehended.

(39) a. Marc est allé à Paris.
Marc be-PR.3RD.SG go-PPT to Paris
'Marc went to Paris.'

b. Marc s'en ira. 'Marc will leave.'
Marc 3.REFL-of.it go-FUT.3RD.SG

c. Ce costume te va bien.
This suit you go-PR.3.SG well
'This suit becomes you.' (lit. goes well to you)

d. Il faut que j'y aille.
It must that I.to.there go-SUBJ.PR.1.SG
'I must go there.'

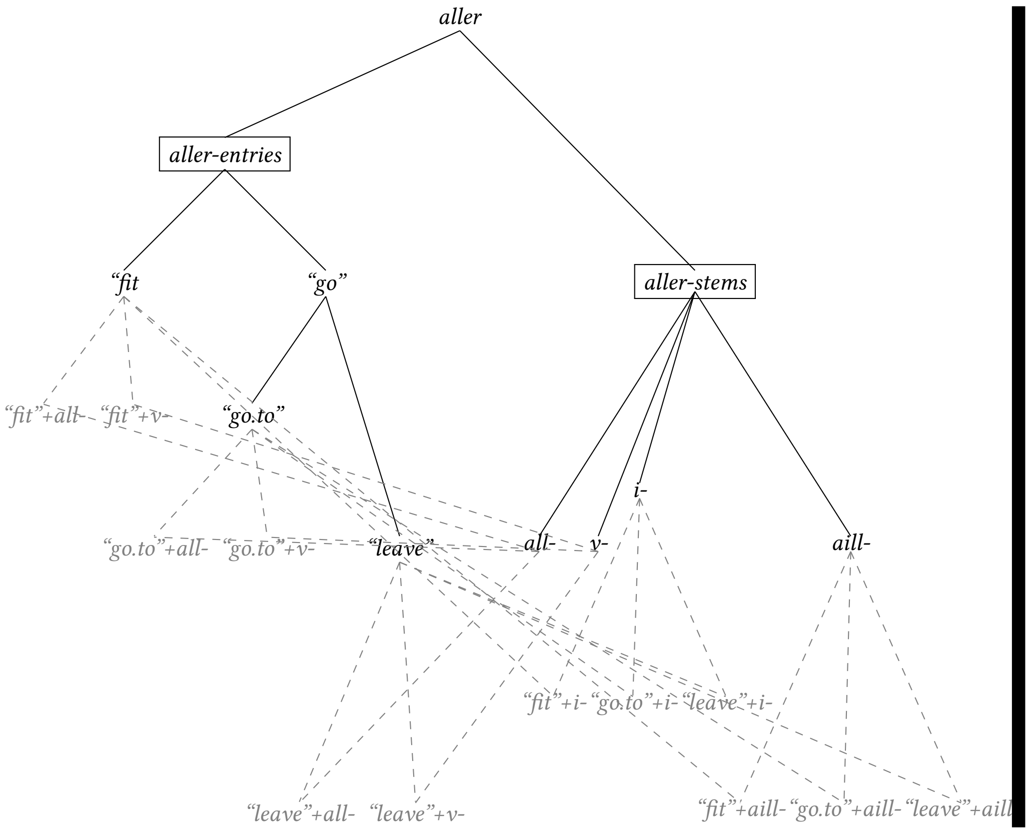


Figure 8: A hierarchy of lexical entries and stem-forms for the French verb *aller*, from Koenig (1999)

Both the distinction between canonical and non-canonical *synsem* and type underspecification avoid conflict between the information specified in the variants of words based on a single lexeme (e.g., conflicts on how syntactic arguments are realized); they abstract over the relevant pieces of conflicting information. Underspecifying information included in lexical entries or lexical types allows a single entry or type to stand for the two distinct entries or types that would be related as input and output by lexical rules.

Lexical rules have played a crucial role in the rise of lexicalist approaches to syntax. But the two alternative analytical tools we discussed in this section (which, of course, can be combined in an analysis) have chipped away at their use in HPSG. Inflectional morphology is now dealt with through lexical types associating morphosyntactic features with forms/positions and constraints on words (ensuring that all morphosyntactic features are realized, see **chapters/chap-morphology** Chapter **??** of this volume). Non-canonical realization of syntactic arguments as affixes or fillers in unbounded dependencies is now modeled by distinguishing kinds of members of the ARG-ST list and constraints on words that relate valence, argument structure, and dependents lists.

So, what remains of the case for lexical rules? Well, first, as we showed above, lexical rules are now simply unary-branching rules within the lexical part of the type hierarchy. As such they are not formally distinct from the rest of the lexical hierarchy or the hierarchy of signs, as they used to be. Second, they are not meant to model just unproductive processes, as they were originally intended to in Jackendoff (1975); Bochner (1993). They can be used to model unproductive processes, but they can also model productive derivational processes (in fact both when a single derivational process is both).

Still, the existence of two distinct ways of dealing with potential conflict of information – underspecification or unary branching rules – raises the issue of which one should be used when. Unfortunately, there is no general guideline; it depends on the nature of the data that needs to be modeled. Müller (2006; 2010) argues that diathesis phenomena, broadly speaking, favor a lexical rules approach over a phrase-structural constructional approach à la Goldberg (1995) or an online type construction approach suggested in Kay (2002). The arguments are convincing, but it should be noted that some of the data involves derivational morphology (e.g., causatives) or passive morphemes, which involves a Type-Based Derivational Morphology of the kind Riehemann argues for (such an approach was suggested in Koenig (1999: Chapter 4)). What remains unclear to us is whether there are instances where lexical rules as unary-branching rules are a better model of "horizontal" generalizations that do not involve morphological processes, i.e. whether the kind of lexical rules Pollard & Sag (1994) proposes (e.g., the Complement Extraction Lexical Rule) are ever motivated over the underspecification treatment of such phenomena proposed in Bouma et al. (2001).

6 Conclusion

Our principal goals in this chapter have been to present the HPSG viewpoint on the structure and content of individual lexical entries, and the organization of the lexicon as a whole. Unsurprisingly, both of these are pervaded by HPSG's lexicalist stance. With regard to lexical entries, this entails informationally rich and sometimes complex representations. A lexical entry models not only a word's idiosyncratic properties, but also its general morphological, distributional, combinatorial, and semantic characteristics. Consequently, HPSG researchers have devoted a great deal of attention to representing all of these in a parsimonious way, so as to avoid massive redundancy in the lexicon. We have surveyed several techniques addressing how to parcel out information shared among entries into descriptions that are true of sets of entries. First, feature geometry plays a key role in organizing portions of this information within a lexical entry in "packages" that tend to recur throughout the lexicon. This in turn allows these recurring portions to be associated with types in a hierarchy. Through inheritance, these common elements can be stated in just one location for the class of words that share them, and multiple inheritance makes it possible to represent numerous crosscutting classifications of words. We have shown two ways in which HPSG scholars have exploited these mechanisms. One is by creating a hierarchy of subtypes of *word*, each with associated constraints. The other, probably more commonly employed in current work, is to posit type hierarchies of various objects within lexical entries, along with implicational statements that constrain the content of a lexical entry containing those types of objects.

This hierarchical character of the HPSG lexicon serves to model the "vertical" relationships among classes of words, based on properties like part of speech, subcategorization, linking, morphological and paradigmatic classes, and so forth. There is also a "horizontal" aspect of lexical relations, however, for which lexical rules explicitly relating one class of lexemes or words to another have been proposed. While their original use was primarily to model systematic sets of, say, forms in an inflectional paradigm, HPSG's lexicalist approach to syntax has also seen them employed in accounts of phenomena such as extraction, traditionally regarded as outside the lexicon. We also presented two alternatives to lexical rules that appear to handle these phenomena equally well. One involves underspecification within lexical entries in a way that permits them to describe the right range of related forms, while the other allows underspecification within type hierarchies, and requires fully specified types to be constructed "online". Both of these alternatives, like lexical rules, avoid massively repetitive specification of properties of families of systematically related words. Lexical rules as well as the two alternatives we outlined are independently needed and, although one can make suggestive remarks as to when to use lexical rules or either alternative, the issue cannot be settled *a priori* and must be argued on a case by case basis. But, the rich and intricate hierarchical lexicon cum lexical rules is a defining, enduring, and pervasive feature of HPSG, more prominent here than in almost any other grammatical framework.

References

Ackerman, Farrell & Gert Webelhuth. 1998. *A theory of predicates*. Stanford, CA: CSLI Publications.

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

Baldridge, Jason. 2002. *Lexically specified derivational control in Combinatory Categorial Grammar*. Edinburgh: University of Edinburgh dissertation.

Bochner, Harry. 1993. *Simplicity in Generative Morphology*. Berlin: Mouton de Gruyter.

Bonami, Olivier & Gilles Boyé. 2001. Suppletion and dependency in inflectional morphology. In Frank van Eyde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the 2001 HPSG conference*, 51–70. Stanford: CSLI Publications.

Bonami, Olivier & Berthold Crysmann. 2016. Morphology in constraint-based lexicalist approaches to grammar. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology*, 609–656. Cambrdgei: Cambridge University Press.

Booij, Gert. 2005. Compounding and derivation: evidence for Construction Morphology. In Wolfgang Dressler, Dieter Kastovsky, Oskar Pfeiffer & Franx Rainer (eds.), *Morphology and its demarcations*, 109–132. Amsterdam: John Benjamins.

Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: syntactic projections and the lexicon. In John Moore & Maria Polinsky (eds.), *The nature of explanation in linguistic theory*, 31–67. Stanford: CSLI Publications.

Bouma, Gosse, Robert Malouf & Ivan Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Brachman, Ronald & James Schmolze. 1985. An overview of the KL-ONE knowledge representation system. *Cognitive Science*. 2nd ser. 9(2). 171–216.

Bresnan, Joan & Sam Mchombo. 1995. The Lexical Integrity Principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13(2). 181–254.

Briscoe, Ted & Ann Copestake. 1999. Lexical rules in constraint-based grammars. *Computational Linguistics* 25(4). 487–526.

Broadwell, George Aaron. 2008. Turkish suspended affixation is lexical sharing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG08 conference*, 198–213. Stanford: CSLI Publications.

Bruening, Benjamin. 2018. The lexicalist hypothesis: both wrong and superfluous. *Language* 94(1). 1–42.

Carpenter, Bob. 1992a. Categorial Grammars, lexical rules, and the English predicative. In Bob Levine (ed.), *Formal grammar: theory and implementation*, 168–242. New York, NY: Oxford University Press.

Carpenter, Bob. 1992b. *The logic of typed feature structures*. Cambridge: Cambridge University Press.

Chaves, Rui. 2008. Linearization-based word-part ellipsis. *Linguistics and Philosophy* 31(3). 261–307.

Chaves, Rui. 2014. On the disunity of right-node raising phenomena: extraposition, ellipsis, and deletion. *Language* 90(4). 834–886.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Boston: Ginn.

Copestake, Ann. 2002. *Implementing typed feature structures*. Stanford: CSLI Publications.

Copestake, Ann & Ted Briscoe. 1991. Lexical operations in a unification-based framework. In James Pustejovsky & Sabine Bergler (eds.), *Lexical semantics and knowledge representation*, 101–119. Berlin: Springer Verlag.

Copestake, Ann & Ted Briscoe. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12(1). 15–67.

Crysmann, Berthold. 2000a. Clitics and coordination in linear structure. In Birgit Gerlach & Janet Grijzenhout (eds.), *Clitics in phonology, morphology and syntax*, 121–159. Amsterdam: John Benjamins Publishing Co.

Crysmann, Berthold. 2000b. On the placement and morphology of Udi subject agreement. In *7th international conference on Head-driven Phrase Structure Grammar*. Berkeley, CA.

Crysmann, Berthold. 2002. *Constraint-based coanalysis: Portuguese cliticisation and morphology-syntax interaction in HPSG*. Saarbrücken: University of Saarland dissertation.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics and Information-based Morphology. *Journal of Linguistics* 52(2). 311–374.

Davis, Anthony. 2001. *Linking by types in the hierarchical lexicon*. Stanford: CSLI Publications.

Davis, Anthony & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.

Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.

Dost, Ascander. 2007. A domain-based approach to second position clitics in Pashto. In Frederick Hoyt, Nikki Seifert, Alexandra Teodorescu & Jessica White (eds.), *Texas Linguistic Society IX: the morphosyntax of underrepresented languages*. Austin, TX.

Dowty, David. 1978. Transformations as lexical rules in a montague grammar. *Linguistic Inquiry* 9. 393–426.

Embick, David. 2015. *The morpheme: a theoretical introduction*. Berlin: de Gruyter Mouton.

Emerson, Guy & Ann Copestake. 2015. Lacking integrity: HPSG as a morphosyntactic theory. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-Driven Phrase Structure Grammar, Nanyang Technological University (NTU), Singapore*, 75–95. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2015/, accessed 2018-2-25.

Flickinger, Dan, Carl Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In *Proceedings of the 23rd annual meeting of the Association for Computational Linguistics*, 262–267. Chicago.

Flickinger, Daniel. 1987. *Lexical rules in the hierarchical lexicon*. Stanford: Stanford University dissertation.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: The University of Chicago Press.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 111–176. Cambridge, MA: MIT Press.

Harris, Alice. 2000. Where in the word is the Udi clitic? *Language* 76(3). 593–616.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80.

Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language*. 3rd ser. 51. 639–671.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Kathol, Andreas. 1999. *Linearization-based syntax*. Oxford: Oxford University Press

Kay, Paul. 2002. An informal sketch of a formal architecture for Construction Grammar. *Grammars* 5. 1–19.

Kim, Jong-Bok & Ivan Sag. 2002. Negation without head movement. *Natural Language and Linguistic Theory* 20(2). 339–412.

King, Paul. 1989. *A logical formalism for head-driven phrase structure grammar*. Manchester: University of Manchester dissertation.

Koenig, Jean-Pierre. 1999. *Lexical relations*. Stanford: CSLI Publications.

Koenig, Jean-Pierre & Anthony Davis. 2003. Semantically transparent linking in HPSG. In Stefan Müller (ed.), *Proceedings of the HPSG 2003 conference*, 222–235. Stanford: CSLI Publications.

Koenig, Jean-Pierre & Daniel Jurafsky. 1994. Type underspecification and on-line type construction in the lexicon. In *Proceedings of WCCFL XIII*, 270–285. Stanford: CSLI Publications.

Lahm, David. 2016. Refining the semantics of lexical rules in hpsg. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar, Polish Academy of Sciences*, 360–379. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2016/headlex2016-lahm.pdf.

Lakoff, George. 1970. *Irregularity in syntax*. New York: Holt, Rinehart, & Winston.

Lascarides, Alex & Ann Copestake. 1999. Default representation in constraint-based frameworks. *Computational Linguistics* 25(1). 55–105.

Malouf, Robert. 2000. *Mixed categories in the hierarchical lexicon*. Stanford, CA: CSLI Publications.

Marantz, Alec. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis, L. Siegel, C. Surek-Clark & A. Williams (eds.), *University of pennsylvania working papers in linguistics*, vol. 4, 201–225. Philadelphia: University of Pennsylvania Linguistics Department.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13(1). 91–106.

Meurers, Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217.

Miller, Philip & Ivan Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Monachesi, Paola. 1993. Object clitics and clitic climbing in Italian HPSG grammar. In *Proceedings of the sixth conference of the European chapter of the Association for Computational Linguistics*, 437–442. Utrecht.

Müller, Stefan. 2006. Phrasal or lexical constructions? *Language* 82(4). 850–883.

Müller, Stefan. 2007. The Grammix CD Rom: A software collection for developing typed feature structure grammars. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar Engineering across Frameworks 2007* (Studies in Computational Linguistics ONLINE), 259–266. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/, accessed 2018-2-25.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655.

Müller, Stefan. 2015. The CoreGram project: theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3. 21–86.

Pollard, Carl & Ivan Sag. 1987. *Information-based syntax and semantics*. Stanford: CSLI Publications.

Pollard, Carl & Ivan Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl & Ivan Sag. 1994. *Head-driven Phrase-Structure Grammar*. Chicago, IL: Chicago University Press.

Przepiórkowski, Adam. 2001. ARG-ST on phrases: evidence from Polish. In Dan Flickinger & Andreas Kathol (eds.), *Proceedings of the 7th international conference on head-driven phrase structure grammar*, 267–284. CSLI Publications.

Reape, Mike. 1994. Domain union and word order variation in german. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head driven Phrase Structure Grammar*, 151–197. Stanford University: CSLI Publications.

Richter, Frank. 2000. *A mathematical formalism for linguistic theories with an application in Head-driven Phrase Structure Grammar*. Tübingen: University of Tübingen dissertation.

Riehemann, Susanne Z. 1998. Type-based derivational morphology. *The Journal of Comparative Germanic Linguistics* 2(1). 49–77.

Roberts, Taylor. 2000. *Clitics and agreement*. Cambridge, MA: MIT dissertation.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Stump, Gregory. 1993. Position classes and morphological theory. In Gert Booij & Jaap van Marle (eds.), *Yearbook of Morphology*, 129–180. Kluwer.

Tegey, Habibullah. 1977. *The grammar of clitics: evidence from Pashto and other languages*. Urbana-Champaign: University of Illinois at Urbana-Champaign dissertation.

Tomasello, Michael. 2003. *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Luc Steels (ed.), *Design patterns in Fluid Construction Grammar*, 115–145. Amsterdam: John Benjamins Publishing.

Wechsler, Stephen. 1999. HPSG, GB, and the Balinese bind. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 179–195. Stanford, CA: CSLI publications.

Chapter 5

HPSG in understudied languages

Douglas L. Ball
Truman State University

Work within HPSG has explored typologically-different and genetically-diverse languages, though the framework is not well-known for such explorations. This chapter details some of that work, focusing on the phenomena of argument indexing (pronoun incorporation or agreement), non-accusative alignment, and VSO constituent order. Examination of proposed and possible analyses within these areas reveals that HPSG can flexibly handle a wide range of languages all while maintaining a certain uniform "underlying structure" within the analyses.

1 Introduction

To date, the most intensely studied language within the HPSG framework has been English; this follows the trend in modern syntactic theorizing at large: English is currently the best described language in the world. Still, there has been plenty of work within HPSG on languages other than English (ISO 639-3 code: eng); in fact, substantial work has occurred within the framework[1] on German (ISO: deu) (Crysmann 2003; Müller 2013), Danish (ISO: dan) (Müller & Ørsnes 2015), Norwegian (ISO: nor) (Hellan & Haugereid 2003), French (ISO: fra) (Abeillé & Godard 2000; 2002; 2004; Abeillé et al. 2006), Spanish (ISO: spa) (Marimon 2013), Portuguese (ISO: por) (Branco & Costa 2008), Mandarin Chinese (ISO: cmn) (Müller & Lipenkova 2013; 2016; Yang & Flickinger 2014), Japanese (ISO: jpn) (Siegel et al. 2016), and Korean (ISO: kor) (Kim et al. 2011), among others. However, work within HPSG is not particularly well-known for exploring a wide range of typologically- and genetically-diverse languages, certainly not to

[1]Citations in this paragraph are to works, if available, whose focus is on the entire morphosyntax of the language in question rather than on particular issues in these languages.

Douglas L. Ball. 2019. HPSG in understudied languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 93–127. Berlin: Language Science Press. DOI:??

the degree of its constraint-based lexicalist cousin, Lexical Functional Grammar. Nevertheless, there has been work within HPSG on such languages; this chapter will discuss some of this work as well as suggesting some further avenues for HPSG work within these languages.

The term I will employ for these typologically- and genetically-diverse languages is *understudied languages*. Which languages qualify as understudied languages, though? Does the term just cover languages that have not previously been investigated, syntactically? Or maybe all the languages without any previous HPSG work? Or maybe the term encompasses any language that is not the most described language, English? Though I reject all these (somewhat jocular) definitions, I do grant that *understudied language* is surely a fuzzy category, with boundaries that are difficult to demarcate and with conditions for inclusion that could be controversial. As a working benchmark for this chapter, I will suppose that the term *understudied languages* includes those languages that have a combined native and non-native speaker population of 1.2 million or fewer (roughly 0.01% of the world's population at present), that are spoken currently or have gone extinct within the last 120 years, that are generally spoken in a smaller, contiguous part of the globe, and that are not usually employed in international diplomacy or commerce. With this benchmark, languages[2] like Tongan (Polynesian, Austronesian; Tonga; ISO: ton), Kimaragang (Dusunic, Austronesian; Sabah, Malaysia; ISO: kqr), Warlpiri (Ngumpin-Yapa, Pama-Nyungan; west central Northern Territory, Australia; ISO: wbp), Burushaski (isolate; Gilgit-Baltistan, Pakistan; ISO: bsk), Lezgian (Lezgic, Nakh-Dagestanian; southern Dagestan, Russia; ISO: lez), Maltese (Semitic, Afro-Asiatic; Malta; ISO: mlt), Basque (isolate; Basque Country, Spain & France; ISO: eus), Welsh (Celtic, Indo-European; Wales, UK; ISO: cym), Oneida (Iroquoian; New York & Wisconsin, USA; Ontario, Canada; ISO: one), Coast Tsimshian (Tsimshianic; NW British Columbia, Canada & SE Alaska, USA; ISO: tsi), Yucatec Maya (Mayan; Yucatan Peninsula, Mexico & Belize; ISO: yua), and Macushi (Cariban; Roraima, Brazil, E Venuzuela, & SE Guyana; ISO: mbc) would all be included, while the ten languages mentioned in the first paragraph would not.[3]

[2] The locations and genetic affiliations of the languages listed here were checked at Hammarström et al. (2018) (accessed 2018-07-31).

[3] Some of the languages listed above will be discussed further in this chapter. Others from the above list are well-known understudied languages from the linguistics literature. A few of these languages have HPSG work that is not mentioned elsewhere in this chapter: on Tongan, see also Dukes (2001); on Warlpiri, see Donohue & Sag (1999); on Maltese, see Müller (2009); on Basque, see also Crowgey & Bender (2011); on Oneida, see also Koenig & Michelson (2010); on Yucatec Maya, see Dąbkowski (2017).

The denotation of the term *understudied languages* is intended to be different from *endangered language*; there is nothing in the above supposition of what an understudied language is that says that an understudied language may (or may not) potentially cease to be spoken within the next five to seventy years (see discussion in Krauss 1992 and Simons & Lewis 2013 for more on endangered languages and the crisis they face). However, the two terms do, in actuality, overlap: many understudied languages are endangered languages. While the use of HPSG (or other formal syntactic framework) or not likely has no direct bearing on the continued viability of a particular language, the practitioners of HPSG join with other linguists in seeing the importance of documenting such languages and supporting the rights of communities of endangered languages to continue to speak these languages.

Understudied languages exhibit a great variety of syntactic behaviors – some of them quite similar to "well-studied languages," some of them quite different – and these languages do not form an obvious natural class, syntactically. Due to space limitations, I will focus on just a very small portion of the syntactic phenomena of understudied languages: argument indexing, non-accusative alignment (chiefly ergativity), and VSO constituent order. These phenomena and their analyses will give the reader a sense of how HPSG has been or could be applied to understudied languages. Unfortunately, this means that a collection of phenomena made famous by understudied languages – including, among others, noun incorporation (but see Malouf 1999; Runner & Aranovich 2003; Ball 2005a,b; 2008), serial verbs (but see Muansuwan 2001; 2002; Kropp Dakubu et al. 2007; Müller & Lipenkova 2009; Lee 2014), clause-chaining, evidentiality systems (but see Lee 2012), object-initial word order, and applicatives (but see Runner & Aranovich 2003; Ball 2008; 2010) – will not be discussed.

In going through the phenomena to be discussed, it will become clear that HPSG can flexibly handle a wide range of languages even while keeping its core characteristics. In fact, in most areas of analytic interest, several different approaches within the framework are equally viable at the outset. Relatedly, the analysis of many areas, especially from a cross-linguistic perspective, is far from settled. This seems to me to be an advantage: it allows competing analyses to be modeled clearly and precisely, while allowing empirical facts to better adjudicate between approaches.[4]

In my discussion, I will move through the three areas of argument indexing (in Section 2), non-accusative alignment (Section 3), and VSO constituent order (Section 4), which corresponds to decreasing pervasiveness – roughly estimated

[4]This point is also made in Fokkens (2014), especially in Chapter 1.

– for each phenomenon across the world's languages.

2 Argument indexing

Widespread among all sorts of natural languages – understudied or not – is what Haspelmath (2013) terms argument indexing. In argument indexing, morphologically dependent elements – that is, affixal elements usually located within or near the verb and with denotations (seemingly) similar to pronouns – either occur in place of arguments of the main semantic predicate of the clause or alongside them.[5] While this phenomenon occurs even a bit in English and throughout other European languages, argument indexing in understudied languages tends to be more "rampant": that is, all (or most) of the verb's arguments are indexed, rather than just the subject being indexed, as is the most common pattern in Europe (Siewierska 2013b). When the argument indexing is "rampant": its treatment within the syntax (and within the morphology-syntax interface) of a language becomes a key question. HPSG analyses offer several possible answers how the syntax of argument indexing works, all while maintaining the framework's surface orientation. Empirically, it is clear that not all argument indexes behave in quite the same way in all languages, so I will explore the analysis of two subtypes of indexes in the sections to follow: first, indexes that do not co-occur with external, role-sharing noun phrases, and, second, indexes that can co-oocur with external, role-sharing noun phrases.[6]

2.1 Indexing in complementary distribution with conominals

Some argument indexes in some languages have the property that they do not – and cannot – appear with a non-pronominal element sharing their same syntactic/semantic role in the same narrow clause (I will refer to these non-pronominal elements as conominals, following Haspelmath 2013). The term that Haspelmath suggests – and I will use – for such argument indexes is pro-index. One language showing pro-index behavior with its argument indexes is Macushi, as revealed from the examples in (1):

[5]Thus, this area includes what has been considered to be predicate-argument agreement as well as what some consider to be "pronoun incorporation", though one of the key points of Haspelmath (2013) is that the pre-existing terminology – if not also the pre-existing analyses – in this domain has been misleading.

[6]See also Saleem (2010) for a similar – though not identical – analysis of the same analytical domain.

(1) Macushi [mbc] (Abbott 1991 via Siewierska 1999 and Corbett 2003)

a. i-koneka-'pî-u-ya
3SG.ABS-make-PAST-1SG-ERG
'I made it'

b. * uurî-ya i-koneka-'pî-u-ya
1SG-ERG 3SG.ABS-make-PAST-1SG-ERG
Intended: 'I made it'

The example in (1a) is just a verb with all its arguments realized as argument indexes. The example in (1b) clearly reveals the pro-index behavior of the argument indexes: the affixed verb is incompatible with an independent pronoun, such as *uurîya* '1SG.ERG'.

The pro-index phenomenon has a straightforward (and, as a result, commonly assumed) analysis within HPSG. The analysis was originally proposed by Miller & Sag (1997) for French "clitics", but could equally be applied to the Macushi case above, among others. Key to this analysis is the idea found in most versions of HPSG (emerging in the mid-to-late 1990s) that there are separate kinds of lists for the combinatorial potential(s) of heads. In fact, not only are there these separate lists, but there can be (principled) mismatches between them (see Borsley & Abeillé 2019, Chapter 1 of this volume and Wechsler, Koenig & Davis 2019, Chapter 9 of this volume). The first of these lists is the ARGUMENT-STRUCTURE (ARG-ST) list. This list handles phenomena related to the syntax-semantic interface, like linking (Davis 2001), case assignment (Przepiórkowski 1999), and binding restrictions (Manning & Sag 1998; Wechsler & Arka 1998). The other lists are the two valence lists, the SUBJECT (SUBJ) list and the COMPLEMENTS (COMPS) list. These are concerned with the "pure" syntax and mediate which syntactic elements can combine with which others.

On the Miller & Sag-style analysis of pro-indexes, the verb's ARG-ST list contains feature descriptions corresponding to all of its arguments. For the examples in (1), the verb's ARG-ST list would include a feature description for both the semantic maker and the semantic element that is made (as will appear in (3)). However, the same verb's SUBJ and COMPS lists would contain no elements corresponding to any affixed arguments. What prompts this disparity? The arguments realized by affixes correspond to a special kind of feature description on the ARG-ST list, typed *non-canonical*.[7] (Intuitively, these arguments are realized

[7] In the version of HPSG of Ginzburg & Sag (2000), *non-canonical* was an immediate subtype of *synsem*; and the relevant feature descriptions were thus seen as syntactico-semantic complexes. In the latter-day version of HPSG known as Sign-Based Construction Grammar (Sag 2012), the *non-canonical* type was re-christened *covert* and was an immediate subtype of *sign*, the

in a non-canonical way.) Feature descriptions of the *non-canonical* type differ from their sibling type *canonical* in how they interact with the SUBJ and COMPS lists. Governing the relationship between the ARG-ST and these valence lists is the Argument Realization Principle, which is stated in (2):

(2) $$\textit{word} \Rightarrow \begin{bmatrix} \text{SS|LOC|CAT} & \begin{bmatrix} \text{SUBJ} & \boxed{B} \\ \text{COMPS} & \boxed{C} \end{bmatrix} \\ \text{ARG-ST} & \boxed{A}\ \textit{list(canonical)}\ (= \boxed{B} \oplus \boxed{C}) \bigcirc \boxed{D}\ \textit{list(non-canonical)} \end{bmatrix}$$

(cf. Ginzburg & Sag 2000: 171)

The constraint in (2) says that only the list of *canonical* feature descriptions on the ARG-ST can (and must) appear on the SUBJ and COMPS lists. Thus, the principle disallows any *non-canonical* argument from being mapped to either of these valence lists. This, then, captures the idea that such affixal arguments are generally inert in the combinatorics of the syntax proper: they saturate an argument slot and that argument slot is no longer available for other (at least direct) syntactic combination.

So, returning to the Macushi word *ikoneka'pîuya* 'I made it' from (1a), it will have the partial lexical description in (3):

(3) $$\begin{bmatrix} \text{SS|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{RELS} & \left\langle \begin{bmatrix} \textit{make-rel} \\ \text{ARG1} & \boxed{1} \\ \text{ARG2} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \langle [\textit{non-canonical}]_{\boxed{1}}, [\textit{non-canonical}]_{\boxed{2}} \rangle \end{bmatrix}$$

In (3), the word's ARG-ST list is comprised of two *non-canonical* feature descriptions (corresponding to the maker argument with index $\boxed{1}$ and the made argument, with index $\boxed{2}$). Yet, by the Argument Realization Principle, the SUBJ and COMPS lists are empty. Consequently, *ikoneka'pîuya* 'I made it', can be a clause by itself – as it is in (1a) – because it requires no other valents (that is, it has a "saturation" level on a par with a clause) and it is headed by a verb (just like a clause is).

The specification of an empty list also contributes to ruling out examples like (1b) with a conominal. On the standard HPSG view of how valence is managed,

relevant feature descriptions being entire signs. In spite of these differences, which may seem significant, the analysis is very similar in both versions of the framework. Several subtypes of *non-canonical*/*covert* have been recognized, the subtype relevant for this example would be *aff*. But I will just use the *non-canonical* type here.

no element with an empty valence list can combine with any possible valence-saturating syntactic entity, like an NP.[8] Thus, the grammar would correctly not license a tree like in Figure 1.[9]

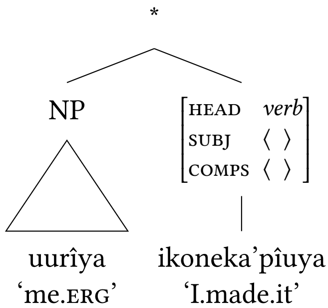


Figure 1: A tree of an illicit conominal–pro-index combination

Overall, the Argument Realization Principle-mandated non-mapping of the *non-canonical* ARG-ST list members to either the SUBJ or the COMPS list captures the key behavior found in the pro-indexing type of argument indexing: the argument indexes occur in complementary distribution with any conominal.

2.2 Indexing co-occurring with conominals

Even though the pro-index type of argument index has a more straightforward analysis, this type is not the most common in the world's languages. Rather, the most common type of argument indexing appears to be the one where the argument indexing affix(es) can co-occur with a conominal, but do(es) not have to. In Haspelmath's (2013) terms, this is the cross-index type.[10] A language exhibiting this type of behavior is Basque, as evident from the example in (8):

[8] To truly rule out the NP from combining with the verb in Figure 1, the NP would also need to not match any non-local requirements of the verb.

[9] The tree in Figure 1, as well as other trees in this chapter, only provides the relevant attribute-value pairs, suppressing the geometry of features found in more articulated feature descriptions.

[10] The behavior of cross-indexes is canonical for so-called "pro-drop" languages, a term arising from the transformational syntax tradition (particularly from Chomsky 1981), but now with wider currency.

(4) Basque [eus] (Laka 1996: 98)

Zuk	niri	liburua	saldu	d-i-da-zu.
2SG.ERG	1SG.DAT	book.DEF	sold	3SG.ABS-AUX-1SG.DAT-2SG.ERG

'You have sold me the book.'

Though *zuk* 'you', *niri* 'me', or even *liburua* 'the book' would not need to be present for the grammaticality of this sentence, this sentence (and language) exhibits cross-index behavior because, even though these conominals are present, the argument indexing affixes on, in this case, the auxiliary *didazu* '3SG.ABS:AUX:-1SG.DAT:2SG.ERG' still occur.

Unlike the pro-indexes, there is no current standard HPSG analysis of cross-indexes.[11] Nevertheless, there are some possible approaches. I detail two in some depth here – what I will call the underspecification analysis and what I refer to as the direct syntax approach – and mention some other options near the end of the section.

2.2.1 Underspecification for Cross-Indexes

On the underspecification analysis, the lexical descriptions of argument index-containing words would have underspecified feature descriptions on their ARG-ST lists, corresponding to their argument indexes. These would then resolve depending on the syntactic context. Which portions of the feature description would be underspecified is a bit flexible (at least, in the abstract) and depends on whether the analyst thought the "agreement" (argument indexing) was more formal or semantic in nature (see Wechsler 2019, Chapter 6 of this volume for a more thorough discussion of what is involved here). For the sake of illustration purposes, I will employ a more semantic approach below.

Let us consider a word, like the Basque auxiliary *dut* 'AUX:3.ABS:1SG.ERG', that has a third-person singular absolutive argument index. Such a word might just be specified, by the constraints on the various lexical types of Basque, as in (5):[12]

[11] This may, in part, be a consequence of the standard way of managing predicate-argument relations in the syntax in HPSG: this management strategy is resource-sensitive – namely, once something is "cancelled" off a SUBJ or COMPS list, it no longer appears on any subsequent (higher) lists and cannot be used for other syntactic purposes. However, Section 2.2.3 will discuss some HPSG approaches where the management strategy is not so resource-sensitive.

[12] The lexical descriptions associated with *dut* in (5)–(7) all have a further argument – the verbal expression associated with the auxiliary – suppressed in these descriptions (with ellipses) because such a verbal argument (and its interaction with the other arguments) is not the focus of the analysis here.

(5) $$\left[\textsc{arg-st}\ \left\langle [\textit{synsem}], \begin{bmatrix}\textit{synsem}\\ \textsc{ss}|\textsc{loc}|\textsc{cont}|\textsc{ind}\ \ \textit{3sg}\end{bmatrix}, \ldots\right\rangle\right]$$

To be consistent with (5), the second ARG-ST list member just needs to be something that is semantically a third person. Therefore, the second ARG-ST list member could ultimately resolve to a *non-canonical* feature description, as in (6):

(6) $$\left[\textsc{arg-st}\ \left\langle [\textit{synsem}], \begin{bmatrix}\textit{non-canonical}\\ \textsc{ss}|\textsc{loc}|\textsc{cont}|\textsc{ind}\ \ \textit{3sg}\end{bmatrix}, \ldots\right\rangle\right]$$

This resolution would be forced when no conominal is present (if this *synsem* on the ARG-ST list resolved to the *canonical* type and a conominal was not present, the COMPS list would illicitly not be emptied). The analysis would, in this condition, be identical to that of the pro-indexes provided in Section 2.1.

However, the second ARG-ST list member could also ultimately resolve to a *canonical* feature description, as in (7):

(7) $$\left[\textsc{arg-st}\ \left\langle [\textit{synsem}], \begin{bmatrix}\textit{canonical}\\ \textsc{ss}|\textsc{loc}|\textsc{cont}|\textsc{ind}\ \ \textit{3sg}\end{bmatrix}, \ldots\right\rangle\right]$$

This resolution would be forced when a conominal is present (otherwise, the conominal could not be syntactically licensed). Thus, the analysis, in this condition, is like an instance of obligatorily co-present conominal and argument index (a gramm-index in Haspelmath's terms).

As the discussion above indicates, there is a certain portion of this analysis that is not lexically mandated: the precise resolution of the argument depends on the specific syntactic expressions appearing in a particular clause. This analysis is also of the dual-nature type discussed by Haspelmath (2013): the argument index is treated as a pro-index when it has no conominal and it is treated as gramm-index when a conominal is present. Other frameworks employ a similar analysis (LFG does, for instance – see Bresnan et al. 2015: Chapter 8). Haspelmath criticizes this approach for positing two distinct structural types for a single kind of affix; though, in the analysis above, it does not seem that the structural types are that radically different (observe that just one underspecified lexical description is associated with forms with affix). Still, we might want to at least consider other options – and, in keeping with the tendency for multiple different approaches to be found within HPSG, there are some.

2.2.2 The "direct syntax" approach to cross-indexes

Another approach to cross-indexes, proposed for Oneida in Koenig & Michelson (2015), takes the view that, in at least some languages, argument indexes always stand for arguments and the combination of a conominal and a verb with argument indexing is more purely semantically mediated and akin to a nominal expression combining with an already saturated narrow clause.[13] This approach Koenig & Michelson have called the "direct syntax approach" (it is direct in the sense that the combinatorics are not mediated by any valence lists, which, arguably, is a bit more "indirect").

As Koenig & Michelson (2015) discuss in detail, it appears that Oneida exhibits some interesting properties that make treating its argument indexing patterns in a (seemingly) rather different way much more plausible. For one, as shown in (8), the verb indexes all its arguments morphologically (except inanimates, like 'his axe' in (9)) – often with portmanteau affixes, as in (8) – making the case that the argument indexes are the actual arguments much stronger.

(8) Oneida [one] (Koenig & Michelson 2015: 5)
wa-hiy-até·kw-a-ht-eʔ
FACT-1SG>3.M.SG-flee-LNK.V-CAUS-PNC
'I chased him away.'

Second, the evidence is equivocal about whether the language has any selection that cannot be treated as semantic selection.

Thus, on Koenig & Michelson's view (and in keeping with the terminology of the previous discussion): all the arguments correspond (at best) to *non-canonical* elements on the ARG-ST list and thus there is never any head-argument combinations in the syntax. Any and all conominals then are licensed via index sharing of a nominal and an element on a NON-LOCAL feature that Koenig & Michelson call DISLOC (see Koenig & Michelson 2015: 39 for discussion of why they consider this the best way to deal with the NON-LOCAL feature), as shown in Figure 2, a tree of (9):

(9) Oneida [one] (Koenig & Michelson 2015: 17)
ʌ-ha-hyoʔthi·yát-eʔ laoto·kʌ́·,
FUT-3M.SG.A-sharpen-PNC his.axe
'he will sharpen his axe,'

Koenig & Michelson's (2015) discussion suggests that the direct syntax type might represent an extreme, occurring only in the most polysynthetic and nonconfigurational of languages, like Oneida and its Iroquoian kin. However, this

[13] This analysis is perhaps the closest any HPSG analysis comes to the so-called Pronominal Argument Hypothesis (Jelinek 1984).

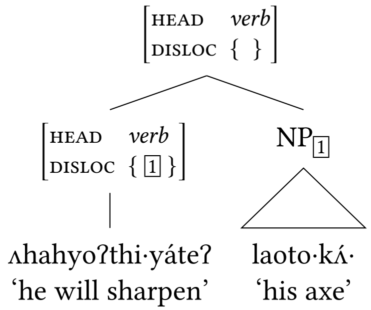


Figure 2: Licensing conominals on the direct syntax approach

claim remains an open question. Perhaps further study will reveal that this sort of analysis could profitably employed in other kinds of languages.

2.2.3 Other possibilities for cross-indexes

In addition to the analyses discussed in the previous two subsubsections, there are a few more conceptual avenues that might be explored for the analysis of cross-indexes, though it is not clear that they have been fully explored in the literature yet (which might raise some questions as to their viability).

One route to explore would be to use lexical rules to create ARG-ST or the valence lists with feature descriptions corresponding to both the argument indexes *and* the conominals (something similar was explored for "clitics" in various Romance varieties by Monachesi 2005). Such a lexical rule might look as in (10):

(10)
$$\begin{bmatrix}\text{ARG-ST} & \boxed{A}\left\langle [\textit{canonical}], \begin{bmatrix}\textit{non-canonical} & \\ \text{SS|LOC|CONT|IND} & \boxed{1}\end{bmatrix}\right\rangle\end{bmatrix} \mapsto \begin{bmatrix}\text{SS|LOC|CAT|COMPS} & \textit{list} \oplus \left\langle\begin{bmatrix}\textit{canonical} & \\ \text{SS|LOC|CONT|IND} & \boxed{1}\end{bmatrix}\right\rangle \\ \text{ARG-ST} & \boxed{A}\end{bmatrix}$$

This approach might be a way to loosen the resource sensitivity of the usual valence regime, though a proposal along these lines would need to take care in considering whether any changes would be needed in the statement of the Argument Realization Principle (and if so, what form they should take) and if there would be any undesirable consequences to allowing single semantic arguments to correspond to more than one syntactico-semantic element.

Another route to consider would be to relax the resource sensitivity in the syntax, instead of in the information associated with single words. Given proposals like the one in Bender (2008) for non-cancellation of arguments (more detailed discussion of this proposal is in Section 6 of Müller 2019, Chapter 10 of this volume[14]) to deal with apparent cases of discontinuous constituency, maybe something similar should also be explored for the cross-index type of argument indexing.

Overall, there seems to be a need for more explorations into cross-index behavior cross-linguistically within HPSG. Certainly, the above discussion shows that, in fact, there is no shortage of possible analyses, but work remains to further determine which of these might be the best analysis, overall, or which of these might be best for which languages.

3 Non-accusative alignments

Another area (in fact, not so distant from argument indexing in function) where understudied languages have enriched the general understanding of natural language morphosyntax is in the area of (morphosyntactic) alignment. Alignment concerns how the morphology of a language (if not also its syntax) groups together (or "aligns") different arguments together into (what seem to be) particular grammatical relations (see Bickel & Nichols 2009 for an overview of alignment).[15] The most widespread alignment is the accusative one – familiar from ancient Indo-European languages and conservative modern day ones – where subjects of transitive and intransitive verbs are treated differently from the objects of transitive verbs. Other recognized alignments include ergative (where subjects of transitive verbs are treated differently from the subjects of intransitives, which, in turn, pattern with the direct objects of transitives) (see Comrie 1978; Plank 1979; Dixon 1979; 1994, among others, for further discussion), split-S/active (where semantic agents and patients are treated differently) (see Klimov 1973; 1974; Dixon 1994: Chapter 4; Mithun 1991; Wichmann & Donohue 2008, among others, for further discussion), tripartite (where subjects of transitive verbs, subjects of intransitive verbs, and objects of transitive verbs are each treated differently) (Dixon 1994: 39–40), Austronesian alignment[16] (where arguments of various semantic roles can flexibly hold a privileged syntactic slot) (see

[14] Also see like-minded proposals in Meurers (1999) and Müller (2008).

[15] Alignment can be explored both in head-marking and dependent-marking (Nichols 1986); however, having already focused on a kind of head-marking strategy in the previous section, I will focus on the corresponding dependent-marking strategy in this section.

[16] This kind of system is known by various different names other than Austronesian alignment, including a symmetrical voice system, a Philippine-type voice system, or an Austronesian focus system.

Schachter 1976; Ross 2002; Himmelmann 2005 for more discussion), and hierarchical alignment (where elements of higher discourse salience are treated differently from elements of lower discourse salience) (see Jacques & Antonov 2014 for a good overview of what is involved).

Surveys from WALS (Comrie 2013a,b; Siewierska 2013a) indicate that accusative alignment is common worldwide,[17] and this seems to be even more true of languages with large numbers of speakers. Of the top 25 most widely spoken languages at present, arguably only a collection of languages from the Indian subcontinent (Hindi-Urdu, Marathi, and Gujarati) have non-accusative alignments, and even those are restricted to certain portions of their respective verbal systems (see Verbeke 2013: Chapter 7 for more on the patterns in these and other Indo-Aryan languages). Impressionistically, it seems that understudied languages do have a much stronger propensity for non-accusative alignments.

Because the non-accusative alignments at least seem to be rather different than accusative alignment, it is an interesting question how a given framework might handle these kinds of systems. In the majority of this section I will focus on the analysis of ergative systems, as a proof of concept (see, however, Drellishak 2009 for analyses of each of the non-accusative alignments, including the hierarchical type).[18]

In dealing with the analysis of ergative systems, it will be useful to divide the discussion into two parts. First, I will consider how particular morphological forms within NPs are licensed in instances when they co-occur with their governing verb – I will call this "the licensing of case in the syntax" (see also Przepiórkowski 2019, Chapter 7 of this volume). Second, I will considered how particular arguments come to be associated with particular morphological forms (whether realized or not) – I will call this "the licensing of case in linking" (see also Wechsler, Koenig & Davis 2019, Chapter 9 of this volume). This division is not commonly recognized in most other frameworks; however, it does present itself as a possible division within HPSG, due to the separate ARG-ST and valence lists.

[17] However, since the surveys focus more on coding patterns rather than behavioral patterns (coding and behavioral in the sense of Keenan 1976 – "coding" related to morphological patterns or function words that signal a particular grammatical relation category; "behavioral" related to reference properties or patterning across clauses), it is possible that they underreport behavioral accusative patterns, even among languages that have so-called "neutral" coding patterns.

[18] There is also some discussion of an HPSG analysis of the ergative-aligned case system of the Caucasian language Archi – similar, in some respects, to the Lezgian examples I consider further on – in Borsley (2016), though the focus of that paper is much more on Archi's argument indexing system rather than its case system.

3.1 The licensing of case in the syntax

The licensing of case in the syntax within HPSG is as straightforward in non-accusative alignments as it is in an accusative alignment system; the fundamentals are the same, regardless of alignment. This comes about due to the use of feature value matching (also known as "feature unification") for case licensing in the syntax.[19] The simple premise of feature value matching is that a value for a particular feature possessed by an argument and a feature value required by its head (for that same argument) must match. The nature of this analysis makes case licensing nearly identical – excepting the different values involved – to the selection of part-of-speech categories.[20]

To actually license case in the syntax with an ergative system, the key elements are (1) a feature for nominal expressions (call it CASE) and (2) appropriate values for CASE, like *ergative* and *absolutive*. Note that the values of *ergative* and *absolutive* are typed, so they can be potentially grouped with other case values into supertypes, like *structural* cases or *semantic* cases, if such groupings are relevant (as was done for the first time in Heinz & Matiasek 1994). With those features in place, the rest of the analysis falls out through the larger theories of syntactic selection, featural identities, and syntactic combination: certain heads will require [CASE *erg(ative)*] and [CASE *abs(olutive)*] of their arguments. If certain potential arguments are just single words, the values for CASE of these words will straightforwardly match or not. If certain potential arguments consist of multiple words, independent constraints on HEAD value identity will ensure that the value for CASE will be identical between the head daughter and overall phrase; constraints on the syntactic combination then ensure that the CASE values of the nominal expressions and the head requirements match.

To see this with an actual example, let us consider the Lezgian sentence in (11):

(11) Lezgian [lez] (Haspelmath 1993: 287)
Aburu zun ajibda.
3PL.ERG 1SG.ABS shame.FUT
'They will shame me.'

The example in (11) could be analyzed with the tree in Figure 3.

[19] Feature value matching does have some conceptual similarity to the "feature checking" approach to case found in more recent Minimalist work (Chomsky 1991; 1993; Adger 2000; 2010; Frampton & Gutmann 2006; Pesetsky & Torrego 2007), though there are notable differences between the approaches, particularly that features are deleted in feature checking, but not in feature value matching.

[20] Thus, to use terms more commonly associated with Mainstream Generative Grammar, HPSG views case licensing as (a specific kind of) c-selection (in the sense of Grimshaw 1979).

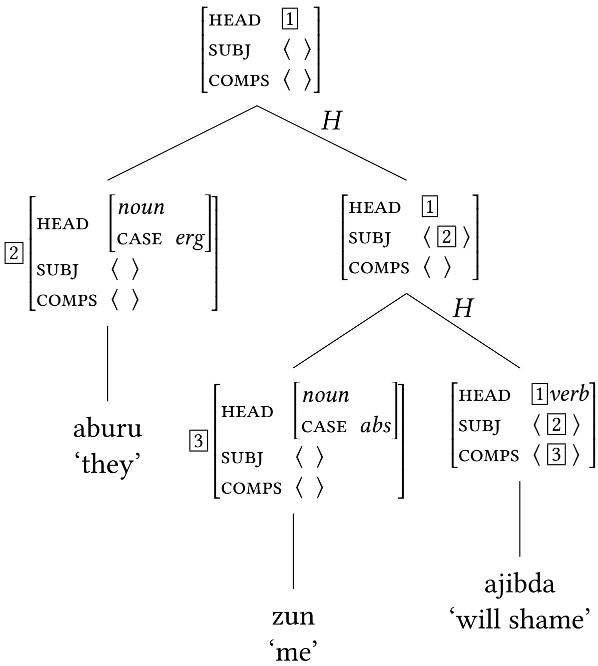


Figure 3: A tree of (11)

The tree in Figure 3 consists of two head-argument combinations, and in fact, the tree has the same geometry as an accusative verb-final language (on standard assumptions about the constituency) – indeed, as the HPSG analysis does not intrinsically tie the analysis of case with constituency, the geometry of clauses would, all else being equal, not differ based on alignment alone. The most idiosyncratic aspect of Figure 3 is that the verb *ajibda* 'will shame' is one that requires an ergative–absolutive combination of arguments. Because the HPSG framework is feature-rich and formally rigorous in how feature values must be constrained within constituent structures, the licensing of case in the syntax in an HPSG analysis is very straightforward.

3.2 The licensing of case in linking

Lurking behind the most idiosyncratic aspect of Figure 3 is the question of how particular heads come to have their particular argument requirements. This is, in fact, the question of how case is licensed in linking. As with other matters of non-accusative alignments, it seems that different alignments need not be treated in a wholly different fashion from each other: thus, the same kinds of analytic moves for accusatively aligned systems could be used for non-accusatively aligned systems. That being so, it is probably too hasty to assume that there is a one-size-fits-all solution for linking of case across all languages (regardless of alignment), as quite a few different factors appear to be important in different languages, among them at least verb class (that is, the classes related to the verbal lexical semantics), the semantic nature of the argument itself, the morphological form of the verb, and the subordination status of the clause headed by the verb (see, for example, discussion in Dixon 1994).

In all known ergative languages, the ergative–absolutive case pattern – clearly indicating that the subjects of transitive verbs are not encoded like the subjects of intransitive verbs – appears with "primary transitive verbs" (a term from Andrews 1985; 2007): predicates with the canonical meaning associated with transitive verbs where an initiating entity causes change in an undergoing entity. Given this basic generalization, a possible analysis of the arguments' case requirements with these "primary transitive verbs" would be through the constraint in (12):

(12)
$$\textit{trans-v-lxm} \Rightarrow \begin{bmatrix} \text{SS|LOC|CONT|KEY} & \begin{bmatrix} \textit{act-und-rel} \\ \text{ARG1}\ \boxed{1} \\ \text{ARG2}\ \boxed{2} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{CASE}\ \textit{erg} \end{bmatrix}_{\boxed{1}}, \begin{bmatrix} \text{CASE}\ \textit{abs} \end{bmatrix}_{\boxed{2}} \right\rangle \oplus \textit{list} \end{bmatrix}$$

In (12), the verb lexeme has an ARG-ST list with both an ergative and an absolutive argument. Key to this result is that the verb lexeme is associated with an *actor-undergoer-relation* (*act-und-rel*), and, in fact, this is the value of the KEY feature in (12), encoding the designated semantic relation relevant for case and linking (see Koenig & Davis 2006 for more on the KEY feature). The *act-und-rel* type designates semantic predicates with precisely the denotation behind the notion of "primary transitive verb" (see Davis 2001: 75–134 for discussion of this type, other related types, and how these types fit into a hierarchy of semantic relations). Provided that the constraint in (12) is the only argument realization constraint to mention the ergative and absolutive cases, the ergative–absolutive collection of arguments would only be available with verbs with this particular sort of meaning.[21] The overall linking constraint is placed on a *transitive-verb-lexeme* (*trans-v-lxm*), so that (a) the case requirements are stated once for all the different inflected forms of the verb and (b) so these case requirements could also

[21] Though to achieve more generality with the licensing of absolutive case, one might follow Ball (2008: Chapter 7) and have separate linking constraints for absolutive and ergative case.

be inherited by other semantically appropriate verbs with even more arguments than the bivalent predicate illustrated in (12).

As alluded to above, in other case "assignment" situations, though, other factors beyond just the semantics of the verb can be relevant (certainly with any instance of "split ergativity", among others). These could be still be treated with constraints with a similar format to (12), but if they needed to refer to, say, just past tense verb forms, the relevant linking constraint would almost assuredly need to reference information from the morphology (perhaps encoded as part of a MORPH(OLOGY) attribute). Given the known claims about what non-accusative alignment can be sensitive to, it seems likely that the sign-based architecture (where all linguistic areas of structure can interact in parallel) would enable the straightforward statement of case constraints based on the previously claimed generalization. And, in fact, having the possibilities of morphological form, semantics, and various syntactic properties easily available for an analysis could be useful as a means of testing and modeling which areas might be relevant in particular examples.

Overall, there is a lot still to be done to better understand the intricate details of case and linking generally, but given the toolbox available in the HPSG framework (again, see Przepiórkowski 2019, Chapter 7 of this volume and Wechsler, Koenig & Davis 2019, Chapter 9 of this volume), it seems like HPSG offers a lot of flexibility for better figuring out what linguistic elements are crucial for particular patterns and for encoding analyses that directly reference the interaction of these elements across different levels of structure.

4 Verb-subject-object constituent order

Let us turn to another interesting phenomenon of understudied languages: Verb-Subject-Object (VSO) constituent order.[22] VSO order appears to be the rarest of the more common orders. Various typology surveys (like Dryer 2013) indicate that it is only found in about 8–10% of the world's languages. Interestingly, a greater number of examples of languages with this order do come from the realm of understudied languages. Of the twelve understudied languages (a non-random sample[23]) mentioned in the introduction, five of them have VSO (or verb-initial

[22] It would probably be clearer to refer to this order as Predicate–Agentive–Patientive order, because, as noted in the previous section, alignment and constituent order are, to a degree, disjoint. However, I will bow to tradition and use the terms verb, subject, and object and their abbreviations, V, S, and O.

[23] See fn. 3 for the rationale behind the choice of those twelve languages.

with no strict ordering of S and O) order. Perhaps a bit more telling of the (apparent) understudied language bias to VSO order is that only one of the top 50 languages by native speakers – Tagalog – reasonably clearly has VSO order.[24] Interestingly, VSO order does occur in a number of languages as a non-dominant word order: for instance, it is found in a great many western European languages (English, German, French, Spanish, among others) as a common order in questions.

In spite of its relative rarity as a dominant order, VSO order (as well as verb-initial order with flexible ordering of verbal dependents) poses some interesting challenges for frameworks that place some importance in constituency (as HPSG has done). Since the V and the O are not (normally) adjacent in VSO clauses, it is less than obvious that there is a constituent that groups them together (as a VP or a V′) in these languages. This contrasts with the more common Subject–Verb–Object and Subject–Object–Verb orders where a constituent that groups the V and O together is much more plausible, on surface adjacencies alone. A long-standing question across constituency-based frameworks is how to best characterize VSO order, both on its own and in the context of the other cross-linguistically attested and common order patterns.

Analyses within Mainstream Generative Grammar have generally analyzed VSO as a derived order; all (or nearly all) of them (especially after the 1970s) have viewed VSO as a derived permutation of some constituent (or more) from a covert SVO order (see Clemens & Polinsky 2017 for an overview of the analyses within this tradition). Some of the suggested HPSG analyses follow a similar line of analysis, and I will briefly touch on those proposals below. However, more HPSG analysts have generally taken VSO order as is, and so I spend more of this section discussing two surface-oriented VSO analyses in HPSG: what I call the flat structure analysis and what I call the binary branching head-initial analysis.

4.1 The analogues of verb movement in HPSG

Interestingly, there is not one, but two styles of HPSG analyses that are roughly analogous to Mainstream Generative Grammar's verb movement, commonly employed to derive VSO order. Both are discussed in greater detail in Müller (2019), Chapter 10 of this volume, so my comments here will be somewhat superficial and will center around verb-initiality. The first of the two uses the DOUBLESLASH

[24] Tagalog's distant Austronesian relatives Indonesian, Javanese, and Sundanese, along with Arabic – all four of these languages are also in the top 50 – had VSO, historically, and each of these languages preserves some instances of VSO order. As is often the case when looking at word orders and languages, things are rarely as cut and dry as they might otherwise seem.

feature (see Section 4.1 of Müller 2019, Chapter 10 of this volume) and so treats the initial verb as involved in a long-distance dependency. Thus, in keeping with other long-distance dependencies in HPSG, information related to the initial verb is passed through the constituent structure to the verb's downstairs position (a trace, with semantic and syntactic structure, though no phonological realization). While this analysis has been explored for some Germanic languages (see Figure 2 in Müller 2019, Chapter 10 of this volume for a pictorial depiction of an analysis of an English verb-initial clause), I am not aware that it has (yet) been seriously explored in the HPSG literature for any particular verb-initial language (let alone for an understudied verb-initial language).

The other verb-movement-like analysis uses constituent order domains and linearization (see Section 5 of Müller 2019, Chapter 10 of this volume. On this analysis, the verb, while combined with its complements at a low level, is constrained at the clausal level to be initial (see Borsley 2006 for more discussion of this in a verb-initial context). This style of analysis has been closely and carefully considered for Welsh in work by Borsley (for example in Borsley 1989; 1995; 2009), but time and again, it seems that Borsley suggests that an analysis (at least for the basics of clausal structure) more in line with what is discussed in Section 3 is to be preferred for Welsh.

Given the rarity (and perhaps reluctance) – noted above – of HPSG researchers to analyze VSO order as covertly SVO (or, more to the point, to recognize a constituent that groups together the V and O within VSO structures), one might wonder why this has (hitherto) been so. Probably, HPSG's surface orientation has played a role, as well as the fact that HPSG-internal considerations do not force or strongly suggest positing a VP constituent. Furthermore, HPSG analysts have also carefully considered how constituency tests might inform such structures. In exploring these, various HPSG researchers (such as Borsley 2006 for Welsh and Ball 2008: Chapter 3 for Tongan) have not found compelling evidence for positing a VP constituent in particular VSO languages.[25] For instance, Ball, in looking at Tongan, found that: putative VP-coordination "over" a subject is not possible; no auxiliary or verb obviously subcategorizes for a verbal constituent that obviously excludes its subject, nor do adverbial elements obviously select for such a constituent; and, while "VP-fronting" and "VP-ellipsis" are possible, they seem to involve NPs rather than VPs. While these facts do not definitively

[25] The undermotivated VP in Welsh is probably just a VP headed by a finite verb, as Welsh does give evidence for non-finite VPs (Borsley et al. 2007). In other languages, like Tongan, the undermotivated VPs might include both finite and non-finite VPs. As has emerged in the study of verb-initial languages in several frameworks, these languages might not be as structurally uniform as the term "verb-initial languages" suggests.

rule out a VP (it is difficult to argue that anything is clearly absent), they suggest that not positing a VP does not complicate the grammar of this kind of language. Undoubtedly, it would be interesting to see what further explorations like these with more verb-initial language might reveal. Still, the VSO-as-covert-SVO analysis may lie on shakier empirically grounds than analyses within Mainstream Generative Grammar have generally acknowledged and this, explicitly or implicitly, has led HPSG analysts to explore other avenues in the analysis of VSO languages.

4.2 The flat structure analysis

The seemingly most common analysis of VSO languages in HPSG is the flat structure analysis.[26] As its name suggests, the proposed structure is flat, with the verb, subject NP, and any complement NPs all being sisters within the same constituent. To license such a structure, one has to depart from rules that put just heads and complements or just heads and subjects together. The flat structure analysis instead makes use of what I call the Head-All-Valents Schema (also sometimes called the Head-Subject-Complements Schema), given in (13):

(13)
$$\textit{head-all-valents-schema} \Rightarrow \left[\begin{array}{ll} \textsc{ss}|\textsc{loc}|\textsc{cat} & \left[\begin{array}{ll} \textsc{subj} & \langle\ \rangle \\ \textsc{comps} & \langle\ \rangle \end{array}\right] \\ \textsc{hd-dtr} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{word}} \\ \textsc{ss}|\textsc{loc}|\textsc{cat} & \left[\begin{array}{ll} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \boxed{2}, \ldots, \boxed{n} \rangle \end{array}\right] \end{array}\right] \\ \textsc{non-hd-dtrs} & \langle \boxed{1}, \boxed{2}, \ldots, \boxed{n} \rangle \end{array}\right]$$

Per its name, it licenses a fully saturated phrase comprising a head – a single word – and all its valents (subject, object, and whatever else). This schema has not just been used for canonical VSO clauses within HPSG, but other clause-level head-initial structures, including polar questions in English. Thus, this schema has a long pedigree in the HPSG literature (compare the schema in (13) with

[26] There are, in fact, several alternative flat structure analyses, differing slightly in how the head's valence features relate to the structure. Besides having the head combine with its subjects and complements simultaneously, as in the main text, one variant has all the arguments as complements and, thus, VSO order arises out of a head-complements structure. Borsley (1995) suggests that different languages might utilize different variants: in particular, Borsley suggests that Syrian Arabic uses the head-subject-complements combination while Welsh uses the head-complements combination. Still other analysts (such as Ball 2008; 2017) utilize a mixture of these two: having all arguments on just one valence feature, but combining all arguments with the head at once.

Schema 3 from Pollard & Sag 1994: 40; *sai-ph* from Ginzburg & Sag 2000: 36; and *aux-initial-cxt* from Sag 2012: 188).

To see an example using the Head-All-Valents Schema, let us consider example (14) from Kimaragang:

(14) Kimaragang [kqr] (Kroeger 2010: 7)
Minangalapak it kogiw do ratu.
PST.AV.TR.split NOM orangutan GEN durian
'The orangutan split (open) a durian.'

By the Head-All Valents Schema (and appropriate inherited constraints concerning the featural identities of HEAD values), a tree for (14) would be as in in Figure 4.

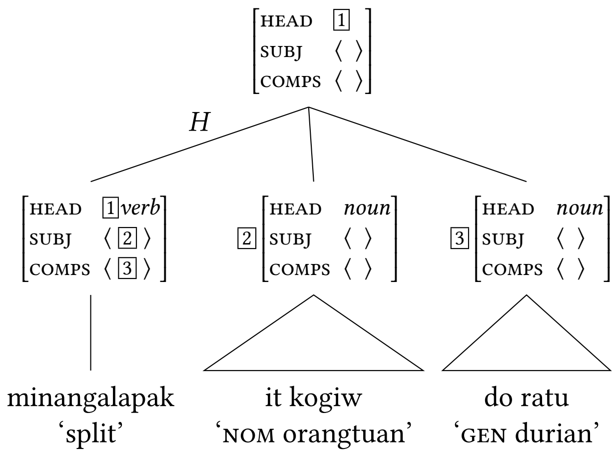


Figure 4: Tree of (14) in the flat structure analysis

To license the tree in Figure 4, we first should observe that the verb *minangalapak* 'split' appears to require both a nominative and a genitive argument. With two such nominal expressions fitting those requirements available, the Head-All-Valents Rule can put all three of these elements – the verb and two NPs – together, and the resulting mother node's SUBJ and COMPS lists would be empty.

In spite of the flatness of Figure 4, the structure is like all head-nexus combinations in HPSG: a head and (at least some of) its dependents. In fact, Figure 4 is identical to certain verb phrases headed by a ditransitive verb (on some HPSG analyses) – just a verb and two NPs. Furthermore, the flat nature of the structure is less of a concern than it would be under c-command-based proposals (which are the off-the-shelf analyses in Mainstream Generative Grammar): binding relations in HPSG are not calculated from the configurations within the tree, but from configurations on the ARG-ST list (see Branco 2019, Chapter 21 of this volume). Other subject-object and agent-patient asymmetries (to the extent they exist) are likewise encoded in HPSG analyses using non-configurational data structures and do not seem to be relevant for determining constituency.

Additionally, assuming a flatter structure for VSO/verb-initial languages eases the analysis of several other phenomena (especially versus a treatment of the same data with a VP constituent). In verb-initial languages where the order of elements following the verb is flexible (as in Tongan, among others), having all arguments together with the verb as part of a single constituent allows for such "scrambling" to be analyzed with simple linear precedence constraints within that constituent, rather than having to deal with different orders across a VP boundary (see the analysis in Ball 2008: Chapter 3 for Tongan and Müller 2019, Chapter 10 of this volume for other HPSG approaches to "scrambling"). There are also a few languages like Coast Tsimshian, where morphological marking (somewhat surprisingly) on one syntactic item refers to the next constituent over. An example of this phenomenon is given in the Coast Tsimshian sentence in (15), where the second line employs brackets to better show which elements are related to which others:

(15) Coast Tsimshian (Sm'algyax̱) [tsi] (Mulder 1994: 32)
Yagwat huumda duusa hoon.
Yagwa-t huum-[da duus]-[a hoon]
CONT-3.ERG smell-[ERG.CN cat]-[ABS.CN fish]
'The cat is sniffing the fish.'

It is far more straightforward to analyze the apparent sideways relationships when the interacting elements are sisters, rather than to manage the relationships across a VP boundary (and possibly other constituent boundaries) (see Ball 2011 for an in-depth look into this syntactic phenomenon in Coast Tsimshian and an analysis of it).

4.3 The binary branching head-initial analysis

Another approach in HPSG to VSO structures takes the view that all verb-headed structures within the clause are maximally binary branching, but strongly head-initial. This approach still has a strong surface-orientation – so it does not take the VSO order to be covertly SVO or SOV – but does posit that more structure is present within a clause than on the flat structure analysis.

On the binary branching head-initial analysis, VSO clauses are built out of several instances of a single rule. The rule, which I call the Head-Valent Schema, is given in (16):[27]

(16)
$$\textit{head-valent-schema} \Rightarrow \begin{bmatrix} \text{SS|LOC|CAT|VAL} & \boxed{\text{A}} \\ \text{HD-DTR} & \left[\text{SS|LOC|CAT|VAL}\ \langle \boxed{1} \rangle \oplus \boxed{\text{A}}\right] \\ \text{NON-HD-DTRS} & \langle \boxed{1} \rangle \end{bmatrix}$$

The rule in (16) allows a head to combine with just one of its valents; in particular, the first one on its VAL list. This aspect of the ordering is crucial to ensure that the subject NP before object NP sequence is licensed.

Returning to the Kimaragang example of (14), we can see how a structure licensed by the Head-Valent Schema in (16) differs from a structure licensed by the Head-All-Valents Schema. The structure licensed by the Head-Valent Schema (and relevant inherited constraints) is given in Figure 5.

Like in Figure 4, in Figure 5, the head verb requires a nominative and a genitive argument. However, instead of combining with both of these at the same time, the verb just combines with the initial nominative argument ($\boxed{2}$), leaving the genitive argument ($\boxed{3}$) to be passed up to the mother. At this second level of structure, the Head-Valent Schema again applies – because there is still at least one element on the relevant head's VAL list – integrating $\boxed{3}$ into the structure. The rule is barred, correctly, from applying to the root node of the tree in Figure 5, as this root node has an empty VAL list and the Head-Valent Schema requires the head daughter to have at least one valent.

A noteworthy feature of the VSO binary branching head-initial analysis is its grouping of verb and the subject NP into a constituent. Exactly this sort of thing has been reported to occur in some verb-initial languages, like Malagasy (Keenan 2000), suggesting the binary branching head-initial analysis might be preferable for such languages. In VSO languages without such evidence, it would seem that either the flat structure analysis or the binary branching head-initial analysis would be possible, all else being equal.

If one is accustomed to seeing the trees from Mainstream Generative Grammar, the structure in Figure 5 may still seem strange (notably, the structural

[27] The Head-Valent Schema here is designed to implement the Categorial Grammar analysis of Keenan (2000) in HPSG terms, and, as such, uses a single VALENCE list, abbreviated VAL. So, for the discussion in this section, I will employ this slightly different feature geometry. Note that the configuration of Figure 5 could also be achieved using the SUBJ and COMPS lists found elsewhere in this chapter (although it requires two rules instead of just one). As has been a recurring theme throughout this chapter, many analyses are possible and more empirical work is needed to see which might be preferred.

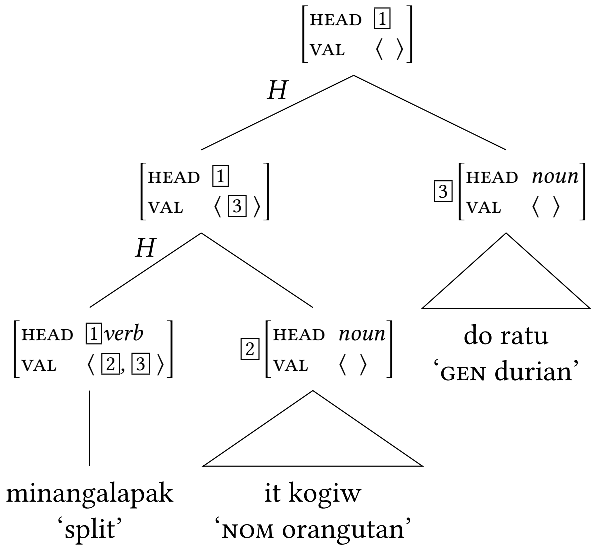


Figure 5: Tree of (14) in the binary branching head-initial analysis

prominence relationships between what seems to be the subject NP and what seems to be the object NP are reversed). Nevertheless, many of the same kinds of comments made for the flat structure analysis hold here as well. The structure in Figure 5 is a just a series of head-argument structures, the most common kind of structure in HPSG. And, once again, the non-configurational approach to binding in HPSG renders any issues related to tree configuration and binding as irrelevant.

Both approaches to VSO order discussed above do raise interesting questions about whether there are any underlying grammatical principles, processing preferences, or historically-driven outcomes behind the patterns. For the binary branching head-initial analysis, there is a question as to why the required order of combination goes from from least oblique to most oblique. A similar set of question can be leveled to the flat structure analysis: what inhibits a more constituent-rich structure? Why are flat structures licensed here and not elsewhere? To my knowledge, these questions have yet to be tackled within the HPSG literature, but they do seem to be reasonable next steps, in addition to better seeing which analyses are appropriate for which verb-initial languages.

5 Wrapping Up

In general, HPSG practitioners have been fairly conservative in what they assume to be universal in syntax: since there is no core assumption in HPSG that particular rich, innate, and universal class of structures help children learn any language (Mainstream Generative Grammar's Universal Grammar), proposals can be (and are) made that are agnostic as to universality. Even so, the brief trip made in this chapter through argument indexing, non-accusative alignments, and verb-initial constituent order found in understudied languages reveals that the more dependency-oriented portions of the framework – in particular, the areas encoded in the SUBJ, COMPS, and ARG-ST lists – are useful for the analysis of all three of these areas, across different languages, and, thus, are candidates for universality[28] (though the current level of understanding does not clearly point to them originating from either a rich language-specific part of cognition or from general cognition). Furthermore, the explorations above show that the rich and precise modeling using attribute-value matrices also allows for uniform sorts of analyses, even though the details may differ.[29] While the precise attributes and feature values may not completely be candidates for universality, they certainly aid in the enterprise of exploring different analyses and determining what precisely must be said to capture certain linguistic phenomena.

In addition to revealing some of the more uniform aspects of HPSG, the above discussion also reveals a certain flexibility in how the framework can be deployed – several analyses might be possible and certain ones might be more appropriate for certain languages and not for others. Thus, on top of a uniform foundation, various languages and phenomena are open to be analyzed in their own terms, dependent on what the specific empirical facts reveal. This mesh of uniformity and parochiality in HPSG analyses seems to strike a good balance as grammarians try to capture the two (somewhat paradoxical) realities one finds when comparing across languages: languages are both surprisingly similar and surprisingly different.

[28] Or in the case of the SUBJ and COMPS lists, a candidate for near-universality, as Koenig & Michelson (2015) argue that Oneida does not require such lists.

[29] Two projects within the HPSG community have explored in-depth how uniform particular HPSG analyses of different languages might be. The Grammar Matrix project (Bender et al. 2010) just starts from a common core and adds language-specific elements as needed; The CoreGram project (Müller 2015) actively tries to use the same sorts of data structures for as many languages as possible within the project.

Acknowledgements

Thanks to Stefan Müller, Robert Borsley, and Emily Bender for comments on this chapter. I alone am responsible for any remaining shortcomings. Thanks also to the editors, especially to Jean-Pierre Koenig (who served as the editors' in-person representative when I was first approached to write this chapter), for suggesting that a chapter like this one exist in this handbook.

References

Abbott, Miriam. 1991. Macushi. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3, 23–160. Berlin: Mouton de Gruyter.

Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng. 2006. The syntax of French *à* and *de*: An HPSG analysis. In Patrick Saint-Dizier (ed.), *Syntax and semantics of prepositions* (Text, Speech and Language Technology 29), 147–162. Springer Verlag.

Abeillé, Anne & Danièle Godard. 2000. French word order and lexical weight. In Robert D. Borsley (ed.), *The nature and function of syntactic categories*, vol. 32 (Syntax and Semantics), 325–360. San Diego: Academic Press.

Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.

Abeillé, Anne & Danièle Godard. 2004. The syntax of French adverbs without functional projections. In Martine Coene, Gretel de Cuyper & Yves d'Hulst (eds.), *Current studies in comparative Romance linguistics*, 1–39. Amsterdam: John Benjamins.

Adger, David. 2000. Feature checking under adjacency and VSO clause structure. In Robert D. Borsley (ed.), *The nature and function of syntactic categories*, 79–100. New York: Academic Press.

Adger, David. 2010. A Minimalist theory of feature structure. In Anna Kibort & Greville G. Corbett (eds.), *Features: perspectives on a key notion in linguistics* (Oxford Linguistics), 185–218. Oxford: Oxford University Press.

Andrews, Avery D. 1985. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description: clause structure*, 1st Edition, 62–154. Cambridge: Cambridge University Press.

Andrews, Avery D. 2007. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description: clause structure*, 2nd, 132–223. Cambridge: Cambridge University Press.

Ball, Douglas L. 2005a. Phrasal incorporation in Tongan. In Jeffrey Heinz & Dimitrios Ntelitheos (eds.), *Proceedings of the Twelfth Annual Conference of Austronesian Formal Linguistics Association (AFLA)*, vol. 12, 19–33. UCLA Working Papers in Linguistics. http://phonetics.linguistics.ucla.edu/wpl/issues/wpl12/papers/Ball_AFLA12.pdf.

Ball, Douglas L. 2005b. Tongan noun incorporation: lexical sharing or argument inheritance? In Stefan Müller (ed.), *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 7–27. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2005/ball.pdf.

Ball, Douglas L. 2008. *Clause structure and argument realization in Tongan*. Stanford University dissertation.

Ball, Douglas L. 2010. A unified analysis of Niuean *Aki*. In Raphael Mercado, Eric Potsdam & Lisa deMena Travis (eds.), *Austronesian and Theoretical Linguistics*, vol. 167 (Linguistik Aktuell/Linguistics Today), 249–270. Amsterdam/Philadelphia: John Benjamins.

Ball, Douglas L. 2011. Morphology in the 'wrong' place: the curious case of Coast Tsimshian connectives. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 25–45. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2011/ball.pdf.

Ball, Douglas L. 2017. ‘‘VP” adverbs without a VP: the syntax of adverbs in Tongan. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 61–81. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-ball.pdf.

Bender, Emily M. 2008. Radical non-configurationality without shuffle operators: An analysis of Wambaya. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 6–24. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/, accessed 2018-2-25.

Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson & Safiyyah Saleem. 2010. Grammar customization. *Research on Language & Computation* 8(1). 23–72. http://dx.doi.org/10.1007/s11168-010-9070-1. 10.1007/s11168-010-9070-1.

Bickel, Balthasar & Johanna Nichols. 2009. Case marking and alignment. In Andrej L. Malchukov & Andrew Spencer (eds.), *The Oxford handbook of case*, 304–321. New York: Oxford University Press.

Borsley, Bob & Anne Abeillé. 2019. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 3–5. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 1989. An HPSG approach to Welsh. *Journal of Linguistics* 25. 333–354.

Borsley, Robert D. 1995. On some similarities and differences between Welsh and Syrian Arabic. *Linguistics* 33(1). 99–122.

Borsley, Robert D. 2006. On the nature of Welsh VSO clauses. *Lingua* 116(4). 462–490.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language and Linguistic Theory* 27(2). 225–265.

Borsley, Robert D. 2016. HSPG and the nature of agreement in Archi. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dustan Brown (eds.), *Archi: complexities of agreement in cross-theoretical perspective*, 118–149. Oxford: Oxford University Press.

Borsley, Robert D., Maggie Tallerman & David Willis. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.

Branco, António. 2019. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 613–665. Berlin: Language Science Press. DOI:??

Branco, António & Francisco Costa. 2008. *A computational grammar for deep linguistic processing of Portuguese: LXGram, version A.4.1*. Tech. rep. TR-2008-17. Universidade de Lisboa, Faculdade de Ciências, Departamento de Informática.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Mark Wechsler. 2015. *Lexical-functional syntax*. 2nd edn. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9781119105664

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin (ed.), *Principles and parameters in Generative Grammar*, 417–454. Cambridge, MA: MIT Press. Reprint as: Chomsky (1995: 129–166).

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 1–52. Cambridge, MA/London: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Clemens, Lauren Eby & Maria Polinsky. 2017. Verb-initial word orders (primarily in Austronesian and Mayan languages). In Martin Everaert & Henk van

Riemsdijk (eds.), *The Wiley-Blackwell companion of syntax*, 2nd. Hoboken, NJ: Wiley-Blackwell.

Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: studies in the phenomenology of language*, 329–294. Austin: University of Texas Press.

Comrie, Bernard. 2013a. Alignment of case marking in full noun phrases. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info/chapter/98, accessed on 2018-08-01.

Comrie, Bernard. 2013b. Alignment of case marking of pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at http://wals.info/chapter/99, accessed on 2018-08-01.

Corbett, Greville G. 2003. Agreement: the range of the phenomenon and the principles of the Surrey Database of Agreement. *Transactions of the Philological Society* 101(2). 155–202.

Crowgey, Joshua & Emily M. Bender. 2011. Analyzing interacting phenomena: word order and negation in Basque. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 46–59. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2011/crowgey-bender.pdf.

Crysmann, Berthold. 2003. On the efficient implementation of German verb placement in HPSG. In *Proceedings of RANLP 2003*, 112–116. Borovets, Bulgaria.

Dąbkowski, Maksymilian. 2017. Yucatecan control and lexical categories in SBCG. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 162–178. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-dabkowski.pdf.

Davis, Anthony R. 2001. *Linking by types in the hierarchical lexicon*. Stanford, CA: CSLI Publications.

Dixon, R.M.W. 1979. Ergativity. *Language* 55(1). 59–138.

Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Donohue, Cathryn & Ivan A. Sag. 1999. Domains in Warlpiri. Manuscript, Stanford University.

Drellishak, Scott. 2009. *Widespread but not universal: Improving the typological coverage of the Grammar Matrix*. University of Washington Doctoral dissertation.

Dryer, Matthew S. 2013. Order of subject, object, and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info/chapter/81, accessed on 2018-08-01.

Dukes, Michael. 2001. The morphosyntax of 2P pronouns in Tongan. In Dan Flickinger & Andreas Kathol (eds.), *The Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, 63–80. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/1/hpsg00dukes.pdf.

Fokkens, Antske. 2014. *Enhancing empirical research for linguistically motivated precision grammars*. Saarland University dissertation.

Frampton, John & Sam Gutmann. 2006. How sentences grow in the mind: agreement and selection inefficient minimalist syntax. In Cedric Boeckx (ed.), *Agreement systems*, 121–157. Amsterdam: John Benjamins.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10(2). 279–326.

Hammarström, Harald, Robert Forkel & Martin Haspelmath. 2018. *Glottolog 3.3*. Jena. Available online at http://glottolog.org, Accessed on 2018-07-31. DOI:10.5281/zenodo.13

Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.

Haspelmath, Martin. 2013. Argument indexing: a conceptual framework for the syntax of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: studies in memory of Anna Siewierska*, 197–226. Berlin: Mouton de Gruyter.

Heinz, Wolfgang & Johannes Matiasek. 1994. Argument structure and case assignment in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 199–236. Stanford, CA: CSLI Publications.

Hellan, Lars & Petter Haugereid. 2003. Norsource – An excercise in the Matrix Grammar building design. In Emily M. Bender, Daniel P. Flickinger, Frederik Fouvry & Melanie Siegel (eds.), *Proceedings of the ESSLLI 2003 Workshop "Ideas and Strategies for Multilingual Grammar Development"*. Vienna, Austria.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian langauges of Asia and Madagascar*, 110–181. London: Routledge.

Jacques, Guillaume & Anton Antonov. 2014. Direct/inverse systems. *Language and Linguistics Compass* 8(7). 301–318.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2(1). 39–78.

Keenan, Edward L. 1976. Towards a universal definition of subject. In Charles N. Li (ed.), *Subject and topic*, 303–333. New York: Academic Press.

Keenan, Edward L. 2000. Morphology is structure: a Malagasy test case. In Ileana Paul, Vivianne Phillips & Lisa Travis (eds.), *Formal issues in Austronesian linguistics*, 27–47. Dordrecht: Kluwer.

Kim, Jong-Bok, Jaehyung Yang, Sanghoun Song & Francis Bond. 2011. Deep processing of Korean and the development of the Korean Resource Grammar. *Linguistic Research* 28(3). 635–672.

Klimov, Georgij A. 1973. *Očerk obščej teorii ergativnosti [Outline of a general theory of ergativity]*. Moscow: Nauka.

Klimov, Georgij A. 1974. On the character of languages of active typology. *Linguistics* 12(131). 11–26.

Koenig, Jean-Pierre & Anthony R. Davis. 2006. The key to lexical semantic representations. *Journal of Linguistics* 42(1). 71–108. DOI:10.1017/S0022226705003695

Koenig, Jean-Pierre & Karin Michelson. 2010. Argument structure of Oneida kinship terms. *International Journal of American Linguistics* 76(2). 169–205.

Koenig, Jean-Pierre & Karin Michelson. 2015. Invariance in argument realization: the case of Iroquoian. *Language* 91(1). 1–47.

Krauss, Michael. 1992. The world's languages in crisis. *Language* 68(1). 4–10.

Kroeger, Paul R. 2010. The grammar of *hitting*, *breaking*, and *cutting* in Kimaragang Dusun. *Oceanic Linguistics* 49(1). 1–20.

Kropp Dakubu, Mary Esther, Lars Hellan & Dorothee Beermann. 2007. Verb sequencing constraints in Ga: Serial verb constructions and the extended verb complex. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 99–119. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/, accessed 2018-2-25.

Laka, Itziar. 1996. *A brief grammar of Euskara, the Basque language*. Manuscript, University of the Basque Country. Available at http://www.ehu.es/grammar, accessed on 2018-08-01.

Lee, Juwon. 2012. The direct evidential *-te* in Korean: its interaction with person and experiencer predicate. In Stefan Müller (ed.), *The Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University, Daejeon*, 251–270. http://cslipublications.stanford.edu/HPSG/2012/lee.pdf.

Lee, Juwon. 2014. Two types of serial verb constructions in Korean: subject-sharing and index-sharing. In Stefan Müller (ed.), *The Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo*, 135–155. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2014/lee.pdf.

Malouf, Robert. 1999. West Greenlandic noun incorporation in a mono-hierarchical theory of grammar. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 47–62. Stanford, CA: CSLI Publications.

Manning, Christopher D. & Ivan A. Sag. 1998. Argument structure, valence, and binding. *Nordic Journal of Linguistics* 21. 107–144.

Marimon, Montserrat. 2013. The Spanish DELPH-IN grammar. *Language Resources and Evaluation* 47(2). 371–397. DOI:10.1007/s10579-012-9199-7

Meurers, Walt Detmar. 1999. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226. http://www.sfs.uni-tuebingen.de/~dm/papers/gagl99.html, accessed 2018-2-25.

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67(3). 510–546.

Monachesi, Paola. 2005. *The verbal complex in Romance: a case study in grammatical interfaces*. Oxford: Oxford University Press.

Muansuwan, Nuttanart. 2001. Directional serial verb constructions in Thai. In Dan Flickinger & Andreas Kathol (eds.), *The proceedings of the 7th international conference on head-driven phrase structure grammar*, 229–246. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/1/.

Muansuwan, Nuttanart. 2002. *Verb complexes in Thai*. University at Buffalo, the State University of New York dissertation.

Mulder, Jean Gail. 1994. *Ergativity in Coast Tsimshian (Sm’algy□x)*. Berkeley: University of California Press.

Müller, Stefan. 2008. Depictive secondary predicates in German and English. In Christoph Schroeder, Gerd Hentschel & Winfried Boeder (eds.), *Secondary predicates in Eastern European languages and beyond* (Studia Slavica Oldenburgensia 16), 255–273. Oldenburg: BIS-Verlag.

Müller, Stefan. 2009. A Head-Driven Phrase Structure Grammar for Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics: Papers from the 1st International Conference on Maltese Linguistics (Bremen/Germany, 18–20 October, 2007)*

(Studies in Language Companion Series 113), 83–112. Amsterdam: John Benjamins Publishing Co.

Müller, Stefan. 2013. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 3rd edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2015. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.91

Müller, Stefan. 2019. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

Müller, Stefan & Janna Lipenkova. 2009. Serial verb constructions in Chinese: An HPSG account. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 234–254. Stanford, CA: CSLI Publications.

Müller, Stefan & Janna Lipenkova. 2013. ChinGram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In Huei-ling Lai & Kawai Chui (eds.), *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, 240–249. Taipei, Taiwan: Department of English, National Chengchi University.

Müller, Stefan & Janna Lipenkova. 2016. *Mandarin Chinese in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In Preparation.

Müller, Stefan & Bjarne Ørsnes. 2015. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In Preparation.

Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62(1). 56–119.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: syntactic derivation and interpretation, in honor of Joseph E. Emonds*, 262–294. Amsterdam: John Benjamins.

Plank, Frans (ed.). 1979. *Ergativity: towards a theory of grammatical relations*. London: Academic Press.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Przepiórkowski, Adam. 1999. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*. Eberhard-Karls-Universität Tübingen dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46147, accessed 2018-2-25.

Przepiórkowski, Adam. 2019. Case. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 155–183. Berlin: Language Science Press. DOI:??

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of western Austronesian voice systems*, 17–62. Canberra: Pacific Linguistics.

Runner, Jeffrey T. & Raúl Aranovich. 2003. Noun incorporation and rule interaction in the lexicon. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University*, 359–379. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2003/runner-aranovich.pdf.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.

Saleem, Safiyyah. 2010. *Argument optionality: a new library for the grammar matrix customization system*. University of Washington MA thesis. http://www.delph-in.net/matrix/saleem-thesis.pdf.

Schachter, Paul. 1976. The subject in Philippine languages: topic, actor, actor-topic, or none of the above. In Charles N. Li (ed.), *Subject and topic*, 317–363. New York: Academic Press.

Siegel, Melanie, Emily M. Bender & Francis Bond. 2016. *Jacy: An implemented grammar of Japanese* (CSLI Studies in Computational Linguistics). Stanford, CA: CSLI Publications.

Siewierska, Anna. 1999. From anaphoric pronoun to grammatical agreement marker: why objects don't make it. *Folia Linguistica* 33(2). 225–251.

Siewierska, Anna. 2013a. Alignment of verbal person marking. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info/chapter/100, accessed on 2018-08-01.

Siewierska, Anna. 2013b. Verbal person marking. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info/chapter/102, accessed on 2018-08-01.

Simons, Gary F. & M. Paul Lewis. 2013. The world's languages in crisis: a 20-year update. In Elena Mihas, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds.), *Responses to language endangerment. in honor of Mickey Noonan*, vol. 142 (Studies in Language Companion Series), 3–19. Amsterdam: John Benjamins.

Verbeke, Saartje. 2013. *Alignment and ergativity in New Indo-Aryan languages.* Berlin: Walter de Gruyter.

Wechsler, Stephen. 2019. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 131–154. Berlin: Language Science Press. DOI:??

Wechsler, Stephen & I. Wayan Arka. 1998. Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory* 16(2). 387–442.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Wichmann, Søren & Mark Donohue (eds.). 2008. *The typology of semantic alignment.* Oxford: Oxford University Press.

Yang, Chunlei & Dan Flickinger. 2014. ManGO: Grammar engineering for deep linguistic processing. *New Technology of Library and Information Service* 30(3). 57–64.

Part II

Syntactic phenomena

Chapter 6

Agreement

Stephen Wechsler
The University of Texas at Austin

Agreement is modeled in HPSG by assigning agreement features such as person, number, and gender ("phi features") to specified positions in the feature structures representing the agreement trigger and target. The locality conditions on agreement follow from the normal operation of the grammar in which those phi features are embedded. In anaphoric agreement, phi features appear on referential indices; in verb agreement, phi features appear on the verb's ARG-ST list items; and in modifier agreement, phi features appear on the MOD value of the modifier. Selective underspecification of agreement features accounts for the alternation between formal and semantic agreement. Within the HPSG framework, long-distance agreement has been analyzed as anaphoric agreement in a special clausal construction, while superficial agreement has been modeled using linearization theory.

1 Introduction

Agreement is the systematic covariation between a semantic or formal property of one element (called the agreement *trigger*) and a formal property of another (called the agreement *target*). In the sentences *I am here* and *They are here*, the subjects (*I* and *they*, respectively) are the triggers; the target verb forms (*am* and *are*, respectively) covary with them. Research on agreement systems within HPSG has been devoted to describing and explaining a number of observed aspects of such systems. Regarding the grammatical relationship between the trigger and the target, we may first of all ask how local that relationship is, and in what grammatical terms it is defined. Having determined the prevailing locality conditions on agreement in a given language, we attempt to explain observed exceptions, that is, cases of apparent "long-distance agreement", as well as cases of superficial agreement defined on string adjacency. Agreement features across

Stephen Wechsler. 2019. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 131–154. Berlin: Language Science Press. DOI:??

languages include person, number, and gender (known as *phi* features), as well as deictic features and case, but various different subsets of those features are involved in particular agreement relations. How can we explain the distribution of features? How are locality and feature distribution related to the diachronic origin of agreement systems? Also, as indicated in the definition of agreement provided in the first sentence of this paper, the features of the target are sometimes determined by the trigger's form and sometimes by its meaning. What regulates this choice? In some cases a single trigger in a sentence determines different features on two different targets. Why does such "mixed agreement" exist, and what does its existence tell us about the grammatical representation of agreement? This chapter reviews HPSG approaches to these questions of locality, grammatical representation, feature distribution, diachrony, semantic versus formal agreement, and mixed agreement. Agreement with coordinate phrases is discussed by Abeillé & Chaves (2019), Chapter 17 of this volume.

HPSG offers an integrated account of these phenomena. In most cases the analysis of agreement phenomena does not involve any special formal devices dedicated for agreement, comparable to the *probe* and *goal*, or the AGREE relation, found in Minimalist accounts (Chomsky 2000). Instead, the observed agreement phenomena arise as a side effect of other grammatical mechanisms responsible for valence saturation, the semantics of modification, and coreference.

2 Agreement as unification

Constraint-based formalisms such as HPSG are uniquely well-suited for modeling agreement. Within such formalisms, agreement occurs when multiple feature sets arising from distinct elements of a sentence specify information about a single abstract object, so that the information must be mutually consistent (Kay 1984). The two forms are said to agree when the values imposed by the two constraints are compatible, while ungrammaticality results when they are incompatible. For example the English verb *is* in (1) specifies that its initial ARG-ST list item, which is identified with the SUBJ list item, has third person, singular features. In the mechanism of valence saturation, the NP list item in the value of SUBJ unifies with the feature description representing the SYNSEM value of the subject NP. The features specified by the verb for its subject and by the subject NP must be compatible; otherwise the representation for the resulting sentence is ill-formed, predicting ungrammaticality as in (3a).

(1) Simplified lexical sign for the verb *is*:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{is} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \boxed{2} \rangle \end{bmatrix} \\ \textsc{arg-st} & \left\langle \boxed{1}\text{NP}\begin{bmatrix} \textsc{pers} & \textit{3rd} \\ \textsc{num} & \textit{sg} \end{bmatrix}, \boxed{2}\text{XP} \right\rangle \end{bmatrix}$$

(2) Simplified lexical signs for *I* and *she*:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{I} \rangle \\ \textsc{head} & \begin{bmatrix} \textit{noun} & \\ \textsc{pers} & \textit{1st} \\ \textsc{num} & \textit{sg} \end{bmatrix} \end{bmatrix} \qquad \begin{bmatrix} \textsc{phon} & \langle \textit{she} \rangle \\ \textsc{head} & \begin{bmatrix} \textit{noun} & \\ \textsc{pers} & \textit{3rd} \\ \textsc{num} & \textit{sg} \\ \textsc{gen} & \textit{fem} \end{bmatrix} \end{bmatrix}$$

(3) a. * I is sober.

b. She is sober.

The features supplied by the trigger and target must be consistent, but there is no general minimum requirement on how many features they specify. Both of them can be, and typically are, underspecified for some agreement features. For example, gender is not specified by the verb in (1) or the first pronoun in (2).

Since unification is commutative, the representation of an agreement construction is the same regardless of whether a feature originates from the trigger or the target. This immediately accounts for common agreement behavior observed when triggers are underspecified (Barlow 1988). For example, Serbo-Croatian is a grammatical gender language, where common nouns are assigned to the masculine, feminine, or neuter gender. The noun *knjiga* 'book' in (4) is feminine, so the modifying determiner and adjective appear in feminine form (Wechsler & Zlatić 2003: 4, ex. (1)).

(4)

Ov-a	star-a	knjig-a	stalno	pad-a.
this-NOM.F.SG	old-NOM.F.SG	book-NOM.SG	always	fall-3SG

'This old book keeps falling.'

However, some nouns are unspecified for gender, such as *sudija* 'judge'. Interestingly, the gender of an agreeing adjective actually adds semantic information, indicating the sex of the judge (Wechsler & Zlatić 2003: 42, ex. (23)).

(5) a.

Taj	stari	sudija	je	dobro	sudio.
that.M	old.M	judge	AUX	well	judged.M

'That old (male) judge judged well.'

b. Ta stara sudija je dobro sudila.
that.F old.F judge AUX well judged.F
'That old (female) judge judged well.'

Here the gender feature comes from the targets instead of the trigger. This illustrates an advantage of constraint-based theories like HPSG over transformational accounts in which a feature is copied from the trigger, where it originates, to the target, where it is then realized. The usual source of the feature (the noun) lacks it in (5), a problem for the feature-copying view.

The same problem occurs even more dramatically in *pro*-drop languages. Many languages allow subject pronouns to drop, and distinguish person, number, and/or gender on the verb. If those features originate from the null subject, then there would have to be distinct null pronouns, one for each verbal and predicate adjective inflection (Pollard & Sag 1994: 64). This would be more complex and stipulative, and moreover the paradigm of putative null pronouns would have to exactly match the set of distinctions drawn in the verb and adjective systems, rather than reflecting the pronoun paradigm. HPSG avoids this suspicious assumption. Null anaphora is modeled by allowing the *pro*-dropped argument to appear on the ARG-ST list but not a VALENCE list (see Wechsler, Koenig & Davis 2019, Chapter 9 of this volume). For example, in the context given in (6) a Serbo-Croatian speaker could omit the subject pronoun.

(6) Context: Speaker comes home to find her bookcase mysteriously empty.
Gde su (one) nestale?
where did (they.F.PL) disappear.F.PL
'Where did they (i.e. the books) go?'

The sign for the inflected participle specifies feminine plural features on the initial item in its ARG-ST list. The SUBJ list item is optional:

(7) Simplified lexical sign for the participle form *nestale*:

$$\begin{bmatrix} \text{PHON} & \langle \textit{nestale} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \left\langle \left(\boxed{1} \right) \right\rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \\ \text{ARG-ST} & \left\langle \boxed{1}\text{NP}\begin{bmatrix} \text{NUM} & \textit{pl} \\ \text{GEN} & \textit{fem} \end{bmatrix} \right\rangle \end{bmatrix}$$

The feminine plural features are specified regardless of whether the subject pronoun appears. When the pronoun is dropped we have the usual underspecification, only in this case the trigger does not exist, so it is effectively fully underspecified, realizing no features at all.

3 Locality in agreement

3.1 Argument and modifier agreement

In HPSG, the grammatical agreement of a predicator with its subject or object, or an adjective, determiner, or other modifier with its head noun, piggy-backs on the mechanism of valence saturation and modification. Agreement is encoded in the grammar by adding features of person, number, gender, case, and deixis to the existing feature descriptions involved in syntactic and semantic composition. This simple assumption is sufficient to explain the broad patterning of distribution of agreement, in contrast to the transformational approach where complex locality conditions must be stipulated.

In HPSG, predicate-argument agreement arises directly from the valence saturation, as illustrated already in (1) above. Thus the locality conditions on the trigger-target relation follow from the conditions on the subject-head or complement-head relation. Similarly, attributive adjectives agree with nouns directly through the composition of the modifier with the head that it selects via the MOD feature. For example, the Serbo-Croatian feminine adjective form *stara* 'old.F' in (5b) specifies feminine singular features for the common noun phrase (N′) that it modifies.

(8) Simplified lexical sign for *stara*:

$$\begin{bmatrix} \text{PHON} & \langle \textit{stara} \rangle \\ \text{MOD} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{NUM} & \textit{sg} \\ \text{GEND} & \textit{fem} \end{bmatrix} \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$$

In head-adjunct phrases, the MOD value of the adjunct daughter is token-identical with the *synsem* value of the head daughter. So *stara*'s feminine singular features cannot conflict with the features of the noun it modifies.

The predicted locality conditions are also affected by the percolation of features from words to phrasal nodes, and this depends on the location of the features within the feature description. Agreement features of the *trigger* appear either within the HEAD value or the semantic CONTENT value (these give rise to CONCORD and INDEX agreement, respectively; see Section 4.2). In either case these features percolate from the trigger's head word to its maximal phrasal projection, due to the Head Feature Principle in the former case and the Semantics Principle in the latter. For example the noun phrase *the books* inherits its [NUM *pl*] feature from the head word *books*. This determines plural agreement on a verb: *These books are/*is interesting*. Apparent exceptions, where a target seems to fail to agree with the head of the trigger, are discussed below.

However, agreement features of the *target* appear in neither the HEAD nor the CONTENT value of the target form, but rather appear embedded in an ARG-ST list item or MOD features. So agreement features of the target do not project to the target's phrasal projection such as VP, S, or AP. This is a welcome consequence. If the subject agreement features of the verb projected to the VP, for example, we would expect to find VP-modifying adverbs that consistently agree with them, but we do not.[1]

4 Varieties of agreement target

4.1 Anaphoric agreement

In anaphoric agreement, an anaphoric pronoun agrees in person, number, and gender with its antecedent. Since Pollard & Sag (1992; 1994), anaphoric agreement has been analyzed in HPSG by assuming that person, number, and gender are formal features of the referential index associated with an NP. Anaphoric binding in HPSG is modeled as coindexation, i.e. sharing of the INDEX value, between the binder and bindee. Thus any specifications for agreement features of the INDEX contributed by the binder and bindee must be mutually consistent. In (9), Principle A of the Binding Theory requires the reflexive pronoun to be coindexed with an o-commanding item, here the subject pronoun:

(9) a. She admires herself.

b. *admire*:

$$\left[\textsc{arg-st}\ \left\langle \text{NP:}\begin{bmatrix} \textit{ppro} & \\ \textsc{index} & \boxed{1}\begin{bmatrix}\textsc{pers} & \textit{3rd}\\ \textsc{num} & \textit{sg}\\ \textsc{gen} & \textit{fem}\end{bmatrix}\end{bmatrix}, \text{NP:}\begin{bmatrix} \textit{ana} & \\ \textsc{index} & \boxed{1}\begin{bmatrix}\textsc{pers} & \textit{3rd}\\ \textsc{num} & \textit{sg}\\ \textsc{gen} & \textit{fem}\end{bmatrix}\end{bmatrix}\right\rangle\right]$$

The agreement features are formal features and not semantic ones, but the semantic correlates of person (speaker, addressee, other), number (cardinality), and

[1]VP-modifying secondary predicates sometimes agree with their own subjects. What we do not find are adjuncts that consistently agree with the subject agreement features of the VP even when the adjunct is not predicated of that subject.

gender (male, female, inanimate, etc.) are invoked under certain conditions (described in Section 5). Thus INDEX agreement is distinct from *pragmatic agreement* whereby semantic features of two coreferential expressions must be semantically consistent in order for them to refer to a single entity. INDEX agreement is enforced only within the syntactic domain defined by binding theory, while pragmatic agreement applies everywhere. For example, feminine pronouns are sometimes used for ships, in addition to neuter pronouns. Whichever gender is chosen, it must be consistent in binding contexts (example based on Pollard & Sag's 1994: 79 example (46a)):

(10) a. The ship lurched, and then it righted itself. She is a fine ship.
b. The ship lurched, and then she righted herself. It is a fine ship.
c. * The ship lurched, and then she righted itself.
d. * The ship lurched, and then it righted herself.

The bound reflexive must agree formally with its antecedent, while other coreferential pronouns need not agree, as they are not coarguments of the antecedent and not subject to the structural binding theory.

In grammatical gender languages, where common nouns are conventionally assigned to a gender, an anaphoric pronoun appearing outside the binding domain of its antecedent can generally agree with that antecedent either formally or, if it is semantically appropriate (such as an animate, sexed entity), it can alternatively agree pragmatically. In most situations pronouns allow either pragmatic or INDEX agreement with their antecedents. For example, pronouns coreferential with the Serbo-Croatian grammatically neuter diminutive noun *devojče* 'girl' can appear in either neuter or feminine gender (from Wechsler & Zlatić 2003: 198):

(11) Ovo malo devojče$_i$ je ušlo.
this.N.SG little.N.SG girl.N.SG AUX.3SG entered.N.SG

a. Ono$_i$ je htelo da telefonira.
it.N.SG AUX.SG wanted.N.SG that telephone

b. Ona$_i$ je htela da telefonira.
she.F.SG AUX.SG wanted.F.SG that telephone

'This little girl$_i$ came in. She$_i$ wanted to use the telephone.'

The neuter pronoun in (11a) reflects INDEX agreement with the antecedent while the feminine pronoun (11b) reflects its reference to a female (pragmatic agreement). But when a reflexive pronoun is locally bound by a nominative subject, agreement in formal INDEX features is preferred:

(12) Devojče je volelo samo/?*samu sebe.
girl.NOM.N.SG AUX3.SG liked.N.SG own.ACC.N.SG/ACC.F.SG self.ACC
'The girl liked herself.'

Again, this illustrates INDEX agreement in the domain defined by the structural binding theory.

4.2 Grammatical agreement: INDEX and CONCORD

As noted above, in HPSG agreement effectively piggy-backs on other independently justified grammatical processes. Anaphoric agreement is a side-effect of binding (Section 4.1) while grammatical agreement is a side-effect of valence saturation and modification (Section 3.1). The formal HPSG analysis of a particular agreement process mainly consists of positing agreement features somewhere in the feature structure; the observed properties follow from the location of those agreement features. With regard to the location of the features, grammatical agreement bifurcates into two types, INDEX and CONCORD.[2] (The attribute name CONCORD was introduced by Wechsler & Zlatić 2000: 799, Wechsler & Zlatić 2003: 14; precursors to the idea were treated as HEAD features in Pollard & Sag 1994, and called AGR by Kathol 1999.) The best way to understand this bifurcation of agreement, and indeed the operation of grammatical agreement systems generally, is by considering their diachronic origin. Although our primary goal is the description of synchronic grammar, a look at diachrony can help explain the forms that the grammar takes, and can also provide clues as to the best formalization of it.

Within the diachronic literature on agreement there are thought to be two different lexical sources for agreement inflections: (i) incorporated pronouns and (ii) incorporated noun classifiers (Greenberg 1978). These two sources, ultimately traced to pronouns and common nouns, give rise to INDEX and CONCORD target inflections, respectively, as explained next.

4.2.1 INDEX agreement

Taking pronouns first, many grammatical agreement systems evolve historically from the incorporation of pronominal arguments into the predicates selecting those arguments, such as verbs and nouns (Bopp 1842; Givón 1976; Wald 1979,

[2]The INDEX/CONCORD theory is sketched in Pollard & Sag (1994: Chapter 2) and Kathol (1999), and developed in detail in Wechsler & Zlatić (2000; 2003), all in the HPSG framework. It has since been adopted into LFG (King & Dalrymple 2004, inter alia) and GB/Minimalism (Danon 2009).

inter alia). When a phrase serving as antecedent of the incorporated pronoun is reanalyzed as the true subject or object of the predicate, the pronominal affix effectively becomes an agreement marker. With this reanalysis the only change in the affix is that it loses its ability to refer: it no longer functions as a pronoun. The affix retains its agreement features, and what was formerly anaphoric agreement with the topic becomes grammatical agreement with the subject or object. This explains why the features of grammatical agreement match those of pronominal anaphora: typically person, number, and gender, with occasional deictic features (Bresnan & Mchombo 1987: 752).

As explained above, structural anaphoric binding involves identifying (structure sharing) the referential indices of the pronoun and its binder. Therefore grammatical agreement derived from it is also INDEX agreement. For example, the signs for English *is* and *I* in (1) and (2) above should be rewritten as follows:

(13) Sign for *is*, illustrating INDEX agreement:

$$\begin{bmatrix} \text{PHON} & \langle \textit{is} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \left\langle \text{NP}\begin{bmatrix}\text{CONTENT}|\text{INDEX} & \begin{bmatrix}\text{PERS} & \textit{3rd}\\ \text{NUM} & \textit{sg}\end{bmatrix}\end{bmatrix}\right\rangle \\ \text{COMPS} & \langle \text{XP} \rangle \end{bmatrix}\end{bmatrix}$$

(14) Sign for *I*, illustrating INDEX features:

$$\begin{bmatrix} \text{PHON} & \langle \textit{I} \rangle \\ \text{CONTENT}|\text{INDEX} & \boxed{1}\begin{bmatrix}\text{PERS} & \textit{1st}\\ \text{NUM} & \textit{sg}\end{bmatrix} \\ \text{CONTEXT} & \textit{speaker}(\boxed{1}) \end{bmatrix}$$

This finite verb form specifies third person singular features of its subject's referential index.

One salient distinguishing characteristic of INDEX agreement is that it includes the PERSON feature. The only known diachronic source of the PERSON feature is from pronouns. Therefore, the other type of agreement, CONCORD, lacks the PERSON feature (as we will see below).

By modeling verb agreement in a way that reflects its historical origin, we are able to explain an array of facts concerning particular agreement systems. Some of these facts and explanations are presented in Section 6 below.

4.2.2 CONCORD

The agreement inflections on modifiers of nouns, such as adjectives and determiners, are thought to derive historically not from pronouns, but from noun classifiers (Greenberg 1978; Reid 1997; Seifart 2009; Grinevald & Seifart 2004, Corbett 2006: 268–269). The classifier morphemes in turn derive historically from lexical common nouns denoting superordinate categories like animal, woman, man, etc. For example Reid (1997) posits the following historical development of Ngan'gityemerri (southern Daly; southwest of Darwin, Australia), a language where the historical stages continue to cooccur in the current synchronic grammar. Originally the language had general-specific pairings of nouns as a common syntactic construction, such as *gagu wamanggal* 'animal wallaby' in (15a) (from Reid 1997: 216, examples (162)–(165)). The specific noun can be omitted when reference to it is established in discourse, leaving the general noun and modifier, to form NPs like *gagu kerre*, literally 'animal big' but functioning roughly like nominal ellipsis 'big one'. Then, where the specific noun is also included, both noun and modifier attract the generic term (15b). The gender markers then reduce phonologically and incorporate, producing modifier gender agreement (15c).

(15) a. Stage I:
Gagu wamanggal kerre ngeben-da.
animal wallaby big 1sg.sb.aux-shoot
'I shot a big wallaby.'

b. Stage II:
Gagu wamanggal gagu kerre ngeben-da.
animal wallaby animal big 1sg.sb.aux-shoot
'I shot a big wallaby.'

c. Stage III:
wa=ngurmumba wa=ngayi darany-fipal-nyine.
male=youth male=mine 3sg.aux-return-foc
'My initiand son has just returned.'

If the same affix is retained on the modifiers and the noun they modify, then the result is symmetrical agreement (also known as alliterative agreement), like the feminine *-a* endings in Spanish *zona rosa* (Corbett 2006: 87–88). But often an asymmetry between the affixes on the noun and the modifiers develops: the noun affix becomes obligatory and is subject to morphophonological processes that do not affect the modifier affix (Reid 1997: 216). This process may further progress to "prefix absorption" into the common noun, as evidenced by "gender prefixed nominal roots being interpreted as stems for further gender marking" (Reid 1997: 217).

Agreement marked with inflections from such nominal sources is called *concord*, which is described using the HPSG CONCORD feature. What is the proper HPSG formalization of this type of agreement, given its provenance? The last stages of the diachronic development, described in the previous paragraph, imply that the *form* of the trigger (the noun) is influenced by the agreement features. That is, noun declension classes tend to correlate with gender assignment (and more generally, phonological and morphological characteristics of nouns correlate with gender assignment); and number is marked on nouns as well. (This close relation between declension class and CONCORD is demonstrated in detail in Wechsler & Zlatić 2003: Chapter 2.) Thus the agreement features must appear both on the head noun (to inform its form and/or its gender selection and number value) and on the phrasal projection of that noun (to trigger agreement via the MOD feature of the agreement targets). Ergo CONCORD is a HEAD feature of the trigger.

Along with the number and gender features, the CONCORD value is assumed to include the case feature when case is a feature of NPs realized on both the head noun and its modifying adjectives or determiner. CONCORD lacks the person feature, since common nouns, from which the agreement inflections on the targets derive, lack the person feature (common nouns do not distinguish person values, since they are all in the third person). Meanwhile, INDEX agreement preserves the pronominal features of person, number, and gender, reflecting its origins. In the usual case the number and gender values found in CONCORD match those found in INDEX. The Serbo-Croatian noun form *knjiga* triggers feminine singular nominative CONCORD on its adjectival possessive specifier and modifier, and third person singular INDEX agreement on the finite auxiliary. (The status of the participle is discussed below.)

(16) Moja stara knjiga je pala.
my.F.NOM.SG old.F.NOM book.NOM.SG AUX.3.SG fall.PPRT.F.SG
'My old book fell.' (Wechsler & Zlatić 2003: 18)

The nominative singular noun form *knjiga* specifies its agreement features in both CONCORD (a HEAD feature) and INDEX, with the respective values for number and gender shared:

(17) Lexical sign for *knjiga* ‘book’ (from Wechsler & Zlatić 2003: 18):

$$\begin{bmatrix} \text{PHONOLOGY} & \langle \textit{knjiga} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{CATEGORY} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{CONCORD} & \boxed{3}\begin{bmatrix} \text{CASE} & \textit{nom} \\ \text{NUM} & \boxed{1}\,\textit{sing} \\ \text{GEN} & \boxed{2}\,\textit{fem} \end{bmatrix} \end{bmatrix} \\ \text{VALENCE|SPR} & \left\langle (\text{AP}[\textit{poss};\ \text{CONCORD}\ \boxed{3}]) \right\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{INDEX} & \boxed{i}\begin{bmatrix} \text{PERS} & \textit{3rd} \\ \text{NUM} & \boxed{1} \\ \text{GEN} & \boxed{2} \end{bmatrix} \\ \text{RESTR} & \{\textit{book}(i)\} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The specifier (SPR) is shown as AP because the possessive phrase is categorically an adjective phrase in Serbo-Croatian. The features in the overlap between CONCORD and INDEX are normally shared as in this example. But with some special nouns, features can be asymmetrically specified in only one of the two values (with no reentrancy linking them, of course). This leads to mismatches between CONCORD and INDEX targets, discussed in Section 6 below.

The phi features also appear within the HEAD value, as shown in (17), so that adjunct APs can agree with those features. For example, concord by the attributive adjective *stara* ‘old’ is guaranteed because its MOD feature is specified for feminine singular features, as shown in (8) in Section 3.1 above.

4.3 Conclusion

To summarize this section, we have seen the two main historical paths to agreement, and shown how HPSG formalizes these two types of agreement so as to capture the syntactic and semantic properties that follow directly from their origins. Agreement that descends from anaphoric agreement of pronouns with their antecedents, through the incorporation of personal pronouns into verbs and other predicators, inherits the INDEX matching process found in the anaphoric agreement from which it descends. Agreement that descends from the incorporation of noun classifiers involves features located in the HEAD value that connect a trigger noun form to its phrasal projection. The feature sets differ for the same reason; PERSON is a feature only of the first type, and CASE only of the second. CONCORD correlates strongly with declension class, while INDEX agreement need not correlate as strongly (for evidence see Wechsler & Zlatić 2003: Chapter 2). The differences in feature sets and morphology further correlate with systematic syntactic differences, described in the following section.

5 Syntactic, semantic, and default agreement

This chapter has so far focused mainly on formal agreement, as opposed to semantic agreement. But this is one of three different ways in which the form of an agreement target may be determined by a grammar:

(18) Formal, semantic, and default determinants of target form.

a. Formal agreement: The target form depends on the trigger's formal phi features.

b. Semantic 'agreement': The target form depends on the trigger's meaning.

c. Failure of agreement: The target fails to agree and hence takes its default form.

In formal agreement, the trigger is grammatically specified for certain features as a consequence of the words making up the trigger phrase: for example a nominal may be marked for a gender as a consequence of the lexical gender of the head noun. In semantic agreement, the target is sensitive to the meaning of the trigger instead of its formal features. English number agreement can be formal as in (19), from Wechsler (2013: 92), or semantic as in (20), from McCloskey (1991: 92):

(19) a. His clothes are/*is dirty.

b. His clothing is/*are dirty.

(20) a. That the position will be funded and that Mary will be hired now seems/??seem likely.

b. That the president will be reelected and that he will be impeached are/??is equally likely at this point.

Regarding (20), McCloskey (1991: 564–565) observes that singular is used for "a single complex state of affairs or situation-type", while plural is possible for "a plurality of distinct states of affairs or situation-types". The latter sort of interpretation is facilitated by the use of the adverb *equally*. Formal and semantic gender agreement are illustrated by the French examples in (21):

(21) a.

La	sentinelle	à la barbe	a	été	{ prise	/ *pris	} en otage.
the.F	sentry	bearded	AUX	been	taken.F.SG	taken.M	hostage

'The bearded sentry was taken hostage.'

b. Dupont est { compétent / compétente }.
Dupont is competent.M.SG competent.F.SG
'Dupont { a man / a woman } is competent.'

The grammatically feminine noun *sentinelle* 'sentry' triggers feminine agreement regardless of the sex of the sentry; but in (21b) feminine agreement indicates that Dupont is female while masculine agreement indicates that Dupont is male.

How does the grammar negotiate between formal and semantic agreement? In HPSG, syntactic and semantic representations are composed in tandem, making the framework well suited to address this question. It was addressed in early HPSG work, including Pollard & Sag (1994: Chapter 1). The specific approach due to Wechsler (2011) exploits the underspecification of agreement features (see Section 2). I posit the Agreement Marking Principle (AMP), which states that target agreement features are semantically interpreted whenever the trigger is underspecified for the formal grammatical features to which the target would normally be sensitive. The subject phrases in (19) are specified for number due to the formal features of the head nouns, but those in (20) are not, as a (coordinate) clause has no grammatical source for those features. Consequently, by the AMP, the verb's number feature is semantically interpreted in (20). Similarly, *sentinelle* in (21a) gives its formal feminine gender feature to the subject, while *Dupont* lacks a gender specification, triggering the semantic interpretation of the target adjectives in (21b): feminine is interpreted as 'female'.

Agreement targets generally have a default form for use when there is no trigger or the normal agreement relation is blocked for some reason. Blocking of agreement comes about in various situations; here we consider a case where the trigger is interpreted metonymically, apparently resulting in a reassignment of the referential index. Swedish predicate adjectives normally agree with their subjects in number (either singular or plural) and grammatical gender, either neuter (NT) or 'common' gender (COM), the gender held in common between masculine and feminine:

(22) a. Hus-et är gott.
house-DEF.N.SG is good.N.SG
'The house is good.'

b. Pannkaka-n är god.
pancake-DEF.COM.SG be.PRES good.COM.SG
'The pancake is good.'

c. { Hus-en / Pannkak-orna } är god-a.
house-PL.DEF pancake-PL.DEF be.PRES good-PL
'The houses / The pancakes are good.'

As shown in (22), a predicate adjective is inflected for number, and, in the singular, for gender, and agrees with its subject. But in sentences like (23), the adjective appears in the neuter singular form, regardless of the number and gender features of the subject. Note that *pannkakor* is the plural form of a common gender noun (Faarlund 1977; Enger 2004; Josefsson 2009):

(23) Pannkak-or är gott.
pancake-PL be.PRES good.N.SG
'Situations involving pancakes are good.' (e.g. 'Eating pancakes is good.')

In general, Swedish predicate adjectives appear in neuter singular when there is no triggering NP, such as with clausal subjects (see (25a) below). Wechsler & Zlatić (2003) posit the index type *unm* ('unmarked') for referential indices that lack phi features, such as those introduced by verbs. So *gott* has a SUBJ list item whose index is disjunctively specified for either neuter singular or type *unm*.

The lack of agreement in (23) then arises because the subject phrase refers, not to the pancakes, but to a situation involving them; hence its referential index is distinct from the one lexically introduced by the noun *pannkakor*. A rule shifts the index and encodes the metonymic relation between the entity and the situation involving it. This is implemented with a non-branching phrasal construction in Wechsler (2013: 82, ex. 20):

(24) *metonymy-ctx:*

$$\left[\begin{array}{ll} \text{MTR} & \left[\begin{array}{ll}\text{SYN} & NP\\ \text{SEM} & \left[\begin{array}{ll}\text{INDEX} & s_{unm}\\ \text{RESTR} & \left\{involve(s,i)\right\} \cup \boxed{1}\end{array}\right]\end{array}\right]\\ \text{DTRS} & \left[\begin{array}{ll}\text{SYN} & NP\\ \text{SEM} & \left[\begin{array}{ll}\text{INDEX} & i\\ \text{RESTR} & \boxed{1}\end{array}\right]\end{array}\right]\end{array}\right]$$

The noun *pannkakor* in (23) has an index marked with the features [PERSON *3rd*], [GENDER *com*], and [NUMBER *pl*], which, by the Semantics Principle, are therefore shared with the index of the daughter NP node in a structure licensed by rule (24). But the construction specifies the mother NP node's index is unmarked for those features, thus explaining the neuter singular adjective.

On the alternative ellipsis analysis, sentence (23) has an elliptical clausal or infinitival subject, with a structure like (25a) except that *att äta* is silent (Faarlund 1977; Enger 2004; Josefsson 2009):

(25) a. Att äta pannkakor är gott.
to eat pancakes be.PRES good.N.SG
'Eating pancakes is good.'

b. Det är gott att äta pannkakor.
it be.PRES good.N.SG to eat pancakes
'It is good to eat pancakes.'

c. * Det är gott pannkakor.
it be.PRES good.N.SG pancakes
Intended: 'It is good to eat pancakes.'

But the metonymic subject behaves in all respects like an NP, and unlike a clause or infinitival phrase. For example, unlike an infinitival it resists extraposition, as shown in (25b, c). The metonymy analysis captures the fact that the subject has a clause-like meaning but not clause-like syntax.

6 Mixed agreement

The two-feature (INDEX/CONCORD) theory of agreement was originally motivated by *mixed agreement*, where a single phrase triggers different features on distinct targets (Pollard & Sag 1994: Chapter 2; Kathol 1999). For example, the French second person plural pronoun *vous* refers to multiple addressees, and also has an honorific or polite use for a single (or multiple) addressee. When used to refer politely to one addressee, *vous* triggers singular on a predicate adjective but plural on the verb, as in (26a):

(26) a. Vous êtes loyal.
you.PL be.2PL loyal.M.SG
'You (singular, formal, male) are loyal.'

b. Vous êtes loyaux.
you.PL be.2PL loyal.PL
'You (plural) are loyal.'

Wechsler (2011) analyzes this by adopting the following suppositions: (i) *vous* has a second person plural marked referential index; (ii) *vous* lacks phi features for

CONCORD; (iii) finite verbs agree with their subjects in INDEX; and (iv) predicate adjectives agree with their subjects in CONCORD. Suppositions (i) and (iii) need not be stipulated, as they follow from the theory: the pronoun must have INDEX phi features since it shows anaphoric agreement (when it serves as binder or bindee); and the verb must agree in INDEX since it includes the PERSON feature. By the Agreement Marking Principle (see Section 5), the (CONCORD) number and gender features of the predicate adjective are interpreted semantically, which is what is shown by example (26).

"Polite plural pronouns" of this kind are found in many languages of the world (Head 1978). The cross-linguistic agreement patterns observed in typological studies (Comrie 1975; Wechsler 2011) confirm the predictions of the theory. Taken together, suppositions (i) and (iii) from the previous paragraph entail that any person agreement targets agreeing with polite pronouns should show formal, rather than semantic, agreement. Targets lacking person, meanwhile, can vary across languages. This pattern is confirmed for all languages with polite plurals that have been surveyed, including Romance languages; Modern Greek; Germanic (Icelandic); West, South and East Slavic; Hindi; Gbaya (Niger-Congo); Kobon and Usan (Papuan); and Sakha (Turkic) (see Comrie 1975 and Wechsler 2011).

The INDEX/CONCORD distinction plays a crucial role in this account of mixed agreement. An earlier hypothesis, proposed by Kathol (1999), is that French predicate adjectives are grammatically specified for semantic agreement with their subjects, while finite verbs show formal agreement. But a plurale tantum noun such as *ciseaux* 'scissors' triggers syntactic agreement on the predicate adjective:

(27) Ces ciseaux sont géniaux! (*génial!)
these.PL scissors(M.PL) are.PL brilliant.M.PL (*brilliant.M.SG)
'These scissors are cool!'

As far as the syntax is concerned, *ciseaux* 'scissors' is an ordinary common noun with masculine plural CONCORD features, so it triggers those features on the adjective. More generally, agreement target types cannot be split into "formal" and "semantic" agreement targets; both formal and semantic agreement are found across all target types. Which of the two is observed for a given agreement feature depends, according to the INDEX/CONCORD theory, on whether the trigger is specified for the grammatical feature, together with the INDEX versus CONCORD status of the target.

7 Agreement defined on other structures

So far our look at grammatical agreement has focused primarily on agreement defined on local grammatical relations like subject, object, and modifier. In this section we look at HPSG analyses of two other types of agreement, namely long-distance and superficial agreement.

7.1 Long-distance agreement

The simple picture of locality in the previous sections is challenged by the phenomenon of long-distance agreement, where the trigger appears within a clause subordinate to the one headed by the target verb. Long-distance agreement has been observed in a number of languages, including Tsez (Nakh-Dagestanian; Polinsky & Potsdam 2001), Hindi-Urdu (Bhatt 2005), and Passamaquoddy (Athabaskan; Bruening 2001; LeSourd 2018).

Passamaquoddy long-distance agreement is illustrated by this sentence (LeSourd 2018: ex. (5)), with the relevant elements indicated in bold:

(28)

N-kosicíy-a-**k**	[eli-	Píyel	-litahási-t
1-know-DIR-PROX.PL	thus-	Peter	-think-3AN

[eli-kis-ankum-í-hti-t	**nìkt**	**ehpíc-ik**
thus-PAST-sell-3/1-PROX.PL-3AN	those.PROX	woman-PROX.PL

posonúti-yil]]
basket-IN.PL

'I know that Peter thinks that those women sold me the baskets.'

The *-k* suffix on the matrix verb *kosicíy* 'know' marks plural, deictically proximate agreement with the phrase *nìkt ehpícik* 'those women' in the doubly embedded subordinate clause. LeSourd (2018) analyzes Passamaquoddy long distance agreement in the HPSG framework. He notes that Passamaquoddy long distance agreement is parallelled by long-distance raising, in which an NP in the matrix clause is coreferential with an implicit argument of a subordinate clause (LeSourd 2018: ex. (4)):

(29)

N-kosicíy-a-**k**	**nìkt**	**ehpíc-ik**$_i$	[eli-	Píyel
1-know-DIR-PROX.PL	those.PROX	woman-PROX.PL	thus-	Peter

-litahási-t	[eli-kis-ankum-í-hti-t	**e**$_i$	posonúti-yil]]
-think-3AN	thus-PAST-sell-3/1-PROX.PL-3AN		basket-IN.PL

'I know about those women$_i$ that Peter thinks that they$_i$ sold me the baskets.'

Passamaquoddy speakers report that sentences (28) and (29) suggest the subject of 'know' (the speaker) is familiar with the women. This provides evidence that the phrase 'those women' in (29) is an argument of the matrix verb 'know', as implied by the translation. Similarly, the matrix clause (28) contains a null argument (cross-referenced by the proximate plural *-k* suffix), which is cataphoric to 'those women'. Hence a more literal translation of (28) is 'I know about them$_i$ that Peter thinks that those women$_i$ sold me the baskets.'[3] What the long-distance agreement and raising constructions share is simply that the matrix object is coreferential with some argument contained in the subordinate clause. The following lexical entry for the verb root *kosicíy* 'know' captures that:

(30) *kosicíy* 'know':

$$\begin{bmatrix} \text{PHON} & \langle \textit{kosicy} \rangle \\ \text{ARG-ST} & \left\langle \text{NP}_i \text{ , NP}_j \text{ , S:}\left[\text{RESTR} \left\langle \ldots \left[\text{PRD}|\text{ARG} \ \ j\right] \ldots \right\rangle\right] \right\rangle \end{bmatrix}$$

LeSourd adopts the version of HPSG described in the Sag et al. (2003) textbook, which uses a simplified Minimal Recursion Semantics. The semantic restrictions feature (RESTR) takes as its value a list of elementary predications (PRD). The list for each node is a concatenation of the restrictions of the daughter nodes. Thus every semantic argument contained within the S complement, whether overt or null, will correspond to some argument (ARG) of an elementary predication (PRD) in S's RESTR list. The lexical entry in (30) stipulates that the matrix object NP corefers with some such argument. In conclusion, Passamaquoddy long-distance agreement is really the anaphoric agreement of a pronoun with an antecedent in a higher clause.

7.2 Superficial agreement

In some languages, string adjacency of the trigger and target, rather than a grammatical relation such as subject or modifier, is a grammatical condition on agreement. This may arise because person agreement derives historically from pronoun incorporation, and a basic syntactic precondition for incorporation is string adjacency between the pronoun and the head into which it incorporates (Givón 1976; Ariel 1999; Wechsler et al. 2010; Fuss 2005). If the trigger occupies the syntactic position that the pronoun occupied prior to incorporation (for example because the trigger is itself a pronoun) then the result is that trigger and target

[3] LeSourd notes that Passamaquoddy lacks Principle C effects, so cataphora of this kind is permitted.

are adjacent. For example, West Flemish complementizers agree with an immediately following subject, even though the complementizer and subject are not related by any grammatical relation (Haegeman 1992). To take another example, Borsley (2009) analyzes Welsh superficial agreement in the HPSG framework, citing examples like the following:

(31) a. Gwelon nhw ddraig.
see.PAST.3PL they dragon
'They saw a dragon.'

b. arno fo
on.3SG.M he
'on him'

c. Gweles i a Megan geffyl.
see.PAST.1SG I and Megan horse
'Megan and I saw a horse.'

The trigger is the subject in (31a), object in (31b), and the first conjunct of a coordinate subject in (31c). But in every case, "An agreeing element agrees with an immediately following noun phrase if and only if the latter is a pronoun" (Borsley 2009: ex. 48). Borsley (2009: ex. 99) expresses this as an HPSG implicational constraint using the DOMain feature from linearization theory (Reape 1994; Müller 1995; 1999; Kathol 2000):

(32) $[\text{DOM} \left\langle [\text{AGR } \boxed{1}], \text{NP: } \textit{ppro}_{\boxed{2}}, \ldots \right\rangle] \Rightarrow \boxed{1} = \boxed{2}$

The DOMain list encodes linear precedence between constituents that are not necessarily sisters. In (32) the AGR value is the set of phi features of the target; the colon following NP represents the semantic CONTENT attribute; and the subscripted tag $\boxed{2}$ is the INDEX value. The rule states that when a constituent bearing the AGR attribute is immediately followed by a personal pronoun (content of type *ppro*), then the AGR value is identified with the pronoun's index (shown here as $\boxed{2}$), that is, it agrees with a right-adjacent pronoun.

8 Conclusion

Agreement is analyzed in HPSG by assigning phi features to specific locations in the feature descriptions representing the grammar. Anaphoric agreement results from phi features appearing on the referential indices of the binder and bindee, together with the assumption that binding consists of the identification of those indices. Verbal agreement with subjects and objects results when phi features appear on the verb's ARG-ST list items that are identified with the SYNSEM values of the subject and object phrases. Modifier agreement with heads occurs when phi features appear within the MOD value of the modifier. According to the INDEX/CONCORD theory, when agreement is historically descended from anaphoric agreement of incorporated pronouns, then those features within the ARG-ST list or MOD items are located on the referential index; while otherwise they are collected in the CONCORD feature and placed within the value of the HEAD features. The locality conditions on agreement follow from the normal operation of the grammar in which those phi features are embedded. Some cases of agreement seem to exist outside those conditions. Long-distance agreement has been analyzed as a kind of anaphoric agreement within a prolepsis construction, and superficial agreement has been defined on string adjacency and precedence, within linearization theory.

References

Abeillé, Anne & Rui P. Chaves. 2019. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 469–516. Berlin: Language Science Press. DOI:??

Ariel, Mira. 1999. The development of person agreement markers: from pronouns to higher accessibility markers. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 197–260. Stanford: CSLI Publications.

Barlow, Michael. 1988. *A Situated Theory of Agreement*. Stanford University PhD thesis.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23(4). 757–807.

Bopp, Franz. 1842. *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Litthauischen, Gothischen und Deutschen [Comparative grammar of Sanskrit, Zend, Greek, Latin, Lithuanian, Old Slavic, Gothic, and German]*. Berlin: Ferdinand Dümmler.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language and Linguistic Theory* 27(2). 225–265.

Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63. 741–82.

Bruening, Benjamin. 2001. *Syntax at the edge: cross-clausal phenomena and the syntax of Passamaquoddy*. MIT dissertation.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, Juan Uriagereka & Samuel Jay Keyser (eds.), *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Comrie, Bernard. 1975. Polite Plurals and Predicate Agreement. *Language* 51(2). 406–418.

Corbett, Greville. 2006. *Agreement*. Cambridge: Cambridge University Press.

Danon, Gabi. 2009. Agreement with quantified nominals: Implications for feature theory. In *Colloque de Syntaxe et Sémantique à Paris*.

Enger, Hans-Olav. 2004. Scandinavian pancake sentences as semantic agreement. *Nordic Journal of Linguistics* 27(01). 5–34.

Faarlund, Jan Terje. 1977. Embedded Clause Reduction and Scandinavian Gender Agreement. *Journal of Linguistics* 13(2). 239–257.

Fuss, Eric. 2005. *The rise of agreement: a formal approach to the syntax and grammaticalization of verbal inflection*. Amsterdam: John Benjamins.

Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 149–88. New York: Academic Press.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Morovcsik (eds.), *Word structure*, vol. III (Universals of Human Language). Stanford: Stanford University Press.

Grinevald, Colette & Frank Seifart. 2004. Noun classes in African and Amazonian languages: Towards a comparison. *Linguistic Typology* 8(2). 243–285.

Haegeman, Liliane. 1992. *Theory and description in generative syntax: a case study in West Flemish*. Cambridge, England: Cambridge University Press.

Head, Brian F. 1978. Respect degrees in pronominal reference. In Joseph H. Greenberg (ed.), *Universals of Human Language: Vol. 3, Word Structure*, 151–212. Stanford, California: Stanford University Press.

Josefsson, Gunlög. 2009. Peas and Pancakes: On apparent disagreement and (null) light verbs in Swedish. *Nordic Journal of Linguistics* 32(01). 35–72.

Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In Robert D. Levine & Georgia M. Green (eds.), *Studies in contemporary Phrase Structure Grammar*, 223–274. Cambridge, UK: Cambridge University Press.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University Press.

Kay, Martin. 1984. *Functional unification grammar: a formalism for machine translation*. Association for Computational Linguistics. 75–78.

King, Tracy Halloway & Mary Dalrymple. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40. 69–104.

LeSourd, Philip S. 2018. Raising and long distance agreement in Passamaquoddy: a unified analysis. *Journal of Linguistics*.

McCloskey, James. 1991. There, it, and agreement. *Linguistic Inquiry* 22(3). 563–567.

Müller, Stefan. 1995. Scrambling in German – Extraction into the *Mittelfeld*. In Benjamin K. T'sou & Tom Bong Yeung Lai (eds.), *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*, 79–83. City University of Hong Kong.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19(19). 583–646.

Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Reape, Mike. 1994. Domain Union and Word Order Variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford: CSLI Publications.

Reid, Nicholas. 1997. Class and classifier in Ngan'gityemerri. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in aboriginal Australia*, 165–228. Amsterdam: John Benjamins.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Seifart, Frank. 2009. Multidimensional typology and Miraña class markers. In Patience Epps & Alexandre Arkhipov (eds.), *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, 365–385. Berlin: Walter de Gruyter.

Wald, Benji. 1979. The development of the Swahili object marker: A study of the interaction of syntax and discourse. In Talmy Givón (ed.), *Discourse and syntax*, vol. 12 (Syntax and Semantics), 505–24. New York: Academic Press.

Wechsler, Stephen. 2011. Mixed agreement, the person feature, and the Index/Concord distinction. *Natural Language and Linguistic Theory* 29(4). 999–1031.

Wechsler, Stephen. 2013. The structure of Swedish *pancakes*. In Elisabeth Norcliffe & Philip Hofmeister (eds.), *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, 71–98. Stanford: CSLI publications.

Wechsler, Stephen, Patience Epps & Elizabeth Coppock. 2010. The origin and distribution of agreement features: a typological investigation. Project description, National Science Foundation grant application.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Wechsler, Stephen & Larisa Zlatić. 2000. A theory of agreement and its application to Serbo-Croatian. *Language* 76. 799–832.

Wechsler, Stephen & Larisa Zlatić. 2003. *The many faces of agreement*. Stanford: CSLI Publications.

Chapter 7

Case

Adam Przepiórkowski
University of Warsaw and Polish Academy of Sciences

The aim of this chapter is to provide an outline of HPSG work on grammatical case. Two issues that attracted much attention of HPSG pracitioners in the 1990s and early 2000s are the locality of case assignment, especially the so-called structural case assignment, as well as case syncretism and underspecification; they are discussed in two separate sections. The third section summarises other work on case carried out within HPSG, including some computational efforts, as well as investigations of case phenomena at the syntax-semantics interface and at the border of syntax and morphology.

HPSG is not widely known for its approach to grammatical case. For example, it is only mentioned in passing in the 2006 monograph *Theories of Case* (Butt 2006: 225) and in the 2009 *Oxford Handbook of Case* (Malchukov & Spencer 2009: 43), which features separate articles on GB/Minimalism, Lexical Functional Grammar, Optimality Theory and other grammatical frameworks. As most of the HPSG work on case was carried out in the 1990s and early 2000s, this perception is unlikely to have changed since the publication of these two volumes.

The aim of this chapter is to provide an overview of HPSG work on grammatical case and to show that it does offer novel solutions to some of the problems related to case. Two main research areas are presented in the two ensuing sections: structural case assignment is discussed in Section 1 and case syncretism and underspecification in Section 2. Some of the other HPSG work on case, including implementational work, is outlined in Section 3.

Adam Przepiórkowski. 2019. Case. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 155–183. Berlin: Language Science Press. DOI:??

1 Structural case assignment

Pollard & Sag (1994) did not envisage a separate theory of case:[1] "Nominative case assignment takes place directly within the lexical entry of the finite verb", while "the subject SUBCAT element of a nonfinite verb […] does not have a case value specified" (p. 30). However, they added in a footnote on the same page that "for languages with more complex case systems, some sort of distinction analogous to the one characterized in GB work as 'inherent' vs. 'structural' is required."

In the transformational Government and Binding theory of the 1980s (GB; Chomsky 1981; 1986), "inherent" – or "lexical" – case is understood as rigidly assigned by the head and independent of syntactic environment, while "structural" case varies with the structural context (e.g., Haider 1985: 70). This difference can be illustrated on the basis of the following examples from German (Przepiórkowski 1999a: 63, based on data from Heinz & Matiasek 1994):

(1) a. Der Mann unterstützt *den Installateur.*
the man.NOM supports the plumber.ACC
'The man is supporting the plumber.'

b. *Der Installateur* wird unterstützt.
the plumber.NOM AUX supported
'The plumber is supported.'

c. das Unterstützen *des Installateurs*
the supporting the plumber.GEN
'the support for/from the plumber'

(2) a. Der Mann hilft *dem Installateur.*
the man.NOM helps the plumber.DAT
'The man is helping the plumber.'

b. *Dem Installateur* wird geholfen.
the plumber.DAT AUX helped
'The plumber is helped.'

c. das Helfen *des Installateurs*
the helping the plumber.GEN
'the help from/*for the plumber'

[1]This section is to some extent based on Przepiórkowski (1999a), Section 3.4 and Chapter 4; see also Müller (2013), Chapter 14.

Note that in what follows, the convention of marking lemmata with small capitals applies. In (1), both arguments of the verb UNTERSTÜTZEN 'support' receive structural case: the patient argument occurs in the accusative in (1a), in the nominative in (1b), and in the genitive in (1c). Similarly, the agent argument is in the nominative in (1a), but it may only occur in the genitive in (1c); hence, the single argument marked as genitive in (1c) is ambiguous between the agent and the patient. In the case of (2), the agent argument of HELFEN 'help' is similarly assigned structural case, but the patient argument receives a rigid inherent case: it is always the dative, so, e.g., the genitive in (2c) may only be understood as marking the agent.

Examples such as those above may still be handled without any general principles of case assignment. For example, lexical rules (Pollard & Sag 1987: 209–218) responsible for forming passive participles (as in the b. examples above) and gerunds (as in the c. examples) might be responsible for manipulating case values of arguments, e.g., for translating nominative and accusative – but not dative – to genitive in the case of gerunds. However, the interaction of the structural/inherent case dichotomy with raising (and – in some languages – with control) motivates a more comprehensive approach to case assignment.

Consider Icelandic raising verbs (all Icelandic data is taken from Sag et al. 1992: 304–305):

(3) a. *Hann* virðist elska hana.
he.NOM seems love.INF her.ACC
'He seems to love her.'

b. Þeir telja *Maríu* hafa skrifað ritgerðina.
they believe Mary.ACC have.INF written the.thesis
'They believe Mary to have written her thesis.'

As in other languages, the subject of the infinitival verb raised to the higher subject position, as in (3a), normally receives the nominative case there, while – in the case it is raised to the object position, as in (3b) – it normally receives the accusative case. This could be easily modelled in accordance with the suggestion of Pollard & Sag (1994: 30) that infinitival verbs do not assign case to their subjects, while finite verbs – in this case finite raising verbs – normally assign the nominative to their subjects and the accusative to their objects. But, as is well known (Andrews 1982; Zaenen & Maling 1983; Zaenen et al. 1985), some Icelandic verbs idiosyncratically assign specific "quirky" cases to their subjects, and when they do, the higher raising verbs must honour this assignment:

(4) a. *Hana* virðist vanta peninga.
her.ACC seems lack.INF money
'She seems to lack money.'

b. Hann telur *mig* vanta peninga.
he.NOM believes me.ACC lack.INF money
'He believes that I lack money.'

(5) a. *Barninu* virðist hafa batnað veikin.
the.child.DAT seems have.INF recovered.from the.disease
'The child seems to have recovered from the disease.'

b. Hann telur *barninu* hafa batnað veikin.
he believes the.child.DAT have.INF recovered.from the.disease
'He believes the child to have recovered from the disease.'

(6) a. *Verkjanna* virðist ekki gæta.
the.pains.GEN seems not be.noticeable.INF
'The pains don't seem to be noticeable.'

b. Hann telur *verkjanna* ekki gæta.
he believes the.pains.GEN not be.noticeable.INF
'He believes the pains to be not noticeable.'

Thus, in (4), the understood subject of the infinitival VANTA 'lack' must be in the accusative, whether it is raised to the object position, as in (4b), where the accusative would be expected anyway, or to the subject position, as in (4a), where normally the nominative case would be expected. This works similarly in the case of verbs idiosyncratically assigning their subject the dative case, as in (5), or the genitive case, as in (6).

The difficulty presented by such examples is this. If the finite raising verbs were assumed to assign case to the raised subjects – nominative in the case of raising to subject and accusative in the case of raising to object – then this would clash with "quirky" cases assigned to their subjects by some verbs: (4a), (5) and (6) would be predicted to be ungrammatical. If, on the other hand, such raising verbs did not assign case to the raised arguments, instead relying on the lower verbs to assign appropriate cases to their subjects, then it is not clear what case should be assigned to their subjects by the usual – not "quirky" – verbs: it cannot always be the nominative, as the accusative case is witnessed when the subject is raised to the object position, as in (3b); similarly, it cannot always be the accusative, as the nominative case surfaces when the subject is raised to the subject position, as in (3a).

The intuition of the analysis proposed in Sag et al. (1992) relies on the distinction between structural and inherent case assignment, although these terms do not appear in that paper. Verbs such as those in (4)–(6) assign their subjects specific inherent cases (accusative in (4), dative in (5) and genitive in (6)), while the usual verbs, as in (3), only mark their subjects as structural, to be assigned case elsewhere. Finite raising verbs are, in a way, sensitive to this distinction, and only assign the nominative (in the case of raising to subject) or accusative (in the case of raising to object) to such structural arguments. While Sag et al. (1992) represent this distinction between structural and inherent case implicitly, via the interaction of two attributes, CASE (realised case) and DCASE (default case), later HPSG work assumes explicit representation of the two kinds of case as two subtypes of *case* in the type hierarchy: *str*(uctural) and *lex*(ical). Such a *case* type hierarchy is, apparently independently, alluded to in Pollard (1994) and introduced in detail in Heinz & Matiasek (1994), to which we turn presently.

On the basis of German examples such as (1)–(2), Heinz & Matiasek (1994) argue that out of four morphological cases in German – nominative, accusative, genitive and dative – the first three (i.e., with the exception of the dative) may be assigned structurally, by general case assignment principles. Similarly, they argue that the last three (i.e., apart from the nominative) may also be assigned lexically, in which case they are stable across various syntactic environments. These empirical observations are translated into the *case* hierarchy in Figure 1.

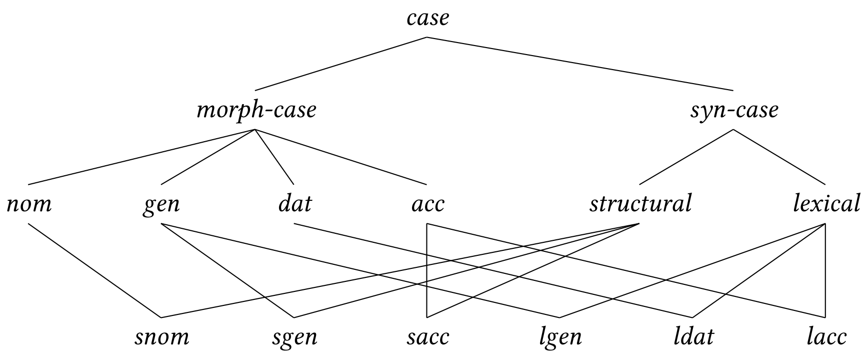


Figure 1: Case hierarchy for German encoding the structural/lexical distinction

Particular verbs may assign specific lexical cases to their arguments, e.g., *ldat*. They may also specify arguments as bearing structural case, in which case only the *str*(*uctural*) supertype is mentioned in the lexicon. For example, the lexical entries for UNTERSTÜTZEN 'support' and HELFEN 'help' contain the following subcategorisation requirements:

(7) a. UNTERSTÜTZEN: [SUBCAT ⟨ NP[*str*], NP[*str*] ⟩]
b. HELFEN: [SUBCAT ⟨ NP[*str*], NP[*ldat*] ⟩]

Assuming a similar *case* hierarchy for Icelandic, the difference between the usual verbs, such as ELSKA 'love' in (3a), and "quirky" subject verbs, such as VANTA 'lack' in (4), could be represented as below (omitting non-initial arguments):

(8) a. ELSKA: [SUBCAT ⟨ NP[*str*], …⟩]
b. VANTA: [SUBCAT ⟨ NP[*lacc*], …⟩]

Since Pollard (1994) and Heinz & Matiasek (1994), such representations of case requirements are generally adopted in HPSG,[2] with the only difference that SUBCAT is currently replaced with ARG-ST. The point where different approaches diverge is how exactly structural case is resolved to a specific morphological case.

The simplest principle would resolve the case of the first *str* argument of a pure (non-gerundial) verb to nominative, i.e., to *snom*, the case of any subsequent *str* argument of a pure verb to accusative, i.e., to *sacc*, and the case of any *str* argument of a gerund to *sgen*. Unfortunately, this simple principle would not work in various cases of raising, e.g., in the case of the Icelandic data above. While the "quirky" cases in (4)–(6) would be properly taken care of by this approach – once the subject is assigned a specific lexical case it is outside of the realm of a principle resolving structural cases – structural subjects raised to a higher verb would be assigned specific case twice (or more times, in the case of longer raising chains): on the SUBCAT (or ARG-ST) of the lower verb and on the SUBCAT (or ARG-ST) of the raising verb. This would not necessarily lead to problems in the case of raising to subject verbs, as in (3a), as the structural argument would be the subject in both subcategorisation frames, so it would be resolved to *snom* twice, but it would create a problem in the case of raising to object verbs, as in (3b), as the raised argument would be resolved to the nominative on the lower subcategorisation frame and to the accusative on the higher frame. So, the problem is not limited to Icelandic, but may be observed in any language with raising to object (also known as Exceptional Case Marking or Accusativus cum Infinitivo or AcI), including German (cf., e.g., Heinz & Matiasek 1994: 231). Obviously, even if a structural argument occurs on a number of SUBCAT or ARG-ST lists, it should

[2] A very recent example being Machicao y Priemer & Fritz-Huechante (2018: 169).

be assigned specific morphological case according to its position on just one of them – the highest one.

Both Pollard (1994) and Heinz & Matiasek (1994) account for such facts via configurational case principles, e.g. (Heinz & Matiasek 1994: 209):

(9) Case Principle (for German):

In a *head-complement-structure* whose head has category
verb[*fin*] the external argument has a case value of *snom*,
verb the internal argument has a case value of *sacc*,
noun the internal argument has a case value of *sgen*.
These are the only saturated or almost saturated *head-complement-structure*s with structural arguments.

(10) *Syntactically External Argument* ('*Subject*'):

If the first element of the subcat list of a sign is an NP[*str*], it is called the (*syntactically*) *external argument* of that sign.

(11) *Syntactically Internal Argument* ('*Direct Object*'):

If the second element of the subcat list of a sign is an NP[*str*], it is called the (*syntactically*) *internal argument* of that sign.

Heinz & Matiasek (1994: 209–210) formalize this Case Principle by giving the following constraints:

(12)
$$\begin{bmatrix} \textit{phrase} & \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{vform} & \textit{fin} \end{bmatrix} \\ \textsc{subcat} & \langle\rangle \end{bmatrix} \\ \textsc{dtrs} & \begin{bmatrix} \textit{h-c-str} & \\ \textsc{head-dtr}|\ldots|\textsc{subcat} & \langle \text{NP}[\textit{str}], \ldots \rangle \end{bmatrix} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} \textsc{dtrs}|\textsc{head-dtr}|\ldots|\textsc{subcat} & \langle \text{NP}[\textit{snom}], \ldots \rangle \end{bmatrix}$$

(13)
$$\begin{bmatrix} \textit{phrase} & \\ \textsc{synsem|loc|cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix}\textit{verb} & \\ \textsc{vform} & \textit{fin}\end{bmatrix} \\ \textsc{subcat} & \langle\rangle \lor \langle \textit{synsem} \rangle \end{bmatrix} \\ \textsc{dtrs} & \begin{bmatrix} \textit{h-c-str} & \\ \textsc{head-dtr|...|subcat} & \langle \textit{synsem}, \text{NP}[\textit{str}], \ldots\rangle \end{bmatrix}\end{bmatrix} \Rightarrow$$
$$\begin{bmatrix}\textsc{dtrs|head-dtr|...|subcat} & \langle \textit{synsem}, \text{NP}[\textit{sacc}]\rangle, \ldots\end{bmatrix}$$

(14)
$$\begin{bmatrix} \textit{phrase} & \\ \textsc{synsem|loc|cat} & \begin{bmatrix} \textsc{head} & \textit{noun} \\ \textsc{subcat} & \langle\rangle \lor \langle \textit{synsem} \rangle \end{bmatrix} \\ \textsc{dtrs} & \begin{bmatrix} \textit{h-c-str} & \\ \textsc{head-dtr|...|subcat} & \langle \textit{synsem}, \text{NP}[\textit{str}], \ldots\rangle \end{bmatrix}\end{bmatrix} \Rightarrow$$
$$\begin{bmatrix}\textsc{dtrs|head-dtr|...|subcat} & \langle \textit{synsem}, \text{NP}[\textit{sgen}]\rangle, \ldots\end{bmatrix}$$

Note that the locus of this Case Principle is *phrase* and that it makes reference to *head-complement-structure* values of the DAUGHTERS (DTRS) attribute. In this sense, this principle is configurational. Similar principles were proposed for Korean (Yoo 1993; Bratt 1996), English (Grover 1995) and Polish (Przepiórkowski 1996a), *inter alia.*

This configurational approach to case assignment is criticised in Przepiórkowski (1996b; 1999a,b) on the basis of conceptual and theory-internal problems. The conceptual problem is that a configurational analysis is employed for what is usually considered an essentially local phenomenon, one concerned with the relation between a head and its dependents (Blake 1994). The – more immediate – theory-internal problem is that such configurational case principles are restricted to locally realised arguments, and are not necessarily compatible with those – dominant since Pollard & Sag (1994: Chapter 9) – HPSG analyses of extraction which do not assume traces and with those HPSG approaches to cliticisation in which the clitic is realised as an affix rather than a tree-configurational constituent (cf., e.g., Miller & Sag 1997 on French and Monachesi 1999 on Italian).

The solution proposed in Przepiórkowski (1996b; 1999a,b) is to resolve structural cases directly within ARG-ST, via local principles operating at the level of *category* of a word (where both head information and argument structure information – but not constituent structure – are available) rather than at the level of *phrase*. This seems to bring back the problem, discussed in connection with the Icelandic data above, of raised arguments, which occur on a number of ARG-ST lists. The innovation of Przepiórkowski (1996b; 1999a,b) is the proposal to mark, within ARG-ST, whether a given argument is realised locally (either tree-configurationally, or as a gap to be extracted higher on, or as an affix) or not. If it is realised locally, it may be assigned appropriate case; if it is not (because it is raised), its structural case must be resolved higher up. On this setup, the above constraints (12)–(13) responsible for the assignment of structural nominative and accusative are replaced with the following two constraints (and similarly for the structural genitive):[3]

(15)
$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{ARG} & \text{NP}[\textit{str}] \\ \text{REALIZED} & + \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \left\langle \begin{bmatrix} \text{ARG} & \text{NP}[\textit{snom}] \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

(16)
$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \boxed{1}\ \textit{nelist} \oplus \left\langle \begin{bmatrix} \text{ARG} & \text{NP}[\textit{str}] \\ \text{REALIZED} & + \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} \text{ARG-ST} & \boxed{1} \oplus \left\langle \begin{bmatrix} \text{ARG} & \text{NP}[\textit{snom}] \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

Obviously, for such constraints to work, values of ARG-ST must be lists of slightly more complex objects than *synsem* (these are now values of ARG within such more complex objects), and additional principles must make sure that values of REALIZED are instantiated properly (see Przepiórkowski 1999a: 78–79 for details).

The analysis of Przepiórkowski (1996b; 1999a,b) assumes that an argument is locally realised – and hence may be assigned structural case – if and only if it is not raised to a higher argument structure. Meurers (1999a,b), on the basis of empirical observations in Haider (1990), Grewendorf (1994) and Müller (1997), shows that this assumption does not always hold in German; rather, structural case should be assigned to arguments on the basis of whether they are raised or not, and not whether they are locally realised or not. Consider the following data (Meurers 1999a: 294):

[3] The antecedents of such principles could be further constrained to apply to *words* only. As usual, '⊕' indicates concatenation of lists.

(17) a. [*Ein* *Außenseiter* gewinnen] wird hier nie.
an.NOM outsider win.INF will here never
'An outsider will never win here.'

b. [*Einen* *Außenseiter* gewinnen] läßt Gott hier nie.
an.ACC outsider win.INF lets god here never
'God never lets an outsider win here.'

Assuming that fronted fragments, marked with square brackets, are single constituents,[4] the subject of *gewinnen* 'win' forms a constituent with this verb, i.e., it has the same configurational realisation in both examples. Hence, configurational case assignment principles should assign it the same case in both instances, contrary to facts: *ein Außenseiter* occurs in the nominative in (17a) and *einen Außenseiter* bears the accusative case in (17b). As argued by Meurers (1999a,b), the reason is that – although the subject is realised locally to its infinitival head – it is in some sense raised further to the subject position of the auxiliary *wird* in (17a) and to the object position of the AcI verb *läßt* in (17b), hence the difference in cases. This suggests that structural case should be assigned not where the argument is realised, but on the highest ARG-ST on which it occurs. A corresponding modification of the non-configurational case assignment approach of Przepiórkowski (1996b; 1999a,b) – replacing the [REALIZED +] with [RAISED −] in constraints such as (15)–(16) and providing appropriate constraints on values of RAISED – is proposed in Przepiórkowski (1999a: 93–95); see also Müller (2013), Section 17.4 (and references therein) for further improvements.

While this non-configurational approach to syntactic case assignment was motivated largely by the need to capture complex interactions in a precise way, it turns out to formalise sometimes apparently contradictory intuitions expressed in various approaches to case. First of all, it preserves the common intuition that case is a local phenomenon, an intimate relation between a head and its dependents. Second, it successfully formalises the distinction between structural and inherent/lexical case known from the transformational literature of the 1980s, and non-configurationally encodes the apparently configurational principles of structural case assignment. Third, while most HPSG literature on case is concerned with syntactic phenomena in European languages, this approach has been extended to case stacking known, e.g., from languages of Australia and case attraction observed, e.g., in Classical Armenian and in Gothic (Malouf 2000). Fourth, by allowing antecedents of implicational constraints such as (15)–(16) to

[4] This assumption is not completely uncontroversial; see Kiss (1994: 100–101) for apparent counterexamples and Müller (2003; 2005; 2017) for a defense of this assumption.

be *local* objects, not just syntactic *categories*, semantic factors influencing case assignment may also be taken into account, as in differential case marking, repeatedly considered in Lexical Functional Grammar (cf., e.g., Butt & King 2003 and references therein), but apparently not (so far) in HPSG. Fifth, as pointed out in Przepiórkowski (1999a,b), the above approach to case formalises the "case tier" intuition of Zaenen et al. (1985), Yip et al. (1987) and Maling (1993) (see also Maling 2009).

Let us illustrate the last point with some Finnish data from Maling (1993):

(18) a.

Liisa	muisti	matkan	vuoden.
Liisa.NOM	remembered	trip.ACC	year.ACC

'Liisa remembered the trip for a year.'

b.

Lapsen	täytyy	lukea	kirja	kolmannen	kerran.
child.GEN	must	read	book.NOM	[third	time].ACC

'The child must read the book for a third time.'

c.

Kekkoseen	luotettiin	yksi	kerta.
Kekkonen.ILL	trust.PASSP	[one	time].NOM

'Kekkonen was trusted once.'

d.

Kekkoseen	luotettiin	yhden	kerran	yksi	vuosi.
Kekkonen.ILL	trust.PASSP	[one	time].ACC	[one	year].NOM

'Kekkonen was trusted for one year once.'

Maling (1993) argues at length that some adjuncts (adverbials of measure, duration and frequency) behave just like objects with respect to case assignment and, in particular, notes the following generalization about syntactic case assignment: only one NP dependent of the verb receives the nominative, namely the one which has the highest grammatical function; other dependents receive the accusative.[5] Thus, if none of the arguments bears inherent case, the subject is in the nominative and other dependents are in the accusative, cf. (18), but if the subject bears an idiosyncratic case, it is the object that gets the nominative, cf. (18b). Furthermore, if all arguments (if any) bear inherent case, the next "available" grammatical function is that of an adjunct, thus one of the adjuncts receives the nominative, cf. (18c)–(18d).

Given such facts, Maling (1993) claims that syntactic case is assigned in Finnish on the basis of grammatical hierarchy and that (at least some) adjuncts belong to this hierarchy. Moreover, as evidenced by (18c)–(18d), adjuncts do not form

[5] See also Zaenen & Maling (1983) and Zaenen et al. (1985) for a similar generalisation with respect to Icelandic.

a single class in this hierarchy: although the multiplicative adverbial *yksi kerta* is nominative in (18c), this case is won over by the duration adverbial in (18d). Taking into consideration also the partitive of negation facts (measure adverbials, but not duration or frequency adverbials, behave like direct objects in the sense that they take partitive case under sentential negation), Maling (1993) extends the grammatical function hierarchy for Finnish in the following way:

(19) SUBJ > OBJ > MEASURE > DURATION > FREQUENCY

While these generalisations are developed in the context of Lexical Functional Grammar, it is not clear how they could be encoded in LFG: there are no formal mechanisms for stating such a hierarchy of grammatical functions and, additionally, all adjuncts are assumed to be elements of an unordered set.[6] On the other hand, given the "adjuncts as complements" approach of Bouma et al. (2001) and others, upon which at least some adjuncts are added to ARG-ST (perhaps renamed to DEPS), and assuming – as is standard in HPSG – that ARG-ST elements satisfy the obliqueness hierarchy, formalisation of the "case tier" approach is easy and consists of two implicational principles similar to (15)–(16). The first principle resolves the first structurally-cased element of extended ARG-ST to nominative, whether this element is the first element of ARG-ST or not (it is not in the case of (18b)–(18d)), and whether it corresponds to the subject, the direct object or an adjunct. The second principle resolves the structural case of all subsequent elements, if any, to accusative.

2 Case syncretism and neutrality

Another important strand of HPSG work on case concerns situations in which a single syncretic form seems to simultaneously bear two (or more) case values, as in the following examples involving coordination, free relatives and parasitic gaps:

(20) Polish coordination (Dyła 1984: 701–702):

a. Kogo who.ACC/GEN Janek Janek.NOM lubi likes(OBJ.ACC) a and Jerzy Jerzy.NOM nienawidzi? hates(OBJ.GEN)

'Who does Janek like and Jerzy hate?'

[6] But see Przepiórkowski (2016) for an attempt to introduce a single ordered list of dependents and formalise the functional hierarchy in LFG.

b. * Co Janek lubi a Jerzy
what.NOM/ACC Janek.NOM likes(OBJ.ACC) and Jerzy.NOM
nienawidzi?
hates(OBJ.GEN)
Intended: 'What does Janek like and Jerzy hate?'

(21) English coordination (Goodall 1987: 70; Levine et al. 2001: 206):

This is the man who$_i$.NOM/ACC Robin saw e_i.ACC and thinks e_i.NOM is handsome.

(22) German coordination (Pullum & Zwicky 1986: 764–765):

a. Er findet und hilft
he.NOM finds(OBJ.ACC) and helps(OBJ.DAT)
Frauen.
women.NOM/ACC/GEN/DAT
'He finds and helps women.'

b. * Sie findet und hilft Männer.
she.NOM finds(OBJ.ACC) and helps(OBJ.DAT) men.NOM/ACC/GEN
Intended: 'She finds and helps men.'

c. * Sie findet und hilft Männern.
she.NOM finds(OBJ.ACC) and helps(OBJ.DAT) men.DAT
Intended: 'She finds and helps men.'

(23) German free relatives (Groos & van Riemsdijk 1981: 212):

Was du mir gegeben hast, ist
what.NOM/ACC you.NOM me.DAT given(OBJ.ACC) have is(SUBJ.NOM)
prächtig.
wonderful
'What you have given to me is wonderful.'

(24) English parasitic gaps (Hukari & Levine 1996: 482; Levine et al. 2001: 205):

Robin is someone who$_i$.NOM/ACC even good friends of e_i.ACC believe e_i.NOM should be closely watched.

In (20a), the fronted syncretic accusative/genitive form *kogo* 'who' satisfies the requirements of the two coordinated verbal constituents: in one, *lubi* 'likes' requires an accusative object, and in the other, *nienawidzi* 'hates' expects a genitive object. A form which is not syncretic between (at least) these two cases cannot occur in the place of *kogo*; this is illustrated in (20b), where the element putatively shared by the two verbal constituents is syncretic between accusative and nominative, rather than accusative and genitive. The English example (21) is similar and involves the relative pronoun *who*, syncretic between accusative and nominative. The well-known example (22) illustrates essentially the same phenomenon in German: the form *Frauen* 'women', which is fully syncretic with respect to case, simultaneously satisfies the accusative requirement of *findet* 'finds' and the dative requirement of *hilft* 'helps'. By contrast, this joint requirement is not satisfied either by *Männer*, which is accusative (among other cases) but not dative, or by *Männern*, which is dative but not accusative. The other two examples show that this phenomenon is not restricted to coordination. In (23), the syncretic form *was* 'what' simultaneously satisfies the constraint that the object of *gegeben* 'given' is accusative and that the subject of *ist* 'is' is nominative. Similarly, the extracted *who* in (24) seems to simultaneously bear the accusative case assigned by the preposition *of* and the nominative case of the subject of *should*.

Such examples were at one point considered problematic not only for HPSG, but for unification-based theories in general (Ingria 1990). The reason is that, on the straightforward approach to case, they should all be ungrammatical. For example, in the case of (22a), the assignment of the accusative to the object of *findet* 'finds' should clash with the assignment of the dative to the object of *hilft* 'helps', as both objects are realised by the same noun *Frauen* 'women'. In other words, the attempt to unify accusative and dative should fail.

The solution first proposed by Levine et al. (2001: 207–208) is to enrich the *case* hierarchy in such a way that the unification of two different morphological cases does not necessarily result in failure.[7] Specifically, assuming that nominative and accusative are structural cases in English, they propose the following part of the structural case hierarchy.[8] The enriched case hierarchy is shown in Figure 2.

Particular nominal forms are specified in the lexicon as either pure accusative (*p-acc*), pure nominative (*p-nom*) or syncretic between the two (*p-nom-acc*):

(25)	*he*	[CASE *p-nom*]
	him	[CASE *p-acc*]
	whom	[CASE *p-acc*]
	who	[CASE *p-nom-acc*]
	Robin	[CASE *p-nom-acc*]

[7] See Ingria (1990: 196) for an earlier implementation of roughly the same idea in the context of unification grammars.

[8] Type names follow the convention in Daniels (2002), for increased uniformity with the remainder of this section.

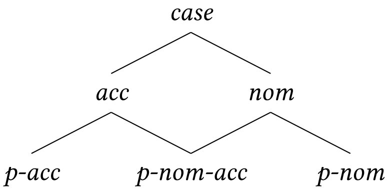


Figure 2: Case hierarchy for English encoding case syncretism

On the other hand, heads – or constraints of a case principle of the kind presented in the previous section – specify particular arguments as *nom* or *acc*. So, in the case of the parasitic gap example (24), the *acc* requirement associated with the preposition *of* and the *nom* requirement on the subject of *should* are not incompatible: their unification results in *p-nom-acc* and the shared dependent may be any form compatible with this case value, e.g., *who* (but not *whom*). Examples (20)–(23) can be handled in a similar way.

A situation often perceived as dual to such case neutrality, sometimes called "case underspecification", occurs when a head specifies the case of its dependent disjunctively and may combine with a coordinate structure containing phrases in both cases, e.g.:

(26) a. Polish (Przepiórkowski 1999a: 175):

Dajcie wina i całą świnię!
give wine.GEN and whole.ACC pig.ACC
'Serve (some) wine and a whole pig!'

b. Russian (Levy 2001: 11):

Včera vec' den' on proždal svoju podrugu Irinu
yesterday all day he expected self's.ACC girlfriend.ACC Irina.ACC
i zvonka ot svoego brata Grigorija.
and call.GEN from self's brother Grigory
'Yesterday he waited all day for his girlfriend Irina and for a call from his brother Grigory.'

In Polish, the object of the verb *dajcie* 'give' is normally in the accusative, but may also be realised as the genitive, when its meaning is partitive; in (26a), the object is a coordination of such a genitive noun *wina* '(some) wine' and the accusative *całą świnię* 'whole pig'. Similarly, according to Levy 2001, the Russian verb *proždal* 'awaited' may combine with accusative or genitive, and in (26b) it happily combines with a coordinate phrase containing both.

If such "accusative and genitive" coordinate phrases bear case at all, the value of this grammatical category must be something like *acc+gen*. Note that this situation differs from case neutrality discussed above: a syncretic case such as *p-acc-gen* intuitively corresponds to intersection: a nominal bearing this case is accusative and genitive at the same time. On the other hand, the intuition behind *acc+gen* is that of union: a (coordinated) nominal with this case value has accusative elements and genitive elements, so it may fill a position disjunctively specified as requiring accusative *or* genitive. However, *acc+gen* coordinate phrases cannot fill either purely accusative positions (because such phrases contain genitive – i.e., non-accusative – conjuncts), or purely genitive positions (because of accusative – i.e., non-genitive – conjuncts), or positions simultaneously specified as accusative *and* genitive, as in (20) above (for both reasons).

This duality is a feature of the Categorial Grammar approach to case and coordination of Bayer (1996) (see also Bayer & Johnson 1995) and the corresponding HPSG analyses were presented in Levy (2001) and Levy & Pollard (2002), as well as in Daniels (2002). As noted in Levy & Pollard (2002: 233), the two HPSG approaches are isomorphic. The main technical difference is that the relevant case hierarchies are construed outside of the usual HPSG type hierarchy in the approach of Levy (2001) and Levy & Pollard (2002), but they are fully integrated in the approach of Daniels (2002). For this reason, and also because it is the basis of some further HPSG work (e.g., Crysmann 2005), this latter approach is presented below.

Intuitively, just as the common subtype of *acc* and *nom*, i.e., *p-nom-acc* in Figure 2, represents forms which are simultaneously accusative and nominative, the common supertype, i.e., *case*, which should perhaps be renamed to *nom+acc*, should represent coordinate structures involving nominative and accusative conjuncts. However, given that all objects are assumed to be sort-resolved in standard HPSG, saying that the case of a coordinate structure is *case* (or *nom+acc*) is paramount to saying that it is either *p-acc* (pure accusative), or *p-nom-acc* (syncretic nominative/accusative), or *p-nom* (pure nominative). One solution is to "make a simple change to the framework's foundational assumptions" (Sag 2003: 268) and to allow linguistic objects to bear non-maximal types. This is proposed and illustrated in detail in Sag (2003). A more conservative solution, proposed in Daniels (2002), is to add dedicated maximal types to all such non-maximal types; for example, the hierarchy in Figure 2 becomes Figure 3. Apart from the trivial renaming of *case* to the more explicit *nom+acc*, a maximal type corresponding to this renamed non-maximal type is added here, namely, *p-nom+acc*.

Let us illustrate this approach with the two Polish examples (20a) and (26a),

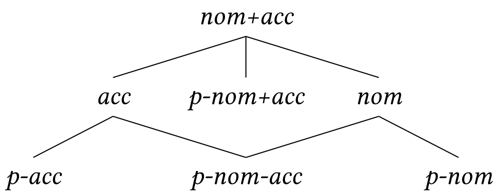


Figure 3: Case (sub)hierarchy encoding nominative/accusative syncretism and underspecification

repeated below as (27a)–(27b):

(27) a. Kogo Janek lubi a Jerzy nienawidzi?
who.ACC/GEN Janek.NOM likes(OBJ.ACC) and Jerzy.NOM hates(OBJ.GEN)
'Who does Janek like and Jerzy hate?'

b. Dajcie wina i całą świnię!
give wine.GEN and whole.ACC pig.ACC
'Serve (some) wine and a whole pig!'

As these examples involve accusative and genitive, I will assume that the complete case hierarchy contains a subhierarchy such as Figure 3 above, but with all occurrences of *nom* replaced by *gen* as in Figure 4.

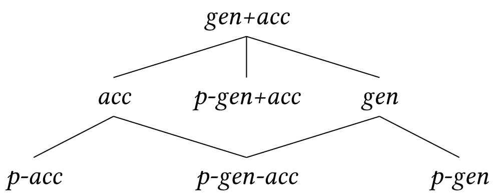


Figure 4: Case (sub)hierarchy encoding accusative/genitive syncretism and underspecification

First of all, heads subcategorise for (or relevant case principles specify) "non-pure" cases, i.e., *acc*, *gen*, *gen+acc*, etc., but not *p-acc*, *p-gen*, *p-gen+acc*, etc. For example, *lubi* 'likes' and *nienawidzi* 'hates' in (27a) expect their objects to have the case values *acc* and *gen*, respectively. Moreover, *dajcie* 'give' in (27b) specifies the case of its object as *gen+acc*. On the other hand, nominal dependents bear "pure" cases. For example, *kogo* 'who' in (27a) is lexically specified as *p-gen-acc*. Similarly to the analysis of the English parasitic gap example above, this neutralised case is compatible with both specifications: *acc* and *gen*.

The analysis of (27b) is a little more complicated, as a new principle is needed to determine the case of a coordinate structure. The two conjuncts, *wina* 'wine' and *całą świnię* 'whole pig', have – by virtue of lexical specifications of their head nouns – the case values *p-gen* and *p-acc*, respectively. Now, the case value of the coordination is determined as follows: take the "non-pure" versions of the cases of all conjuncts (here: *gen* and *acc*), find their (lowest) common supertype (here: *gen+acc*), and assign to the coordinate structure the "pure" type corresponding to this common supertype (here: *p-gen+acc*). This way the coordinate structure in (27b) ends up with the case value *p-gen+acc*, which is compatible with the *gen+acc* requirement posited by the verb *dajcie* (or by an appropriate principle of structural case assignment). Obviously, a purely accusative, purely genitive or accusative/genitive neutralised object would also satisfy this requirement.

One often-perceived – both within and outside of HPSG – problem with this approach is that it leads to very complex type hierarchies for *case* and rather inelegant constraints (Sag 2003: 272, Dalrymple et al. 2009: 63–66). Let us, following Daniels (2002), simplify the presentation of type hierarchies such as Figure 3, by removing all those "pure" types which are only needed to represent some non-maximal types as maximal as in Figure 5. Hence, the representation in this figure

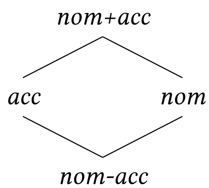


Figure 5: Simplified case (sub)hierarchy encoding nominative/accusative syncretism and underspecification

corresponds to seven types shown explicitly in Figure 3 (each non-maximal type in Figure 5 has an additional *p-* type, while the maximal *nom-acc* in Figure 5 is the same as *p-nom-acc* in Figure 3). What would a similar hierarchy for three morphological cases look like? Daniels (2002: 143) provides the following visualisation, involving 18 nodes corresponding to 35 types in the full type hierarchy in Figure 6. As mentioned in Levy & Pollard (2002: 225), the size of such a type hierarchy grows double exponentially with the number of grammatical cases, so it would already be next to impossible to visualise such a hierarchy for German, with its four cases, not to mention Polish with its seven cases or Finno-Ugric languages with around 15 cases. And matters are further complicated by the fact that sometimes form syncretism simultaneously involves a number of grammati-

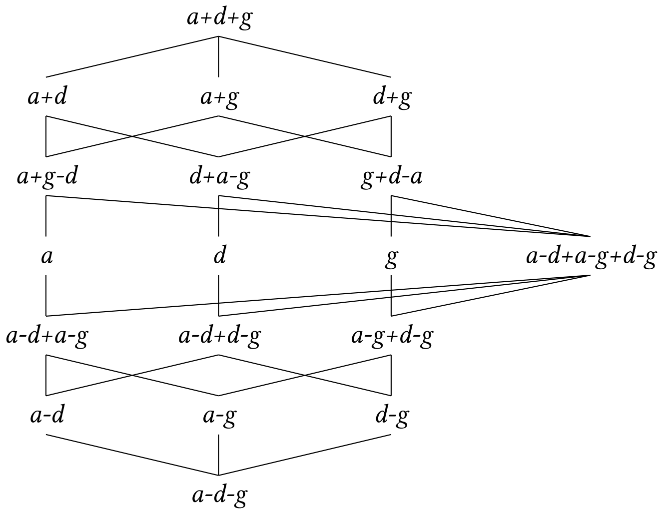


Figure 6: Simplified case (sub)hierarchy encoding accusative/dative/genitive syncretism and underspecification

cal categories, so perhaps such type hierarchies should combine case information with person, gender and number (Daniels 2002: 145, Crysmann 2005), and by the fact that coordinated elements may be specified for different categories (e.g., an NP specified for case may be coordinated with a sentence), in which case it is not clear what categories should be borne by the coordinate structure as a whole (see, e.g., the inconclusive fn. 10 in Sag 2003: 277).

After the early 2000s, such complex *case* hierarchies do not appear in HPSG work. A possible reason for this is the increasing popularity of ellipsis-based accounts of various coordinate constructions, including unlike category coordination cases, of which the "case underspecification" examples (26a)–(26b) may be seen as special cases.[9] Such ellipsis accounts are usually formulated within the linearisation approach of Reape (1992; 1994) and Kathol (1995), and they have been claimed to deal with some of the cases discussed in this section, e.g., by Crysmann (2008), Beavers & Sag (2004), and Chaves (2006; 2008). However, such linearisation-based approaches to coordination have more recently come under attack: see Levine (2011) and Kubota & Levine (2015) (as well as Yatabe 2012;

[9] Another HPSG approach to unlike category coordination which obviates the need for such complex hierarchies is that of Yatabe (2004), according to which the – perhaps disjunctive or underspecified – requirements of the head independently distribute to all conjuncts, in a manner similar to (but more general than) distributivity within coordinate structures assumed in LFG (Dalrymple & Kaplan 2000; Dalrymple et al. 2009; Przepiórkowski & Patejuk 2012).

2016 and, especially, Yatabe & Tam 2018 for a defence of ellipsis-based accounts of some cases of coordination). Hence, it is difficult to predict at the moment whether ellipsis-based analyses will permanently remove the need for complex type hierarchies modelling neutralisation and underspecification in coordination. But even if they do, some of the examples given at the beginning of this section, namely (23)–(24), demonstrated that feature neutrality is not limited to coordinate structures, but also occurs at least in free relatives and multiple gapping, so case hierarchies of the kind illustrated in Figure 2, with separate types representing syncretic cases, are still needed in contemporary HPSG, regardless of the analysis of coordination; an example of a more recent analysis which does assume such a case hierarchy (to account for gapping and resumptive pronouns in Modern Standard Arabic) is Alotaibi & Borsley (2013).[10]

3 Other HPSG work on case

Apart from the two clearly identifiable strands of HPSG work described in the two preceding sections, there are also single papers concerned with various theoretical and implementational aspects of grammatical case. Of these, the report by Drellishak (2008) on modelling complex case phenomena in the Grammar Matrix (Bender et al. 2002) has the widest typological scope. It describes the treatment of various case systems in the multilingual platform for implementing HPSG grammars: not only the pure nominative-accusative, ergative-absolutive and tripartite systems, but also systems with various types of split ergativitity, systems – known from Austronesian languages, including Tagalog – in which case marking interacts with focus marking, and so-called "direct-inverse" systems, exemplified by Algonquian languages, in which case marking partially depends on the hierarchies – or scales – of nominal phrases, e.g., based on person and/or animacy. Similarly to the non-configurational case assignment principles discussed in Section 1 above, such systems are described – via constraints on specific lexical types – by specifying case values of elements on ARG-ST. Also, a typologically very interesting language, Nias, usually assumed to display the ergative-absolutive alignment but with the typologically exceptional property of marking the absolutive – rather than the ergative – case, is reanalysed as a nominative-accusative language in Crysmann (2009), with the sole argument of intransitive verbs mapped to the grammatical function of object, rather than subject.

[10]But see Crysmann (2017) for a reanalysis which does not need to refer to such a case hierarchy.

Two other works mentioned here are concerned with two very different aspects of case systems of particular languages. Ryu (2013) investigates the issue of case spreading from an argument of a verb to certain nominal dependents of this argument in Korean. He investigates the semantic relations that must hold between the two nominals for such "case copying" to occur and proposes a repertoire of 16 semantic relations (collected in five coherent groups, further classified into two general classes) which make the spreading of the nominative possible, 10 of which (three of the five groups, one of the two classes) license the spreading of the accusative. On the syntactic side, the dependents of such nominal arguments are raised to become valency elements of the governing verbs. In particular, dependents of the subject are raised to the VAL|SUBJ list, resulting in multiple valency subjects. Configurational case assignment rules constrain the value of case of each valency subject to nominative, and of each valency complement to accusative. The paper does not discuss the (im)possibility of formulating such case assignment rules non-configurationally, within local ARG-ST (or DEPS), but the challenge for the non-configurational case assignment seems to be the fact that multiple argument structure elements may correspond to valency subjects (and multiple to valency complements), so – looking at the argument structure alone – it is not immediately clear how many initial elements of this list should be assigned the nominative case, and which final elements should get the accusative.

Finally, a very different aspect of Hungarian case is investigated in Thuilier (2011), namely, whether case affixes should be distinguished from postpositions and, if so, where to draw the line. In Hungarian, postpositions behave in some respect just like case affixes (e.g., they do not allow any intervening material between them and the nominal phrase) which led some researches to deny the existence of this distinction. Thuilier (2011) shows that, in this case, the traditional received wisdom is right, and that case affixes and postpositions differ in a number of morphological and syntactic ways. The proposed tests suggest that the essive element *ként*, normally considered to be a case affix, should be reanalysed as a postposition, thus establishing the number of Hungarian cases as 16. The resulting analysis of Hungarian case affixes and postpositions is couched within Sign-Based Construction Grammar (Boas & Sag 2012).

In summary, while perhaps HPSG is not best known for its approach to grammatical case, it does offer a range of interesting accounts of a variety of case-related phenomena in diverse languages ranging from German, Icelandic and Polish through Finnish and Hungarian to Korean and Nias; it provides perhaps the only formal implementation of the influential "case tier" idea; and it successfully captures somewhat conflicting intuitions concerning the locality of case assignment.

Acknowledgements

I would like to thank the following colleagues for their comments on a previous version of this chapter: Rui Chaves, Tony Davis, Detmar Meurers, Stefan Müller and Shûichi Yatabe. I wish I could blame them for any remaining errors and omissions.

References

Alotaibi, Mansour & Robert D. Borsley. 2013. Gaps and resumptive pronouns in Modern Standard Arabic. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 6–26. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2013/alotaibi-borsley.pdf, accessed 2018-2-25.

Andrews, Avery D. 1982. The representation of case in Modern Icelandic. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 427–503. Cambridge, MA/London: MIT Press.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72(3). 579–616.

Bayer, Samuel & Mark Johnson. 1995. Features and agreement. In Hans Uszkoreit (ed.), *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 70–76. Cambridge, MA: Association for Computational Linguistics.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 48–69. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

Bender, Emily M., Daniel P. Flickinger & Stephan Oepen. 2002. The Grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the Workshop on Grammar*

Engineering and Evaluation at the 19th International Conference on Computational Linguistics, 8–14. Taipei, Taiwan.

Blake, Barry J. 1994. *Case*. Cambridge, UK: Cambridge University Press.

Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based Construction Grammar* (CSLI Lecture Notes 193). Stanford, CA: CSLI Publications.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bratt, Elizabeth Owen. 1996. *Argument composition and the lexicon: Lexical and periphrastic causatives in Korean*. Stanford University dissertation.

Butt, Miriam. 2006. *Theories of case*. Cambridge, UK: Cambridge University Press.

Butt, Miriam & Tracy Holloway King. 2003. Case systems: beyond structural distinctions. In Ellen Brandner & Heike Zinsmeister (eds.), *New perspectives on Case Theory* (CSLI Lecture Notes 156), 53–87. Stanford, CA: CSLI Publications.

Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, 102–122. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2006/, accessed 2018-2-25.

Chaves, Rui P. 2008. Linearization-based word-part ellipsis. *Linguistics and Philosophy* 31(3). 261–307. DOI:10.1007/s10988-008-9040-3

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of language: its nature, origin, and use*. Westport: Praeger.

Crysmann, Berthold. 2005. Syncretism in German: A unified approach to underspecification, indeterminacy, and likeness of case. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 91–107. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2005/, accessed 2018-2-25.

Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, 47–62. Stanford, CA: CSLI Publications.

Crysmann, Berthold. 2009. Deriving superficial ergativity in Nias. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 68–88. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2009/, accessed 2018-2-25.

Crysmann, Berthold. 2017. Resumption and case: A new take on Modern Standard Arabic. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 120–240. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/.

Dalrymple, Mary & Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 74(4). 759–798. DOI:10.2307/417199

Dalrymple, Mary, Tracy Holloway King & Louisa Sadler. 2009. Indeterminacy by underspecification. *Journal of Linguistics* 45(1). 31–68. DOI:10.1017/S0022226708005513

Daniels, Michael W. 2002. On a type-based analysis of feature neutrality and the coordination of unlikes. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 137–147. Stanford, CA: CSLI Publications.

Drellishak, Scott. 2008. Complex case phenomena in the Grammar Matrix. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 67–86. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/, accessed 2018-2-25.

Dyła, Stefan. 1984. Across-the-board dependencies and case in Polish. *Linguistic Inquiry* 15(4). 701–705.

Goodall, Grant. 1987. *Parallel structures in syntax: coordination, causatives, and restructuring* (Cambridge Studies in Linguistics 46). Cambridge, UK: Cambridge University Press.

Grewendorf, Günther. 1994. Kohärente Infinitive und Inkorporation. In Anita Steube & Gerhild Zybatow (eds.), *Zur Satzwertigkeit von Infinitiven und Small Clauses* (Linguistische Arbeiten 315), 31–50. Tübingen: Max Niemeyer Verlag.

Groos, Anneke & Henk van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar. In A. Belletti, L. Brandi & L. Rizzi (eds.), *Theory of markedness in Generative Grammar*, 171–216. Pisa: Scuola Normale Superiore.

Grover, Claire. 1995. *Rethinking some empty categories: Missing objects and parasitic gaps in HPSG*. Department of Language & Linguistics University of Essex dissertation.

Haider, Hubert. 1985. The case of German. In Jindřich Toman (ed.), *Studies in German grammar* (Studies in Generative Grammar 21), 23–64. Dordrecht: Foris Publications.

Haider, Hubert. 1990. Topicalization and other puzzles of German syntax. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and Barriers* (Linguistik Aktuell/Linguistics Today 5), 93–112. Amsterdam: John Benjamins Publishing Co. DOI:10.1075/la.5.06hai

Heinz, Wolfgang & Johannes Matiasek. 1994. Argument structure and case assignment in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 199–236. Stanford, CA: CSLI Publications.

Hukari, Thomas E. & Robert D. Levine. 1996. Phrase Structure Grammar: The next generation. *Journal of Linguistics* 32(2). 465–496. DOI:10.1017/S0022226700015978

Ingria, Robert J. P. 1990. The limits of unification. In *28th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 194–204. Pittsburgh, Pennsylvania: Association for Computational Linguistics.

Kathol, Andreas. 1995. *Linearization-based German syntax*. Ohio State University dissertation.

Kiss, Tibor. 1994. Obligatory coherence: The structure of German modal verb constructions. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 71–108. Stanford, CA: CSLI Publications.

Kubota, Yusuke & Robert Levine. 2015. Against ellipsis: Arguments for the direct licensing of 'noncanonical' coordinations. *Linguistics and Philosophy* 38. 521–576.

Levine, Robert D. 2011. Linearization and its discontents. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 126–146. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2011/, accessed 2018-2-25.

Levine, Robert D., Thomas E. Hukari & Mike Calcagno. 2001. Parasitic gaps in English: Some overlooked cases and their theoretical consequences. In Peter W. Culicover & Paul M. Postal (eds.), *Parasitic gaps* (Current Studies in Linguistics 35), 181–222. London, England/Cambridge, MA: MIT Press.

Levy, Roger. 2001. Feature indeterminacy and the coordination of unlikes in a totally well-typed HPSG. http://www.mit.edu/~rplevy/papers/feature-indet.pdf, accessed 2018-7-29. Unpublished manuscript.

Levy, Roger & Carl Pollard. 2002. Coordination and neutralization in HPSG. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 221–234. Stanford, CA: CSLI Publications.

Machicao y Priemer, Antonio & Paola Fritz-Huechante. 2018. Korean and Spanish psych-verbs: Interaction of case, theta-roles, linearization, and event structure in HPSG. In Stefan Müller & Frank Richter (eds.), *Proceedings of the 25th In-*

ternational Conference on Head-Driven Phrase Structure Grammar, University of Tokyo, 155–175. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2018/hpsg2018-machicaoypriemer-fritz-huechante.pdf.

Malchukov, Andrej & Andrew Spencer (eds.). 2009. *The Oxford handbook of case*. Oxford: Oxford University Press.

Maling, Joan. 1993. Of nominative and accusative: The hierarchical assignment of grammatical case in Finnish. In Anders Holmberg & Urpo Nikanne (eds.), *Case and other functional categories in Finnish syntax*, 51–76. Dordrecht: Mouton.

Maling, Joan. 2009. The case tier: A hierarchical approach to morphological case. In Andrej Malchukov & Andrew Spencer (eds.), 72–87. Oxford: Oxford University Press.

Malouf, Robert. 2000. A head-driven account of long-distance case assignment. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism 8), 201–214. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar. 1999a. *Lexical generalizations in the syntax of German non-finite constructions*. Tübingen: Eberhard-Karls-Universität dissertation.

Meurers, Walt Detmar. 1999b. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226. http://www.sfs.uni-tuebingen.de/~dm/papers/gagl99.html, accessed 2018-2-25.

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Monachesi, Paola. 1999. *A lexical approach to Italian cliticization* (CSLI Lecture Notes 84). Stanford, CA: CSLI Publications.

Müller, Stefan. 1996. Yet another paper about partial verb phrase fronting in German. In *Proceedings of COLING 96*, 800–805. Copenhagen.

Müller, Stefan. 1997. *Yet another paper about partial verb phrase fronting in German*. Research Report RR-97-07. Saarbrücken: DFKI. A shorter version appeared as Müller (1996).

Müller, Stefan. 2003. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330.

Müller, Stefan. 2013. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 3rd edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2017. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Revise and resubmit.

Pollard, Carl J. 1994. Toward a unified account of passive in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 273–296. Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Przepiórkowski, Adam. 1996a. Case assignment in Polish: Towards an HPSG analysis. In Claire Grover & Enric Vallduví (eds.), *Studies in HPSG*, vol. 12 (Edinburgh Working Papers in Cognitive Science), 191–228. Centre for Cognitive Science, University of Edinburgh.

Przepiórkowski, Adam. 1996b. Non-configurational case assignment in HPSG. Paper delivered at the 3rd International Conference on HPSG, 20–22 May 1996, Marseille, France.

Przepiórkowski, Adam. 1999a. *Case assignment and the complement-adjunct dichotomy: A non-configurational constraint-based approach*. Eberhard-Karls-Universität Tübingen dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46147, accessed 2018-2-25.

Przepiórkowski, Adam. 1999b. On case assignment and "adjuncts as complements". In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 231–245. Stanford, CA: CSLI Publications.

Przepiórkowski, Adam. 2016. How *not* to distinguish arguments from adjuncts in LFG. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 560–580. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2016/headlex2016-przepiorkowski.pdf.

Przepiórkowski, Adam & Agnieszka Patejuk. 2012. On case assignment and the coordination of unlikes: The limits of distributive features. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG'12 conference*, 479–489. Stanford, CA: Center for the Study of Language & Information. http://cslipublications.stanford.edu/LFG/17/lfg12.html.

Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 62(4). 751–773.

Reape, Mike. 1992. *A formal theory of word order: A case study in West Germanic*. University of Edinburgh dissertation.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Ryu, Byong-Rae. 2013. Multiple case marking as case copying: A unified approach to multiple nominative and accusative constructions in Korean. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 182–202. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2013/, accessed 2018-8-17.

Sag, Ivan A. 2003. Coordination and underspecification. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the HPSG 2002 conference*, 267–291. Stanford, CA: Center for the Study of Language & Information.

Sag, Ivan A., Lauri Karttunen & Jeffrey Goldberg. 1992. A lexical analysis of Icelandic case. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters* (CSLI Lecture Notes 24), 301–318. Stanford, CA: CSLI Publications.

Thuilier, Juliette. 2011. Case suffixes and pospositions in Hungarian. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 209–226. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2011/, accessed 2018-2-25.

Yatabe, Shûichi. 2004. A comprehensive theory of coordination of unlikes. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 335–355. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

Yatabe, Shûichi. 2012. Comparison of the ellipsis-based theory of non-constituent coordination with its alternatives. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*, 454–474. Stanford, CA: CSLI Publications.

Yatabe, Shûichi. 2016. Medial left-node raising in Japanese. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 681–701. Stanford, CA: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2016/headlex2016-yatabe.pdf.

Yatabe, Shûichi & Wai Lok Tam. 2018. In defense of an HPSG-based theory of non-constituent coordination: A reply to Kubota and Levine. http://ling.auf.

net/lingbuzz/003152, accessed 2018-7-29. Unpublished manuscript, version of April 16, 2018.

Yip, Moira, Joan Maling & Ray S. Jackendoff. 1987. Case in tiers. *Language* 63(2). 217–250. DOI:10.2307/415655

Yoo, Eun Jung. 1993. Subcategorization and case marking in Korean. In Andreas Kathol & Carl J. Pollard (eds.), *Papers in syntax* (OSU Working Papers in Linguistics 42), 178–198. Ohio State University: Department of Linguistics.

Zaenen, Annie & Joan Maling. 1983. Passive and oblique case. In Lori S. Levin, Malka Rappaport & Annie Zaenen (eds.), *Papers in Lexical-Functional Grammar*, 159–191. Bloomington, IN: Indiana University Linguistics Club.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3(4). 441–483.

Chapter 8

Nominal structures

Frank Van Eynde
University of Leuven

This chapter shows how nominal structures are treated in HPSG. The introduction puts the discussion in the broader framework of the NP vs. DP debate and differentiates three HPSG treatments, i.e. the specifier treatment, the DP treatment and the functor treatment. They are each presented in some detail and applied to the analysis of ordinary nominals, covering both syntactic and semantic aspects. A comparison reveals that the DP treatment does not mesh as well with the monostratal surface-oriented nature of the HPSG framework as the other treatments. Then it is shown how the specifier treatment and the functor treatment deal with nominals that have idiosyncratic properties, such as the Big Mess Construction, gerunds, nominalized infinitives and irregular P+NP combinations. Their analysis typically requires an appeal to phrasal constraints and a modicum of constructionalism.

1 Introduction

We use the term 'nominal' in a theory-neutral way as standing for a noun and its phrasal projection. In this broad sense all of the bracketed strings in (1) are nominals.

(1) [the [red [box]]] is empty

For the analysis of nominals there are two main approaches in generative grammar. One treats the noun (N) as the head all the way through. In that analysis the largest bracketed string in (1) is an NP. The other makes a distinction between the nominal core, consisting of a noun with its complements and modifiers, if any, and a functional outer layer, comprising determiners, quantifiers and numerals. In that analysis the noun is the head of *red box*, while the determiner is the head of *the red box*, so that the category of the latter is DP.

Frank Van Eynde. 2019. Nominal structures. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 185–223. Berlin: Language Science Press. DOI:??

The first approach, henceforth called the **NP approach**, prevailed in generative grammar up to and including the Government and Binding model (Chomsky 1981). One of its modules, the categorial component, consists of phrase structure rules, such as those in (2).

(2) a. VP → V NP
    b. NP → Det Nom

They are required to "meet some variety of X-bar theory" (Chomsky 1981: 5). The original variety is that of Chomsky (1970). It consists of the following cross-categorial rule schemata:

(3) a. X′ → X …
    b. X″ → [Spec, X′] X′

X′ stands for the combination of a head X and its complements, where X is N, A or V, and X″ stands for the combination of X′ and its specifier "where [Spec, N′] will be analyzed as the determiner" (Chomsky 1970: 52). X-bar theory was further developed in Jackendoff (1977), who added a schema for the addition of adjuncts and who extended the range of X with P, the category of adpositions. Generalized Phrase Structure Grammar developed a monostratal version of it, exemplified by the analysis in Figure 1, quoted from Gazdar et al. (1985: 126). The top node is the double-bar category N″, which consists of the determiner and the single-bar category N′. The AP and the relative clause are adjoined to N′, and the lowest N′ consists of the noun and its PP complement.

The second approach, henceforth called the **DP approach**, results from an extension of the range of X in (3) to the functional categories. This was motivated by the fact that some of the phrase structure rules, such as (4), do not fit the X-bar mould.

(4) S → NP Aux VP

To repair this, the category Aux, which contained both auxiliaries and inflectional verbal affixes (Chomsky 1957), was renamed as I(nfl) and treated as the head of S. More specifically, I(nfl) was claimed to combine with a VP complement, yielding I′, and the latter was claimed to combine with an NP specifier (the subject), yielding I″ (formerly S). For the analysis of nominals such an overhaul did not at first seem necessary, since the relevant PS rules did fit the X-bar mould, but it took place nonetheless, mainly in order to capture similarities between nominal and clausal structures. These are especially conspicuous in

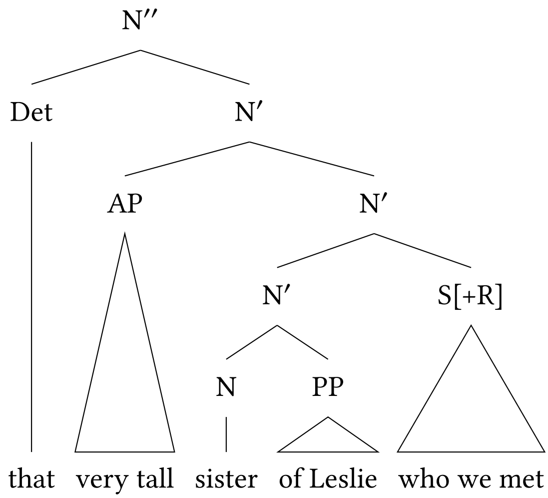


Figure 1: The NP approach

gerunds, nominalized infinitives and nominals with a deverbal head, and were seen as evidence for the claim that determiners have their own phrasal projection, just like the members of I(nfl) (Abney 1987). More specifically, members of D were claimed to take an N″ complement (formerly Nom), yielding D′, and the latter was claimed to have a specifier sister, as in Figure 2. The DP approach was also taken on board in other frameworks, such as Word Grammar (Hudson 1990) and Lexical Functional Grammar (Bresnan 2001: 99).

Turning now to Head-driven Phrase Structure Grammar we find three different treatments. The first and oldest can be characterized as a lexicalist version of the NP approach, more specifically of its monostratal formulation in GPSG. It is first proposed in Pollard & Sag (1987) and further developed in Pollard & Sag (1994) and Ginzburg & Sag (2000). We henceforth call it the **specifier treatment**, after the role which it assigns to the determiner. The second is a lexicalist version of the DP approach. It is first proposed in Netter (1994) and further developed in Netter (1996), and Nerbonne & Mullen (2000). We will call it the **DP treatment**. The third adopts the NP approach, but neutralizes the distinction between adjuncts and specifiers, treating them both as functors. It is first proposed in Van Eynde (1998) and Allegranza (1998) and further developed in Van Eynde (2003), Van Eynde (2006) and Allegranza (2006). It is also adopted in Sign-Based Construction Grammar (Sag 2012). We will call it the **functor treatment**. This chapter presents the three treatments and compares them wherever this seems

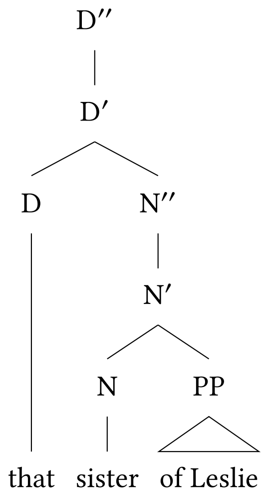


Figure 2: The DP approach

appropriate.

We first focus on ordinary nominals (Section 2) and then on nominals with idiosyncratic properties (Section 3). For exemplification we use English and a number of other Germanic and Romance languages, including Dutch, German, Italian and French. We assume familiarity with the typed feature structure notation and with such basic notions as unification, inheritance and token-identity.

2 Ordinary nominals

We use the term 'ordinary nominal' for a nominal that contains a noun, any number of complements and/or adjuncts and at most one determiner. This section shows how such nominals are analyzed in respectively the specifier treatment (2.1), the functor treatment (2.2) and the DP treatment (2.3).

2.1 The specifier treatment

The specifier treatment adopts the same distinction between heads, complements, specifiers and adjuncts as X-bar syntax, but the integration of these notions in a monostratal lexicalist framework inevitably leads to various differences. The presentation is mainly based on Pollard & Sag (1994) and Ginzburg & Sag (2000). We first discuss the syntactic structure (2.1.1) and the semantic composition (2.1.2) of nominals, and then turn to nominals with a phrasal specifier (2.1.3).

2.1.1 Syntactic structure

Continuing with the same example as in Figures 1 and 2, a relational noun, such as *sister*, selects a PP as its complement and a determiner as its specifier, as spelled out in its CATEGORY value (5).

(5) $\begin{bmatrix} \textit{category} & \\ \text{HEAD} & \textit{noun} \\ \text{SPR} & \langle \text{Det} \rangle \\ \text{COMPS} & \langle \text{PP}[\textit{of}\,] \rangle \end{bmatrix}$

The combination with a matching PP, as in *sister of Leslie*, is subsumed by the *head-comp(lement)s-phr* type, as defined in Borsley & Abeillé (2019), Chapter 1 of this volume, and yields a nominal with an empty COMPS list. Similarly, the combination of this nominal with a matching determiner, as in *that sister of Leslie*, is subsumed by the *head-sp(ecifie)r-phr* type, as defined in Borsley & Abeillé (2019), Chapter 1 of this volume. and yields a nominal with an empty SPR list. This is spelled out in Figure 3.

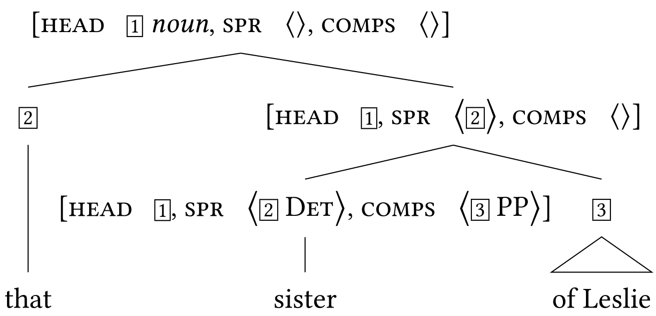


Figure 3: Adnominal complements and specifiers

Since the noun is the head of *sister of Leslie* and since the latter is the head of *that sister of Leslie*, the Head Feature Principle implies that the phrase as a whole shares the HEAD value of the noun ($\boxed{1}$). The valence features, COMPS and SPR, have a double role. On the one hand, they register the degree of saturation of the nominal; in this role they supersede the bar levels of X-bar theory. On the other hand, they capture co-occurrence restrictions, such as the fact that the complement of *sister* be a PP, rather than an NP or a clause, and that its specifier be a singular determiner, rather than a plural one.[1]

In contrast to complements and specifiers, adjuncts are not included in valence lists, since their addition has no effect on the degree of saturation. At the same time, since the combination of an adjunct with its head is subject to co-occurrence restrictions, one needs a way to capture those. For that purpose Pollard & Sag (1994: 55–57) employs the feature MOD(IFIED). It is part of the HEAD value of the substantive parts-of-speech, i.e. noun, verb, adjective and preposition. Its value is of type *synsem* in the case of modifying items and of type *none* otherwise.

(6) $\textit{substantive}\text{:}\ \begin{bmatrix}\textsc{mod} & \textit{synsem} \lor \textit{none}\end{bmatrix}$

As an example, let us take the number and gender agreement between nouns and adjectives in the Romance languages.[2] The Italian *grossa* 'big', for instance, is compatible with singular feminine nouns, such as *scatola* 'box', but not with plural feminine nouns, such as *scatole* 'boxes', nor with masculine nouns, such as *libro* 'book' or *libri* 'books'. This is made explicit in its MOD value, spelled out in (7).

(7) $$\begin{bmatrix}\textit{category} & \\ \textsc{head} & \begin{bmatrix}\textit{adjective} & \\ \textsc{mod}|\textsc{loc}|\textsc{cat} & \begin{bmatrix}\textsc{head} & \begin{bmatrix}\textit{noun} & \\ \textsc{number} & \textit{sing} \\ \textsc{gender} & \textit{fem}\end{bmatrix} \\ \textsc{spr} & \langle \textrm{D}\textsc{et} \rangle\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

The token-identity of the MOD(IFIED) value of the adjective with the SYNSEM value of its head sister is part of the definition of type *head-adj(unct)-phr*, as defined in Borsley & Abeillé (2019), Chapter 1 of this volume. Since the MOD(IFIED) feature is part of the HEAD value, it follows from the Head Feature Principle that it is shared between an adjective and the AP which it projects. As a consequence, the MOD(IFIED) value of *molto grossa* 'very big' is shared with that of *grossa* 'big'.

Besides its role in modeling agreement, the MOD value is also instrumental in capturing constraints on linear order. The fact, for instance, that the AP in *that very tall bridge* appears after the determiner and not before, as in **very tall*

[1] The value of the third valence feature (SUBJ) is invariably the empty list for nouns, except in the case of predicative nouns, whose SUBJ list contains an NP that is identified with the target of the predication, see Ginzburg & Sag (2000: 409).

[2] This is an instance of concord, as defined in Wechsler (2019), Chapter 6 of this volume.

that bridge, is captured by the stipulation that the MOD value of the adjective is a nominal with a determiner on its SPR list. This blocks the combination with a fully saturated NP, as in **very tall that bridge*.[3] Also here, the MOD value of the adjective is shared with that of the AP, as shown in Figure 4.

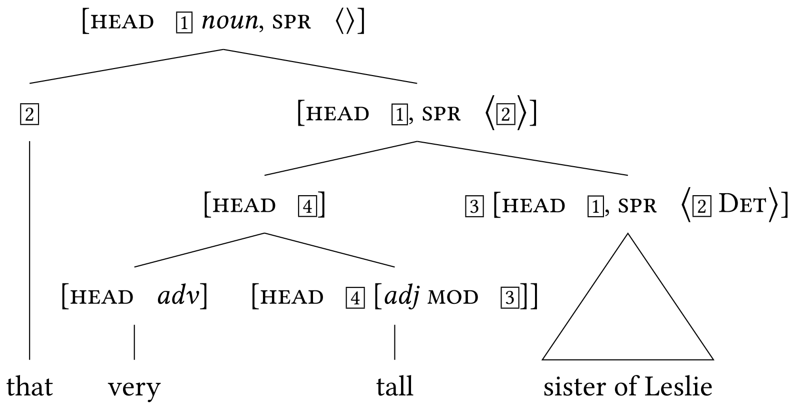


Figure 4: Adnominal modifiers

2.1.2 Semantic composition

Given the monostratal nature of HPSG, semantic representations do not constitute a separate level of representation, but take the form of attribute value pairs that are added to the syntactic representations. Phrase formation and semantic composition are, hence, modeled in tandem. Technically, the CONTENT feature is declared for the same type of objects as the CATEGORY feature.

(8) $local\text{:} \begin{bmatrix} \text{CAT} & \textit{category} \\ \text{CONTENT} & \textit{semantic-object} \end{bmatrix}$

For nominals the value of the CONTENT feature is of type *scope-obj(ect)* in Ginzburg & Sag (2000). A scope-object is an index-restriction pair in which the index stands for entities and in which the restriction is a set of facts which constrain the denotation of the index, as in the CONTENT value of the noun *box*:

(9) $\begin{bmatrix} \textit{scope-obj} & \\ \text{INDEX} & \boxed{1}\ \textit{index} \\ \text{RESTR} & \left\{ \begin{bmatrix} \textit{box} & \\ \text{ARG} & \boxed{1} \end{bmatrix} \right\} \end{bmatrix}$

[3]This order constraint is overruled in the Big Mess Construction, see Section 3.2.

This is comparable to the representations which are canonically used in Predicate Logic (PL), such as {x|*box*(x)}, where x stands for the entities that the predicate *box* applies to. In contrast to PL variables, HPSG indices are sorted with respect to person, number and gender. This provides the means to model the type of agreement that is called index agreement in Wechsler (2019), Chapter 6 of this volume.

(10) $$\textit{index}\text{:} \begin{bmatrix} \textsc{person} & \textit{person} \\ \textsc{number} & \textit{number} \\ \textsc{gender} & \textit{gender} \end{bmatrix}$$

CONTENT values of attributive adjectives are also of type *scope-object*. When combined with a noun, as in *red box*, the resulting representation is one in which the indices of the adjective and the noun are identical, as in (11).[4]

(11) $$\begin{bmatrix} \textit{scope-obj} & \\ \textsc{index} & \boxed{1} \\ \textsc{restr} & \left\{ \begin{bmatrix} \textit{red} & \\ \textsc{arg} & \boxed{1} \end{bmatrix}, \begin{bmatrix} \textit{box} & \\ \textsc{arg} & \boxed{1} \end{bmatrix} \right\} \end{bmatrix}$$

Also this is comparable to the PL practice of representing such combinations with one variable to which both predicates apply, as in {x|*red*(x) & *box*(x)}. What triggers the index sharing is the MOD(IFIED) value of the adjective, as illustrated by the AVM of *red* in (12) (Pollard & Sag 1994: 55).[5]

(12) $$\begin{bmatrix} \textsc{cat}|\textsc{head} & \begin{bmatrix} \textit{adjective} & \\ \textsc{mod}|\textsc{loc}|\textsc{content} & \begin{bmatrix} \textit{scope-obj} & \\ \textsc{index} & \boxed{1} \\ \textsc{restr} & \boxed{\Sigma} \end{bmatrix} \end{bmatrix} \\ \textsc{content} & \begin{bmatrix} \textsc{index} & \boxed{1} \\ \textsc{restr} & \left\{ \begin{bmatrix} \textit{red} & \\ \textsc{arg} & \boxed{1} \end{bmatrix} \right\} \cup \boxed{\Sigma} \end{bmatrix} \end{bmatrix}$$

The adjective selects a scope-object, shares its index and adds its own restriction to those that are already present. The resulting CONTENT value is then shared with the mother.

To model the semantic contribution of determiners, Ginzburg & Sag (2000) makes a distinction between scope-objects that contain a quantifier (*quant-rel*),

[4] This is an example of intersective modification. The semantic contribution of other types of adjectives, such as *alleged* and *fake*, are modeled differently (Pollard & Sag 1994: 330–331).

[5] $\boxed{\Sigma}$ stands for an object of type *set*, as in Ginzburg & Sag (2000).

and those that do not (*parameter*). In terms of this distinction, the addition of a determiner to a nominal, as in *every red box*, triggers a shift from *parameter* to *quant-rel*. To capture this the specifier treatment employs the feature SPEC(IFIED). It is part of the HEAD value of the determiners, and its value is of type *semantic-obj(ect)*.[6]

(13) *determiner*: $\begin{bmatrix}\text{SPEC} & \textit{semantic-obj}\end{bmatrix}$

In the case of *every*, the SPEC value is an object of type *parameter*, but its own CONTENT value is a subtype of *quant-rel* and this quantifier is put in store, to be retrieved at the place where its scope is determined, as illustrated by the AVM of *every* in (14) (Ginzburg & Sag 2000: 204).

(14)
$$\begin{bmatrix} \text{CAT}|\text{HEAD} & \begin{bmatrix} \textit{det} & \\ \text{SPEC} & \begin{bmatrix} \textit{parameter} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \boxed{\Sigma} \end{bmatrix} \end{bmatrix} \\ \text{CONTENT} & \boxed{2} \begin{bmatrix} \textit{every-rel} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \boxed{\Sigma} \end{bmatrix} \\ \text{STORE} & \{\boxed{2}\} \end{bmatrix}$$

Notice that the addition of the SPEC feature yields an analysis in which the determiner and the nominal select each other: The nominal selects its specifier by means of the valence feature SPR and the determiner selects the nominal by means of SPEC.

2.1.3 Nominals with a phrasal specifier

Specifiers of nominals tend to be single words, but they can also take the form of a phrase. The bracketed phrase in [*the Queen of England's*] *sister*, for instance, is in complementary distribution with the possessive determiner in *her sister* and has a comparable semantic contribution. For this reason it is treated along the same lines. More specifically, the possessive marker *'s* is treated as a determiner that takes an NP as its specifier, as shown in Figure 5 (Pollard & Sag 1994: 51–54) and (Ginzburg & Sag 2000: 193).[7]

[6] In Pollard & Sag (1994: 45) the SPEC(IFIED) feature was also assigned to other function words, such as complementizers, and its value was of type *synsem*.

[7] The treatment of the phonologically reduced *'s* as the head of a phrase is comparable to the treatment of the homophonous word in *he's ill* as the head of a VP. Notice that the possessive *'s* is not a genitive affix, for if it were, it would be affixed to the head noun *Queen*, as in **the Queen's of England sister*, see Sag et al. (2003: 199).

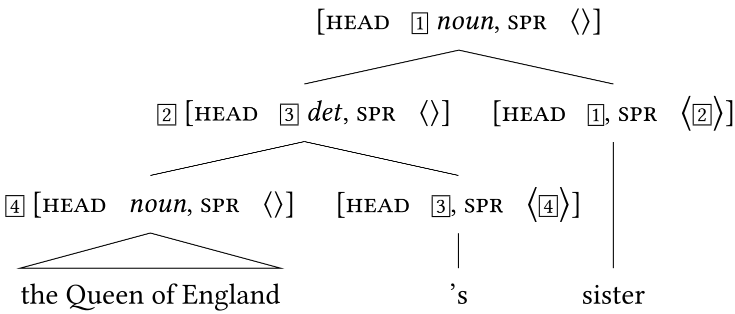


Figure 5: Phrasal specifiers

In this anaysis the specifier of *sister* is a DetP that is headed by *’s* and the latter takes the NP *The Queen of England* as its specifier.[8] Semantically, *’s* relates the index of its specifier (the possessor) to the index of the nominal that it selects (the possessed), as spelled out in (15).[9]

(15)
$$\begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{det} & \\ \text{SPEC} & \begin{bmatrix}\textit{parameter} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \boxed{\Sigma}\end{bmatrix}\end{bmatrix} \\ \text{SPR} & \left\langle \begin{bmatrix}\text{INDEX} & \boxed{3}\end{bmatrix}\right\rangle \end{bmatrix} \\ \text{CONTENT} & \boxed{2} \begin{bmatrix} \textit{the-rel} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \left\{ \begin{bmatrix}\textit{poss-rel} & \\ \text{POSSESSOR} & \boxed{3} \\ \text{POSSESSED} & \boxed{1}\end{bmatrix}\right\} \cup \boxed{\Sigma} \end{bmatrix} \\ \text{STORE} & \{\boxed{2}\} \end{bmatrix}$$

The assignment of *the-rel* as the CONTENT value captures the definiteness of the resulting NP. Notice that this analysis contains a DetP, but in spite of that, it is not an instance of the DP approach, since the determiner does not head the nominal, but only its specifier.

[8] Since the specifier of *’s* is an NP, it may in turn contain a specifier that is headed by *’s*, as in *John’s uncle’s car.*

[9] The terms ‘possessor’ and ‘possessed’ are meant to be understood in a broad not-too-literal sense.

2.2 The functor treatment

The functor treatment also adopts the NP approach, but –in contrast to the specifier treatment– it does not draw a distinction between specifiers and adjuncts, nor between lexical (or substantive) and functional categories.[10] The presentation in this section is mainly based on Van Eynde (2006) and Allegranza (2006). We first introduce the basics (Section 2.2.1) and then turn to nominals with a phrasal specifier (Section 2.2.2) and to nominals without specifier (Section 2.2.3).

2.2.1 Basics

The distinction between specifiers and adjuncts is usually motivated by the assumption that the former are obligatory and non-stackable, while the latter are optional and stackable. In practice, though, this distinction is blurred by the fact that many nominals are well-formed without specifier. To accommodate this one can employ phonetically empty determiners, optionality in the SPR list or a non-branching phrase type, but the functor treatment provides a more radical response and abandons the distinction between specifiers and adjuncts. Technically, this implies that the SPR feature is dropped.[11]

Besides, the functor treatment abandons the distinction between lexical and functional categories. The words which are commonly treated as determiners are not treated as members of a separate functional category 'Det', but are claimed to belong to independently needed lexical categories, such as Adjective, Adverb, Pronoun and Noun. The argumentation is mainly –but not only– based on matters of NP-internal agreement and inflectional variation (Van Eynde 2006). In Dutch, for instance, prenominal adjectives show agreement with the nouns they modify: They take the affix *-e* in combination with plural and singular non-neuter nominals, but not in combination with singular neuter nominals.[12]

(16) zwarte muren, zwarte verf, zwart zand
black wall.PL, black paint.SG.FEM, black sand.SG.NEU

[10] The term *functor* is also used in Categorial (Unification) Grammar, where it has a very broad meaning, subsuming the nonhead daughter in combinations of a head with a specifier or an adjunct, and the head daughter otherwise, see Bouma (1988). This broad notion is also adopted in Reape (1994). We adopt a more restrictive version in which functors are nonhead daughters which lexically select their head sister.

[11] Intriguingly, Chomsky has recently argued that there is no need for the notion of specifier in Transformational Grammar either, see Chomsky (2013: 43).

[12] If the adjective is preceded by a definite determiner, it also takes the affix in singular neuter nominals. This phenomenon is treated in Section 2.2.2.

The same holds for several of the words that are claimed to be determiners, such as the possessive *ons* 'our' and the interrogative *welk* 'which'.

(17) onze ouders, onze muur, ons huis
our parent.PL, our wall.SG.MAS, our house.SG.NEU

(18) welke boeken, welke man, welk boek
which book.PL, which man.SG.MAS, which book.SG.NEU

By contrast, prenominal (pro)nouns do not show NP-internal agreement and never take the *-e* affix. This not only holds for genitive nouns, such as *Jans*, but also for several of the words that are claimed to be determiners, such as the interrogative *wiens* 'whose' and the quantifying *wat* 'some'.

(19) Jans ouders, Jans fiets, Jans huis
Jan.GEN parent.PL, Jan.GEN bike.SG.MAS, Jan.GEN house.SG.NEU

(20) wiens ouders, wiens muur, wiens huis
whose parent.PL, whose wall.SG.MAS, whose house.SG.NEU

(21) wat boeken, wat verf, wat zand
some book.PL, some paint.SG.FEM, some sand.SG.NEU

To model this the functor treatment assigns adjectival status to determiners like *ons* 'our' and *welk* 'which', and pronominal status to determiners like *wiens* 'whose' and *wat* 'some'. This distinction is also relevant for other languages. The Italian possessives of the first and second person, for instance, show the same alternation for number and gender as the adjectives and are subject to the same constraints on NP-internal agreement.

(22) il nostro futuro, la nostra scuola, i nostri
the our future.SG.MAS, the our school.SG.FEM, the our
genitori, le nostre scatole
parent.PL.MAS, the our box.PL.FEM

The possessive of the third person plural, by contrast, does not show any inflectional variation and does not show agreement with the noun.[13]

(23) il loro futuro, la loro scuola, i loro genitori,
the their future.SG.MAS, the their school.SG.FEM, the their parent.PL.MAS,
le loro scatole
the their box.PL.FEM

[13] Confirming evidence for the pronominal status of *loro* is provided by the fact that it is also used as a personal pronoun, as in *l'ho dato a loro* 'I gave it to them'.

An example from French is provided in Abeillé et al. (2004), which assigns adverbial status to the specifier in *beaucoup de farine* 'much flour'. Technically, the elimination of the distinction between lexical and functional categories implies that there is no longer any need for drawing a distinction between the selection features MOD(IFIED) and SPEC(IFIED).

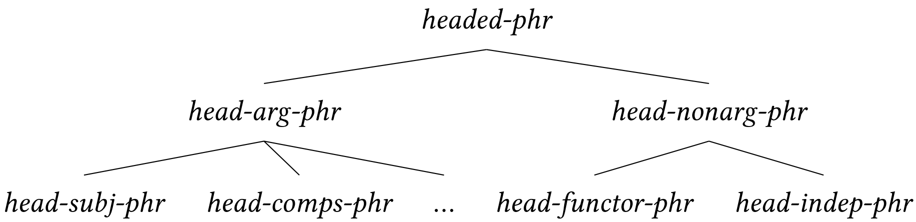


Figure 6: Hierarchy of headed phrases

To spell out the functor treatment we start from the hierarchy of headed phrases in Figure 6. The basic distinction is that between *head-argument-phr* and *head-nonargument-phr*. In the former the head daughter selects its non-head sister(s) by means of valence features, such as COMPS and SUBJ (but not SPR!), and it is their values that register the degree of saturation of the phrase, as shown for COMPS in Section 2.1.1. In head-nonargument phrases the degree of saturation is registered by the MARK(ING) feature. It is declared for objects of type *category*, along with the HEAD and valence features.[14] Its value is shared with the head daughter in head-argument phrases and with the non-head daughter in head-nonargument phrases, as spelled out in (24) and (25) respectively.

(24)
$$\begin{bmatrix} \textit{head-arg-phr} \\ \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{MARK} \;\; \boxed{1}\; \textit{marking} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{MARK} \;\; \boxed{1} \end{bmatrix}$$

(25)
$$\begin{bmatrix} \textit{head-nonarg-phr} \\ \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{MARK} \;\; \boxed{1}\; \textit{marking} \\ \text{DTRS} \;\; \left\langle \begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{MARK} \;\; \boxed{1} \end{bmatrix}, \boxed{2} \right\rangle \\ \text{HEAD-DTR} \;\; \boxed{2}\; \textit{sign} \end{bmatrix}$$

Besides, instead of MOD and SPEC, the functor treatment employs one feature to model the selection of a head by its non-head sister. It is called SEL(ECT), and its value is an object of type *synsem* that is shared with the head daughter, as spelled out in (26).

[14] The MARKING feature is introduced in Pollard & Sag (1994: 46) to model the combination of a complementizer and a clause.

(26) $$\begin{bmatrix} \textit{head-functor-phr} \\ \text{DTRS} \quad \left\langle \left[\text{SYNSEM|LOC|CAT|HEAD|SEL} \quad \boxed{1}\right], \boxed{2}\ \textit{sign} \right\rangle \\ \text{HEAD-DTR} \quad \boxed{2}\left[\text{SYNSEM} \quad \boxed{1}\ \textit{synsem}\right] \end{bmatrix}$$

This is a subtype of *head-nonargument-phr*. It subsumes the phrases in which the non-head daughter selects its head sister. As such it contrasts with phrases of type *head-independent-phr*, whose defining characteristic is that the nonhead daughter does not select its head sister: Its SEL value is hence *none*.[15]

(27) $$\begin{bmatrix} \textit{head-independent-phr} \\ \text{DTRS} \quad \left\langle \left[\text{SYNSEM|LOC|CAT|HEAD|SEL} \quad \textit{none}\right], \boxed{1} \right\rangle \\ \text{HEAD-DTR} \quad \boxed{1}\ \textit{sign} \end{bmatrix}$$

An illustration of the functor treatment is given in Figure 7. The combination of the noun with the adjective is an instance of *head-functor-phr*, in which the adjectival functor selects an unmarked nominal ($\boxed{3}$), shares its MARKING value ($\boxed{5}$), and, being a non-argument, shares it with the mother as well. The combination of the resulting nominal with the demonstrative is also an instance of *head-functor-phr*, in which the demonstrative functor selects an unmarked nominal ($\boxed{4}$), but –differently from the adjective– its MARKING value is of type *marked*, and this value is shared with the mother ($\boxed{2}$). This accounts for the fact that adjectives can be stacked, while determiners cannot. It also accounts for the ill-formedness of combinations like **long the bridge*, since attributive adjectives are not compatible with a marked nominal. Notice that the demonstrative does not belong to a functional part of speech, such as Det. Instead, it is claimed to be a pronoun. The only difference between adnominal *that* and nominal *that* concerns the SEL(ECT) value: While the former selects an unmarked singular nominal, the latter does not select anything.

Whether an adnominal functor is marked or unmarked is subject to cross-linguistic variation. The Italian possessives, for instance, are unmarked and can, hence, be preceded by an article, as in *il suo cane* 'the his dog' and *un mio amico* 'a my friend', while the French possessives are marked: *(*le) mon chien* '(*the) my dog' and *(*un) mon ami* '(*a) my friend'.

The treatment of postnominal dependents is similar to that in the specifier treatment. A relational noun like *sister*, for instance, takes an optional PP[*of*] complement. Similarly, a deverbal noun like *description* takes two optional PP complements. If realized, as in *a description of the Hungarian NP by Ivan*, their

[15]This type is introduced in Van Eynde (1998: 130). It will be used in Section 3 to deal with idiosyncratic nominals.

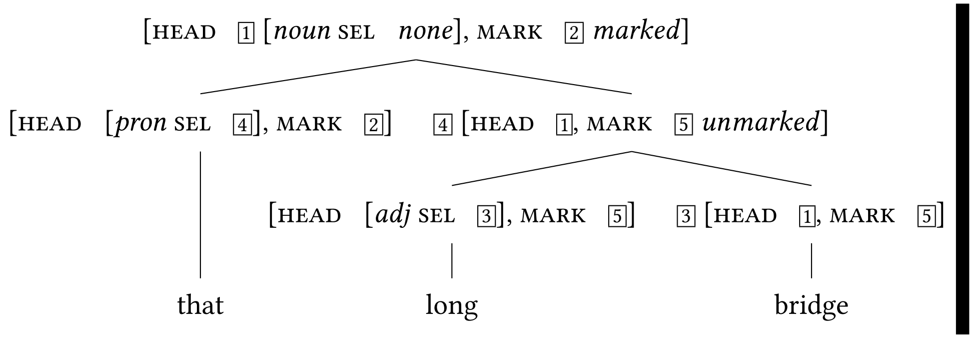


Figure 7: Adnominal functors

indices are identified with those of the second and the first argument respectively. There is a difference, though, in the treatment of nominals whose first argument is realized as a possessive, as in *his description of the Hungarian NP*. In this case, the first argument is not put on the SPR list of the noun. Instead, the possessive selects an unmarked nominal with an optional PP[*by*] complement on its COMPS list, and shares the index of that PP, as spelled out in Figure 8.

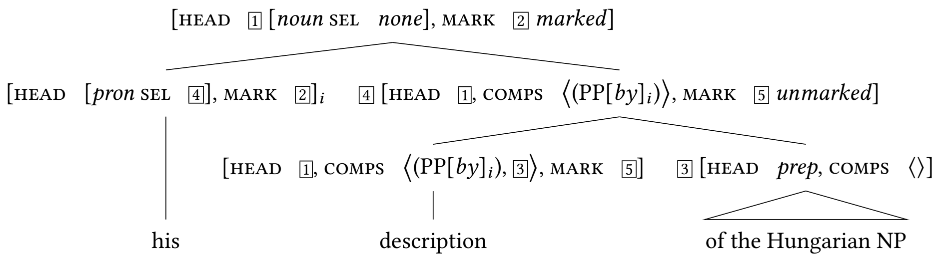


Figure 8: Deverbal nominals

2.2.2 Nominals with a phrasal functor

To illustrate how the functor treatment deals with phrasal specifiers we take the nominal *a hundred pages*. It has a left branching structure in which the indefinite article selects the unmarked singular noun *hundred* –its plural counterpart is *hundreds*– and in which the resulting NP selects the unmarked plural noun *pages*, as spelled out in Figure 9. The MARKING value of the article is shared with its mother ($\boxed{1}$), yielding a marked NP, and since the latter is a functor it is also shared with the NP as a whole ($\boxed{1}$). This accounts for the fact that it cannot be preceded by a determiner, as in **those a hundred pages*. By contrast, if the place of the article is taken by a numeral, as in *two hundred pages*, the addition of a determiner is possible, as in *those two hundred pages*, since the numeral is unmarked. Notice that the HEAD value of the entire NP is identified with that of *pages*. This accounts for the fact that it is plural.

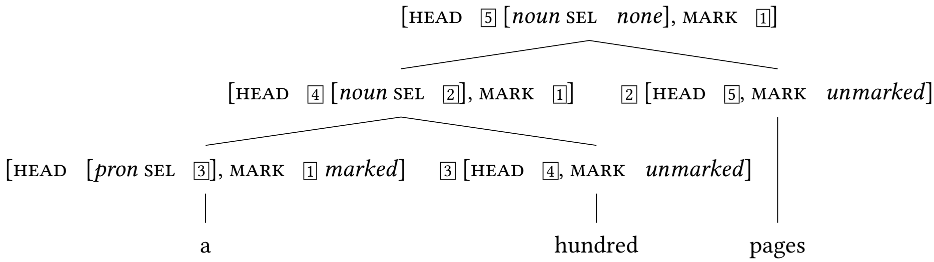


Figure 9: Phrasal functors

2.2.3 The hierarchy of MARKING values

Nominals without specifier have a MARKING value of type *unmarked*. Whether they need to be marked, depends on a multitude of factors and is largely language specific. Latin, for instance, does not require its nominals to contain a specifier, also if they are singular and count. Yet, there are cases in which an unmarked nominal is incomplete. To illustrate this let us take the attributive adjectives of Dutch again. As already pointed out in Section 2.2.1, they take the form without affix in singular neuter nominals, as in *zwart zand* 'black sand'. A complication, though, is that they canonically take the form with the affix if the nominal is introduced by a definite determiner, as in *het zwarte zand* 'the black sand'. This has consequences for the status of nominals with a singular neuter head: *zwart zand* and **zwarte zand*, for instance, are both unmarked, but while the former is well-formed as it is, the latter is only well-formed if it is preceded by a definite determiner. To model this Van Eynde (2006) differentiates between two types of *unmarked* nominals, as shown in Figure 10.

Employing the more specific subtypes, the adjectives without affix which select a singular neuter nominal have the value *bare*, while their declined counterparts which select a singular neuter nominal have the value *incomplete*. Since this MARKING value is shared with the mother, the MARKING value of *zwart zand* is *bare*, while that of **zwarte zand* is *incomplete*. The fact that the latter must be preceded by a definite determiner is modeled in the SEL(ECT) value of the deter-

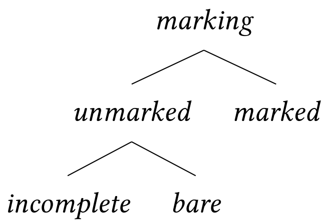


Figure 10: Hierarchy of MARKING values

miner: While definite determiners select an unmarked nominal, which implies that they are compatible with both bare and incomplete nominals, non-definite determiners select a bare nominal and are, hence, not compatible with an incomplete one, as in **een zwarte huis* 'a black house'. The MARKING feature is, hence, useful to differentiate bare nominals from incomplete nominals. In a similar way, one can make finer-grained distinctions in the hierarchy of *marked* values to capture co-occurrence restrictions between determiners and nominals, as in the functor treatment of the Italian determiner system of Allegranza (2006). See also the treatment of nominals with idiosyncratic properties in Section 3.

2.3 The DP treatment

In contrast to the specifier treatment and the functor treatment, the DP treatment identifies the determiner as the head of an NP. The presentation in this section is mainly based on Netter (1994). We first sketch its main characteristics (2.3.1) and then list some problems for it (2.3.2).

2.3.1 Two types of complementation

As in the transformational DP treatment, nominal projections are divided in a nominal core, consisting of a noun and its complements and/or adjuncts, on the one hand, and an external functional layer, comprising the determiner, on the other hand. The former is analyzed in much the same way as in the specifier treatment. The relational noun *sister* in Figure 11, for instance, is combined with its PP complement in the usual way. The addition of the determiner, by contrast, is modeled differently. It is not the nominal which selects its specifier by means of the valence feature SPR; instead, it is the determiner which selects a nominal by means of the valence feature COMPS, yielding a DP with an empty COMPS list. To account for the fact that this DP inherits many of its properties from its

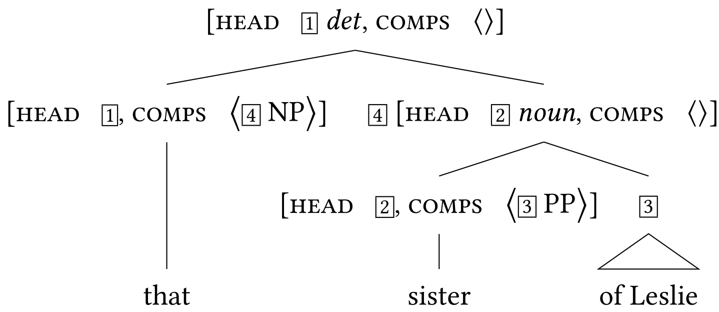


Figure 11: Lexicalist DP treatment

nominal non-head daughter, Netter makes a distinction between functional complementation and ordinary complementation, and differentiates between major and minor HEAD features:

(28) *part-of-speech*:
$$\begin{bmatrix} \text{MAJOR} & \begin{bmatrix} \text{N} & \textit{boolean} \\ \text{V} & \textit{boolean} \end{bmatrix} \\ \text{MINOR} & \begin{bmatrix} \text{FCOMPL} & \textit{boolean} \end{bmatrix} \end{bmatrix}$$

The MAJOR attribute includes the boolean features N and V, where nouns are [+N, −V], adjectives [+N, +V], verbs [−N, +V] and prepositions [−N, −V]. Besides, [+N] categories also have the features CASE, NUMBER and GENDER. Typical of functional complementation is that the functional head shares the MAJOR value of its complement (Netter 1994: 311–312).

(29) In a lexical category of type *func-cat* the value of its MAJOR attribute is token identical with the MAJOR value of its complement.

Since a determiner shares the MAJOR value of its nominal complement and since this value is also shared with the DP (given the Head Feature Principle), it follows that the resulting DP is [+N,−V] and that its CASE, NUMBER and GENDER values are identical to those of its nominal non-head daughter. This differentiates functional complementation from ordinary complementation, where a head and its complement do not share their MAJOR value. The noun *sister* in Figure 11, for instance, does not share the part-of-speech of its PP complement.

The MINOR attribute is used to model properties which a functional head does *not* share with its complement. It includes FCOMPL, a feature which registers whether a projection is functionally complete or not. Its value is positive for determiners, negative for singular count nouns and underspecified for plurals and mass nouns. Determiners take a nominal complement with a negative FCOMPL value, but their own FCOMPL value is positive and since they are the head, they share this value with the mother, as in Figure 12. In this analysis, a nominal is complete, if it is both saturated (empty COMPS list) and functionally complete (positive FCOMPL).

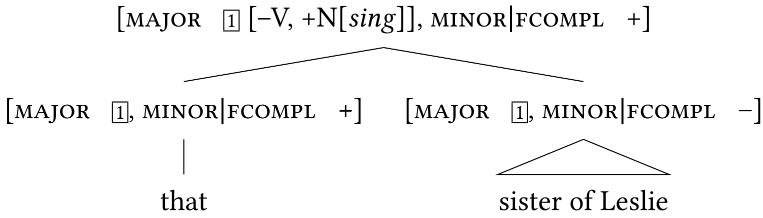


Figure 12: Functional completeness

2.3.2 Two problems for the DP treatment

A problem for the DP treatment concerns the notion of functional complementation, as defined in (29). If determiners share the MAJOR value of the nominals which they select, then it follows that they are nominal themselves, i.e. [+N, −V]. However, while this makes sense for determiners with (pro)nominal properties, such as the English *that*, the Dutch *wat* 'some' and the Italian *loro* 'their', it is rather implausible for determiners with adjectival properties, such as the Italian *mio* 'my', the Dutch *ons* 'our' and the German *welch* 'which'. The problem also affects the associated agreement features, i.e. CASE, NUMBER and GENDER. If a determiner is required to share the values of these features with its nominal complement, as spelled out in (29), then we get implausible results for nominals in which the determiner and the noun do not show agreement. In the Dutch *wiens huizen* 'whose houses' and *'s lands hoogste bergen* 'the highest mountains of the country', for instance, the selected nominals are plural and non-genitive, and so are the entire DPs, but the selecting determiner (phrase) is singular and genitive.

Another problem concerns the assumption "that all substantive categories will require the complement they combine with to be both saturated and functionally complete" (Netter 1994: 311). Complements of verbs and prepositions must, hence, be positively specified for FCOMPL. This, however, is contradicted by the existence of prepositions which require their complement to be functionally incomplete. The German prepositions *am, im, vom, beim* and *zum*, for instance, which diachronically result from the incorporation of a dative form of the definite article (*dem*) in respectively *an* 'at', *in* 'in', *von* 'of', *bei* 'at' and *zu* 'to', count as syntactic atoms in HPSG, just like the forms without incorporated article, but in contrast to the latter they require their complement to lack a determiner: While *in dem/diesem Zimmer* 'in the/this room' and *im Zimmer* 'in.the room' are both well-formed, **im dem/diesem Zimmer* 'in.the the/this room' is not.[16] Moreover, there are prepositions which do not contain an incorporated article, but nonetheless require a determinerless nominal. The Dutch *te* and *per*, for instance, are not compatible with nominals that contain a determiner, even if the nominal is singular and count, as in *te (*het/een) paard* 'on horse' and *per (*de/een) trein* 'by train'. This must be due to a lexical property of these prepositions, since most other prepositions require such nominals to have a determiner, as in *ze viel van *(het/een) paard* 'she fell from *(the/a) horse' and *ze zit op *(de/een) trein naar Londen* 'she is on *(the/a) train to London'. The requirement that complements of prepositions must be functionally complete is, hence, untenable.

2.4 Conclusion

This section has presented the three main treatments of nominal structure in HPSG. Two adopt the NP approach and one the DP approach. Overall, the former turn out to be more amenable to integration in a monostratal surface-oriented framework like HPSG than the latter. See also Müller (2018). Of the two NP treatments the specifier treatment is closer to early versions of $\overline{\mathrm{X}}$ syntax and to GPSG. The functor treatment is closer to versions of Categorial (Unification) Grammar, and has also been adopted in Sign-Based Construction Grammar (Sag 2012).

3 Idiosyncratic nominals

This section focusses on the analysis of nominals with idiosyncratic properties. Since their analysis often requires a relaxation of the strictly lexicalist approach

[16] A similar phenomenon exists in Italian, where the prepositions *a* 'at', *da* 'from', *su* 'on', *di* 'of' and *in* 'in' have alternate forms with an incorporated form of the definite article, as in *al, dal, sul, del* and *nel*, each with feminine and plural counterparts (*alla, agli, alle, dalla, dagli, dalle,* …). The nominals which combine with such forms may not be introduced by a determiner: *di/*della questa scatola* 'of/*of.the this box' and *in/*nel ogni palazzo* 'in/*in.the every palace'. The same holds m.m. for the French portemanteau forms *du, des, au* and *aux*.

of early HPSG we first introduce some basic notions of constructional HPSG.

The lexicalist approach of early HPSG can be characterized as one in which the properties of phrases are mainly determined by properties of the constituent words and only to a small extent by properties of the combinatory operations as such. Pollard & Sag (1994: 391), for instance, employs no more than seven types of combinatory operations, including those which were exemplified in Section 2.1, i.e. head-specifier, head-complements and head-adjunct.[17] Over time, though, this radical lexicalism gave way to an approach in which the properties of the combinatory operations play a larger role. The small inventory of highly abstract phrase types got replaced by a more fine-grained hierarchy in which the types contain more specific and –if need be– idiosyncratic constraints. This development started in Sag (1997), was elaborated in Ginzburg & Sag (2000), and gained momentum afterward, leading to what is now known as constructional HPSG. Characteristic of this approach is the use of a bi-dimensional hierarchy of phrasal signs. In such a hierarchy the phrases are not only partitioned in terms of HEADEDNESS, but also in terms of a second dimension, called CLAUSALITY, as in Figure 13.

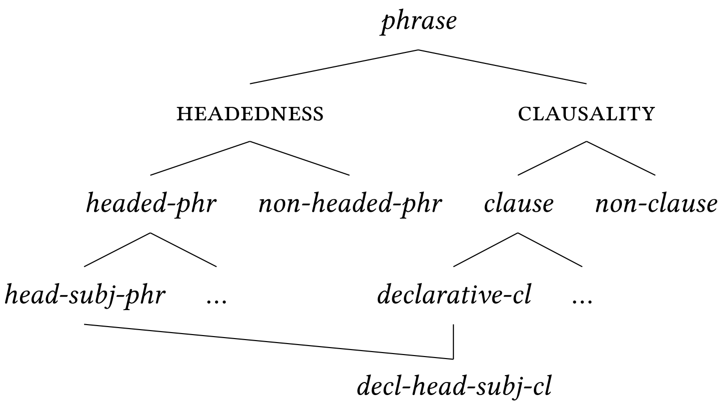


Figure 13: Bi-dimensional hierarchy of clauses

The types in the CLAUSALITY dimension are associated with constraints, in much the same way as the types in the HEADEDNESS dimension. Clauses, for instance, are required to denote an object of type *message* (Ginzburg & Sag 2000: 41).

[17] The remaining four are head-subject, head-subject-complements, head-marker and head-filler.

(30) $\begin{bmatrix} \textit{clause} \\ \textsc{synsem}|\textsc{loc}|\textsc{content} \quad \textit{message} \end{bmatrix}$

At a finer-grained level, the clauses are partitioned into declarative, interrogative, imperative, exclamative and relative clauses, each with their own constraints. Interrogative clauses, for instance, denote a question, while indicative declarative clauses denote a proposition.

Exploiting the possibilities of multiple inheritance one can define types which inherit properties from more than one supertype. The type *decl(arative)-head-subj(ect)-cl(ause)*, for instance, inherits the properties of *head-subject-phr*, on the one hand, and *declarative-cl*, on the other hand. Besides, it may have properties of its own, such as the fact that its head daughter is a finite non-inverted verb (Ginzburg & Sag 2000: 43). This combination of multiple inheritance and specific constraints on maximal phrase types is also useful for the analysis of nominals with idiosyncratic properties, as will be shown presently.

3.1 The Big Mess Construction

In ordinary nominals determiners precede attributive adjectives. Changing the order yields ill-formed combinations, such as **long that bridge* and **very tall every man*. However, this otherwise illegitimate order is precisely what we find in the Big Mess Construction (BMC), a term coined by Berman (1974).

(31) a. It's *so good a bargain* I can't resist buying it.
b. *How serious a problem* is this?

The idiosyncratic order in (31) is required if the nominal is introduced by the indefinite article, and if the preceding AP is introduced by one of a small set of degree markers, including *so, as, how, this, that* and *too*.

A specifier treatment of the BMC is provided in Ginzburg & Sag (2000: 201). It adopts a left branching structure, as in [[[*so good*] *a*] *bargain*], in which *so good* is the specifier of the indefinite article and in which *so good a* is the specifier of *bargain*. This is comparable to the treatment of the possessive in [[[*the Queen of England*] *'s*] *sister*], see Section 2.1.3. However, while there is evidence that *the Queen of England's* is a constituent, since it may occur independently as in *this crown is the Queen of England's*, there is no evidence that *so good a* is a constituent. Instead, there is evidence that the article forms a constituent with the following noun, since it also precedes the noun when the AP is in postnominal position, as in *we never had a bargain so good as this one*.

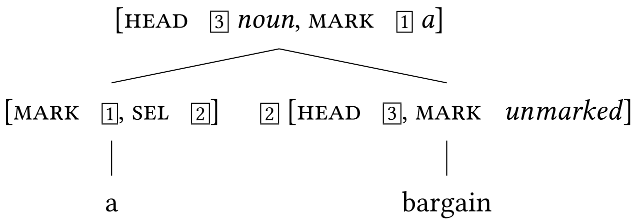


Figure 14: The lower NP

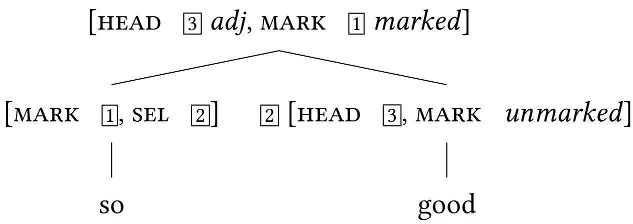


Figure 15: The marked AP

A structure in which the article sides with the noun, as in [[*so good*] [*a bargain*]], is adopted in Van Eynde (2007), Kim & Sells (2011), Kay & Sag (2012), Arnold & Sadler (2014) and Van Eynde (2018), all of which are functor treatments. The structure of the lower NP is spelled out in Figure 14. The article has a MARKING value of type *a* which is a subtype of *marked* and which it shares with the mother.[18] The AP is also treated as an instance of the head-functor type in Van Eynde (2007), Kim & Sells (2011) and Van Eynde (2018). The adverb has a MARKING value of type *marked*, so that the AP is marked as well, as shown in Figure 15. In combination with the fact that the article selects an unmarked nominal, this accounts for the ill-formedness of (32).

(32) a. * It's a so good bargain I can't resist buying it.
b. * A how serious problem is it?

By contrast, adverbs like *very* and *extremely* are unmarked, so that the APs which they introduce are admissible in this position, as in (33).

(33) a. This is a very serious problem.

[18] The MARKING value of the article looks similar to its PHONOLOGY value, but it is not the same. The PHONOLOGY values of *a* and *an*, for instance, are different, but their MARKING value is not.

b. We struck an extremely good bargain.

To model the combination of the AP with the lower NP it may at first seem plausible to treat the AP as a functor which selects an NP that is introduced by the indefinite article. This, however, has unwanted consequences: Given that SEL(ECT) is a HEAD feature, its value is shared between the AP and the adjective, so that the latter has the same SEL(ECT) value as the AP, erroneously licensing such combinations as **good a bargain*. To avoid this Van Eynde (2018) exploits the possibilities of the bi-dimensional hierarchy of phrase types in Figure 16.

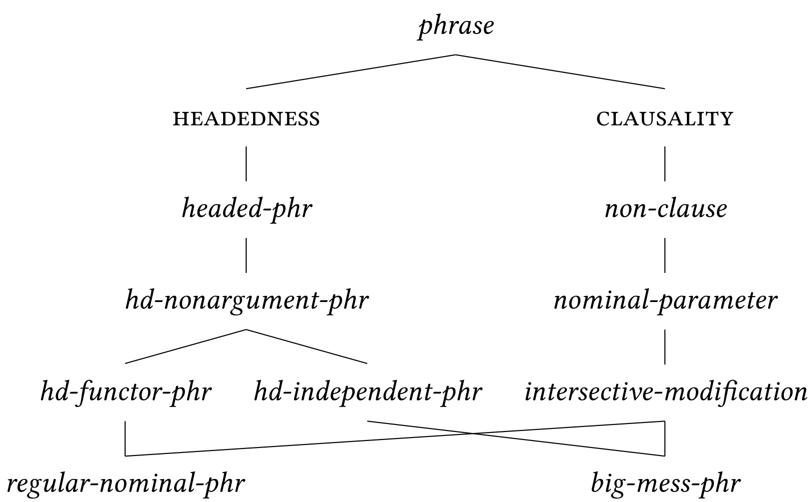


Figure 16: Bi-dimensional hierarchy of nominals

The types in the HEADEDNESS dimension are a subset of those in Figure 6. The types in the CLAUSALITY dimension mainly capture semantic and category-specific properties, in analogy with the hierarchy of clausal phrases in Ginzburg & Sag (2000). One of the non-clausal phrase types is *nominal-parameter*:

(34)
$$\begin{bmatrix} \textit{nominal-parameter} & \\ \textsc{synsem|loc} & \begin{bmatrix} \textsc{cat|head} & \textit{noun} \\ \textsc{content} & \begin{bmatrix} \textit{parameter} & \\ \textsc{index} & \boxed{1}\ \textit{index} \\ \textsc{restr} & \boxed{\Sigma_1} \cup \boxed{\Sigma_2} \end{bmatrix} \end{bmatrix} \\ \textsc{dtrs} & \left\langle \begin{bmatrix} \textsc{synsem|loc|content|restr} & \boxed{\Sigma_1} \end{bmatrix}, \boxed{2} \right\rangle \\ \textsc{head-dtr} & \boxed{2} \begin{bmatrix} \textsc{synsem|loc|content} & \begin{bmatrix} \textit{parameter} & \\ \textsc{index} & \boxed{1} \\ \textsc{restr} & \boxed{\Sigma_2} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The mother shares its index with the head daughter ($\boxed{1}$) and its RESTR(ICTION) value is the union of the RESTR values of the daughters ($\boxed{\Sigma_1}$ and $\boxed{\Sigma_2}$). In the hierarchy of non-clausal phrases, this type contrasts amongst others with the quantified nominals, which have a CONTENT value of type *quant-rel* (Ginzburg & Sag 2000: 203–205). A subtype of *nominal-parameter* is *intersective-modification*, as defined in (35).

(35)
$$\begin{bmatrix} \textit{intersective-modification} & \\ \textsc{synsem|loc|content|index} & \boxed{1}\ \textit{index} \\ \textsc{dtrs} & \left\langle \begin{bmatrix} \textsc{synsem|loc|content|index} & \boxed{1} \end{bmatrix}, \boxed{2} \right\rangle \\ \textsc{head-dtr} & \boxed{2}\ \textit{sign} \end{bmatrix}$$

This constraint requires the mother to share its index also with the non-head daughter. It captures the intuition that the noun and its non-head sister apply to the same entities, as in the case of *red box*.[19]

The maximally specific types inherit properties of one of the types of headed phrases, on the one hand, and of one of the non-clausal phrase types, on the other hand. Regular nominal phrases, for instance, such as *red box*, are subsumed by a type, called *regular-nominal-phr*, that inherits the constraints of *head-functor-phrase*, on the one hand, and *intersective-modification*, on the other hand. Another maximal type is *big-mess-phrase*. Its immediate supertype in the CLAUSALITY hierarchy is the same as for the regular nominal phrases, i.e. *intersective-modification*, but the one in the HEADEDNESS hierarchy is different: Being a subtype of *head-independent-phrase*, its non-head daughter does not select the head-daughter. Beside the inherited properties the BMC has some properties of its own. They are spelled out in (36).

[19] Another subtype of *nominal-parameter* is *inverted-predication*, which subsumes amongst others the binominal noun phrase construction and certain types of apposition (see Section 3.4).

(36)
$$\begin{bmatrix} \textit{big-mess-phr} & \\ \textsc{dtrs} & \left\langle \begin{bmatrix} \textit{hd-functor-phr} & \\ \textsc{synsem|loc|cat} & \begin{bmatrix} \textsc{head} & \textit{adjective} \\ \textsc{mark} & \textit{marked} \end{bmatrix} \end{bmatrix}, \boxed{1} \right\rangle \\ \textsc{head-dtr} & \boxed{1} \begin{bmatrix} \textit{regular-nominal-phr} & \\ \textsc{synsem|loc|cat|mark} & \textit{a} \end{bmatrix} \end{bmatrix}$$

The head daughter is required to be a regular nominal phrase whose MARKING value is *a*, and the non-head daughter is required to be an adjectival head-functor phrase with a MARKING value of type *marked*. This licenses APs which are introduced by a marked adverb, as in *so good a bargain* and *how serious a problem*, while it excludes unmarked APs, as in **good a bargain* and **very big a house*. Iterative application is not licensed, since (36) requires the head daughter to be of type *regular-nominal-phr*, which is incompatible with the type *big-mess-phrase*. This accounts for the fact that a big mess phrase cannot contain another big mess phrase, as in **that splendid so good a bargain*. A reviewer remarked that this analysis allows combinations like *so big a red expensive house*, suggesting that it should not. We are not sure, though, that this combination is ill-formed, and are anyway not inclined to exclude the presence of adjectives in the lower NP, since that would erroneously block the formation of the well-formed combinations in (37).

(37) a. How big a new shrub from France were you thinking of buying?
  b. That's as beautiful a little black dress as I've ever seen.

These are quoted from Zwicky (1995: 116) and Troseth (2009: 42) respectively.

3.2 Nominals with a verbal core

Ordinary nominals have a nominal core, but there are also nominals with a verbal core, such as gerunds and nominalized infinitives. They are of special interest, since they figure prominently in the argumentation that triggered the shift from the NP approach to the DP approach in Transformational Grammar. We first present a specifier treatment of the English gerund and then a functor treatment of the Dutch nominalized infinitive.

3.2.1 The English gerund: a mixed category

Some examples of gerunds are given in (38–40), quoted from Quirk et al. (1985: 1290).

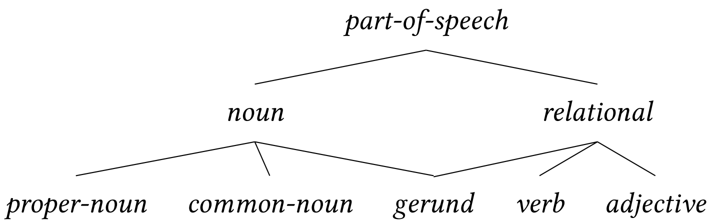


Figure 17: Gerunds as a mixed category

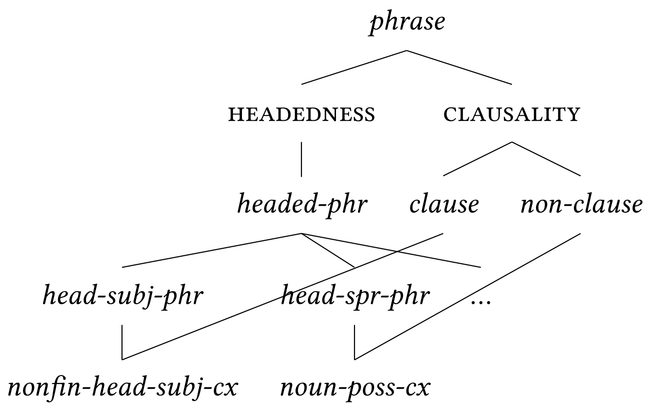


Figure 18: Bi-dimensional hierarchy of gerundial phrases

(38) [Brown's deftly painting his daughter] is a delight to watch.

(39) I dislike [Brown painting his daughter].

(40) Brown is well known for [painting his daughter].

The bracketed phrases have the external distribution of an NP, taking the subject position in (38), the complement position of a transitive verb in (39) and the complement position of a preposition in (40). The internal structure of these phrases, though, shows a mixture of nominal and verbal characteristics. Typically nominal is the presence of the possessive in (38). Typically verbal are the presence of the accusative subject in (39), of the adverbial modifier in (38) and of the NP complements in (38–40). To model this Rob Malouf treats the gerund as a mixed category, introducing a separate part-of-speech for it, which is a subtype of *noun*, on the one hand, and *relational*, on the other hand, see Figure 17 (Malouf 2000: 65). The distinctive properties of this mixed category are spelled out in a lexical rule which derives gerunds from the homophonous present participles (Malouf 2000: 66).[20]

(41)
$$\begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{vform} & \textit{prp} \end{bmatrix} \\ \textsc{subj} & \langle \boxed{1}\,\text{NP} \rangle \\ \textsc{comps} & \boxed{A} \\ \textsc{spr} & \langle\ \rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \textsc{head} & \textit{gerund} \\ \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \boxed{A} \\ \textsc{spr} & \langle \boxed{1} \rangle \end{bmatrix}$$

The gerunds are claimed to take the same kind of complements as the present participles from which they are derived ($\boxed{A}$), and the compatibility with adverbial modifiers follows from the fact that adverbs typically modify objects of type *relational*. The availability of different options for realizing the subject is captured by the inclusion of the subject requirement of the present participle in both the SUBJ list and the SPR list of the gerund ($\boxed{1}$). To model the two options Malouf (2000: 15) employs the bi-dimensional hierarchy of phrase types in Figure 18. The combination with an accusative subject is subsumed by *nonfin-head-subj-cx*, which is a subtype of *head-subject-phr* and *clause*. Its defining properties are spelled out in (42) (Malouf 2000: 16).

(42)
$$\begin{bmatrix} \textit{nonfin-head-subj-cx} & \\ \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head}|\textsc{root} & - \\ \textsc{dtrs} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head} & \begin{bmatrix} \textit{noun} & \\ \textsc{case} & \textit{acc} \end{bmatrix} \end{bmatrix}, \boxed{1} \right\rangle \\ \textsc{head-dtr} & \boxed{1} \end{bmatrix}$$

This construction type subsumes combinations of a non-finite head with an accusative subject, as in (39). When the non-finite head is a gerund, the HEAD value of the resulting clause is *gerund* and since that is a subtype of *noun*, the clause is also a nominal phrase. This accounts for the fact that its external distribution is that of an NP. By contrast, the combination with a possessive subject is subsumed by *noun-poss-cx*, which is a subtype of *head-spr-phr* and *non-clause* (Malouf 2000: 16).

(43)
$$\begin{bmatrix} \textit{noun-poss-cx} & \\ \textsc{synsem}|\textsc{loc} & \begin{bmatrix} \textsc{cat}|\textsc{head} & \textit{noun} \\ \textsc{content} & \textit{nom-obj} \end{bmatrix} \\ \textsc{dtrs} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head} & \begin{bmatrix} \textit{noun} & \\ \textsc{case} & \textit{gen} \end{bmatrix} \end{bmatrix}, \boxed{1} \right\rangle \\ \textsc{head-dtr} & \boxed{1} \end{bmatrix}$$

[20] Boxed capitals stand for objects of type *list*, as in Ginzburg & Sag (2000).

This construction subsumes combinations of a nominal and a possessive specifier, as in *Brown's house*.[21] It also subsumes combinations with the gerund, as in (38), since *gerund* is a subtype of *noun*.

In sum, Malouf's analysis of the gerund involves the postulation of a mixed type in the part-of-speech hierarchy, which allows the gerund to simultaneously function as a nominal and a relational category.

3.2.2 The Dutch nominalized infinitive: phrasal conversion

The closest equivalent of the gerund in Dutch is the nominalized infinitive. Some examples are given in (44–45).

(44) [geld wegschenken] maakt vrouwen gelukkig
[money donate.INF] makes women happy
'Donating money makes women happy'

(45) voor [het op diepte houden van de Vlaamse kusthavens] dient
for [the on depth keep.INF of the Flemish coast.ports] needs
gebaggerd te worden
dredged to be
'Dredging is necessary to keep the Flemish coastal harbors accessible'

Also here the bracketed phrases have the external distribution of an NP, taking the subject position in (44) and the complement position of a transitive verb in (45). And also here, the internal structure shows a mix of nominal and verbal characteristics. Typically nominal are the presence of the article and the post-nominal PP complement in (45). Typically verbal are the presence of the direct object complement in (44) and the predicative PP complement in (45). To model this Van Eynde (2019) makes a distinction between the verbal core and the nominal fringe of a nominalized infinitive, as in the structure of (45), spelled out in Figure 19.

The infinitive is treated as unambiguously verbal at the lexical level and remains verbal when combined with its predicative PP complement, but then it is converted into a nominal projection and combined with a postnominal PP[*van*] complement and the definite article.[22] The conversion is modelled in terms of a non-headed phrase type, since it does not comply with the Head Feature Principle: The HEAD value of the mother is not shared with the daughter. More specif-

[21] Malouf treats the English possessive as a genitive, differently from Sag et al. (2003).

[22] Since the article is a pronominal determiner, it is assigned the category N in the functor treatment.

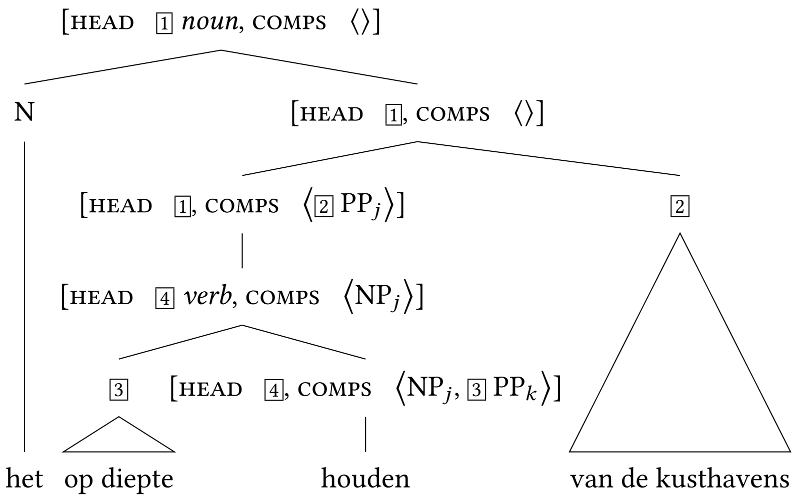


Figure 19: Dutch nominalized infinitive

ically, there is a subtype of the non-headed phrases, called *convert-phr*, whose defining characteristic is that they have a single daughter. In that respect they differ from coordinate phrases, which have at least two daughters.

(46)
$$\begin{bmatrix} \textit{convert-phr} & \\ \text{DTRS} & \langle \text{X} \rangle \end{bmatrix}$$

The conversion which we observe in the Dutch nominalized infinitive is modeled in terms of a phrase type that inherits properties of *convert-phr* and *non-clause*. Its properties are spelled out in (47).

(47)
$$\begin{bmatrix} \textit{nom-inf-phr} & \\ \text{SYNSEM|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{NUMBER} & \textit{singular} \\ \text{GENDER} & \textit{neuter} \end{bmatrix} \\ \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \left\langle \left(\text{PP}_j \right) \right\rangle \oplus \left\langle \left(\text{PP}_i \right) \right\rangle \oplus \boxed{\text{A}} \\ \text{MARKING} & \textit{unmarked} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM|CAT} & \begin{bmatrix} \text{HEAD|VFORM} & \textit{inf} \\ \text{SUBJ} & \langle \text{NP}_j \rangle \\ \text{COMPS} & \left\langle \left(\text{NP}_i \right) \right\rangle \oplus \boxed{\text{A}} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

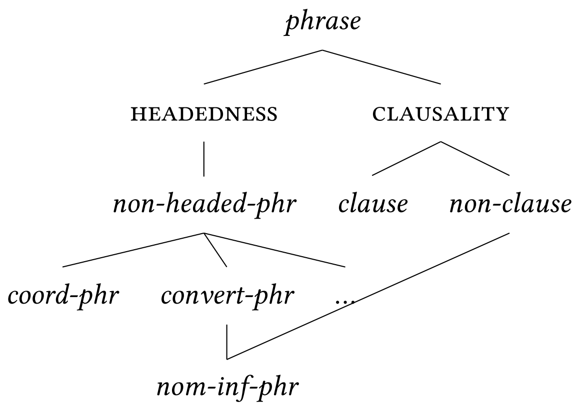


Figure 20: Nominalized infinitives in the bi-dimensional hierarchy

The daughter is an infinitive with a non-empty SUBJ list and a possibly empty COMPS list. The mother is an unmarked singular neuter nominal that inherits the unsaturated complement requirements of its daughter, albeit with the twist that NP complements become PP complements. The subject requirement of the infinitive is added to the COMPS list of the nominal and becomes a PP too. It is made optional, since it is often left unexpressed, as in (44–45). If present, it is introduced by *van* 'of' or *door* 'by', as in (48) and (49) respectively.

(48) het trage afsterven van de koraalriffen
the slow die.INF of the coral.reefs
'the slow dying of the coral reefs'

(49) het uitschakelen van Chelsea door Real Madrid
the eliminate.INF of Chelsea by Real Madrid
'the elimination of Chelsea by Real Madrid'

For a full treatment that also covers the semantic type shift that accompanies the syntactic conversion, we refer to Van Eynde (2019).

3.3 Idiosyncratic P+NP combinations

When a nominal combines with a preposition the result is usually a PP, but not always. The French *de* 'of', for instance, heads a PP in *je viens de Roubaix* 'I come from Roubaix', but not in *beaucoup de farine* 'much flour'. As argued in Abeillé et al. (2004), there are many differences between these two uses of *de*, both syntactic and semantic ones. To model them they treat the former as the ordinary head of a PP and the latter as a weak head. Typical of a weak head is that it shares nearly all properties of its complement, as spelled out in (50).

(50)
$$\begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{subj} & \boxed{A} \\ \textsc{spr} & \boxed{B} \\ \textsc{comps} & \left\langle \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{subj} & \boxed{A} \\ \textsc{spr} & \boxed{B} \\ \textsc{comps} & \langle\ \rangle \end{bmatrix} \\ \textsc{mark} & \textit{unmarked} \\ \textsc{content} & \boxed{4} \end{bmatrix} \right\rangle \\ \textsc{mark} & \textit{de} \end{bmatrix} \\ \textsc{content} & \boxed{4} \end{bmatrix}$$

de has the same values for HEAD, SUBJ, SPR and CONTENT as its complement. When combined with *farine* 'flour', it is, hence, a noun that selects a specifier and that denotes a parameter.[23] The only difference between the weak head and its complement concerns the MARKING value: *de* requires an unmarked complement, but its own MARKING value is of type *de*, and this value is shared with the mother. Specifiers that require the presence of *de*, such as *beaucoup* 'much/many', select a nominal with that MARKING value, as shown in (51).[24]

(51)
$$\begin{bmatrix} \textsc{cat}|\textsc{head} & \begin{bmatrix} \textit{adverb} \\ \textsc{spec}|\textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \textit{noun} \\ \textsc{spr} & \langle \mathrm{X} \rangle \\ \textsc{mark} & \textit{de} \end{bmatrix} \\ \textsc{content} & \begin{bmatrix} \textsc{index} & \boxed{2} \\ \textsc{restr} & \boxed{\Sigma} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textsc{content} & \boxed{1} \begin{bmatrix} \textit{beaucoup-rel} \\ \textsc{index} & \boxed{2} \\ \textsc{restr} & \boxed{\Sigma} \end{bmatrix} \\ \textsc{store} & \{\boxed{1}\} \end{bmatrix}$$

Besides, the selected nominal is required to be unsaturated for SPR and the quantifier is required to share the index and restrictions of its head sister.

[23] The weak *de* also combines with infinitival VPs, as in *de sortir un peu te ferait du bien* 'getting out a bit would do you some good'. In that combination, *de* is a verb that selects a subject and that denotes a state-of-affairs.

[24] In this AVM, quoted from Abeillé et al. (2004: 18), the value of SPEC is of type *synsem*, as in Pollard & Sag (1994), and not of type *semantic-object*, as in Ginzburg & Sag (2000).

The weak head analysis of Abeillé et al. (2004) fits the mold of the specifier treatment. The same data can be analyzed in a way that fits the mold of the functor treatment, as shown in Figure 21. In this analysis the preposition in *beaucoup de farine* 'much flour' is a functor that selects a nominal of type *bare* and that shares its own MARKING value (*de*) with the mother. The functor *beaucoup* in its turn selects a nominal with the MARKING value *de* and shares its own MARKING value with the NP as a whole. In this analysis *de* does not share the part-of-speech, valence and meaning of its nominal sister. Instead, those properties are shared directly between *farine* and *de farine*.

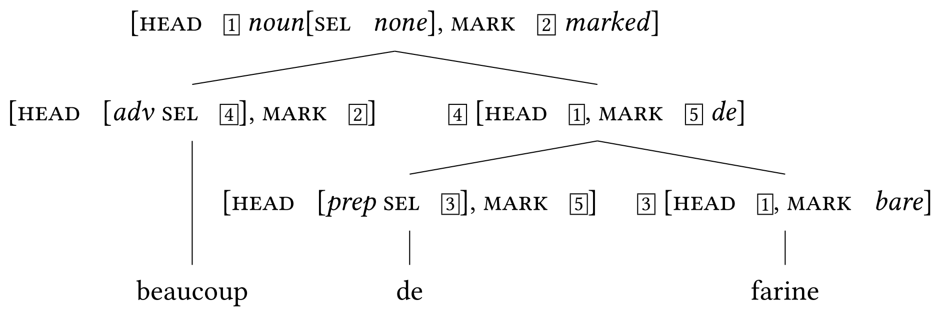


Figure 21: A prepositional functor

While the differences between the weak head treatment and the functor treatment are small and subtle, they may have empirical consequences. This is shown in Maekawa (2015), which offers an analysis of English nominals of the *kind*/*type*/*sort* variety. To account for the fact that the determiner shows agreement with the rightmost noun in *these sort of problems* and *those kind of pitch changes* Maekawa considers the option of treating *of* and the immediately preceding noun as weak heads, but considers this unsatisfactory, since it has the unwanted effect of treating *sort/kind/type* as plural. As an alternative, he develops an analysis in which *of* and the preceding noun are functors (Maekawa 2015: 149). This avoids the unwanted side-effect, since functors do not share the HEAD value of their head sister. Further evidence is provided in Maekawa (2016), which analyzes combinations of a singular numeral with a plural noun, as in *these thousand pages*. If *thousand* is treated as a weak head, it inadvertantly inherits the number value of *pages*, i.e. plural, whereas it is in fact singular, its plural counterpart being *thousands*. This unwanted effect is avoided if *thousand* is treated as a functor.

3.4 Other nominals with idiosyncratic properties

A much studied nominal with idiosyncratic traits is the binominal noun phrase construction (BNPC), exemplified in (52).

(52) a. She blames it on [her nitwit of a husband].
b. She had [a skullcracker of a headache].

In contrast to ordinary [NP–*of*–NP] sequences, as in *an employee of a Japanese company*, where the first nominal is the head of the entire NP, and where the second nominal is part of its PP adjunct, the BNPC shows a pattern in which the relation between the nominals is a predicative one: her husband is claimed to be a nitwit, and the headache is claimed to be like a skullcracker. HPSG treatments of the BNPC are provided in Kim & Sells (2014) and Van Eynde (2018). The latter uses the phrase type hierarchy in (16), defining the BNPC as a maximal type that inherits from *head-independent-phr* and *inverted-predication*. To capture the intuition that the second nominal is the head of the entire NP, the preposition *of* is not treated as the head of a PP, but as a functor that selects a nominal head.

Comparable to NPs with a verbal core are NPs with an adjectival core, such as *the very poor* and *the merely skeptical*. They are described and provided with an HPSG analysis in Arnold & Spencer (2015). Interestingly, it employs a device for phrasal conversion that is similar to the one for the Dutch nominalized infinitive in Section 3.2.2.

Idiosyncratic are also the nominals with an extracted determiner, as in the French (53) and the Dutch (54).

(53) Combien as-tu lu [__ de livres en latin] ?
how.many have-you read [__ of books in Latin]
'How many books have you read in Latin?'

(54) Wat zijn dat [__ voor vreemde geluiden] ?
what are that [__ for strange noises]
'What kind of strange noises are those?'

The French example is analyzed in Abeillé et al. (2004: 20–21) and the Dutch one in Van Eynde (2004: 47–50).

Another special case is apposition. It comes in (at least) two types, known as close apposition and loose apposition. Relevant examples are respectively *my brother Richard is a soldier* and *Sarajevo, the capital of Bosnia, is where WW I began.* Both are compared and analyzed in Kim (2012) and Kim (2014). Van Eynde & Kim (2016) provides an analysis of loose apposition in Sign-Based Construction Grammar.

4 Conclusion

This chapter has provided a survey of how nominals are analyzed in HPSG. Over time three treatments have taken shape, i.e. the specifier treatment, the functor treatment and the DP treatment. They were presented and applied to ordinary nominals in Section 2. Nominals with idiosyncratic properties were discussed in Section 3. Modeling them requires a relaxation of the strictly lexicalist stance of early HPSG. The more flexible tools of constructional HPSG have been put to use in the analysis of the Big Mess Construction and of nominals with a verbal core, such as gerunds and nominalized infinitives. For other nominals with idiosyncratic properties pointers have been given to relevant literature.

In terms of the NP vs. DP debate, it is clear that the NP approach has been more successful in HPSG than the DP approach. This fits in with the tendency to avoid the postulation of functional categories with their respective projections. Clauses, for instance, are not analyzed as IPs either. The finer-grained differences between the specifier treatment and the functor treatment are a topic of ongoing debate. They both have their advocates and for the analysis of ordinary nominals there does not seem to be any evidence that decisively tilts the balance. For nominals with idiosyncratic properties, however, there is a clear tendency in the recent literature to adopt the functor treatment, usually in combination with phrasal constraints.

Acknowledgements

For their comments on earlier versions of this chapter I would like to thank Bob Borsley, Stefan Müller, Liesbeth Augustinus and two anonymous reviewers. My contributions to the HPSG framework span three decades by now, starting with work on a EU-financed project in the early nineties, to which also Valerio Allegranza and Louisa Sadler contributed. Decisive for my conversion to HPSG at the time was not only the EU-money, but also conversations with Danièle Godard and Ivan Sag. I became a regular at the annual HPSG conferences by 1998 and started teaching it in Leuven around the same time. A heart-felt thanks goes to the colleagues who over the years have created such a stimulating environment to work in.

References

Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng. 2004. The syntax of French N' phrases. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-driven Phrase Structure Grammar*, 6–26. Stanford: CSLI Publications.

Abney, Stephen. 1987. *The English noun phrase in its sentential aspects*. MIT dissertation.

Allegranza, Valerio. 1998. Determiners as functors: NP structure in Italian. In Sergio Balari & Luca Dini (eds.), *Romance in HPSG*, 55–107. Stanford: CSLI Publications.

Allegranza, Valerio. 2006. *The signs of determination. Constraint-based modelling across languages*. Peter Lang.

Arnold, Doug & Louisa Sadler. 2014. The big mess construction. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG14 conference*. Stanford: CSLI Publications.

Arnold, Doug & Andrew Spencer. 2015. A constructional analysis for the skeptical. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-driven Phrase Structure Grammar*, 41–60. Stanford: CSLI Publications.

Berman, Judith. 1974. *Adjectives and adjective complement constructions in English*. Harvard University dissertation.

Borsley, Bob & Anne Abeillé. 2019. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 3–5. Berlin: Language Science Press. DOI:??

Bouma, Gosse. 1988. Modifiers and specifiers in Categorial Unification Grammar. *Linguistics* 26. 21–46.

Bresnan, Joan. 2001. *Lexical Functional Syntax*. Blackwell.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In R.A. Jacobs & P.S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, Mass.: Ginn & Company.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Gazdar, Gerald, Ewan Klein, Geoff Pullum & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

Ginzburg, Jonathan & Ivan Sag. 2000. *Interrogative investigations*. Stanford: CSLI Publications.

Hudson, Richard. 1990. *English Word Grammar*. Oxford: Blackwell.

Jackendoff, Ray. 1977. *X-bar Syntax. A study of phrase structure*. Cambridge: MIT Press.

Kay, Paul & Ivan Sag. 2012. Cleaning up the big mess: discontinuous dependencies and complex determiners. In Hans Boas & Ivan Sag (eds.), *Sign-Based Construction Grammar*, 229–256. Stanford: CSLI Publications.

Kim, Jong-Bok. 2012. Form and function mismatch in the English appositional construction. *Language Research* 48. 1–23.

Kim, Jong-Bok. 2014. Similarities and differences between the close and loose apposition in English. *English Language and Linguistics* 20. 79–98.

Kim, Jong-Bok & Peter Sells. 2011. The big mess construction: interactions between the lexicon and constructions. *English Language and Linguistics* 15. 335–362.

Kim, Jong-Bok & Peter Sells. 2014. English binominal NPs: a construction-based perspective. *Journal of Linguistics* 15. 1–33.

Maekawa, Takafumi. 2015. Agreement mismatch between sort/kind/type and the determiner. In Stefan Müller (ed.), *Proceedings of the 22nd International Conference on Head-driven Phrase Structure Grammar*, 136–156. Stanford: CSLI Publications.

Maekawa, Takafumi. 2016. Seminumerals, determiners and nouns in English. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the Joint 2016 Conference on HPSG and LFG*, 422–441. Stanford: CSLI Publications.

Malouf, Robert. 2000. *Mixed categories in the hierarchical lexicon*. Stanford: CSLI Publications.

Müller, Stefan. 2018. Headless in Berlin: Headless structures in Head-Driven Phrase Structure Grammar. In Ulrike Freywald & Horst Simon (eds.), *Headedness and/or grammatical anarchy?* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. To appear.

Nerbonne, John & Tony Mullen. 2000. Null-headed nominals in German and English. In Frank Van Eynde, Ineke Schuurman & Ness Schelkens (eds.), *Computational Linguistics in the Netherlands 1998*, 125–146. Amsterdam: Rodopi.

Netter, Klaus. 1994. Towards a theory of functional heads: German nominal phrases. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 297–340. Stanford, CA: CSLI Publications.

Netter, Klaus. 1996. *Functional categories in an HPSG for German*. Saarbrücken: Universität des Saarlandes Dissertation.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London & New York: Longman.

Reape, Michael. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-driven Phrase Structure Grammar*, 151–197. Stanford: CSLI Publications.

Sag, Ivan. 1997. English relative clause constructions. *Journal of Linguistics* 33. 431–484.

Sag, Ivan. 2012. Sign-Based Construction Grammar: an informal synopsis. In Hans Boas & Ivan Sag (eds.), *Sign-Based Construction Grammar*, 69–202. Stanford: CSLI Publications.

Sag, Ivan, Thomas Wasow & Emily Bender. 2003. *Syntactic Theory. a Formal Introduction. Second Edition*. Stanford: CSLI Publications.

Troseth, Erika. 2009. Degree inversion and negative intensifier inversion in the English DP. *The Linguistic Review* 26. 37–65.

Van Eynde, Frank. 1998. The immediate dominance schemata of HPSG. In Peter-Arno Coppen, Hans van Halteren & Lisanne Teunissen (eds.), *Computational Linguistics in the Netherlands 1997*, 119–133. Amsterdam/Atlanta: Rodopi.

Van Eynde, Frank. 2003. Prenominals in Dutch. In Jong-Bok Kim & Steve Wechsler (eds.), *Proceedings of the 9th International Conference on Head-driven Phrase Structure Grammar*, 333–356. Stanford: CSLI Publications.

Van Eynde, Frank. 2004. Minor Adpositions in Dutch. *Journal of Comparative Germanic Linguistics* 7. 1–58.

Van Eynde, Frank. 2006. NP-internal agreement and the structure of the noun phrase. *Journal of Linguistics* 42. 139–186.

Van Eynde, Frank. 2007. The big mess construction. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-driven Phrase Structure Grammar*, 415–433. Stanford University: CSLI Publications.

Van Eynde, Frank. 2018. Regularity and idiosyncracy in the formation of nominals. *Journal of Linguistics* 54. 823–858.

Van Eynde, Frank. 2019. Nominalized infinitives in Dutch. In Stefan Müller (ed.), *Proceedings of the 26th International Conference on Head-driven Phrase Structure Grammar*. Stanford: CSLI Publications.

Van Eynde, Frank & Jong-Bok Kim. 2016. Loose apposition. A construction-based analysis. *Functions of Language* 23. 17–39.

Wechsler, Stephen. 2019. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 131–154. Berlin: Language Science Press. DOI:??

Zwicky, Arnold. 1995. Exceptional degree markers : a puzzle in internal and external syntax. *OSU Working Papers in Linguistics* 47. 111–123.

Chapter 9

Argument structure and linking

Stephen Wechsler
The University of Texas

Jean-Pierre Koenig
University at Buffalo

Anthony Davis
Southern Oregon University

What we cannot speak about, we must pass over in silence.

1 Introduction

When a verb or other predicator is composed with the phrases or pronominal affixes expressing its semantic arguments, the grammar must specify the mapping between the semantic participant roles and syntactic dependents of that verb. For example, the grammar of English indicates that the subject of *eat* fills the eater role and the object of *eat* fills the role of the thing eaten. In HPSG this mapping is usually broken down into two simpler mappings by positing an intermediate representation called ARG-ST ('argument structure'). The first mapping connects the participant roles within the semantic CONTENT with the elements of the ARG-ST feature; here we will call the theory of this mapping *linking theory* (see Section 4). The second mapping connects those ARG-ST list elements to the elements of the VALlists, namely COMPS ('complements') and SUBJ ('subject'; or SPR, 'specifier'); we will refer to this second mapping as *argument realization* (see Section 2).[1] These two mappings are illustrated with the simplified lexical sign

[1]Some linguists, such as Levin & Rappaport Hovav (2005), use the term 'argument realization' more broadly, to encompass linking as well.

Stephen Wechsler, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

for the verb *eat* in (1).

(1) Lexical sign for the verb *eat*

$$\begin{bmatrix} \text{PHON} & \langle \text{EAT} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\text{NP}_i \text{ , } \boxed{2}\text{NP}_j \rangle \\ \text{CONTENT} & \textbf{\textit{eat}}(i, j) \end{bmatrix}$$

In (1), 'NP' abbreviates a feature structure representing syntactic and semantic information about a nominal phrase. The variables i and j are the referential indices for the eater and eaten arguments, respectively, of the **eat** relation. The semantic information in NP_i semantically restricts the value or referent of i.

The ARG-ST feature plays an important role in HPSG grammatical theory. In addition to regulating the mapping from semantic arguments to grammatical relations, ARG-ST is the locus of the theories of anaphoric binding and other construal relations such as control and raising. (This chapter focuses on the function of ARG-ST in semantic mapping, with some discussion of binding and other construal relations only insofar as they interact with that mapping. A more detailed look at binding is presented in Chapter **??**.)

In HPSG, verb diathesis alternations, voice alternations, and derivational processes such as category conversions are all captured within the lexicon (see Section 5 and Chapter **??**). The different variants of a word are grammatically related either through lexical rules or by means of the lexical type hierarchy. HPSG grammars explicitly capture paradigmatic relations between word variants, making HPSG a *lexical approach to argument structure*, in the sense of Müller & Wechsler (2014a). This fundamental property of lexicalist theories contrasts with many transformational approaches, where such relationships are treated as syntagmatically related through operations on phrasal structures representing sentences and other syntactic constituents. Arguments for the lexical approach are reviewed in Section 8.

Within the HPSG framework presented here, we will formulate and address a number of empirical questions:

- We know that a verb's meaning influences its valence requirements, (via the ARG-ST list, on this theory). What are the principles governing the mapping from CONTENT to ARG-ST? Are some aspects of ARG-ST idiosyncratically stipulated for individual verbs? What aspects of the semantic CONTENT bear on the value of ARG-ST, and what aspects do not? (For example, what is the role of modality?)

- How are argument alternations defined with respect to our formal system? For each alternation we may ask which of the following it involves: a shuffling of the ARG-ST list; a change in the mapping from ARG-ST to VAL; or a change in the CONTENT, with a concomitant change in the ARG-ST?

These questions will be addressed below in the course of presenting the theory. We begin by considering ARG-ST itself (Section 2), followed by the mapping from ARG-ST to VAL(Section 3) and the mapping from CONTENT to ARG-ST (Section 4). The remaining sections address further issues relating to argument structure: the nature of argument alternations, extending the ARG-ST attribute to include additional elements, whether ARG-ST is a universal feature of languages, and a comparison of the lexicalist view of argument structure presented here with phrasal approaches.

2 The representation of argument structure in HPSG

In the earliest versions of HPSG, the selection of dependent phrases was specified in the SUBCAT feature of the head word (Pollard & Sag (1987), Pollard & Sag (1994: ch. 1–8)). The value of SUBCAT is a list of items, each of which corresponds to the SYNSEM value of a complement or subject. Following are SUBCAT features for an intransitive verb, a transitive verb, and a transitive verb with obligatory PP complement:

(2) a. *laugh*: [SUBCAT ⟨ NP ⟩]
    b. *eat*: [SUBCAT ⟨ NP, NP ⟩]
    c. *put*: [SUBCAT ⟨ NP, NP, PP ⟩]

Phrase structure rules in the form of immediate dominance schemata identify a certain daughter node as the head daughter (HEAD-DTR) and others as complement daughters (COMP-DTRS). In keeping with the *Subcategorization Principle*, here paraphrased from Pollard & Sag (1994: 34), list items are effectively 'cancelled' from the SUBCAT list as complement phrases are joined with the selecting head:

(3) Subcategorization Principle: In a headed phrase, the SUBCAT value of the HEAD-DTR ('head daughter') is the concatenation of the phrase's SUBCAT list with the list of SYNSEM values of the COMPS-DTRS ('complement daughters').

Phrasal positions are distinguished by their saturation level: 'VP' is defined as a verbal projection whose SUBCAT list contains a single item, corresponding to the subject; and 'S' is defined as a verbal projection whose SUBCAT list is empty.

The 'subject' of a verb, a distinguished dependent with respect to construal processes such as binding, control, and raising, was then defined as the first item in the SUBCAT list, hence the last item with which the verb combines. However, defining 'subject' as the last item to combine with the head proved inadequate (Pollard & Sag 1994: Ch. 9). There are many cases where the dependent displaying subject properties need not be the last item added to the head projection. For example, in German the construal subject is a nominal in nominative case (Reis 1982), but the language allows subjectless clauses containing only a dative or genitive non-subject NP. If that oblique NP is the only NP dependent to combine with the verb then it is *ipso facto* the last NP to combine.

Consequently, the SUBCAT list was split into two valence lists, a SUBJ list of length zero or one for subjects, and a COMPS list for complements. Nonetheless, certain grammatical phenomena, such as binding and other construal processes, must still be defined on a single list comprising both subject and complements (Manning & Sag 1999). Additionally, some syntactic arguments are unexpressed or realized by affixal pronouns, rather than as subject or complement phrases. The new list containing all the syntactic arguments of a predicator was named ARG-ST ('argument structure').

In clauses without implicit or affixal arguments, the ARG-ST is the concatenation of SUBJ and COMPS respectively. For example, the SUBCAT list for *put* in (2c) is replaced with the following:

(4)
$$\begin{bmatrix} \text{PHON} & \langle \text{PUT} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\text{NP}, \boxed{2}\text{NP}, \boxed{3}\text{PP} \rangle \end{bmatrix}$$

The idealization according to which ARG-ST is the concatenation of SUBJ and COMPS is canonized as the *Argument Realization Principle* (ARP) (Sag et al. 2003: PAGE). Exceptions to the ARP, that is, dissociations between VALENCE and ARG-ST, are discussed in Section 3.2 below.

A predicator's VALENCE lists indicate its requirements for syntactic concatenation with phrasal dependents (Section 3). ARG-ST, meanwhile, provides syntactic information about the expression of semantic roles and is related, via linking theory, to the lexical semantics of the word (Section 3.2). The ARG-ST list contains specifications for the union of the verb's syntactic local phrasal dependents (the subject and complements, whether they are semantic arguments, raised phrases, or expletives) and its syntactic arguments that are not realized locally, whether they are unbounded dependents, affixal, or unexpressed.

Figure 1 provides a schematic representation of linking and argument realization in HPSG. Linking principles govern the mapping of of participant roles in a predicator's CONTENT to syntactic arguments on ARG-ST. Argument realization is shown in this figure only for mapping to VAL; affixal and null arguments are not depicted. Here, the semantic roles are just arbitrary labels, but we discuss in Section (4) how they can be systematically related to lexical entailments of predicators. The ARG-ST and VALlists in this figure contain only arguments linked to participant roles, but in Section ?? we note motivations for extending ARG-ST to include additional elements. And in Section (3), we examine cases where the relationship between ARG-ST and VALviolates the ARP.

Semantics (CONTENT)	$\begin{bmatrix} \textit{sem-rel-1} & \\ \textsc{sem-role-1} & \boxed{1} \\ \textsc{sem-role-2} & \boxed{2} \\ \textsc{soa} & \begin{bmatrix} \textit{sem-rel-2} & \\ \textsc{sem-role-3} & \boxed{3} \\ \textsc{sem-role-4} & \boxed{4} \end{bmatrix} \end{bmatrix}$	Structured representation of semantic relations, their roles and the participants filling them
	$\Updownarrow$ *Linking principles*	
Argument structure (ARG-ST)	$\begin{bmatrix} \textsc{arg-st} \ \langle \textsc{xp}_1, \textsc{xp}_2, \ldots \textsc{xp}_n \rangle \end{bmatrix}$	List of syntactic arguments
	$\Updownarrow$ *Argument Realization Principle*	
Syntax (VALENCE)	$\begin{bmatrix} \textsc{spr} & \langle\ \rangle \\ \textsc{subj} & \langle \textsc{xp}_1 \rangle \\ \textsc{comps} & \langle \textsc{xp}_2, \ldots \textsc{xp}_n \rangle \end{bmatrix}$	Lists of locally realized dependents

Figure 1: How linking works in HPSG

3 Argument realization: The mapping from ARG-ST to VALlists

3.1 Variation in the expression of arguments

The VALfeature is responsible for composing a verb with its phrasal dependents, but this is just one of the ways that semantic arguments of a verb are expressed in natural language. Semantic arguments can be expressed in various linguistic forms: as local syntactic dependents (SUBJ and COMPS), as affixes, or displaced in unbounded dependency constructions (SLASH).

Affixal arguments can be illustrated with the first person singular Spanish verb *hablo* ‘speak.1SG’, as in 5.

(5) a. Habl-o español.
speak-1SG Spanish
‘I speak Spanish.’

b. *hablo*‘speak.1SG’:

$$\begin{bmatrix} \text{PHON} & \langle \text{HABLO} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle\ \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{bmatrix} \\ \text{ARG-ST} & \left\langle \text{NP:}\begin{bmatrix} \textit{ppro} & \\ \text{INDEX} & \begin{bmatrix} \text{PERS} & \textit{1st} \\ \text{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix}, \boxed{2}\text{NP} \right\rangle \end{bmatrix}$$

The *-o* suffix contributes the first person singular pronominal subject content to the verb form (the morphological process is not shown here; see Chapter MORPHOLOGY-CHAPTER). The pronominal subject appears on the ARG-ST list and hence is subject to the binding theory. But it does not appear in SUBJ, if no subject NP appears in construction with the verb.

A lexical sign whose ARG-ST list that is just the concatenation of its SUBJ and COMPS lists conforms to the Argument Realization Principle (ARP); such signs are called *canonical signs* by Bouma et al. (2001). Non-canonical signs, which violate the ARP, have been approached in two ways. In one approach, a lexical rule takes as input a canonical entry and derives a non-canonical one by removing items from the VALlists, while adding an affix or designating an item as an unbounded dependent by placement on the SLASH list. In the other approach, a feature of each ARG-ST list item specifies whether the item is subject to the ARP (hence mapped to a VALlist), or ignored by it (hence expressed in some other way). See the chapter on the lexicon for more detail and Miller & Sag (1997) for a treatment of French clitics as affixes.

A final case to consider is null anaphora, in which a semantic argument is simply left unexpressed and receives a definite pronoun-like interpretation. Japanese *mi-* 'see' is transitive but the object NP can be omitted as in (6).

(6) Naoki-ga mi-ta.
Naoki-NOM see-PAST
'Naoki saw it/him/her/*himself.'

Null anaphors of this kind typically arise in discourse contexts similar to those that license ordinary weak pronouns, and the unexpressed object often has the (Principle B) obviation effects characteristic of overt pronouns, as shown in (6). But HPSG eschews the use of silent formatives like 'small *pro*' when there is no evidence for such items, such as local interactions with the phrase structure. Instead, null anaphors of this kind are present in ARG-ST but absent from VALlists. ARG-ST is directly linked to the semantic CONTENT and is the locus of binding theory, so the presence of a syntactic argument on the ARG-ST list but not a VALlist accounts for null anaphora. To account for obviation, the ARG-ST list item, when unexpressed, receives the binding feature of ordinary (non-reflexive) pronouns, usually *ppro*. This language-specific option can be captured in a general way by VALand ARG-ST defaults in the lexical hierarchy for verbs.

3.2 The syntax of ARG-ST and its relation to VALENCE

The ARG-ST ordering represents a preliminary syntactic structuring of the set of argument roles. In that sense it functions as an interface between the lexical semantics of the verb, and the expressions of dependents as described in Section 3. Its role thus bears some relation to the initial stratum in Relational Grammar, *argument structure* (including intrinsic classifications) in LFG Lexical Mapping Theory, D-structure in Government/Binding theory, and the Merge positions of arguments in Minimalism, assuming in the last two cases the Uniform Thematic Alignment Hypothesis (Baker 1988) or something similar. However, it also differs from all of those in important ways.

Semantic constraints on ARG-ST are explored in Section 4 below. But ARG-ST is not only structured by semantic distinctions between the arguments but also by syntactic ones. Specifically, the list ordering represents relative syntactic *obliqueness* of arguments. The least oblique argument is the subject (SUBJ), followed by the complements (COMPS). Following Manning (1996) term arguments (direct arguments, i.e. objects) are assumed to be less oblique than 'oblique' arguments (adpositional and oblique case marked phrases), followed finally by predicate and clausal complements. The transitive ordering relation on the ARG-ST list is called *o-command* ('obliqueness command'): the subject list item o-commands those of the complements; an object list item o-commands those of any obliques; and so on.

Voice alternations like the passive, which are defined on the ARG-ST list, illustrate the ordering of terms before obliques on the ARG-ST list. Passivization alters the syntactic properties of ARG-ST list items: the initial item of the active, normally mapped to SUBJ of the active, is an oblique (*by* phrase) or unexpressed argument in the passive. Given that terms precede obliques in the list order, any term arguments must o-command the passive oblique, so passive effectively reorders the initial item in ARG-ST to a list position following any terms.

(7) a. Susan gave Mary a book.
  b. Mary was given a book by Susan.

(8) a. *give* (active): $[\textsc{arg-st} < \textsc{np}_i, \textsc{np}_j, \textsc{np}_k >]$
  b. *given* (passive): $[\textsc{arg-st} < \textsc{np}_j, \textsc{np}_k, \textsc{pp}[by]_i >]$

Relative obliqueness conditions a number of syntactic processes and phenomena, including anaphoric binding. The o-command relation replaces the c-command in the Principles A, B, and C of Chomsky's (1981) configurational theory of binding. For example, HPSG's Principle B states that an ordinary pronoun cannot be o-commanded by its coargument antecedent, which accounts for the pronoun obviation observed in the English sentence *Naoki*$_i$ *saw him*$_{*i/j}$, and also accounts for obviation in the Japanese sentence (6) above.

Relative obliqueness also conditions the accessibility hierarchy of Keenan & Comrie (1977), according to which a language allowing relativization of some type of dependent also allows relativization of any less oblique than it. Hence if a language has relative clauses at all, it has subject relatives; if it allows obliques to relativize then it also allows subject and object relatives; and so on. Similar implicational universals apply to verb agreement with subjects, objects, and obliques Greenberg (1966).

Returning now to argument realization, we saw above that the rules for the selection of the subject from among the verb's arguments are also stated on the ARG-ST list. In a 'canonical' realization the subject is the first list item, o-commanding all of its coarguments.

3.3 Syntactic ergativity

The autonomy of ARG-ST from the VALlists is further illustrated by cross-linguistic variation in the mapping between them. As just noted, in English and many other languages the initial item in ARG-ST maps to the subject. However, languages with so-called *syntactically ergative* clauses have been analyzed as following a different mapping rule. Crucially, the ARG-ST ordering in those languages is still supported by independent evidence from properties such as binding and NP versus PP categorial status of arguments. Balinese (Austronesian), as analyzed by Wechsler & Arka (1998), is such a language. In the morphologically unmarked, and most common voice, called 'Objective voice' (OV), the subject is any term *except* the ARG-ST-initial one.

Balinese canonically has SVO order, regardless of the verb's voice form (Artawa 1994; Wechsler & Arka 1998). The preverbal NPs in (9) are the surface subjects and the postverbal ones are complements. When the verb appears in the unmarked OV verb, a non-initial term is the subject, as in (9a). But verbs in 'Agentive Voice' (AV) select as their subject the ARG-ST-initial item, as in (9b).

(9) a. Bawi adol ida.
pig OV.sell 3sg
'He/She sold a pig.'

b. Ida ng-adol bawi.
3sg AV-sell pig
'He/She sold a pig.'

A ditransitive verb such as the benefactive applied form of *beli* 'buy' in (10), has three term arguments on its ARG-ST list. The subject can be either term that is non-initial in ARG-ST:

(10) a. Potlote ento beli-ang=a I Wayan.
pencil-DEF that OV.buy-APPL=3 Art Wayan
'(s)he bought Wayan the pencil.'

b. I Wayan beli-ang=a potlote ento.
Art Wayan OV.buy-APPL=3 pencil-DEF that
'(s)he bought Wayan the pencil.'

Unlike the passive voice, which reorders the ARG-ST list, the Balinese OV does not affect ARG-ST list order. Thus the agent argument can bind a coargument reflexive pronoun (but not vice versa), regardless of whether the verb is in OV or AV form:

(11) a. Ida ny-ingakin ragan idane.
3sg AV-see self
'(s)he saw himself/herself'

b. Ragan idane cingakin ida.
self OV.see 3sg
'(s)he saw himself/herself'

The 'seer' argument o-commands the 'seen', with the AV versus OV voice forms regulating subject selection:

(12) Agentive Voice form of 'see':

$$\begin{bmatrix} \textsc{phon} & \langle \textit{nyinkagin} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \boxed{2} \rangle \end{bmatrix} \\ \textsc{arg-st} & \langle \boxed{1}\text{NP}_i , \boxed{2}\text{NP}_j \rangle \\ \textsc{content} & \begin{bmatrix} \textit{see-rel} & \\ \textsc{seer} & i \\ \textsc{seen} & j \end{bmatrix} \end{bmatrix}$$

(13) Objective Voice form of 'see':

$$\begin{bmatrix} \textsc{phon} & \langle \textit{cinkagin} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{2} \rangle \\ \textsc{comps} & \langle \boxed{1} \rangle \end{bmatrix} \\ \textsc{arg-st} & \langle \boxed{1}\text{NP}_i , \boxed{2}\text{NP}_j \rangle \\ \textsc{content} & \begin{bmatrix} \textit{see-rel} & \\ \textsc{seer} & i \\ \textsc{seen} & j \end{bmatrix} \end{bmatrix}$$

Languages like Balinese illustrate the autonomy of ARG-ST. Although the agent binds the patient in both (11)a and b, the binding conditions cannot be stated directly on the thematic hierarchy. For example, in HPSG a raised argument appears on the ARG-ST list of the raising verb even though that verb assigns no thematic role to that list item. But a raised subject can bind a coargument reflexive in Balinese (this is comparable to English *John seems to himself to be ugly*). Anaphoric binding in Balinese raising constructions thus behaves as predicted by the ARG-ST based theory (Wechsler 1999). In conclusion, neither VALNOR CONTENT provides the right representation for defining binding conditions, but ARG-ST fills the bill.

Syntactically ergative languages that have been analyzed as using an alternative mapping between ARG-ST and VALinclude Tagalog, Inuit, some Mayan languages, Chukchi, Toba Batak, Tsimshian languages, and Nadëb (Manning 1996; Manning & Sag 1999).

Interestingly, while the GB/Minimalist configurational binding theory may be defined on analogues of VALor CONTENT, those theories lack any analogue of ARG-ST. This leads to special problems for such theories in accounting for binding in many Austronesian languages like Balinese. In transformational theories since Chomsky (1981), anaphoric binding conditions are usually stated with respect to the A-positions ('argument positions'). A-positions are analogous to HPSG VALlist items, with relative c-command in the configurational structure corresponding to relative list ordering in HPSG, in the simplest cases. Meanwhile, to account for data similar to (11), where agents asymmetrically bind patients, Austronesian languages like Balinese were said to define binding on the 'thematic structure' encoded in d-structure or Merge positions, where agents asymmetrically c-command patients regardless of their surface positions (Guilfoyle et al. 1992). But the interaction with raising shows that neither of those levels is appropriate as the locus of binding theory (Wechsler 1999).[2]

3.4 Symmetrical objects

We have thus far tacitly assumed a total ordering of elements on the ARG-ST list, but Ackerman et al. (2018), Ackerman et al. (2017) propose a partial ordering for certain so-called 'symmetrical object' languages. In Moro (Kordofanian), the two term complements of a ditransitive verb have exactly the same object properties. Relative linear order of the theme and goal arguments is free, as shown by the two translations of (14) (from Ackerman et al. 2017: 9):

(14) é-g-a-natʃ-ó ... óráŋ ŋerá
1SB.SM-CLg-MAIN-give-PFV CLg.man CLŋ.girl
'I gave the girl to the man.' / 'I gave the man to the girl.'

More generally, the two objects have identical object properties with respect to occurrence in post-predicate position, case marking, realization by an object marker, and ability to undergo passivization (Ackerman et al. 2017: 9).

[2] To account for (11b) under the configurational binding theory, the subject position must be an A-bar position; but to account for binding by a raised subject, it must be an A-position. See Wechsler (1999).

Ackerman et al. (2017) propose that the two objects are unordered on the ARG-ST list. This allows for two different mappings to the COMPS list, as shown here:

(15) a. Goal argument as primary object:

$$\begin{bmatrix} \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix} \\ \text{ARG-ST} & \left\langle \boxed{1}\text{NP}_i , \left\{ \boxed{2}\text{NP}_j , \boxed{3}\text{NP}_k \right\} \right\rangle \\ \text{CONTENT} & \begin{bmatrix} \textit{give-rel} & \\ \text{AGENT} & i \\ \text{GOAL} & j \\ \text{THEME} & k \end{bmatrix} \end{bmatrix}$$

b. Theme argument as primary object:

$$\begin{bmatrix} \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{3}, \boxed{2} \rangle \end{bmatrix} \\ \text{ARG-ST} & \left\langle \boxed{1}\text{NP}_i , \left\{ \boxed{2}\text{NP}_j , \boxed{3}\text{NP}_k \right\} \right\rangle \\ \text{CONTENT} & \begin{bmatrix} \textit{give-rel} & \\ \text{AGENT} & i \\ \text{GOAL} & j \\ \text{THEME} & k \end{bmatrix} \end{bmatrix}$$

The primary object properties, which are associated with the initial term argument of COMPS, can go with either the goal or theme argument.

To summarize this section, while the relationship between ARG-ST, SUBJ, and COMPS lists was originally conceived as a straightforward one, enabling binding principles to maintain their simple form by defining ARG-ST as the concatenation of the other two, the relationship was soon loosened. Non-canonical relationships between ARG-ST and the VALlists are invoked in accounts of several core syntactic phenomena. Arguments not realized overtly in their canonical positions, due to extraction, cliticization, or pro-drop (null anaphora), appear on ARG-ST but not in any VALlist. Accounts of syntactic ergativity in HPSG involve variations in the mapping between ARG-ST and VALlists; in particular, the element of SUBJ is not, in such languages, the first element of ARG-ST. Modifications of ARG-ST play a role in treatments of passivization, where its expected first element is suppressed, and in languages with multiple, symmetric objects, where a partial rather than total ordering of ARG-ST elements has been postulated. Thus ARG-ST has now acquired an autonomous status within HPSG, and is not merely a predictable rearrangement of information present elsewhere in lexical entries.

4 Linking: the mapping between semantics and ARG-ST

4.1 HPSG approaches to linking

The term *linking* refers to the mapping specified in a lexical entry between participant roles in the semantics and their syntactic representations on the ARG-ST list. Early HPSG grammars stipulated the linking of each verb: semantic CONTENT values with predicator-specific attributes like DEVOURER and DEVOURED were mapped to the subject and object, respectively, of the verb *devour*. But linking observes prevailing patterns, e.g. if one argument of a transitive verb in active voice has an agentive role, it will map to the subject, not the object. Thus these early accounts were unsatisfying, as they lead to purely stipulative accounts of linking, specified verb by verb. Beginning with Wechsler (1995b) and Davis (1996), researchers formulated linking principles stated on more general semantic properties holding across verbs.

Within the history of linguistics there have been three general approaches to modeling the lexico-semantic side of linking: thematic role types (Pāṇini ca 400 B.C., Fillmore 1968); lexical decomposition (Foley & Van Valin 1984; Rappaport Hovav & Levin 1998); and the proto-roles approach (Dowty 1991). In developing linking theories within the HPSG framework Wechsler (1995b) and Davis (1996) employed a kind of lexical decomposition that also incorporated some elements of the proto-roles approach. The reasons for preferring this over the alternatives are discussed in Section 4.4 below.

Wechsler's 1995b linking theory constrains the relative order of pairs arguments on the ARG-ST list according to semantic relations entailed between them. For example, his *notion rule* states that if one participant in an event is entailed to have a mental notion of another, then the first must precede the second on the ARG-ST list. The *conceive-pred* type is defined by the following type declaration (based on Wechsler (1995b: 127), with formal details adjusted for consistency with current usage):

(16) *conceive-pred:* $\begin{bmatrix} \textsc{arg-st} & \langle \mathrm{NP}_i, \mathrm{NP}_j \rangle \\ \textsc{content} & \begin{bmatrix} \textit{conceive-rel} & \\ \textsc{conceiver} & i \\ \textsc{conceived} & j \end{bmatrix} \end{bmatrix}$

This accounts for a host of linking facts in verbs as varied as *like*, *enjoy*, *invent*, *claim*, and *murder*, assuming these verbs belong to the type *conceive-pred*. It explains the well known contrast between experiencer-subject *fear* and experiencer-object *frighten* verbs: *fear* entails that its subject has some notion of its object, so

The tourists feared the lumberjacks entails that the tourists are aware of the lumberjacks. But the object of *frighten* need not have a notion of its subject: in *The lumberjacks frightened the tourists (by cutting down a large tree that crashed right in front of them*, the tourists may not be aware of the lumberjacks' existence.

Two other linking rules appeared in Wechsler (1995). One stated that 'affected themes,' that is, participants that are entailed to undergo a change, map to the object, rather than subject, of a transitive verb. Another pertained to stative transitive verbs entailing a part-whole relation between the two participants, such as *include* and *contain*: the whole maps to the subject and the part to the object.

The linking constraints do not rely on a total ordering of thematic roles, nor on an exhaustive asssignment of thematic role types to every semantic role in a predicator. Instead, a small set of partial orderings of semantic roles, based on lexical entailments, suffices to account for the linking patterns of a wide range of verbs. This insight was adopted in a slightly different guise in work by Davis (1996; 2001) and Davis & Koenig (2000), who develop a more elaborated representation of lexical semantics, with which simple linking constraints can be stated. The essence of this approach is to posit a small number of dyadic semantic relations such as *act-und-rel* ('actor-undergoer relation') with attributes ACT(OR) and UND(ERGOER) that serve as intermediaries between semantic roles and syntactic arguments (akin to the notion of Generalized Semantic Roles discussed in Van Valin 1999).

What are the truth conditions of *act-und-rel*? Following Fillmore (1977), Dowty (1991), and Wechsler (1995b), Davis and Koenig note that many of the pertinent lexical entailments come in related pairs. For instance, one of Dowty's entailments is that one participant causally affects another, and of course the other is entailed to be causally affected. Another involves the entailments in Wechsler's notion rule (16); one participant is entailed to have a notion of another. These entailments of paired participant types characterize classes of verbs (or other predicators), and can then be naturally represented as dyadic relations in CONTENT. Collecting those entailments we arrive at a disjunctive statement of truth conditions:

(17) **act-und-rel**(x, y) is true iff x causes a change in y, or x has a notion of y.

We can designate the x participant in the pair as the value of ACTOR (or ACT) and y as the value of UNDERGOER (or UND), in a relation of type *act-und-rel*. Semantic arguments that are ACTOR or UNDERGOER will then bear at least one of the entailments characteristic of ACTORS or UNDERGOERS (Davis & Koenig 2000: 72). This then simplifies the statement of linking constraints for all of these paired participant types. Davis (1996) and Koenig & Davis (2001) argue that this obviates counting the relative number of proto-agent and proto-patient entailments, as advocated by Dowty (1991).

The linking constraints 18 and 19 state that a verb whose semantic CONTENT is of type *act-und-rel* will be constrained to link the ACT participant to the the first element of the verb's ARG-ST list (its subject), and the UND participant to the second element of the verb's ARG-ST list (this is analogous to Wechsler's constraints based on partial orderings).

These linking constraints can be viewed as parts of the definition of lexical types, as in Davis (2001), where (18) defines a particular class of lexemes (or words).[3]

(18) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix}\textsc{actor} & \boxed{1}\end{bmatrix} \\ \textsc{arg-st} & \left\langle \text{NP}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$$

(19) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix}\textsc{undergoer} & \boxed{2}\end{bmatrix} \\ \textsc{arg-st} & \left\langle \ldots, \text{NP}_{\boxed{2}}, \ldots \right\rangle \end{bmatrix}$$

(20) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix} \textit{cause-possess-rel} \\ \textsc{soa} \; \begin{bmatrix}\textsc{actor} & \boxed{3}\end{bmatrix} \end{bmatrix} \\ \textsc{arg-st} & \left\langle \textit{synsem} \right\rangle \oplus \left\langle \text{NP}_{\boxed{3}}, \ldots \right\rangle \end{bmatrix}$$

The first constraint, in (18), links the value of ACT (when not embedded within another attribute) to the first element of ARG-ST. The second, in (19), merely links the value of UND (again, when not embedded within another attribute) to some NP on ARG-ST. Given this understanding of how the values of ACTOR and UNDERGOER are determined, these constraints cover the linking patterns of a wide range of transitive verbs: *throw* (ACT causes motion of UND), *slice* (ACT causes change of state in UND), *frighten* (ACT causes emotion in UND), *imagine* (ACT has a notion of UND), *traverse* (ACT "measures out" UND as an incrememtnal theme), and *outnumber* (ACT is superior to UND on a scale).

[3] Alternatively, (18) (and other linking constraints) can be recast as implicational constraints on lexemes or words (Koenig & Davis 2003). (i) is an implicational constraint indicating that a word whose semantic content includes an ACTOR role must map that role to the initial item in the ARG-ST list.

(i) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix}\textsc{actor} & \boxed{1}\end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \textsc{arg-st} & \left\langle \text{NP}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$$

The third constraint, in (20), links the value of an ACT attribute embedded within a SOA attribute to an NP that is second on ARG-ST. This constraint accounts for the linking of the (primary) object of ditransitives. In English, these verbs (*give*, *hand*, *send*, *earn*, *owe*, etc.) involve (prospective) causing of possession (Pinker 1989; Goldberg 1995), and the possessor is represented as the value of the embedded ACT in (20). There could be additional constraints of a similar form in langages with a wider range of ditransitive constructions; conversely, such a constraint might be absent in languages that lack ditransitives entirely. As mentioned earlier in this section, the range of subcategorization options varies somewhat from one language to another.

The KEY attribute in (18) – (20) also requires explanation. The formulation of linking constraints here employs the architecture used in Koenig & Davis (2006), in which the semantics represented in CONTENT values is expressed as a set of *elementary predications*, formalized within Minimal Recursion Semantics (Copestake et al. 2001; 2005). Each elementary predication is a simple relation, but the relationships among them may be left unspecified. For linking, one of the elementary predications is designated the KEY, and it serves as the locus of linking. This allows us to indicate the linking of participants that play multiple roles in the denoted situation. The KEY selects one relation as the "focal point," and the other elementary predications are then irrelevant as far as linking is concerned. The choice of KEY then becomes an issue demanding consideration; we will see in the discussion of argument alternations in Section 5 how this choice might account for some alternation phenomena.

Note too that these linking constraints are treated as constraints on classes in the lexical hierarchy (see Chapter ??). One consequence of this fact merits brief mention. Constraint (19), which links the value of UND to some NP on ARG-ST, is a specification of one class of verbs. Not all verbs (and certainly not all other predicators, such as nominalizations) with a CONTENT value containing an UND value realize it as an NP. Verbs obeying this constraint include the transitive verbs noted above, and intransitive "unaccusative" verbs such as *fall* and *persist*. But some verbs with both ACT and UND attributes in their CONTENT are intransitive, such as *impinge (on)*, *prevail (on*, and *tinker (with)*. Interactions with other constraints, such as the requirement that verbs (in English, at least) have an NP subject, determine the range of observed linking patterns.

These linking constraints also assume that the proto-role attributes ACTOR, UNDERGOER, and SOA are appropriately matched to entailments, as described above. Other formulations are possible, such as that of Koenig & Davis (2003), where the participant roles pertinent to each lexical entailment are represented in CONTENT by corresponding, distinct attributes.

In addition to the linking constraints, there may be some very general well-formedness conditions on linking. We rarely find verbs that obligatorily map one semantic role to two distinct members of the ARG-ST list that are both expressed overtly. A verb meaning 'eat', but with that disallowed property, could appear in a ditransitive sentence like (4.1), with the meaning that Pat ate dinner, and his dinner was a large steak.

(21) *Pat ate dinner a large steak.

Typically semantic arguments map to at most one (overtly expressed) ARG-ST list item (Davis 2001: 262-268).

4.2 Linking oblique arguments

In this section we discuss linking of oblique arguments, that is, PP's and oblique case marked NP's. In some instances, a verb's selection of a particular preposition appears at least partly arbitrary; it is hard to explain why we *hanker after* and *yearn for*, but we don't **yearn after*. In these cases, the choice of preposition may be stipulated by the individual lexical entry. But as Gawron (1986) and Wechsler (1995a) have shown, many prepositions are semantically meaningful. *For* in the above-mentioned cases, and in *look for*, *wait for*, and *aim for* is surely not a lexical accident. And in the cases like *cut with*, *with* is used in an instrumental sense, denoting a *use-rel* relation, as with verbs that either allow (*eat*) or require (*cut*) an instrument. Davis (1996; 2001) adopts the position of Gawron and Wechsler in his treatment of linking to PPs As an example of this kind of account, the linking type in (22) characterizes a verb selecting a *with*-PP. The PP argument is linked from the RELS list rather the KEY.

(22)
$$\left[\begin{array}{ll} \text{CONTENT} & \left[\begin{array}{ll} \text{KEY} & \boxed{1} \\ \text{RELATIONS} & \left\langle \boxed{1}, \boxed{2}\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{use-rel}} \\ \text{ACT} & a \\ \text{UND} & u \\ \text{SOA} & s \end{array}\right], \ldots \right\rangle \end{array}\right] \\ \text{ARG-ST} & \left\langle \ldots, \text{PP}_{with}{:}\boxed{2} \ldots \right\rangle \end{array}\right]$$

Apart from the details of individual linking constraints, we have endeavored here to describe how linking can be modeled in HPSG using the same kinds of constraints used ubiquitously in the framework. Within the hierarchical lexicon, constraints between semantically defined classes and syntactically defined ones, can furnish an account of linking patterns, and there is no resort to additional mechanisms such as a thematic hierarchy or numerical comparison of entailments.

4.3 To what extent does meaning predict linking?

The framework outlined above allows us to address the following question: how much of linking is strictly determined by semantic factors, and how much is left open to lexically arbitrary subcategorization specifications, or perhaps subject to other factors?

Subcategorization— the position and nature of ARG-ST elements, in HPSG terms— is evidently driven to a great extent by semantics, but debate continues about how much, and which components of semantics are involved. Views have ranged from the strict, highly constrained relationship in which lexical semantics essentially determines syntactic argument structure, to a looser one in which some elements of subcategorization may be stipulated. Among the first camp are those who espouse the Uniformity of Theta Assignment Hypothesis proposed in Baker (1988: 46) or Baker (1997), which maintains that "identical thematic relationships between items are represented by identical structural relationships" in the syntax. With regard to the source of diathesis alternations, Levin (1993: 12-13) notes that "studies of these properties suggest that argument structures might in turn be derivable to a large extent from the meaning of words", and accordingly "pursues the hypothesis of semantic determinism seriously to see just how far it can be taken."

Others, including Pollard & Sag (1987) (Section 5.3) and Davis (2001) (Section 5.1), have expressed caution, pointing out cases where subcategorization and diathesis alternations seem to be at least partly arbitrary. Pollard & Sag note contrasts like these:

(23) a. Sandy spared/*deprived Kim a second helping.
     b. Sandy *spared/deprived Kim of a second helping. (Pollard & Sag 1987: ex. 214–215)

And Davis provides these pairs of semantically similar verbs with differing subcategorization requirements:

(24) a. Few passengers waited for/awaited the train.
     b. Homer opted for/chose a chocolate frosted donut.
     c. The music grated on/irritated the critics. (Davis 2001: ex. 5.4)

Other cases where argument structure seems not to mirror semantics precisely include raising constructions, in which one of a verb's direct arguments bears no semantic role to it at all. Similarly, overt expletive arguments cannot be seen as deriving from some participant role in a predicator's semantics. Like the examples above, these phenomena suggest that some aspects of subcategorization are specified independently of semantics.

Another point against strict semantic determination of argument structure comes from cross-linguistic observations of subcategorization possibilities. It is evident, for example, that not all languages display the same range of direct argument mappings. Some lack ditransitive constructions entirely (Halkomelem), some allow them across a limited semantic range (English), some quite generally (Georgian), and a few permit tritransitives (Kinyarwanda and Moro). Gerdts (1992) surveys about twenty languages and describes consistent patterns like these. The range of phenomena such as causative and applicative formation in a language is constrained by what she terms its "relational profile;" this includes, in HPSG terms, the number of direct NP arguments permitted on its ARG-ST lists. Again, it is unclear that underlying semantic differences across languages in the semantics of verbs meaning 'give' or 'write' would be responsible for these general patterns.

Summarizing, there is much evidence tempting us to derive the contents of ARG-ST solely from lexical semantics. If this ultimately proves feasible, then ARG-ST serves more as a convenient interface notion with little possibility of independently expressing strictly syntactic aspects of subcategorization. This view, however satisfying it might be, does not accord with our current best understanding of the syntactic and semantic evidence. In the following sections we delve into some of the nuances that make linking more than a simple rendering of lexical semantics. We begin by noting a point on which HPSG accounts of linking differ from many others— the absence of traditional thematic roles.

4.4 HPSG and thematic roles

The ARG-ST list constitutes the syntactic side of the mapping between semantic roles and syntactic dependents. As ARG-ST is merely an ordered list of arguments, without any semantic "labels," it contains no counterparts to thematic roles, such as AGENT, PATIENT, THEME, or GOAL. Thematic roles like these, however, have been a mainstay of linking in generative grammar since Fillmore (1968) and have antecedents going back to (Pāṇini. Ranking them in a *thematic hierarchy*, and labeling each of a predicator's semantic roles with a unique thematic role, then yields an ordering of roles analogous to the ordering on the ARG-ST list. Indeed, it would not be difficult to import this kind of system into HPSG, as a means of determining the order of elements on the ARG-ST list. However, HPSG researchers have generally avoided using a thematic hierarchy, for reasons we now briefly set out.

Fillmore (1968) and many others thereafter have posited a small set of disjoint, thematic roles, with each semantic role of a predicator assigned exactly one thematic role. Thematic hierarchies depend on these properties for a consistent linking theory. But they do not hold up well to formal scrutiny. Jackendoff (1987) and Dowty (1991) note (from somewhat different perspectives) that numerous verbs have arguments not easily assigned a thematic role from the typically posited inventory (e.g., the objects of *risk*, *blame*, and *avoid*), that more than one argument might sensibly be assigned the same role (e.g., the subjects and objects of *resemble*, *border*, and some alternants of commercial transaction verbs), and that multiple roles can be sensibly assigned to a single argument (the subjects of verbs of volitional motion are like both an AGENT and a THEME). In addition, consensus on the inventory of thematic roles has proven elusive, and some, notoriously THEME, have resisted clear definition. Work in formal semantics, including Ladusaw & Dowty (1988), Dowty (1989), Landman (2000), and Schein (2002), casts doubt on the prospects of assigning formally defined thematic roles to all of a predicator's arguments, at least in a manner that would allow them to play a crucial part in linking. Thematic role types seem to pose problems, and there are alternatives that avoids those problems. As Carlson (1998) notes about thematic roles, "It is easy to conceive of how to write a lexicon, a syntax, a morphology, a semantics, or a pragmatics without them."

4.5 CONTENT decomposition and ARG-ST

Instead of thematic role types, lexical decomposition is typically used in HPSG to model the semantic side of the linking relation. The word meaning represented by the CONTENT value is decomposed into a set of elementary predications that share arguments, as described in Section 4 above. Lexical decompositions cannot be directly observed, but the decompositions are justified indirectly by the roles they play in the grammar. Decompositions play a role in at least the following processes:

- *Linking*. As described in Section 4, linking constraints are stated on semantic relations like *act-und-rel* ('actor-undergoer relation'), so those relations must be called out in the CONTENT field.

- *Sublexical scope.* Certain modifiers can scope over a part of the situation denoted by a verb (Dowty 1979).

  (25) John sold the car, and then he bought it again.

  In this sentence the adverb *again* either adds the presupposition that John bought it before, or, in the more probable interpretation, it adds the presupposition that *the result of buying the car* obtained previously. The result of buying a car is owning it, so this sentence presupposes that John previously owned the car. Thus the decomposition of the verb *buy* in (36) below includes the *possess-rel* ('possession relation') holding between the buyer and the goods. This is available for modification by adverbials like *again*.

- *Argument alternations.* Some argument alternations can be modeled as highlighting of different portions of a single lexical decomposition. See Section 5.

In general, sublexical decompositions are included in the CONTENT field only insofar as they are visible to the grammar for processes like these.

The ARG-ST feature lies at the syntax side of the linking relation. Much like the CONTENT field, the ARG-ST items are justified only insofar as they are visible to the syntax. Many ARG-ST list items are obviously justified by being explicitly expressed as subject and complement phrases or as affixal pronouns. Certain implicit arguments appear if they are subject to the binding theory as applied to the ARG-ST list (as discussed in Section 3.1 above).

Implicit arguments can also participate in the syntax, and therefore appear on the ARG-ST list, by acting as controllers of adjunct clauses. For example, English rationale clauses like the infinitival phrase in (26a) are controlled by the agent argument in the clause, *the hunter* in this example. The implicit agent of a short passive can likewise control the rationale clause as shown in(26b). But the middle in (26c) lacks an implicit agent that is capable of controlling, even though native speakers assume that some agent must have caused the gun to load. This contrast was observed by Keyser & Roeper (1984) and confirmed in experimental work by Mauner & Koenig (2000).

(26) a. The shotgun was loaded quietly by the hunter to avoid the possibility of frightening off the deer.

b. The shotgun was loaded quietly to avoid the possibility of frightening off the deer.

c. *The shotgun had loaded quietly to avoid the possibility of frightening off the deer.

If the syntax of control is specified such that the controller of the rationale clause is an (agent) argument on the ARG-ST list of the verb, then this contrast is captured by assuming that the agent appears on the ARG-ST list of the passive verb but not the middle.

4.6 Modal transparency

Another observation concerning lexical entailments and linking was developed by Koenig & Davis (2001), who point out that linking appears to ignore modal elements of lexical semantics, even when those elements invalidate entailments (expanding on an observation implicit in Goldberg 1995). For instance, there are various English verbs displaying linking patterns like the ditransitive verbs of possession transfer *give* and *hand*, but which denote situations in which the transfer need not, or does not, take place. Thus, *offer* describes a situation where the transferor is willing to effect the transfer, *owe* one in which the transferor should effect the transfer but has not yet, *promise* describes a situation where the transferor commits to effect the transfer, and *deny* one in which the transferor does not effect the contemplated transfer. Koenig and Davis argue that modal elements should be clearly separated in CONTENT values from the representations of predicators and their arguments. (27) exemplifies this factoring out of sublexical modal information from core situational information. This pattern of linking functioning independently of sublexical modal information applies not only to these ditransitive cases, but also to verbs involving possession (cf. *own* and *obtain*, vs. *lack*, *covet* and *lose*), perception (*see* vs. *ignore* and *overlook*), and carrying out an action (*manage* vs. *fail* and *try*). Whatever the role of lexical entailments in linking, then, the modal information should be factored out, since the entailments canonically driving, e.g., the ditransitive linking patterns of verbs like *give* and *hand*, do not hold of *offer*, *owe*, or *deny*. The constraints in (18)-(20) need only been minimally altered to target the value of SIT-CORE within the representation of relation.

(27) The lexical semantic representation of *promise* (Koenig & Davis 2001: 101)

$$\begin{bmatrix} \textit{promise-sem} \wedge \textit{cause-possess-sem} & & \\ \textsc{sit-core} & \boxed{3} & \begin{bmatrix} \textit{cause-possess-rel} & \\ \textsc{actor} & \boxed{1} \\ \textsc{undergoer} & \boxed{2} \\ \textsc{soa} & \begin{bmatrix} \text{sit-core} & \boxed{5} \begin{bmatrix} \textit{have-rel} & \\ \textsc{actor} & \boxed{2} \\ \textsc{undergoer} & \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textsc{modal-base} & & \left\langle \begin{bmatrix} \textit{deontic-mb} \wedge \textit{condit-satis-mb} \\ \textsc{soa}\ \boxed{3} \end{bmatrix} \right\rangle \end{bmatrix}$$

5 The semantics and linking of argument alternations

A verb can often occur in varied syntactic contexts, as *find* does in (28); these are termed *valence alternations* or *diathesis alternations*, in reference to their different argument structuress. Levin (1993) lists around 50 kinds of alternations in English, and there are still more, including the alternation illustrated in (28).

(28) a. I found that the chair was comfortable
  b. I found the chair to be comfortable
  c. I found the chair comfortable

Another well studied alternation, the locative alternation, is exemplified by the two uses of *spray* in (29).

(29) a. $spray_{loc}$: Joan sprayed the paint onto the statue.
  b. $spray_{with}$: Joan sprayed the statue with paint.

It is typically assumed that these two different uses of *spray* in (29) have slightly different meanings, with the statue being in some sense more affected in the *with* alternant. This exemplifies the "holistic" effect of direct objecthood, which we will return to. Here, we will examine how semantic differences between alternant relate to their linking patterns. The semantic side of linking has often been devised with an eye to syntax (e.g., Pinker (1989), and see Koenig & Davis (2006) for more examples). There is a risk of stipulation here, without independent evidence for these semantic differences. In the case of locative alternations, though, the meaning difference between (29a) and (29b) is easily stated (and Pinker had the right intuition), as (29b) entails (29a), bot not conversely. Informally, (29a) describes a particular kind of caused motion situation, while (29b) describes a situation in which this kind of caused motion additionally results in a caused change of state. The difference is depicted in the two structures in (30).

(30) a. cause (Joan, go (paint, to (statue)))

b. act-on (Joan, statue, by (cause (Joan, go (paint, to (statue)))))

This description of the semantic difference between sentences (29a) and (29b) provides a strong basis for predicting their different argument structures. But we still need to explain how linking principles give rise to this difference. Pinker's account rests on semantic structures like (30), in which depth of embedding reflects sequence of causation, with ordering on arg-st stemming from depth of semantic embedding, a strategy adopted in Davis (1996; 2001). This is one reasonable alternative, although the resulting complexity of some of the semantic rpresentations raises valid questions about what independent evidence supports them. An alternative appears in Koenig & Davis (2006), who borrow from Minimal Recursion Semantics (see the chapter on Semantics for an introduction to MRS). MRS "flattens" semantic relationss, rather than embedding them in one another, so the arrangment of these *elementary predications*, as they are temed, is of less import. They posit a relations (or rels) attribute that collects a set of elementary predications, each representing some part of the predicator's semantics. A key attribute specifies a particular member of rels as the relevant one for linking (of direct syntactic arguments). In the case of (29b) the key is the caused change of state description. These MRS-style representations of the two alternants of *spray*, with different key values, are shown in (31) and (32).

(31)
$$\begin{bmatrix} \text{KEY} & \boxed{5} \\ \text{RELATIONS} & \left\langle \boxed{5} \begin{bmatrix} \textit{spray-ch-of-loc-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-loc-rel} & \\ \text{FIG} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

(32)
$$\begin{bmatrix} \text{KEY} & \boxed{3} \begin{bmatrix} \textit{spray-ch-of-st-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-st-rel} & \\ \text{UND} & \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{RELATIONS} & \left\langle \boxed{3}, \begin{bmatrix} \textit{use-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \boxed{3} \end{bmatrix}, \begin{bmatrix} \textit{spray-ch-of-loc-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-loc-rel} & \\ \text{FIG} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Generalizing from this example, one possible characterization of valence alternations, implicit in Koenig & Davis (2006), is as systematic relations between two sets of lexical entries in which the RELS of any pair of related entries are in a subset/subset relation (a weaker version of that definition would merely require an overlap between the RELS values of the two entries). Consider another case; (33) illustrates the causative-inchoative alternation, where the intransitive alternant describes only the change of state, while the transitive one ascribes a explicit causing agent.

(33) a. John broke the window.
 b. The window broke.

Under a MRS representation, the change of state relation is a separate member of RELS; it is also included in the RELS of the transitive alternant, which contains a cause relation as well. Again, the RELS value of one member of each pair of related entries is a subset of the RELS value of the other.

Many other alternations linvolve one argument shifting from direct to oblique. Some English examples include conative, locative preposition drop, and *with* preposition drop alternations, as shown in (34):

(34) a. Rover clawed (at) Spot.
 b. Bill hiked (along) the Appalachian Trail.
 c. Burns debated (with) Smithers.

It is well known that the direct objects in these alternations seem to be "affected" more than their oblique counterparts. So if Rover clawed Spot, we infer that Spot was subjected to direct contact with Rover's claws and may have been injured by them, while if Rover merely clawed *at* Spot, no such inference can be made. Similarly, to say the one has hiked the Appalachian Trail suggests that one has hiked its entire length, not merely hiked along some portion of it. This holistic effect is not so evident in cases like (34c), though the direct object variant suggests that a formally organized debate took place, while the *with* variant could just describe an informal discussion. How might these varying intuitions related to "affectedness" relate to lexical semantic representations like those in (31) and (32)? Beavers (2010) provides one analytical advance in this direction, similar to the subset relationship between RELS values described above. He generalizes from affectedness to strength of entailments, where one semantic role's entailments are stronger than another's if and only if the set of entailments characterizing the second role also hold of the first. That is, what is true for any particpant that bears the first role will be true for any particpant that bears the second, but not necessarily the converse. His *Morphosyntactic Alignment Principle* then relates this to linking, as stated in (35), where an "L-thematic role" is a linguistically relevant semantic role:

(35) When participant x may be realized as either a direct or oblique argument of verb V, it bears L-thematic role R as a direct argument and L-thematic role $Q \subseteq_M R$ as an oblique. (Beavers 2010: 848)

Here, $Q \subseteq_M R$ means that Q is a "minimally weaker" role than R; in other words, there is no role P in the predicator such that $Q \subset P \subset R$. Thus, the substantive claim is essentially that the MAP rules out "verbs where the alternating participant has MORE lexical entailments as an oblique than the corresponding object realization" (Beavers 2010: 849).

The entailments Beavers employs differ somewhat from those we have discussed here, involving quantized change, nonquantized change, potential for change (where change can refer to change in location, possession, state, or something more abstract), furnishing the clear ordering by strength that is central to his proposal. But they do resemble entailments of semantic relations we have represented as elementary predications, such as incremental theme, change of state, and possession, along with the modal effects described in Koenig & Davis (2001). Thus the notion of a stronger role in Beavers' analysis has a rough analog in terms of whether a particular elementary predication is present in the semantics of a particular alternant. And only if an elementary predication is present, can it be designated as the KEY, and its roles linked directly. For example, in (31), there is nothing representing affectedness of the location, while in (32), there is, and it is designated as the KEY. As noted earlier, the semantics in (32) represents this additional entailment borne by the location argument. However, we are not aware of any simple, general way to represent Beavers' MAP within the EP-based model of Koenig & Davis (2006). Indeed, there is an aspect of Beavers' view that seems more in accord with numerical comparison approaches such as those of Dowty (1991) and Ackerman & Moore (2001), in that role strength is determined by the number of entailments that hold of it relative to others.

Having outlined the semantic basis of the different linking patterns of alternating verbs, we briefly take up three other issues. First is the question of how the alternants are related to one another. Second is how KEY selection has been used to account not just for alternants of the same verb, but for (nearly) synonymous verbs whose semantics contain the same set of elementary predications. Third is whether passives, which arguably do not differ semantically from their active counterparts, should be assimilated with other alternations or treated distinctly, as a kind of non-canonical lexical item.

The hypothesis pursued in Davis (1996; 2001) is that most alternations are the consequence of classes of lexical entries having two related meanings. This follows researchers such as Pinker (1989) and Levin (1993) in modeling subcategorization alternations as underlyingly meaning alternations. This change in meaning is crucial in Koenig & Davis (2006) KEY shifts as well. In some cases, the value of the RELS attribute of the two valence alternates differ (as in the two alternates of *spray* in *spray/load* alternation we discussed earlier). In some cases, the alternation might be different construals of the same event for some verbs, but not others, as Rappaport Hovav & Levin (2008) claim for the English ditransitive alternations, which adds the meaning of transfer for verbs like *send*, but not for verbs like *promise*; a KEY change would be involved (with the addition of a *cause-possess-rel*) for the first verb only. But KEY shifts and diathesis alternations do not always involve a change in meaning. The same elementary predications can be present in as the CONTENT values of two alternants, with each alternant designating a different elementary predication as the KEY.

Koenig & Davis propose this not only for cases in which there is no obvious meaning difference betwen two alternants of a verb, but also for different verbs that appear to be truth-conditionally equivalent, one famous example being the verbs of commercial exchange *buy* and *sell* (but see Van Valin (1999: 387-388), Levin & Rappaport Hovav (2005: 20), and Wechsler (2005a) for arguments that *buy* and *sell* are not equivalent). Koenig and Davis argue that a commercial event involves two reciprocal actions, an exchange of goods (which involves giving goods and obtaining goods) and an exchange of money (which involves giving money and obtaining money). Individual verbs might select one or the other these four relations, thus accounting for the differences in subject and object selection. As shown in (36) and (37), each of these verbs contains four elementary predications: one *exch-give-rel* and one *obtain-rel* for the transfer of goods, and one of each for the counter-transfer of money. *Buy* designates the *obtain-rel* representing the transfer of goods as the KEY, while *sell* designates the *exch-give-rel* representing the transfer of goods as the KEY. Other verbs, such as *pay* or *charge*, choose elementary predications representing the counter-transfer as the KEY. In all cases, the same linking constraints apply between the KEY and the ARG-ST list, yielding the different argument realizations of these verbs while preserving their underlying semantic commonality. The relevant portions of the entries for *buy* and *sell* in (36) and (37) below illustrate: critically, the KEY relation for *buy* is not the same as that for *sell*.

(36) A representation of the relevant parts of the lexical entry for *buy*:

$$
\begin{bmatrix}
\textsc{content} & \begin{bmatrix}
\textsc{key} & \boxed{7} \\
\textsc{relations} & \left\langle
\begin{array}{l}
\boxed{7}\begin{bmatrix}
\textsc{act} & \boxed{1}\ \textit{(buyer)} \\
\textsc{und} & \boxed{2} \\
\textsc{soa} & \boxed{5}\begin{bmatrix}\textit{possess-rel} & \\ \textsc{act} & \boxed{1} \\ \textsc{und} & \boxed{2}\ \textit{(goods)}\end{bmatrix}
\end{bmatrix},
\begin{bmatrix}
\textsc{act} & \boxed{1}\ \textit{(buyer)} \\
\textsc{und} & \boxed{4} \\
\textsc{soa} & \boxed{6}\begin{bmatrix}\textit{possess-rel} & \\ \textsc{act} & \boxed{3}\ \textit{(seller)} \\ \textsc{und} & \boxed{4}\ \textit{(money)}\end{bmatrix}
\end{bmatrix}, \\
\begin{bmatrix}
\textit{obtain-rel} & \\
\textsc{act} & \boxed{3}\ \textit{(seller)} \\
\textsc{und} & \boxed{4} \\
\textsc{soa} & \boxed{6}\begin{bmatrix}\textit{possess-rel} & \\ \textsc{act} & \boxed{3}\ \textit{(seller)} \\ \textsc{und} & \boxed{4}\ \textit{(money)}\end{bmatrix}
\end{bmatrix},
\boxed{8}\begin{bmatrix}
\textit{exch-give-rel} & \\
\textsc{act} & \boxed{3}\ \textit{(seller)} \\
\textsc{und} & \boxed{2} \\
\textsc{soa} & \boxed{5}\begin{bmatrix}\textit{possess-rel} & \\ \textsc{act} & \boxed{1} \\ \textsc{und} & \boxed{2}\ \textit{(goods)}\end{bmatrix}
\end{bmatrix}
\end{array}
\right\rangle
\end{bmatrix} \\
\textsc{arg-st} & \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2}, \text{PP(}\textsc{from}\text{):}\boxed{8} \rangle
\end{bmatrix}
$$

(37) A representation of the relevant parts of the lexical entry for *sell*:

$$
\begin{bmatrix}
\textsc{content} & \begin{bmatrix}
\textsc{key} & \boxed{7} \\
\textsc{relations} & \left\langle
\boxed{7}\begin{bmatrix}
\textsc{act} & \boxed{1}\ \textit{(seller)} \\
\textsc{und} & \boxed{2} \\
\textsc{soa} & \boxed{6}\begin{bmatrix}\textit{POSSESS-REL} & \\ \textsc{act} & \boxed{3}\textit{(buyer)} \\ \textsc{und} & \boxed{2}\ \textit{(goods)}\end{bmatrix}
\end{bmatrix},
\begin{bmatrix}
\textsc{act} & \boxed{1}\ \textit{(seller)} \\
\textsc{und} & \boxed{4} \\
\textsc{soa} & \boxed{6}\begin{bmatrix}\textit{POSSESS-REL} & \\ \textsc{act} & \boxed{1} \\ \textsc{und} & \boxed{4}\ \textit{(money)}\end{bmatrix}
\end{bmatrix}
\right\rangle
\end{bmatrix} \\
\textsc{arg-st} & \langle \text{NP:}\boxed{1}, \text{NP:}\boxed{2}, \text{PP(}\textsc{to}\text{):}\boxed{7} \rangle
\end{bmatrix}
$$

As a final example of semantic alternations in the fine-grained meaning possibilities of verbs, we consider here the source-final product alternation exemplified in (38) where the direct object can be either the final product or the material source of the final product. Davis proposes that the (38a) sentences involve an alternation between the two meanings of entries represented in (39). We adapt Davis (2001) to make it consistent with Koenig & Davis (2006) and also treat the alternation as an alternation of meaning of *entries*. Note that in the meaning alternation described in (39), we use, informally, a double-headed arrow. One of the potential drawbacks of a lexical rule approach to valence alternations is that it requires selecting one or the other alternant as basic and the other as derived (e.g., is the inchoative or the causative basic?). This is not always an easy decision, as Goldberg (1995) or Levin & Rappaport Hovav (1994) have pointed out. Sometimes, morphology provides a clue, although in different languages the clues may point in different directions. French, and other Romance languages, use a "reflexive" clitic as a detransitivizing affix. In English, though, there is no obvious "basic" form or directionality. It is to avoid committing ourselves to a directionality in the meaning relation described in (39) that we eschews treating it as a lexical rule.

(38) a. Kim made/carved/sculpted/crafted a toy (out of the wood).

b. Kim made/carved/sculpted/crafted the wood into a toy.

(39)
$$\left[\begin{array}{ll} \text{CONTENT} & \left[\begin{array}{ll} \text{KEY} & \boxed{3}\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{affect-incr-th-rel}} \\ \text{ACT} & \boxed{1} \\ \text{SOA} & \boxed{2}\textit{(final product)} \end{array}\right] \\ \text{RELS} & \langle \boxed{3}, \boxed{4} \rangle \end{array}\right] \end{array}\right]$$

$$\longleftrightarrow$$

$$\left[\begin{array}{ll} \text{CONTENT} & \left[\begin{array}{ll} \text{KEY} & \boxed{4}\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{affect-incr-th-rel}} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \textit{(source material)} \\ \text{SOA} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{affect-incr-th-rel}} \\ \text{ACT} & \boxed{1} \\ \text{SOA} & \boxed{2}\textit{(final product)} \end{array}\right] \end{array}\right] \\ \text{RELS} & \langle \boxed{3}, \boxed{4} \rangle \end{array}\right] \end{array}\right]$$

Although most diathesis alternations can be modeled as alternations in meaning and as KEY shifts, some arguably cannot. We discuss the active/passive alternation here, but impersonals, as well as raising structures exemplified in (28) are good candidates too. The semantic relations of actives and long passives, as in (40), are practically identical and the difference between the two alternates is pragmatic in nature. Arguably, then, passives are a degenerate case of the subset relationship between RELS attributes, where the RELS values of the two entries are identical and so are the two entries' KEY. But this raises the question of whether linking in passives violates the constraints in (18)–(20), especially (18, which links the value of ACT to the first element of ARG-ST.

(40) a. Fido dug a couple of holes.

b. A couple of holes were dug by Fido.

One typical HPSG method for modeling valence alternations like passives is through lexical rules (see Chapter ??) with one alternant serving as input and the other as output; the main effect of the lexical rule in such an approach is to alter the ARG-ST of the input, going from the ARG-ST of *give* to that of *given* in (8). Critically, we must assume that the output cannot be subject to linking constraint (18), since the actor argument is not linked to the first member of the ARG-ST list. A simplified representation of what such a rule would look like is provided in (41) where we assume that the input to the rule must be transitive.

(41)
$$\begin{bmatrix} \textit{trans-vb} & \\ \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}}, \text{NP}_{\boxed{2}} \right\rangle \oplus \boxed{3}\,\textit{list} \end{bmatrix} \mapsto \begin{bmatrix} \text{HEAD} & \left[\text{VFORM} \ \ \textit{pass}\right] \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{2}} \right\rangle \oplus \boxed{3} \oplus \left\langle \text{PP}[\textit{by}]_{\boxed{1}} \right\rangle \end{bmatrix}$$

To sum up, in contrast to most meaning-driven alternations, valence alternations like the active/passive are modeled through the use of lexical rules that alter the ARG-ST of "base" entries. Which alternations pattern with active/passive and require positing lexical rules that alter a "base" entry's ARG-ST list is, as of yet, not settled. We turn now to roles that are putatively present semantically, but not realized syntactically at all.

6 Extended ARG-ST

Most of this chapter focuses on cases where semantic roles linked to the ARG-ST list are arguments of the verb's core meaning. But in quite a few cases, complements (or even subjects) of a verb are not part of this basic meaning; consequently, the ARG-ST list must be extended to include elements beyond the basic meaning. We consider three cases here, illustrated in (42)–(44).

Resultatives, illustrated in (42), express an effect, which is caused by an action of the type denoted by the basic meaning of the verb. The verb *fischen* 'to fish' is a simple intransitive verb (42a) that does not entail that any fish were caught, or any other specific effect of the fishing.

(42) a. dass er fischt
that he fishes
'that he is fishing'

b. dass er ihn leer fischt
that he it empty fishes
'that he is fishing it empty'

c. wegen der Leerfischung der Nordsee
because.of the empty.fishing of.the North.See.GEN
'because of the North Sea being fished empty'

In (42b) we see a resultative construction, with an object NP and a secondary predicate AP. The meaning is that he is fishing, causing it (the body of water) to become empty of fish. Müller (2002) posits a German lexical rule applying to the verb that augments the ARG-ST list with an NP and AP, and adds the causal semantics to the CONTENT (see Wechsler (2005b) for a similar analysis of English resultatives). The existence of deverbal nouns like *Leerfischung* 'fishing empty', which takes the body of water as an argument in genitive case (see 42c) confirms that the addition of the object is a lexical process, as noted by Müller (2002).

Romance clause-union structures as in (43) have long been analyzed as cases where the complements of the complements of a clause-union verb (*faire* in (43)) are complements of the clause-union verb itself (Aissen 1979).

(43) Johanna a fait manger les enfants.
Johanna have made eat the children
'Johanna had the children eat.'

Within HPSG, the "union" of the two verbs' dependents is modeled via the composition of ARG-ST lists of the clause union verb, following Hinrichs & Nakazawa (1994) (this is a slight simplification, see Chapter ?? for details).

Sentence (44) illustrates a slightly different point, namely that some semantic modifiers, such as *souvent* in (44), can be realized as complements, and thus should be added as members of ARG-ST (or members of the DEPS list if one countenances such an additional list).

(44) Mes amis m' ont souvent aidé.
My friends me have often helped
'My friends often helped me.'

Abeillé & Godard (1997) have argued that many adverbs including negative adverbs and negation in French are complements of the verb and Kim & Sag (2002) extended that view to some uses of negation in English. In contrast to resultatives, which affect the meaning of the verb, or to clause union, where one verb co-opts the argument structure of another verb, what is added to the ARG-ST list in these cases is typically considered a semantic adjunct and a modifier in HPSG (thus it selects the verb or VP via the MOD attribute).

Another case of an adjunct that behaves like a complement is found in (45). The clitic *en* expressing the cause of death is not normally an argument of the verb *mourir* 'die', but rather an adjunct (Koenig & Davis 2006):

(45) Il en est mort
He of.it is dead.PERF.PAST
'He died of it' (Koenig & Davis 2006, ex. 12a)

On the widespread assumption (at least within HPSG) that pronominal clitics are verbal affixes (see Miller & Sag 1997), the adjunct cause of the verb *mourir* must be represented within the entry for *mourir*, so as to trigger affixation by *en*. Bouma, Malouf & Sag (2001) discuss such cases and other cases where "adverbials" as they call them, can be part of a verb's lexical entry. To avoid mixing those adverbials with the argument structure list (and have to address their relative obliqueness with syntactic arguments of verbs), they introduce yet an additional list, the dependents list (abbreviated as DEPS) which includes the ARG-ST list but also a list of adverbials. Each adverbial selects for the verb on whose DEPS list it appears as argument, as shown in (46). But, of course, not all verb modifiers can be part of the DEPS list and Bouma, Malouf & Sag discuss at length some of the differences between the two kinds of "adverbials."

(46) $$verb \Rightarrow \begin{bmatrix} \text{CONT}|\text{KEY} & \boxed{2} \\ \text{HEAD} & \boxed{3} \\ \text{DEPS} & \boxed{1} \oplus \mathit{list}\left(\begin{bmatrix} \text{MOD} & \begin{bmatrix} \text{HEAD} & \boxed{3} \\ \text{key} & \boxed{2} \end{bmatrix} \end{bmatrix}\right) \\ \text{ARG-ST} & \boxed{1} \end{bmatrix}$$

Although the three cases we have outlined result in an extended ARG-ST, the ways in which this extension arises differ. In the case of resultatives, the extension results partly or wholly from changing the meaning in a way similar to Rappaport Hovav & Levin (1998): by adding a causal relation, the effect argument of this causal relation is added to the membership in the base ARG-ST list (see Section 5 for a definition of the attributes KEY and RELS; suffice it to say for now that a *cause-rel* is added to the list of relations that are the input of the rule). The entries of the clause union verbs are simply stipulated to include on their ARG-ST lists the syntatic arguments of their (lexical) verbal arguments. Finally, (negative) adverbs that select for a verb (VP) are added to the ARG-ST of the verb they select. A simplified representation of all three processes is provided in (47)-(49).

(47) $$\begin{bmatrix} \text{KEY} & \boxed{2} \\ \text{RELS} & \boxed{1}\langle \ldots \boxed{2} \ldots \rangle \end{bmatrix} \mapsto \begin{bmatrix} \text{KEY} & \boxed{3}\,\mathit{cause\text{-}rel} \\ \text{RELS} & \boxed{1} \oplus \boxed{3} \end{bmatrix}$$

(48) $$\begin{bmatrix} \text{ARG-ST} & \left\langle \ldots, \begin{bmatrix} \text{HEAD} & \mathit{verb} \\ \text{ARG-ST} & \boxed{1} \end{bmatrix} \right\rangle \bigcirc \boxed{1} \end{bmatrix}$$

(49) $$\begin{bmatrix} \text{ARG-ST} & \boxed{1} \end{bmatrix} \mapsto \begin{bmatrix} \text{ARG-ST} & \boxed{1} \bigcirc \langle \text{ADV}_{neg} \rangle \end{bmatrix}$$

7 Is ARG-ST universal?

In this section, we briefly consider the question of whether something akin to Figure 1 offers a satisfying account of the grammatical encoding of semantic arguments across all languages. Because of its role in accounting for the syntax of basic clauses, the presence of an ARG-ST list on lexical entries comes with expectations about the syntactic realization of semantic roles. In recent work, Koenig & Michelson (2015) argue these expectations are not universally borne out, based on data from Oneida (Norther Iroquoian). The only grammatical reflex of semantic arguments in Oneida, they argue, is inflectional: the referencing of semantic arguments by so-called pronominal prefixes, which are better thought of as agreement markers à la Evans (2002). Koenig and Michelson distinguish between grammatical and syntactic arguments. Grammatical arguments include not only syntactic arguments (that is, those on ARG-ST) but also inflectional referencing of semantic roles. Some ordering analogous to linking in other languages is present in Oneida, because the semantic roles are not arbitrarily associated with agreement morphemes, but this can be captured in an ordered list of semantic indices , called INFL-STR in this proposal. INFL-STR is part of the morphological information relevant to word-internal inflectional processes, what Anderson (1992) calls 'Morphosyntactic Representation' in his treatment of Georgian agreement markers. The ordering of semantic indices on INFL-STR insures that the predicator is properly inflected. For example, the prefix *lak-* occurs if a third singular masculine proto-agent argument is acting on a first singular proto-patient argument as in *lak-hlo·lí-heʔ* 'he tells me' (habitual aspect), whereas the prefix *li-* occurs if a first singular proto-agent argument is acting on a third masculine singular argument, as in *li-hlo·lí-heʔ* 'I tell him' (habitual aspect).

In (50) and (51) we show the distinction between grammatical arguments realized as syntactic arguments (as in most languages) and those that are not (as in Oneida). Here, Koenig and Michelson follow the encoding of linking constraints as implicational constraints, as in Koenig & Davis (2003), although nothing critical hinges on that choice.

(50) $\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} \ \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \langle \text{NP}, \ldots \rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \text{ARG-ST} \ \langle \text{NP:}\boxed{1}, \ldots \rangle \end{bmatrix}$

(51) $\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} \ \boxed{1}\textit{anim} \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{INFL-STR} \ \langle \boxed{1}, \ldots \rangle \end{bmatrix}$

(50) constrains the association between a cause and a *synsem* member of the ARG-ST list; (51) constrains the semantic *index* of the cause to be the first member of INFL-STR.

If a language like Oneida (Northern Iroquoian) only includes an ordering of semantic indices for inflectional purposes and constraints such as (51) and no

ARG-ST list, a number of predictions follow, which Koenig and Michelson claim are borne out. Briefly summarizing their evidence, the relation between semantic arguments and external phrases, when they occur, is not necessarily one of co-indexing, no binding constraints exist between external phrases (e.g., condition C violations can be found), there are no valence alternations, and no syntactic constraints on extraction. In other words, Oneida contains none of the evidence supporting the presence of an ARG-ST list and an ordering of syntactic arguments along an obliqueness hierarchy we have discussed in this chapter. The ARG-ST list may thus not a universal attribute of words, though present in the overwhelming majority of languages. Linking, understood as constraints between semantic roles and members of the ARG-ST list, is then but one possibility; constraints that relate semantic roles to an INFL-STR list of semantic indices is also an option. In languages that exclusively exploit that latter possibility, syntax is indeed simpler.

8 The lexical approach to argument structure

We end this chapter with a necessarily brief comparison between the approach to argument structure we describe in this chapter with other approaches to argument structure that have developed since the 1990's. This chapter describes a *lexical approach to argument structure*, which is typical of research in HPSG. The basic tenet of such approaches is that lexical items include argument structures, which represent essential information about potential argument selection and expression, but abstract away from the actual local phrasal structure. In contrast, *phrasal approaches*, which are common both in Construction Grammar and in transformational approaches such as Distributed Morphology, reject such lexical argument structures. Let us briefly review the reasons for preferring a lexical approach. (This section is drawn from Müller & Wechsler (2014b), which may be consulted for more detailed and extensive argumentation).

In phrasal approaches to argument structure, components of a verb's apparent meaning are actually 'constructional meaning' contributed directly by the phrasal structure. The linking constraints of the sort discussed above are then said to arise from the interaction of the verb meaning with the constructional meaning. For example, agentive arguments tend to be realized as subjects, not objects, of transitive verbs. On the theory presented above, that generalization is captured by the linking constraint (18), which states that the ACTOR argument of an *act-und-rel* ('actor-undergoer relation') is mapped to the initial item in the ARG-ST list. In a phrasal approach, the agentive semantics is directly associated with the subject position in the phrase structure. In transformational theories, a silent 'light verb' (usually called 'little *v*') heads a projection in the phrase structure and assigns the agent role to its specifier (the subject). In constructional theories, the phrase structure itself assigns the agent role. In either type of phrasal approach, the agentive component of the verb meaning is actually expressed by the phrasal structure into which the verb is inserted.

The lexicalist's predicate argument structure provides essential information for a verb's potential combination with argument phrases. If a given lexical entry could only combine with the particular set of phrases specified in a single VALfeature, then the lexical and phrasal approaches would be difficult to distinguish: whatever information the lexicalist specifies for each VALlist item could, on the phrasal view, be specified instead for the phrases realizing those list items. But crucially, the verb need not immediately combine with its specified arguments. Alternatively it can meet other fates: it can serve as the input to a lexical rule; it can combine first with a modifier in an adjunction structure; it can be coordinated with another word with the same predicate argument structure; instead of being realized locally, one or more of its arguments can be effectively transferred to another head's valence feature (raising or argument transfer); or arguments can be saved for expression in some other syntactic position (partial fronting). Here we consider two of these, lexical rules and coordination.

The predicate argument structure is abstract: it does not directly encode the phrase structure or precedence relations between this verb and its arguments. This abstraction captures the commonality across different syntactic expressions of the arguments of a given root.

(52) a. The rabbits were nibbling the carrots.
  b. The carrots were being nibbled (by the rabbits).
  c. a large, partly nibbled, orange carrot
  d. the quiet, nibbling, old rabbits
  e. the rabbit's nibbling of the carrots
  f. The rabbit gave the carrot a nibble.
  g. The rabbit wants a nibble (on the carrot).
  h. The rabbit nibbled the carrot smooth.

Verbs undergo morpholexical operations like passive (52d), as well as antipassive, causative, and applicative in other languages. They have cognates in other parts of speech such as adjectives (52c,d) and nouns (52e,f,g). Verbs have been argued to form complex predicates with resultative secondary predicates (52h), and with serial verbs in other languages.

The same root lexical entry *nibble*, with the same meaning, appears in all of these contexts. The effects of lexical rules together with the rules of syntax dictate the proper argument expression in each context. For example, if we call the first two arguments in an ARG-ST list (such as the one in (52) above) Arg1 and Arg2, respectively, then in an active transitive sentence Arg1 is the subject and Arg2 the object; in the passive, Arg2 is the subject and the referential index of Arg1 is optionally assigned to a *by*-phrase. The same rules of syntax dictate the position of the subject, whether the verb is active or passive. When adjectives are derived from verbal participles, whether active (*a nibbling rabbit*) or passive (*a nibbled carrot*), the rule is that whichever role would have been expressed as the subject of the verb is assigned by the participial adjective to the referent of the noun that it modifies, see Bresnan (1982) and Bresnan et al. (2015: Chapter 3). The phrasal approach, in which the agent role is assigned to the subject position, is too rigid.

Nor could this be solved by associating each syntactic environment with a different meaningful phrasal construction: an active construction with agent role in the subject position; a passive construction with agent in the *by*-phrase position; etc. The problem for that view is that that one lexical rule can feed another. In the example above, the output of the verbal passive rule (see (52d)) feeds the adjective formation rule (see (52e)).

A verb can also be coordinated with another verb with the same valence requirements. The two verbs then share their dependents. This causes problems for the phrasal view, especially when a given dependent receives different semantic roles from the two verbs. For example, in an influential phrasal analysis, Hale & Keyser (1993) derived denominal verbs like *to saddle* through noun incorporation out of a structure akin to [PUT a saddle ON x]. Verbs with this putative derivation routinely coordinate and share dependents with verbs of other types:

(53) Realizing the dire results of such a capture and that he was the only one to prevent it, he quickly [saddled and mounted] his trusted horse and with a grim determination began a journey that would become legendary.[4]

Under the phrasal analysis the two verbs place contradictory demands on a single phrase structure. But on the lexical analysis, this is simple V^0 coordination.

To summarize, a lexical argument structure is an abstraction or generalization over various occurrences of the verb in syntactic contexts. To be sure, one key use of that argument structure is simply to indicate what sort of phrases the verb must (or can) combine with, and the result of semantic composition; if that were

[4] http://www.jouetthouse.org/index.php?option=com_content&view=article&id=56&Itemid=63, 21.07.2012

the whole story then the phrasal theory would be viable. But it is not. As it turns out, this lexical valence structure, once abstracted, can alternatively be used in other ways: among other possibilities, the verb (crucially including its valence structure) can be coordinated with other verbs that have a similar valence structure; or it can serve as the input to lexical rules specifying a new word bearing a systematic relation to the input word. The phrasal approach prematurely commits to a single phrasal position for the realization of a semantic argument. In contrast, a lexical argument structure gives a word the appropriate flexibility to account for the full range of expressions found in natural language.

References

Abeillé, Anne & Danièle Godard. 1997. The syntax of French negative adverbs. In Paul Hirschbühler & F. Marineau (eds.), *Negation and polarity: syntax and semantics*, 1–17. Amsterdam: John Benjamins.

Ackerman, Farrell, Robert Malouf & John Moore. 2017. Symmetrical objects in Moro: challenges and solutions. *Journal of Linguistics* 53. 3–50.

Ackerman, Farrell, Robert Malouf & John Moore. 2018. Symmetric objects in Moro. http://csli-publications.stanford.edu/HPSG/2013/, accessed 2018-7-5. Paper presentend at the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin.

Ackerman, Farrell & John Moore. 2001. *Proto-properties and grammatical encoding: a correspondence theory of argument selection*. Stanford: CSLI Publications.

Aissen, Judith. 1979. *The syntax of causative constructions*. New York: Garland.

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Artawa, Ketut. 1994. *Ergativity and Balinese syntax*. Bundoora, Australia: La Trobe University dissertation.

Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press.

Baker, Mark. 1997. Thematic roles and syntactic structures. In Liliane Haegeman (ed.), *Elements of grammar*, 73–137. Kluwer.

Beavers, John. 2010. The structure of lexical meaning: why semantics really matters. *Language* 86. 821–864.

Bouma, Gosse, Robert Malouf & Ivan Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 3–86. Cambridge, MA/London: MIT Press.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Mark Wechsler. 2015. *Lexical-functional syntax*. 2nd edn. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9781119105664

Carlson, Greg. 1998. Thematic roles and the individuation of events. In Susan Rothstein (ed.), *Events and grammar*, 35–51. Dordrecht: Kluwer.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3. 281–332. DOI:10.1007/s11168-006-6327-9

Copestake, Ann, Alex Lascarides & Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th ACL meeting*. Toulouse, France.

Davis, Anthony. 1996. *Lexical semantics and linking in the hierarchical lexicon*. Stanford University dissertation.

Davis, Anthony. 2001. *Linking by types in the hierarchical lexicon*. Stanford: CSLI Publications.

Davis, Anthony & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.

Dowty, David. 1989. On the semantic content of the notion of 'thematic role'. In Gennaro Chierchia, Barbara Partee & Raymond Turner (eds.), *Properties, types, and meaning*, vol. 2, 69–129. Dordrecht: Kluwer.

Dowty, David R. 1979. *Word meaning and montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ*. Vol. 7 (Studies in Linguistics and Philosophy). Dordrecht, Netherlands: Reidel. DOI:10.1007/978-94-009-9473-7

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Evans, Nicholas. 2002. The true status of grammatical object affixes: evidence from Bininj Gun-Wok. In Nicholas Evans & Hans-Jürgen Sasse (eds.), *Problems of polysynthesis*, 15–50. Berlin: Akademie Verlag.

Fillmore, Charles. 1968. The case for case. In Emmon Bach & Robert Harms (eds.), *Universals in linguistic theory*, 1–87. New York: Holt, Rinehart & Winston.

Fillmore, Charles. 1977. The case for case reopened. In Peter Cole & Jerold Sadock (eds.), *Grammatical relations*, 59–81. New York: Academic Press.

Foley, William & Robert Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Gawron, Jean Mark. 1986. Situations and prepositions. *Linguistics and Philosophy* 9. 327–382.

Gerdts, Donna. 1992. Morphologically-mediated relational profiles. In Laura Buszard-Welcher et al. (ed.), *Proceedings of the 18th annual meeting of the berkeley linguistics society*, 322–337. Berkeley, CA: Berkeley Linguistics Society.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: The University of Chicago Press.

Greenberg, Joseph H. 1966. *Language universals*. The Netherlands: Mouton.

Guilfoyle, Eithne, Henrietta Hung & Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10(3). 375–414.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 53–109. Cambridge, MA: MIT Press.

Hinrichs, Erhard & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-driven Phrase-Structure Grammar*, 11–37. Stanford: CSLI Publications.

Jackendoff, Ray. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18. 369–411.

Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.

Keyser, Samuel & T. Roeper. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry*. 2nd ser. 15. 381–416.

Kim, Jong-Bok & Ivan Sag. 2002. Negation without head movement. *Natural Language and Linguistic Theory* 20(2). 339–412.

Koenig, Jean-Pierre & Anthony Davis. 2001. Sublexical modality and the structure of lexical semantic representations. *Linguistics and Philosophy* 24. 71–124.

Koenig, Jean-Pierre & Anthony Davis. 2003. Semantically transparent linking in HPSG. In Stefan Müller (ed.), *Proceedings of the HPSG 2003 conference*, 222–235. Stanford: CSLI Publications.

Koenig, Jean-Pierre & Anthony Davis. 2006. The KEY to lexical semantic representations. *Journal of Linguistics* 42. 71–108.

Koenig, Jean-Pierre & Karin Michelson. 2015. Invariance in argument realization: the case of Iroquoian. *Language* 91. 1–47.

Ladusaw, William & David Dowty. 1988. Toward a nongrammatical account of thematic roles. In Wendy Wilkins (ed.), *Syntax and semantics 21: thematic relations*, 62–74. New York: Academic Press.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures*. Dordrecht: Kluwer.

Levin, Beth. 1993. *English verb classes and alternations*. Chicago, IL: Chicago University Press.

Levin, Beth & Malka Rappaport Hovav. 1994. A preliminary analysis of causative verbs in English. In Lila Gleitman & Barbara Landau (eds.), *The acquisition of the lexicon*, 35–77. Cambridge, MA: MIT Press.

Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge, UK: Cambridge University Press.

Manning, Christopher. 1996. *Ergativity: argument structure and grammatical relations*. Stanford, CA: CSLI publications.

Manning, Christopher & Ivan A Sag. 1999. Dissociations between argument structure and grammatical relations. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation*, 63–78. Stanford: CSLI Publications.

Mauner, Gail & Jean-Pierre Koenig. 2000. Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language* 43. 110–134.

Miller, Philip & Ivan Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan & Stephen Mark Wechsler. 2014a. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Müller, Stefan & Stephen Mark Wechsler. 2014b. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Pinker, Steven. 1989. *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments*, 97–134. Stanford: CSLI Publications.

Rappaport Hovav, Malka & Beth Levin. 2008. The English dative alternation: the case for verb sensitivity. *Journal of Linguistics* 44. 129–167.

Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und prag-*

matischen Fundierung (Studien zur deutschen Grammatik 15), 171–211. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Schein, Barry. 2002. Events and the semantic content of thematic relations. In Gerhard Preyer & Georg Peter (eds.), *Logical form and language*, 263–344. Oxford: Oxford University Press.

Van Valin, Robert. 1999. Generalized semantic roles and the syntax-semantics interface. In *Empirical issues in formal syntax and semantics 2*, 373–388. The Hague, The Netherlands: Thesus.

Wechsler, Stephen. 1995a. Preposition selection outside the lexicon. In Raul Aranovich, William Byrne, Susanne Preuss & Matha Senturia (eds.), *Proceedings of WCCFL XIII*, 416–431. Stanford: CSLI Publications.

Wechsler, Stephen. 1995b. *The semantic basis of argument structure*. Stanford, CA: CSLI Publications.

Wechsler, Stephen. 1999. HPSG, GB, and the Balinese bind. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 179–195. Stanford, CA: CSLI publications.

Wechsler, Stephen. 2005a. What is right and wrong about little v. In Mila Vulchanova & Tor A. Åfarli (eds.), *Grammar and beyond – essays in honour of Lars Hellan*. Novus Press.

Wechsler, Stephen Mark. 2005b. Resultatives under the event-argument homomorphism model of telicity. In Nomi Erteschik-Shir & Tova Rapoport (eds.), *The syntax of aspect—Deriving thematic and aspectual interpretation*, 255–273. Oxford University Press.

Wechsler, Stephen & I Wayan Arka. 1998. Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory* 16. 387–442.

Chapter 10

Constituent order

Stefan Müller
Humboldt-Universität zu Berlin

This chapter discusses local ordering variants and how they can be analyzed in HPSG. So-called scrambling, the local reordering of arguments of a head, can be accounted for by assuming flat rules or binary branching rules with arbitrary order of saturation. The difference between SVO and SOV is explained by assuming different mappings between the argument structure list (a list containing all arguments of a head) and valence features for subjects and complements. The position of the finite verb in initial or final position in languages like German can be accounted for by flat rules and a separation between immediate dominance and linear precedence information or by something analogous to head-movement in transformational approaches. The chapter also addresses the analysis of languages allowing even more freedom than just scrambling arguments. It is shown how one such langauge, namely Warlpiri, can be analyzed with so-called constituent order domains allowing for discontinuous constituents. I discuss problems of domain-based approaches and provide an alternative account of Warlpiri that does not rely on discontinuous constituents.

1 Introduction

This chapter deals with constituent order, with a focus on local order variants. English is the language that is treated most thoroughly in theoretical linguistics but is probably also a rather uninteresting language as far as the possibilities of reordering constituents is concerned: the order of subject, verb, and object is fixed in sentences like (1):

(1) Kim likes bagels.

Of course, there is the possibility to front the object as in (2) but this is a special, non-local construction that is not the topic of this chapter but is treated in Borsley & Crysmann (2019), Chapter 14 of this volume.

Stefan Müller. 2019. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

(2) Bagels, Kim likes.

This chapter deals with scrambling (the local reordering of arguments) and with alternative placements of heads (called *head movement* in some theories). Examples of the former are the subordinate clauses in (3) and an example of the latter is given in (4):

(3) a. [weil] der Mann dem Kind das Buch gibt (German)
because the.NOM man the.DAT child the.ACC book gives

b. [weil] der Mann das Buch dem Kind gibt
because the.NOM man the.ACC book the.DAT child gives

c. [weil] das Buch der Mann dem Kind gibt
because the.ACC book the.NOM man the.DAT child gives

d. [weil] das Buch dem Kind der Mann gibt
because the.ACC book the.DAT child the.NOM man gives

e. [weil] dem Kind der Mann das Buch gibt
because the.DAT child the.NOM man the.ACC book gives

f. [weil] dem Kind das Buch der Mann gibt
because the.DAT child the.ACC book the.NOM man gives

(4) Gibt der Mann dem Kind das Buch? (German)
gives the.NOM man the.DAT child the.ACC book
'Does the man give the child the book?'

(3) shows that in addition to the unmarked order in (3a) (see Höhle (1982) on the notion of unmarked order), five other argument orders are possible in sentences with three-place verbs. As with the examples just given, I will use German if a phenomenon does not exist in English. Section 6.2 discusses examples from Warlpiri, a language having even freer constituent order.

(4) shows that the verb is placed in initial position in questions in German. This contrasts with the verb-final order in the subordinate clause in (3a), which has the same order as far as the arguments are concerned. This alternation of verb placement is usually treated as head movement in the transformational literature (Bach 1962; Bierwisch 1963: 34; Reis 1974; Thiersch 1978: Chapter 1). Declarative main clauses in German are V2 clauses and the respective fronting of the preverbal constituent is usually treated as a non-local dependency (see Borsley & Crysmann (2019), Chapter 14 of this volume). Hence, V2 sentences will not be handled here.

The following sections explore the theoretical options within the HPSG framework for dealing with these phenomena. I first discuss the separation of grammar rules into an immediate dominance part and a linear precedence component in Section 2 and then flat vs. binary branching structures (Section 3). While flat structures allow verbs to be ordered clause-finally or clause-initially, this is not the case for binary branching structures, since only sisters can be ordered. So, for (3a) one would get the bracketing in (5a). If *das Buch* 'the book' and *gibt* 'gives' are ordered in a different order, (5b) results.

(5) a. [weil] [der Mann [dem Kind [das Buch gibt]]]
    because the.NOM man the.DAT child the.ACC book gives

  b. * [weil] [der Mann [dem Kind [gibt das Buch]]]
    because the.NOM man the.DAT child gives the.ACC book

Hence, local reordering is not sufficient to get clause-initial verb order and therefore, proposals with binary branching structures are usually paired with HPSG's analogue of what is head-movement in transformational theories. These are explained in Section 5. Section 6 introduces an extension to standard HPSG developed by Reape (1994): constituent order domains. Such constituent order domains allow for discontinuous constituents and have been used to account for languages like Warlipri (Donohue & Sag 1999). In contrast, Section 7 shows how such languages can be analyzed without admitting discontinuous constituents.

2 ID/LP format

HPSG was developed out of Generalized Phrase Structure Grammar (GPSG) and Categorial Grammar. The ideas concerning linearization of daughters in a local tree were taken over from GPSG (Gazdar, Klein, Pullum & Sag 1985). In GPSG a separation between immediate dominance and linear precedence is assumed. So, while in classical phrase structure grammar, a phrase structure rule like (6) states that the NP[nom], NP[dat] and NP[acc] have to appear in exactly this order, this is not the case in GPSG and HPSG:

(6) S → NP[nom] NP[dat] NP[acc] V

The HPSG schemata corresponding to the immediate dominance rule (ID rule) in (6) do not express information about ordering. Instead, there are separate linear precedence (LP) rules (also called linearization rules). A schema like (6) licenses 24 different orders: the six permutations of the three arguments that were shown in (3) and all possible placements of the verb (to the right of NP[acc], between

NP[dat] and NP[acc], between NP[nom] and NP[dat], to the left of NP[nom]). Orders like NP[nom], NP[dat], V, NP[acc] are not attested in German and hence these orderings have to be filtered out. This is done by linearization rules, which can refer to features or to the function of a daughter in a schema. (7) shows some examples of linearization rules:

(7) a. X < V
  b. X < V[INI−]
  c. X < Head [INI−]

The first rule says that all constituents have to precede a V in the local tree. The second rule says that all constituents have to precede a V that has the INITIAL value −. One option to analyze German would be the one that was suggested by Uszkoreit (1987) within the framework of GPSG: one could allow for two linearization variants of finite verbs. So in addition to the INI− variant of verbs there could be a INI+ variant and this variant would be linearized initially. This reduces the number of permutations licensed by (6) and LP rules to 12: verb-initial placement and 6 permutations of the NPs and verb-final placement with 6 permutations of the arguments. The ID rule in (6) together with the two linearization rules linearizing the verb in initial or final position therefore licenses the same orders as the following twelve phrase structure rules would do:

(8) a. S → NP[nom] NP[acc] NP[dat] V
  S → NP[nom] NP[dat] NP[acc] V
  S → NP[acc] NP[nom] NP[dat] V
  S → NP[acc] NP[dat] NP[nom] V
  S → NP[dat] NP[nom] NP[acc] V
  S → NP[dat] NP[acc] NP[nom] V
  b. S → V NP[nom] NP[acc] NP[dat]
  S → V NP[nom] NP[dat] NP[acc]
  S → V NP[acc] NP[nom] NP[dat]
  S → V NP[acc] NP[dat] NP[nom]
  S → V NP[dat] NP[nom] NP[acc]
  S → V NP[dat] NP[acc] NP[nom]

Note that we do not need a linearization rule for every ID rule. For example, in a grammar with rules for intransitive, transitive, and ditransitive verbs, head ordering is taken care of by general LP rules of the type in (7b) applying to the respective ID rules. The LP rule in (7c) is even more general than (7b) in that it does not mention the part of speech but instead refers to the function of the constituent. The rule says that a head that has the INI value '–' has to be linearized to the right of all other elements in the local tree. Hence, it also applies to adjectives, nouns, and prepositions and their dependents.

This treatment of constraints on linearization has an advantage that was already pointed out by researchers working in GPSG: it captures the generalizations regarding linearization. For instance, the order of verbs with respect to their arguments is the same in embedded sentences in German, independent of the finiteness of the verb:

(9) a. dass er dem Mann das Buch gab
that he.NOM the.DAT man the.ACC book gave
'that he gave the man the book'

b. dass er versucht, [dem Mann das Buch zu geben]
that he.NOM tried the.DAT man the.ACC book to give
'that he tried to give the man the book'

This is also true for the relative order of dative and accusative object in (9). The constraints regarding linearization hold across rules. By factoring these constraints out, the generalizations can be captured. See Uszkoreit (1987) for weighted constraints for the ordering of constituents in the *Mittelfeld*.

Furthermore, cross-linguistic generalizations about constituent structure can be captured. For example, the two phrase structure rules in (10) would be needed for head-initial and head-final languages, respectively:

(10) a. VP → V NP NP
b. VP → NP NP V

In an ID/LP framework only one ID rule is needed to describe both sorts of languages. The linearization of the head is factored out of the rules.

Similarly, HPSG has just one schema for Head-Adjunct structures, although languages like English allow adjuncts to precede or follow the head they modify. Schema 1 does not say anything about the order of the daughters.

Schema 1 (Head-Adjunct Schema)

head-adjunct-phrase ⇒

$$\begin{bmatrix} \text{HEAD-DTR} & \begin{bmatrix} \text{SYNSEM } \boxed{1} \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD|MOD } \boxed{1} \\ \text{SPR } \langle\rangle \\ \text{COMPS } \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

There is a head daughter and a list of non-head daughters. The respective daughters are specified as the value of a feature or as an element in a list but they are not ordered with respect to each other in the schema. Ordering is taken care of by two LP rules saying that adjuncts marked as pre-modifiers (e.g., attributive adjectives) have to precede their head while those that are marked as post-modifiers (noun-modifying prepositions) follow it:

(11) a. Adjunct[PRE-MODIFIER +] < Head

b. Head < Adjunct[PRE-MODIFIER −]

In general, there are two options for two daughters: head-initial and head-final order. Examples are given in (12):[1]

(12) a. head-initial:

$$\begin{bmatrix} \text{PHON} & \boxed{1} \oplus \boxed{2} \\ \text{HEAD-DTR} & \begin{bmatrix}\text{PHON} & \boxed{1}\end{bmatrix} \\ \text{NH-DTRS} & \left\langle \begin{bmatrix}\text{PHON} & \boxed{2}\end{bmatrix} \right\rangle \end{bmatrix}$$

example:

$$\begin{bmatrix} \text{PHON} & \langle \textit{ squirrel, from, America } \rangle \\ \text{HEAD-DTR} & \begin{bmatrix}\text{PHON} & \langle \textit{ squirrel } \rangle\end{bmatrix} \\ \text{NH-DTRS} & \left\langle \begin{bmatrix}\text{PHON} & \langle \textit{ from, America } \rangle\end{bmatrix} \right\rangle \end{bmatrix}$$

b. head-final:

$$\begin{bmatrix} \text{PHON} & \boxed{2} \oplus \boxed{1} \\ \text{HEAD-DTR} & \begin{bmatrix}\text{PHON} & \boxed{1}\end{bmatrix} \\ \text{NH-DTRS} & \left\langle \begin{bmatrix}\text{PHON} & \boxed{2}\end{bmatrix} \right\rangle \end{bmatrix}$$

example:

$$\begin{bmatrix} \text{PHON} & \langle \textit{ grey, squirrel } \rangle \\ \text{HEAD-DTR} & \begin{bmatrix}\text{PHON} & \langle \textit{ squirrel } \rangle\end{bmatrix} \\ \text{NH-DTRS} & \left\langle \begin{bmatrix}\text{PHON} & \langle \textit{ grey } \rangle\end{bmatrix} \right\rangle \end{bmatrix}$$

When linearization rules enforce head-initial order, as in the case of modification by a PP in English, the PHON value of the head daughter is concatenated with the PHON value of the non-head daughter, and if the order has to be the other way around as in the case of adjectives modifying nouns, the non-head daughter is concatenated with the head daughter. An adjective is specified as PRE-MODIFIER + and a preposition as PRE-MODIFIER −. Since these features are head-features (see **chapters/wherever** Chapter ?? of this volume on head features), they are also accessible at the level of adjective phrases and prepositional phrases.

For languages with free variation in head-adjunct order, it would suffice to not state any LP rule and one would get both orders with the same Head-Adjunct schema. So, the separation of immediate dominance and linear precedence allows for an underspecification of order. Therefore HPSG grammarians are not forced

[1] $\oplus$ (append) is a relational constraint that concatenates two lists.

to assume several different constructions for attested patterns or derivational processes that derive one order from another more basic one.

3 Flat and binary branching structures

The previous section discussed LP rules and used flat phrase structure rules for illustration. The corresponding flat structures are also used in HPSG. Schema 2 shows a Head-Complement schema that combines a head with all the complements selected via the COMPS list.[2]

Schema 2 (Head-Complement Schema)

$$
\textit{head-complement-phrase} \Rightarrow
\begin{bmatrix}
\textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{comps}\ \langle\rangle \\
\textsc{head-dtr}\ \begin{bmatrix}\textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{comps}\ \boxed{1}\end{bmatrix} \\
\textsc{non-head-dtrs}\ \text{synsem2signs}(\boxed{1})
\end{bmatrix}
$$

`synsem2signs` is a relational constraint mapping a list of *synsem* objects as they are contained in the COMPS list onto a list of objects of type *sign* as they are contained in daughters (see Ginzburg & Sag 2000: 34 for a similar proposal).[3] How this schema can be used to analyze VPs like the one in (13) is shown in Figure 1.

(13) Kim gave Sandy a book.

HPSG differs from purely phrase structure-based approaches in that the form of a linguistic object is not simply the concatenation of the forms associated with the terminal symbols in a tree (words or morphemes). Every linguistic object has its own phonology representation. So in principle one could design theories in which the combination of *Mickey Mouse* and *sleeps* is pronounced as *Donald*

[2] Ginzburg & Sag (2000: 4) assume a list called DTRS for all daughters including the head daughter. It is useful to be able to refer to specific non-head daughters without having to know a position in a list. For example in head-adjunct structures the adjunct is the selector. So I keep DTRS for a list of ordered daughters and HEAD-DTR and NON-HEAD-DTRS for material that is not necessarily ordered.

[3] In Sign-Based Construction Grammar the objects in valence lists are of the same type as the daughters. A relational constraint would not be needed in this variant of the HPSG theory. Theories working with a binary branching Head-Complement Schema as Schema 3 on page 275 would not need the relational constraint either, since the *synsem* object in the COMPS list can be shared with the SYNSEM value of the element in the list of non-head daughters directly.

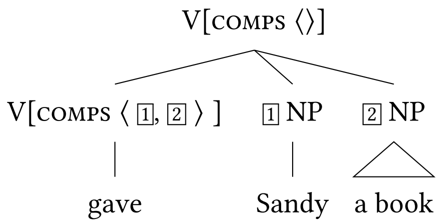


Figure 1: Analysis of the VP *gave Sandy a book* with a flat structure

Duck laughs. Of course, this is not done. The computation of the PHON value of the mother is dependent of the PHON values of the daughters. But the fact that the PHON values of a linguistic sign are not necessarily a strict concatenation of the PHON values of the daughters can be used to model languages having a less strict order than English. Pollard & Sag (1987: 168) formulate the Constituent Order Principle, which is given as Principle 1 in adapted form:

Principle 1 (Constituent Order Principle)

$$\textit{phrase} \Rightarrow \begin{bmatrix} \textsc{phon} & \text{order-constituents}(\boxed{1}) \\ \textsc{dtrs} & \boxed{1} \end{bmatrix}$$

DTRS is a list of all daughters including the head daughter (if there is one). This setting makes it possible to have the daughters in the order in which the elements are ordered in the COMPS list (primary object, secondary object, and obliques) and then compute a PHON value in which the secondary object precedes the primary object. French is a language with freer constituent order than English and such flat structures with appropriate reorderings are suggested by Abeillé & Godard (2000). For English the function `order-constituents` would just return a concatenation of the PHON values of the daughters, but for other languages it would be much more complicated. In fact this function and its interaction with linear precedence constraints was never worked out in detail.

Researchers working on English and French usually assume a flat structure (Pollard & Sag 1994: 39–40, 362; Sag 1997: 479; Ginzburg & Sag 2000: 34; Abeillé & Godard 2000) but assuming binary branching structures would be possible as well, as is clear from analyses in Categorial Grammar, where binary combinatory rules are assumed (Ajdukiewicz 1935; Steedman 2000). For languages like German it is usually assumed that structures are binary branching (but see Reape 1994: 156 and Bouma & van Noord 1998: 51). The reason for this is that adverbials can be placed anywhere between the arguments, as the following example from Uszkoreit (1987: 145) shows:

(14) *Gestern* hatte *in der Mittagspause* der Vorarbeiter *in der*
yesterday had during the lunch.break the foreman in the
Werkzeugkammer dem Lehrling *aus Boshaftigkeit langsam* zehn
tool.shop the apprentice maliciously slowly ten
schmierige Gußeisenscheiben *unbemerkt* in die Hosentasche gesteckt.
greasy cast.iron.disks unnoticed in the pocket put
'Yesterday during the lunch break, the foreman maliciously put ten greasy cast iron disks slowly into the apprentice's pocket unnoticed.'

A way to straightforwardly analyze adjunct placement in German and Dutch is to assume that adjuncts can attach to any verbal projection. For example, Figure 2 shows the analysis of (15):

(15) weil deshalb jemand gestern dem Kind schnell das Buch gab
because therefore somebody yesteray the child quickly the book gave
'because somebody quickly gave the child the book yesterday'

The adverbials *deshalb* 'therefore', *gestern* 'yesterday' and *schnell* 'quickly' may attach to any verbal projection. For example, *gestern* could also be placed at the other adjunct positions in the clause.

Binary branching structures with attachment of adjuncts to any verbal projection also account for recursion and hence the fact that arbitrarily many adjuncts can attach to a verbal projection. Of course it is possible to formulate analyses with flat structures that involve arbitrarily many adjuncts (Kasper 1994; van Noord & Bouma 1994; Abeillé & Godard 2000; Bouma et al. 2001), but these analyses involve relational constraints in schemata or in lexical items or an infinite lexicon. In Kasper's analysis, the relational constraints walk through lists of daughters of unbounded length in order to compute the semantics. In the other three analyses, (some) adjuncts are treated as valents, which may be problematic because of scope issues. This cannot be dealt with in detail here, but see Levine & Hukari (2006) and Chaves (2009) for discussion.

The following schema licenses binary branching head-complement phrases:

Schema 3 (Head-Complement Schema (binary branching))

head-complement-phrase ⇒

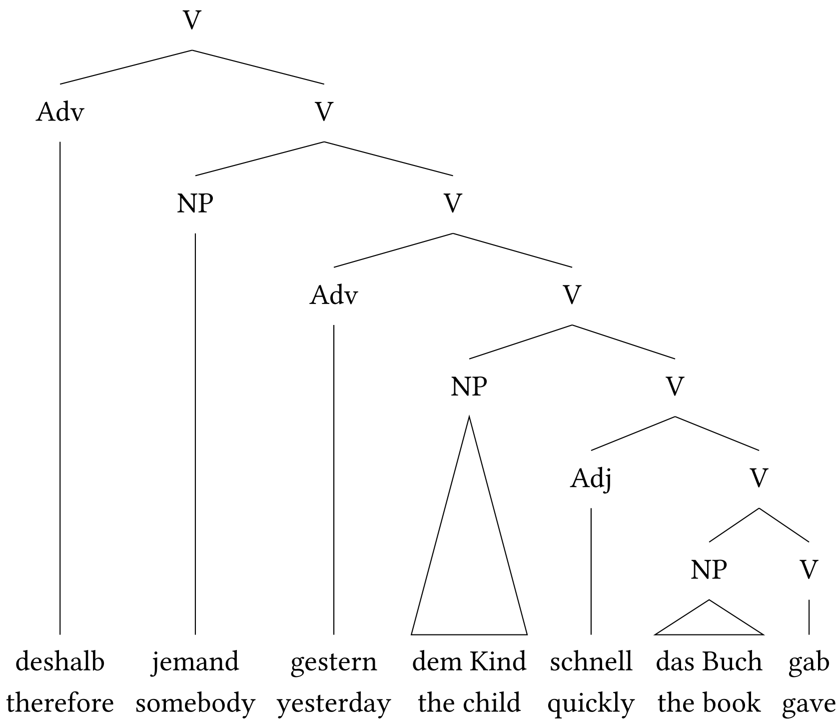


Figure 2: Analysis of [*weil*] *deshalb jemand gestern dem Kind schnell das Buch gab* 'because somebody quickly gave the child the book yesterday' with binary branching structures

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{COMPS} & \boxed{1} \oplus \boxed{2} \\ \text{HEAD-DTR} & \begin{bmatrix}\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{COMPS}\ \boxed{1} \oplus \langle\ \boxed{3}\ \rangle \oplus \boxed{2}\end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix}\text{SYNSEM}\ \boxed{3}\end{bmatrix} \right\rangle \end{bmatrix}$$

The COMPS list of the head daughter is split into three lists: a beginning ($\boxed{1}$), a list containing $\boxed{3}$ and a rest ($\boxed{2}$). $\boxed{3}$ is identified with the SYNSEM value of the non-head daughter. All other elements of the COMPS list of the head daughter are concatenated and the result of this concatenation ($\boxed{1} \oplus \boxed{2}$) is the COMPS list of the mother node. This schema is very general. It works for languages that allow for scrambling, since it allows an arbitrary element to be taken out of the COMPS list of the head daughter and realize it in a local tree. The schema can also be "parameterized" to account for languages with fixed word order. For head-final languages with fixed order, $\boxed{2}$ would be the empty list (= combination with the last element in the list) and for head-initial languages with fixed order (e.g., English), $\boxed{1}$ would be the empty list (= combination with the first element in the list). Since the elements in the COMPS list are ordered in the order of Obliqueness (Keenan & Comrie 1977; Pullum 1977) and since this order corresponds to the order in which the complements are serialized in English, the example in (13) can be analyzed as in Figure 3.[4] The second tree in the figure is the German

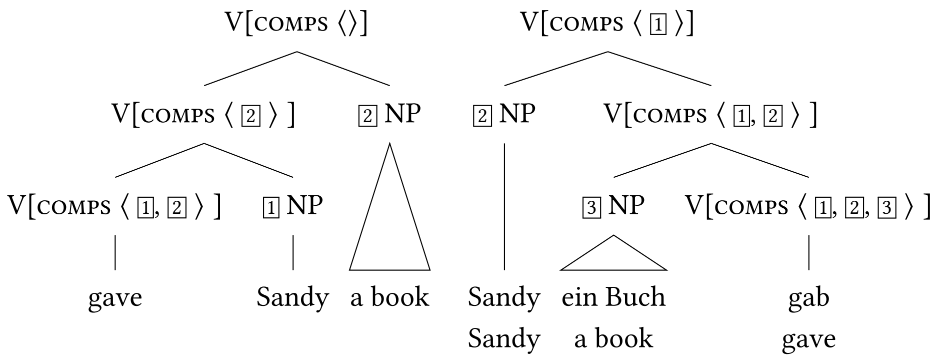


Figure 3: Analysis of the English VP *gave Sandy a book* and the corresponding German verbal projection *Sandy ein Buch gab* with binary branching structures

counterpart of *gave Sandy a book*: the finite verb in final position with its two objects in normal order. Section 4 explains why SOV languages like German and Japanese contain their subject in the COMPS list while SVO languages like English and Romance languages do not.

The alternative to using relational constraints as in Schema 3 is to use sets rather than lists for the representation of valence information (Gunji 1986; Hinrichs & Nakazawa 1989; Pollard 1996; Engelkamp, Erbach & Uszkoreit 1992). The Head-Complement Schema would combine the head with one of its complements. Since the elements of a set are not ordered, any complement can be taken and hence all permutations of complements are accounted for.

The disadvantage of set-based approaches is that sets do not impose an order on their members, but an order is needed for various subtheories of HPSG (see

[4]This structure may seem strange to those working in Mainstream Generative Grammar (MGG, GB/Minimalism). In MGG, different branchings are assumed, since the form of the tree plays a role in Binding Theory. This is not the case in HPSG: Binding is done on the ARG-ST list. See Branco (2019), Chapter 21 of this volume for a discussion of HPSG's Binding Theory and Borsley & Müller (2019a), Chapter L of this volume for a comparison between HPSG and Minimalism.

Przepiórkowski (2019), Chapter 7 of this volume on case assignment, and Branco (2019), Chapter 21 of this volume on Binding Theory). In the approach proposed above and in Müller (2003b; 2015a,b), the valence lists are ordered but the schema allows for combination with any element of the list. For valence representation and the order of elements in valence lists see Wechsler, Koenig & Davis (2019), Chapter 9 of this volume.

4 SVO vs. SOV

The careful reader will have noticed that the COMPS list of *gave* in Figure 3 contains the two objects, while its German counterpart *gab* has three elements in the COMPS list. The rationale behind this is explained in this section.

In principle, one could assume a rule like (6) for SVO languages like English as well. The SVO order would then be accounted for by linearization rules stating that NP[*nom*] precedes the finite verb while other arguments follow it. This would get the facts about simple sentences like (16a) right but leaves the analysis of (16c) open.

(16) a. Peter reads books.
b. Peter often read books.

The generalization about languages like English is that adverbials can attach to the left or to the right of verbs with their objects, that is, to the left or to the right of the VP. Researchers like Borsley (1987) argued that subjects, specifiers, and complements differ in crucial ways and should be represented by special (valence) features. For example, the subject of the VP *to read more books* is not realized but is referred to in Control Theory (Abeillé 2019, Chapter 13 of this volume).

(17) Peter tries to read more books.

The subject in English main clauses is similar to the determiner in nominal structures, so one way of expressing this similarity is by using the same valence features and the same schema for subject-VP combinations as for determiner-noun combinations.[5] The schema is given here as Schema 4:

[5]This is non-standard in HPSG. Usually the SUBJECT feature is used for subjects and SPR for determiners. I follow the German HPSG tradition and use SUBJ for unexpressed subjects. See also Eynde (2019), Chapter 8 of this volume for alternative analyses of nominal structures that do not assume a selection of the determiner by the noun.

Schema 4 (Specifier-Head Schema)

specifier-head-phrase ⇒

$$\left[\begin{array}{l} \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{SPR}\ \boxed{1} \\ \text{HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT} \left[\begin{array}{ll}\text{SPR} & \boxed{1} \oplus \langle \boxed{2} \rangle \\ \text{COMPS} & \langle\rangle\end{array}\right] \\ \text{NON-HEAD-DTRS} \left\langle \left[\text{SYNSEM}\ \boxed{2}\right] \right\rangle \end{array}\right]$$

The last element of the SPR list is realized as the non-head daughter. The remaining list is passed up to the mother node. Note that the non-head daughter is taken from the end of the SPR list. For heads that have exactly one specifier this difference is irrelevant, but in the analysis of object shift in Danish suggested by Müller & Ørsnes (2013), the authors assume multiple specifiers and hence the difference in order of combination is relevant. The head-daughter must have an empty COMPS list. This way it is ensured that verbs form a unit with their objects (the VP) and the subject is combined with the VP, rather than the subject combining with a lexical verb and this combination combining with objects later.

The analysis of the sentence in (18) including the analysis of the NP is given in Figure 4.

(18) Kim gave Sandy a book.

For German, it is standardly assumed that the subjects of finite verbs are treated like complements (Pollard 1996: 295–296, Kiss 1995: 80) and hence are represented at the COMPS list (as in Figure 3). The assumption that arguments of finite verbs are complements is also made by researchers working in different research traditions (e.g. Eisenberg 1994: 376). By assuming that the subject is listed among the complements of a verb it is explained why it can be placed in any position before, between, and after them. So in summary, German differs from English in the way the arguments are distributed on the valence lists, on the one hand in order to capture the similarity in English between combinations of subjects with VPs and determiners with nouns, and on the other hand to allow German the flexible constituent order it needs. However, HPSG has a more basic representation in which the languages do behave the same: the argument structure represented at the ARG-ST list. The ARG-ST list contains *synsem* objects and is used for linking (Wechsler, Koenig & Davis 2019, Chapter 9 of this volume), case assignment (Przepiórkowski 2019, Chapter 7 of this volume), and binding (Branco 2019, Chapter 21 of this volume). Ditransitive verbs in German and English have

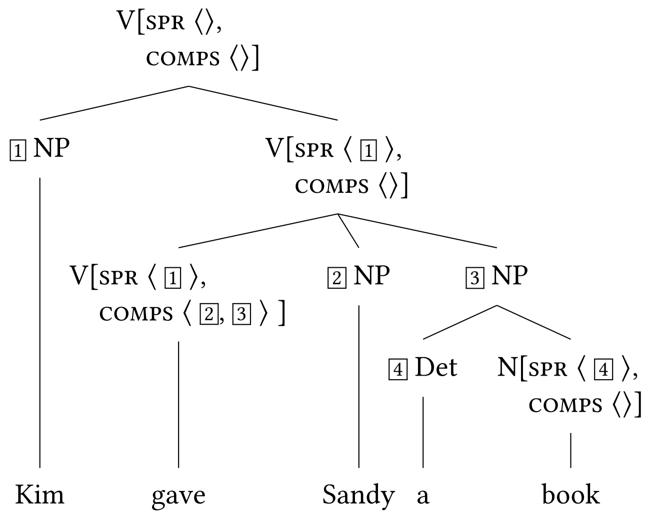


Figure 4: Analysis of *Kim gave Sandy a book* with SPR and COMPS feature and a flat VP structure

three NP arguments on their ARG-ST and they are linked in the same way to the semantic representation (Müller 2018). (19) shows the mapping from ARG-ST to SPR and COMPS:

(19) *gives* (English, SVO language):
$$\begin{bmatrix} \textsc{spr} & \langle \boxed{1} \rangle \\ \textsc{comps} & \boxed{2} \\ \textsc{arg-st} & \langle \boxed{1}\,\text{NP} \rangle \oplus \boxed{2} \langle \text{NP, NP} \rangle \end{bmatrix}$$

gibt (German, SOV language):
$$\begin{bmatrix} \textsc{spr} & \langle\rangle \\ \textsc{comps} & \boxed{1} \\ \textsc{arg-st} & \boxed{1} \langle \text{NP, NP, NP} \rangle \end{bmatrix}$$

In SVO languages, the first element of the ARG-ST list is mapped to SPR and all others to COMPS and in languages without designated subject positions all ARG-ST elements are mapped to COMPS.

Having explained scrambling in HPSG and the order of subjects in SVO languages, I now turn to “head movement”.

5 Head movement vs. constructional approaches assuming flat structures

The Germanic languages signal the clause type by verb position. All Germanic languages with the exception of English are V2 languages: the finite verb is in second position in declarative main clauses. The first position can be filled by any other constituent, for example a subject, objects, or adverbials. (20) shows an example from the V2 language German and its English translation.

(20) Eigentlich mag ich Katzen sehr. (German)
actually like I cats really
'I actually really like cats.'

The fronted material is not necessarily from the matrix clause, clause boundary crossing non-local dependencies are possible. The same holds for questions with *w* phrases.

Yes/no questions are formed by putting the verb in initial position:

(21) Magst du Katzen? (German)
like you cats
'Do you like cats?'

English is a so-called *residual V2 language*, that is, there are some constructions that are parallel to what is known from V2 languages. For example, while declarative clauses are in base order (SVO), questions follow the pattern that is known from other Germanic languages with the finite verb in second position.[6]

(22) What$_i$ will Kim read $\_{}_i$?

Analyses assuming flat structures (or flat linearization domains, see Section 6) usually treat alternative orders of verbs in Germanic languages as linearization variants (Reape 1994; Kathol 2001; Müller 1995; 2003b; Bjerre 2006), but this is not necessarily so, as Bouma and van Noord's analysis of Dutch clauses show (Bouma & van Noord 1998: 62, 71). The alternative to verb placement as linearization is something that is similar to verb movement in Government & Binding: an empty element takes the position of the verb in its canonical position and the verb is realized in initial or – if something is realized before the finite verb – in second position. The following subsection deals with such approaches in more detail. Subsection 5.2 deals with a constructional approach.

[6] SVO is not V2 although the verb is in second position in SVO sentences. Languages can be categorized into SOV, SVO, VSO, OSV, OVS, and VOS languages and into V2 or non-V2 languages. These two dimensions are independent. For example, Danish is an SVO language that is V2, while German is SOV and V2 (Haider 2019). See Müller (2019) for discussion and the analysis of this variation in HPSG.

5.1 Head movement approaches

Building on work by Jacobson (1987) in the framework of Categorial Grammar, Borsley (1989) showed that in addition to the analysis of auxiliary inversion in English that was suggested in GPSG (Gazdar et al. 1985), an analysis that is similar to the movement-based analysis in GB is possible in HPSG as well. Head movement analyses in GPSG and HPSG are concerned with the verb placement in pairs such as the one in (23) rather than with adverb placement as in GB analyses by Pollock (1989) and Cinque (1999).

(23) a. Will Kim get the job?
b. Kim will get the job.

The technique that is used in Borsley's analysis is basically the same that was developed by **Gazdar81** for the treatment of nonlocal dependencies in GPSG. An empty category is assumed and the information about the missing element is passed up the tree until it is bound off at an appropriate place (that is, by the fronted verb). Note that the heading of this section contains the term *head movement* and I talk about traces, but it is not the case that something is actually moved. There is no underlying structure with a verb that is transformed into one with the verb fronted and a remaining trace in the original position of the verb. Instead, the empty element is a normal element in the lexicon and can function as the verb in the respective position. The analysis of (23a) is shown in Figure 5. A special variant of the auxiliary is licensed by a unary rule. The unary rule has as a daughter the auxiliary as it appears in canonical SVO order as in (23b). It licenses an auxiliary selecting a full clause in which the daughter auxiliary (with the LOCAL value $\boxed{2}$) is missing. The fact that the auxiliary is missing is represented as the value of DOUBLE SLASH (DSL). The value of DSL is a *local* object, that is, something that contains syntactic and semantic information ($\boxed{2}$ in Figure 5). DSL is a head feature and hence available everywhere along a projection path (see Borsley & Abeillé (2019), Chapter 1 of this volume for the Head Feature Principle). The empty element for head movement is rather simple:

check reference to Chapter 1.

(24) empty element for head movement:

$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & \langle\rangle \\ \text{SYNSEM|LOC} & \boxed{1}\begin{bmatrix}\text{CAT|HEAD|DSL}\ \boxed{1}\end{bmatrix} \end{bmatrix}$$

It states that there is an empty element that has the local requirements that correspond to its DSL value. For cases of verb movement it says: I am a verb that is

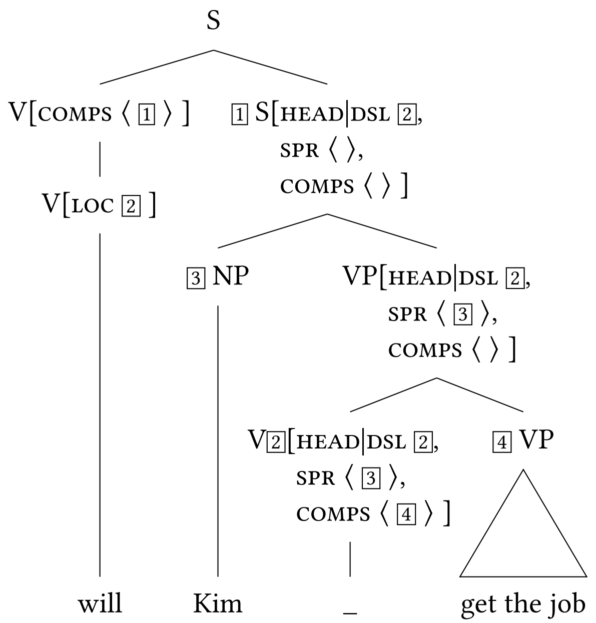


Figure 5: Analysis of English auxiliary constructions as head-movement following Borsley (1989)

missing itself.

Such head-movement analyses are assumed by most researchers working on German (Kiss & Wesche 1991: Section 4.7; Oliva 1992; Netter 1992; Kiss 1993; Frank 1994; Kiss 1995; Feldhaus 1997, Meurers 2000; Müller 2005a; 2017) and also by Bouma & van Noord (1998: 62, 71) in their work on Dutch, by Müller & Ørsnes (2015) in their grammar of Danish and by Müller (2019) for Germanic in general.

5.2 Constructional approaches

The alternative to head-movement-based approaches is a flat analysis with an alternative serialization of the verb. This was already discussed with respect to German, but I want to discuss English auxiliary constructions here, since they have figured prominently in linguistic discussions.[7] In the analysis of (25) shown in Figure 6, the auxiliary *did* selects for the subject *Kim* and a VP *get the job*.

(25) Did Kim get the job?

[7] For a discussion including French verb placement see Abeillé & Godard (1997) and Kim & Sag (2002).

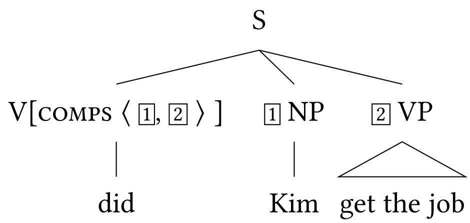


Figure 6: Analysis of English auxiliary constructions according to Sag et al. (2020)

The tree in Figure 6 is licensed by a schema combining a head with its subject ($\boxed{1}$) and its VP complement ($\boxed{2}$) in one go.[8] As has been common in HPSG since the mid-1990s (Sag 1997), phrasal schemata are organized in type hierarchies and the general schema for auxiliary-initial constructions has the type *aux-initial-cxt*. Fillmore (1999) and Sag et al. (2020) argue that there are various usages of auxiliary-initial constructions and assign the respective usages to subconstructions of the general auxiliary-initial construction. Technically this amounts to stating subtypes of *aux-initial-cxt*. For example, Sag et al. (2020) posit a subtype *polar-int-cl* for polar interrogatives like (26a) and another subtype *auxinitial-excl-cl* for exclamatives like (26b).

(26) a. Are they crazy?
 b. Are they crazy!

Chomsky (2010) compared the various clause types used in HPSG with the – according to him – much simpler Merge-based analysis in Minimalism. Minimalism assumes just one very general schema for combination (External Merge is basically equivalent to our Schema 3 above, see Müller (2013)), so this rule for combining linguistic objects is very simple, but this does not help in any way when considering the facts: there are at least five different meanings associated with auxiliary initial clauses (polar interrogative, blesses/curses, negative imperative, exclamatives, conditionals) and these have to be captured somewhere in a grammar. One way is to state them in a type hierarchy as is done in some HPSG analyses and in Sign-Based Construction Grammar (SBCG), another way is to use implicational constraints that assign meaning with respect to actual configurations (see Section 5.3), and a third way is to do everything lexically. The only

[8] An alternative is to assume a separate valence feature for the subject (SUBJ) and a schema that combines the head with the element in the SUBJ list and the elements in the COMPS list (Ginzburg & Sag 2000: 36).

option for Minimalism is the lexical one. This means that Minimalism has to either assume as many lexical items for auxiliaries as there are types in HPSG or to assume empty heads that contribute the meaning that is contributed by the phrasal schemata in HPSG (Borsley 2006: Section 5; Borsley & Müller 2019b). The latter proposal is generally assumed in Cartographic approaches (Rizzi 1997). Since there is a fixed configuration of functional projections that contribute semantics, one could term these Rizzi-style analyses *Crypto-Constructional*.

Having discussed a lexical approach involving an empty element and a phrasal approach that can account for the various meanings of auxiliary inversion constructions, I turn now to a mixed approach in the next section and show how the various meanings associated with certain patterns can be integrated into accounts with rather abstract schemata for combinations like the one described in Section 5.1.

5.3 Mixed approaches

The situation with respect to clause types is similar in German. Verb first sentences can be yes/no questions (27a), imperatives (27b), conditional clauses (27c), and declarative sentences with topic drop (27d).

(27) a. Kommt Peter? (German)
comes Peter
'Is Peter coming?'

b. Komm!
come

c. Kommt Peter, komme ich nicht.
comes Peter come I not
'If Peter comes, I won't come.'

d. Kommt. (Was ist mit Peter?)
comes what is with Peter
'What about Peter?' 'He is coming.'

Verb second sentences can be questions (28a), declarative sentences (28b), or imperatives (28c).

(28) a. Wer kommt? (German)
who comes

b. Peter kommt.
Peter comes

c. Jetzt komm!
   now come
   'Come now!'

While one could try and capture this situation by assuming surface order-related clause types, such approaches are rarely used in HPSG (but see Kathol (2001) and Wetta (2011), and see Section 6.4.2 on why such approaches are doomed to failure). Rather, researchers assumed binary branching head-complement structures together with verb movement (for references see the end of Section 5.1).[9]

As was explained in Section 5.1, the head movement approaches are based on lexical rules or unary projections. These license new linguistic objects that could contribute the respective semantics. In analogy to what Borsley (2006) has discussed with respect to extraction structures, this would mean that one needs seven versions of fronted verbs to handle the seven cases in (27) and (28), which would correspond to the seven phrasal types that would have to be stipulated in phrasal approaches. But there is a way out of this: one can assume one lexical item with underspecified semantics. HPSG makes it possible to use implicational constraints referring to a structure in which an item occurs. Depending on the context, the semantics contributed by a specific item can be further specified. Figure 7 shows the construction-based and the lexical-rule-based analysis in the abstract for comparison. In the construction-based analysis, the daughters

(a) Phrasal construction

(b) Unary construction and implication

Figure 7: Construction-based, phrasal approach and approach with implicational constraint

contribute x and y as semantic values and the whole construction adds the construction meaning f. In the lexical-rule- or unary-projection-based analysis, the lexical rule/unary projection adds the f and the output of the rule is combined compositionally with the other daughter. Now, implicational constraints can be

[9] I assumed linearization domains (see Section 6) for ten years and then switched to the head-movement approach (Müller 2005a,b; 2017).

used to determine the exact contribution of the lexical item (Müller 2016). This is shown with the example of a question in Figure 8. The implication says: when

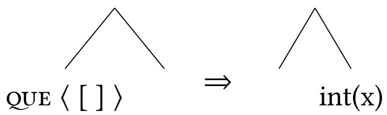


Figure 8: Implication for interrogative sentences

the configuration has the form that there is a question pronoun in the left daughter, the output of the lexical rule gets question semantics. Since HPSG represents all linguistic information in the same attribute value matrix (AVM), such implicational constraints can refer to intonation as well and hence, implications for establishing the right semantics for V1 questions (27a) vs. V1 conditionals (27c) can be formulated.

6 Constituent order domains and linearization

There is an interesting extension to standard HPSG that opens up possibilities for analyses that are quite different from what is usually done in theoretical linguistics: Mike Reape (1991; 1992; 1994) working on German suggested formal tools that allow for the modeling of discontinuous constituents.[10] His original motivation was to account for scrambling of arguments of verbs forming verbal complexes, but this analysis was superseded by Hinrichs and Nakazawa's analysis (Hinrichs & Nakazawa 1989; 1994) since purely linearization-based approaches are unable to account for agreement and the so-called remote passive (Kathol 1998: Section 5.1, Section 5.2; Müller 1999: Chapter 21.1). Nevertheless, Reape's work was taken up by others and was used for analyzing German (Kathol & Pollard 1995; Kathol 2000; Müller 1995; 1996; 2004; Wetta 2011; 2014). As will be discussed below in Section 6.4, there were reasons for abandoning linearization-based analyses of German assuming discontinuous constituents (Müller 2005b; 2017) but constituent order domains still play a major role in analyzing ellipsis (Nykiel & Kim 2019, Chapter 20 of this volume) and coordination (Abeillé & Chaves 2019, Chapter 17 of this volume). Bonami, Godard & Marandin (1999) show that complex predicate formation does not account for subject-verb inversion in French and suggest a domain-based approach. Bonami & Godard (2007),

[10] See also Wells (1947), Dowty (1996), and Blevins (1994) for proposals assuming discontinuous constituents in other frameworks.

also working on French, treat sentential adverbs in a domain-based approach.

6.1 A special representational layer for constituent order

The technique that is used to model discontinuous constituents in frameworks like HPSG goes back to Mike Reape's work on German (1991; 1992; 1994). Reape uses a list called DOMAIN to represent the daughters of a sign in the order in which they appear at the surface of an utterance. (29) shows an example in which the DOM value of a headed-phrase is computed from the DOM value of the head and the list of non-head daughters.

(29) $$\textit{headed-phrase} \Rightarrow \begin{bmatrix} \text{HEAD-DTR}|\text{DOM} & \boxed{1} \\ \text{NON-HEAD-DTRS} & \boxed{2} \\ \text{DOM} & \boxed{1} \bigcirc \boxed{2} \end{bmatrix}$$

The symbol '$\bigcirc$' stands for the *shuffle* relation. *shuffle* relates three lists A, B and C iff C contains all elements from A and B and the order of the elements in A and the order of the elements of B is preserved in C. (30) shows the combination of two sets with two elements each:

(30) $$\begin{aligned} \langle\, a, b \,\rangle \bigcirc \langle\, c, d \,\rangle = {} & \langle\, a, b, c, d \,\rangle \vee \\ & \langle\, a, c, b, d \,\rangle \vee \\ & \langle\, a, c, d, b \,\rangle \vee \\ & \langle\, c, a, b, d \,\rangle \vee \\ & \langle\, c, a, d, b \,\rangle \vee \\ & \langle\, c, d, a, b \,\rangle \end{aligned}$$

The result is a disjunction of six lists. *a* is ordered before *b* and *c* before *d* in all of these lists, since this is also the case in the two lists $\langle\, a, b \,\rangle$ and $\langle\, c, d \,\rangle$ that have been combined. But apart from this, *b* can be placed before, between, or after *c* and *d*.

On the linearization-based approach, every word comes with a domain value that is a list that contains the word itself:

(31) Domain contribution of single words, here *gibt* 'gives':

$$\boxed{1} \begin{bmatrix} \text{PHON} & \langle\, \textit{gibt} \,\rangle \\ \text{SYNSEM} & \ldots \\ \text{DOM} & \langle\, \boxed{1} \,\rangle \end{bmatrix}$$

The description in (31) may seem strange at first glance, since it is cyclic, but it can be understood as a statement saying that *gibt* contributes itself to the items that occur in linearization domains.

The constraint in (32) is responsible for the determination of the PHON values of phrases:

(32) $$phrase \Rightarrow \begin{bmatrix} \text{PHON} & \boxed{1} \oplus \ldots \oplus \boxed{n} \\ \text{DOM} & \left\langle \begin{bmatrix} sign \\ \text{PHON} \ \boxed{1} \end{bmatrix}, \ldots, \begin{bmatrix} sign \\ \text{PHON} \ \boxed{n} \end{bmatrix} \right\rangle \end{bmatrix}$$

It states that the PHON value of a sign is the concatenation of the PHON values of its DOMAIN elements. Since the order of the DOMAIN elements corresponds to their surface order, this is the obvious way to determine the PHON value of the whole linguistic object.

Figure 9 shows how this machinery can be used to license binary branching structures with discontinuous constituents in the sentence *dass dem Kind ein Mann das Buch gibt* 'that a man gives the child the book'. Words or word se-

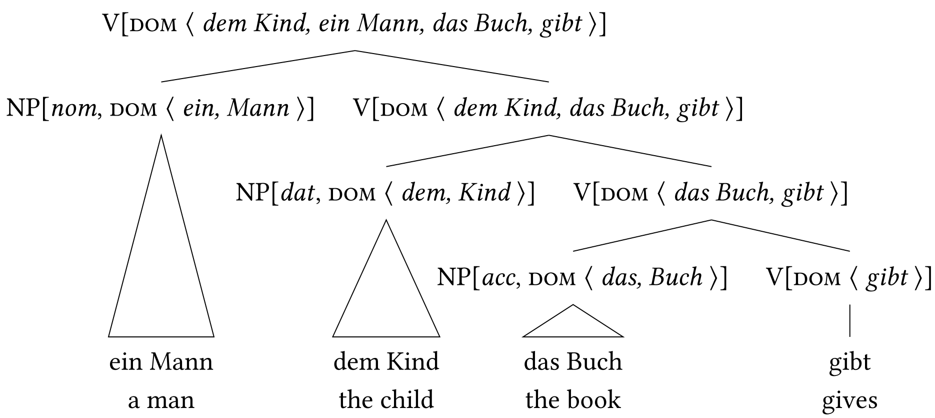


Figure 9: Analysis of *dass dem Kind ein Mann das Buch gibt* 'that a man gives the child the book' with binary branching structures and discontinuous constituents. The tree shows the order of combination, which does not correspond to the linearization of the DOMAIN objects.

quences that are separated by commas stand for separate domain objects, that is, ⟨ *das, Buch* ⟩ contains the two objects *das* and *Buch* and ⟨ *das Buch, gibt* ⟩ contains the two objects *das Buch* and *gibt*. The important point to note here is that the arguments in the tree are combined with the head in the order accusative, dative, nominative, although the elements in the constituent order domain (i.e. in the list of DOMAIN elements and in the surface sentence) are realized in the order dative, nominative, accusative, rather than nominative, dative, accusative, which is what one might expect based on the order in which they are combined in the tree. This is possible since the formulation of the computation of the DOM value using the shuffle operator allows for discontinuous constituents. The node for *dem Kind das Buch gibt* 'the child the book gives' is discontinuous: *ein Mann* 'a man' is inserted into the domain between *dem Kind* 'the child' and *das Buch* 'the book'. This is more obvious in Figure 10, which has a serialization of NPs that corresponds to their order.

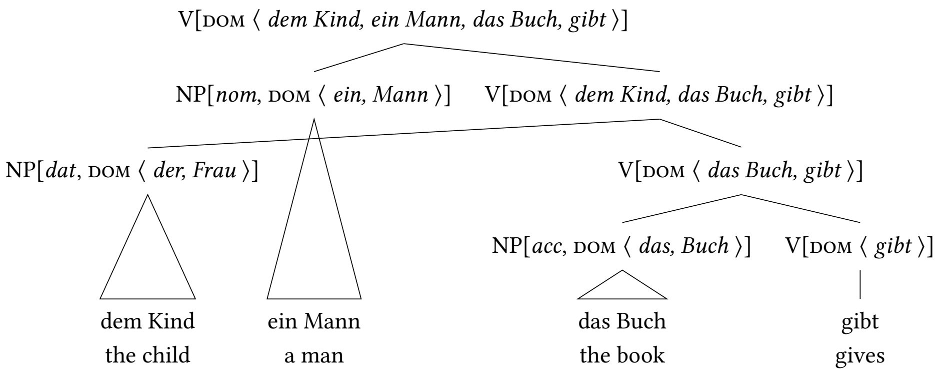


Figure 10: Analysis of *dass dem Kind ein Mann das Buch gibt* 'that a man gives the child the book' with binary branching structures and discontinuous constituents, more clearly showing the discontinuity

6.2 Absolutely free

While German is more striking than English in terms of constituent order, languages like Warlpiri are even more so, since they have much freer constituent order. In Warlpiri the auxiliary has to be in first or in second position (Donohue & Sag 1999: 8), but apart from this, even parts of what are noun phrases in German and English can appear separated from each other. For example, the two parts of the NP *Kurdujarrarlu witajarrarlu* 'child small' may appear discontinuously since they are marked with the same case:

(33) Kurdu-jarra-rlu ka-pala maliki wajili.pi-nyi wita-jarra-rlu.
child-DU-ERG PRES-3DU.SBJ dog.ABS chase-NPST small-DU-ERG
'Two small children are chasing the dog.' or
'Two children are chasing the dog and they are small.'

Donohue & Sag (1999) develop an analysis for this that simply liberates domain elements and inserts them into the next higher domain. (34) shows how this is formalized:

(34) *liberating-phrase* ⇒

$$\begin{bmatrix} \textsc{dom} & \delta_0 \bigcirc \delta_1 \bigcirc \ldots \bigcirc \delta_n \\ \textsc{head-dtr} & \begin{bmatrix} \textsc{dom} & \delta_0 \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle \begin{bmatrix} \textsc{dom} & \delta_1 \end{bmatrix}, \ldots, \begin{bmatrix} \textsc{dom} & \delta_n \end{bmatrix} \right\rangle \end{bmatrix}$$

Rather than inserting the complete daughters into the domain of the mother as in (29), the DOM values of the daughters are shuffled into the domain of the mothers. So instead of having the NPs in the same domain as the verb as in the German example in the previous section, one has all the parts of NPs in the next higher domain. Hence, a single nominal element being placed in front of the auxiliary in second position is explained without difficulty. Figure 11 shows Donohue & Sag's (1999) analysis of a version of (33) with the VP constituents *maliki wajilipinyi* 'dog chase' serialized after *witajarrarlu* 'small'. Here *Kurdujarrarlu* 'child' and

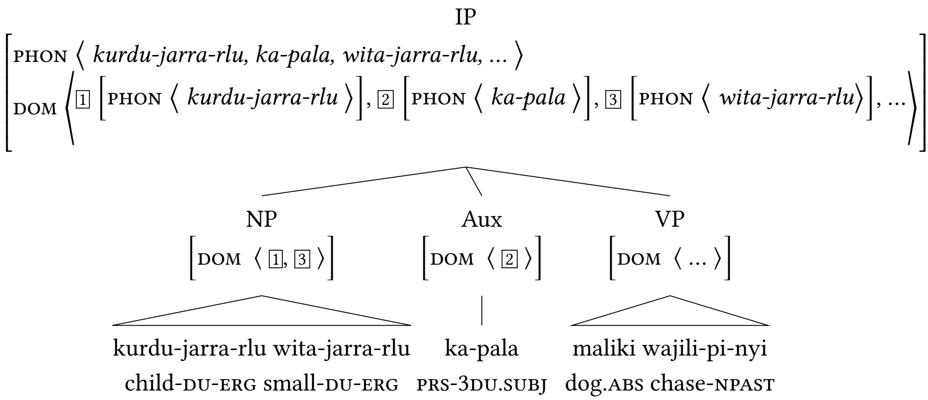


Figure 11: Analysis of free constituent order in Warlpiri according to Donohue & Sag (1999)

witajarrarlu 'small' form an NP. They contribute two independent domain objects

($\boxed{1}$ and $\boxed{3}$) to the domain of the mother. The second element in this domain has to be the auxiliary ($\boxed{2}$), $\boxed{1}$ is realized initially and $\boxed{3}$ follows the auxiliary.

We have seen so far an analysis that inserts complete objects into the domain of the mother (the analysis of German) and an analysis that inserts all domain objects of objects into the domain of the mother (the analysis of Warlpiri), and in the next subsection I want to look at an intermediate case, so-called *partial compaction*.

6.3 Partial compaction (extraposition)

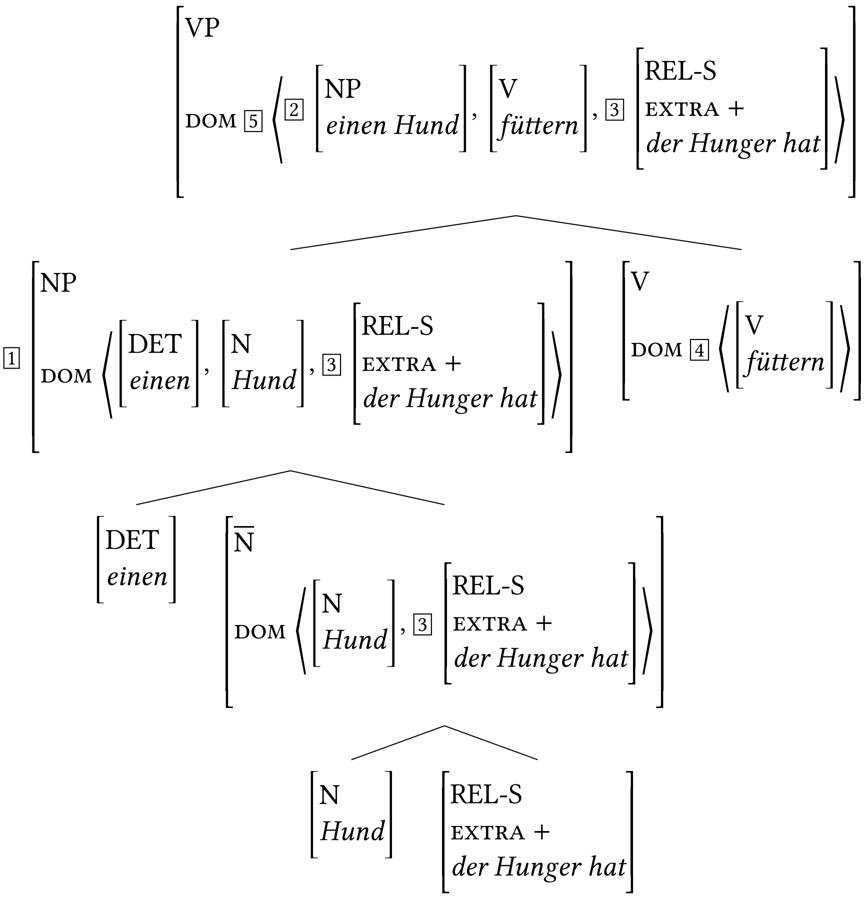
Kathol & Pollard (1995) develop an analysis of extraposition that is a mix of the strategies discussed in the two previous subsections: most of one NP object is inserted into the domain of the mother as a single object, while only those parts that are extraposed are liberated and inserted as individual domain objects into the domain of the mother.[11] Kathol & Pollard's analysis of (35) is given in Figure 12.[12]

(35) einen Hund füttern, der Hunger hat (German)
a dog feed that hunger has
'feed a dog that is hungry'

einen Hund, der Hunger hat 'a dog who is hungry' consists of three domain objects: *einen* 'a', *Hund* 'dog', and *der Hunger hat* 'who is hungry'. The two initial ones are inserted as one object (the NP *ein Hund* 'a dog') into the higher domain and the relative clause is liberated. While the formation of the new domain at the mother node is relatively straightforward in the cases discussed so far, a complex relational constraint is needed to split the relative clause ($\boxed{3}$) from the other domain objects and construct a new domain object that has the determiner and the noun as constituents ($\boxed{2}$). Kathol and Pollard have a relational constraint called *compaction* that builds new domain objects for insertion into higher domains.

[11] This analysis of extraposition is not the only option available in HPSG. I explain it here since it shows the flexibility of the domain approach. The more common analysis of extraposition is one that is parallel to the SLASH-based approach to extraction that is explained in Borsley & Crysmann (2019), Chapter 14 of this volume. Since constraints regarding locality differ for fronting to the left and extraposition to the right, a different feature is used (EXTRA). See Keller (1995) and Müller (1999: Section 13.2) for discussion. More recent approaches assume the projection of semantic indices (Kiss 2005) to be able to solve puzzles like Link's (1984) hydra sentences and even more recent proposals mix index projection and EXTRA projection (Crysmann 2013).

[12] The figure is taken over from Kathol & Pollard. Words in italics are the object language. Part of speech or category labels are provided at the top of AVMs.



p-compaction($\boxed{1}$, $\boxed{2}$, $\langle\ \boxed{3}\ \rangle$)

$\boxed{5} = \langle\ \boxed{2}\ \rangle \bigcirc \langle\ \boxed{3}\ \rangle \bigcirc \boxed{4}$

Figure 12: Analysis of extraposition via partial compaction of domain objects according to Kathol & Pollard (1995)

partial compaction takes an initial part of a domain and forms a new domain object from this, returning the remaining domain objects for separate insertion into the higher domain. Due to space limitations, this constraint will not be discussed here, but see Müller (1999: 244) for a refined version of Kathol and Pollard's constraint. The effect of partial compaction in Figure 12 is that there is a new object $\boxed{2}$ and a list containing the remaining objects, in the example $\langle \boxed{3} \rangle$. A list containing the new object $\langle \boxed{2} \rangle$ and the list containing the remaining objects $\langle \boxed{3} \rangle$ are shuffled with the domain list of the head $\boxed{4}$. Since the relative clause is now in the same domain as the verb, it can be serialized to the right of the verb.

So, this subsection showed how examples like (35) can be analyzed by allowing for a discontinuous constituent consisting of an NP and a relative clause. Rather than liberating all daughters and inserting them into the domain of the mother node as in the Warlpiri example, determiner and noun form a new object, an NP, and the newly created NP and the relative clause are inserted into the domain of the mother node. This explains why determiner and noun have to stay together while the relative clause may be serialized further to the right.

6.4 Problems with order domains

Constituent order domains may seem rather straightforward since linearization facts can be handled easily. I assumed constituent order domains and discontinuous constituents for German myself for over a decade (Müller 1995; 2004). However, there are some problems that seem to suggest that a traditional GB-like head-movement approach is the better alternative. In what follows I want to discuss just two problematic aspects of linearization approaches: spurious ambiguities and apparently multiple frontings.

6.4.1 Partial fronting and spurious ambiguities

Kathol (2000) suggests an analysis with binary branching structures in which all arguments are inserted into a linearization domain and can be serialized there in any order, provided no LP rule is violated. Normally one would have the elements of the COMPS list in a fixed order, combine the head with one element from the COMPS list after another, and let the freedom in the DOM list be responsible for the various attested orders. So, both sentences in (36) would have analyses in which the verb *erzählt* 'tells' is combined with *Geschichten* 'stories' first and then *Geschichten erzählt* 'stories tells' is combined with *den Wählern* 'the voters'. Since the verb and all its arguments are in the same linearization domain they can be ordered in any way, including the two possibilities in (36):

(36) a. weil er den Wählern Geschichten erzählt (German)
because he the voters stories tells
'because he tells the voters stories'

b. weil er Geschichten den Wählern erzählt
because he stories the voters tells

The problem with this approach is that examples like (37) show that grammars have to account for fronted combinations of any of the verb and any of its objects to the exclusion of the other:

(37) a. Geschichten erzählen sollte man den Wählern nicht. (German)
stories tell should one the voters not
'One should not tell the voters such stories.'

b. Den Wählern erzählen sollte man diese Geschichten nicht.
the voters tell should one these stories not

Kathol (2000: Section 8.9) accounts for examples like (37) by relaxing the order of the objects in the valence list. He uses the shuffle operator in the valence representation:

(38) $\langle$ NP[*nom*] $\rangle \oplus (\langle$ NP[*dat*] $\rangle \bigcirc \langle$ NP[*acc*] $\rangle)$

This solves the problem with examples like (37) but it introduces a new one: sentences like (36) now have two analyses each. One is the analysis we had before and another one is the one in which *den Wählern* 'the voters' is combined with *erzählt* 'tells' first and the result is then combined with *Geschichten* 'stories'. Since both objects are inserted into the same linearization domain, both orders can be derived. So we have too much freedom: freedom in linearization and freedom in the order of combination. The proposal that I suggested in Müller (2005a: Section 2.1; 2017: Section 2.2.1) and which is implemented in Schema 3 above has just the freedom in the order of combination and hence can account for both (36) and (37) without spurious ambiguities.

6.4.2 Surface order, clause types, fields within fields, and empty elements

Kathol (2001) develops an analysis of German that uses constituent order domains and determines the clause types on the basis of the order of elements in such domains. He suggests the topological fields *1*, *2*, *3*, and *4*, which correspond to the traditional topological fields *Vorfeld* 'prefield', *linke Satzklammer* 'left sentence bracket', *Mittelfeld* 'middle field', *rechte Satzklammer* 'right sentence bracket'. Domain objects may be assigned to these fields, and they are then ordered by linearization constraints stating that objects assigned to *1* have to precede objects of type *2*, type *3*, and type *4*. Objects of type *2* have to precede type *3*, and type *4* and so on. For the *Vorfeld* and the left sentence bracket, he stipulates uniqueness constraints saying that at most one constituent may be of this type. This can be stated in a nice way by using the linearization constraints in (39):

(39) a. *1* < *1*
     b. *2* < *2*

This trick was first suggested by Gazdar et al. (1985: 55, Fn. 3) in the framework of GPSG and it works because, if there were two objects of type *1*, then each one would be required to precede the other one, resulting in a violation of the linearization constraint. So in order to avoid such constraint violation there must not be more than one *1*.

Kathol (2001) assumes the following definition for V2 clauses:

(40) $$\textit{V2-clause} \Rightarrow \begin{bmatrix} \text{S}[\textit{fin}] \\ \text{DOM} \quad \left\langle [1], \begin{bmatrix} 2 \\ \text{V}[\textit{fin}] \end{bmatrix}, \ldots \right\rangle \end{bmatrix}$$

This says that the constituent order domain starts with one element assigned to field *1*, followed by another domain object assigned to field *2*. While this is in accordance with general wisdom about German, which is a V2 language, there are problems for entirely surface-based theories: German allows for multiple constituents in front of the finite verb. (41) shows some examples:

(41) a. [Zum zweiten Mal] [die Weltmeisterschaft] errang Clark 1965 … [13]
        to.the second time the world.championship won Clark 1965
        'Clark won the world championship for the second time in 1965.'
     b. [Dem Saft] [eine kräftige Farbe] geben Blutorangen.[14]
        the.DAT juice a.ACC strong color give blood.oranges
        'Blood oranges give the juice a strong color.'

Müller (2003a) extensively documents this phenomenon. The categories that can appear before the finite verb are almost unrestricted. Even subjects can be fronted

[13] (Beneš 1971: 162)

[14] Bildhauer & Cook (2010) found this example in the *Deutsches Referenzkorpus* (DeReKo), hosted at Institut für Deutsche Sprache, Mannheim: http://www.ids-mannheim.de/kl/projekte/korpora, 2018-09-13.

together with other material (Bildhauer & Cook 2010: 72; Bildhauer 2011: 371). The empirical side of these apparent multiple frontings was further examined in the Collective Research Center 632, Project A6, and the claim that only constituents depending on the same verb can be fronted together (Fanselow 1993; Hoberg 1997: 1634) was confirmed (Müller 2017: Chapter 3). A further insight is that the linearization properties of the fronted material (NPs, PPs, adverbs, adjectives) correspond to the linearization properties they would have in the *Mittelfeld*. The example in (42) is even more interesting. It shows that there can be a right sentence bracket (the particle *los*) and an extraposed constituent (something following the particle: *damit*) before the finite verb (*geht* 'goes'):

(42) *Los* damit *geht* es schon am 15. April.[15] (German)
off there.with goes it PRT on 15. April
4 5 2 3 3 3
'The whole thing starts on April 15th.'

As far as topology is concerned, this sentence corresponds to sentences with VP fronting and extraposition like the one in (43) discussed in Reis (1980: 82).

(43) [Gewußt, daß du kommst,] haben wir schon seit langem. (German)
known that you come have we PART since long
'We have known for a while that you are coming.'

In (43) *gewußt, dass du kommst* 'known that you come' forms a VP in which *gewußt* is the right sentence bracket and *daß du kommst* 'that you come' is extraposed. We have the same situation in (42) with *los* 'off' and *damit* 'there.with', except that one would not want to claim that *damit* 'there.with' depends on *los* 'off'.

In Kathol's system, *los* would be of type *4* and *damit* would have to be of type *5* (an additional type for extraposed items). Without any modification of the general system, we would get a *4* and a *5* ordered before a *2* (a right sentence bracket and a postfield preceding the left sentence bracket), something that is ruled out by Kathol's linearization constraints.

Müller (2002), still working in a domain-based framework, developed an analysis assuming an empty verbal head to explain the fact that the fronted constituents have to depend on the same verb and that there is a separate topological area that is independent of the remaining clause. So, *los* and *damit* are domain objects within a larger domain object placed in the prefield. Wetta (2011) suggests an analysis in which two or more constituents are compacted into one

[15] taz, 01.03.2002, p. 8.

domain object, so *los* and *damit* would form one object that is inserted into the domain containing the finite verb. However, this begs the question of what kind of object it is that is formed. Section 6.3 dealt with partial compaction of NPs. Some of the elements from an NP domain were liberated and other elements were fused into a new object that had the same category as the object containing all material, namely NP. But the situation with examples like (41) and (43) is quite different. We have a particle and a pronominal adverb in (43) and various other combinations of categories in the examples collected by Müller (2003a) and Bildhauer (2011). It would not make sense to claim that the fronted object is a particle or a pronominal adverb. Note that it is not an option to leave the category of the fronted object unspecified, since HPSG comes with the assumption that models of linguistic objects are total, that is, maximally specific (King 1999, see also Richter (2019), Chapter 3 of this volume). Leaving the category and valence properties of the item in the prefield unspecified would make such sentences infinitely ambiguous. Of course Wetta could state that the newly created object is a verbal projection, but this would just be stating the effect of the empty verbal head within a relational constraint, which I consider less principled than stating the empty element.

However, the empty verbal head that I stated as part of a linearization grammar in 2002 comes as a stipulation, since its only purpose in the grammar of German was to account for apparent multiple frontings. Müller (2005b; 2017) drops the linearization approach and assumes head-movement instead. The empty head that is used for accounting for the verb position in German can also be used to account for apparent multiple frontings. The analysis is sketched in (44):

(44) a. [$_{\text{VP}}$ [Zum zweiten Mal] [die Weltmeisterschaft] $\__{\text{V}}$]$_i$ errang$_j$
to.the second time the world.championship won
Clark 1965 $\__i$ $\__j$.
Clark 1965

b. [$_{\text{VP}}$ *Los* $\__{\text{V}}$ damit]$_i$ *geht*$_j$ es schon am 15. April $\__i$ $\__j$.
off there.with goes it PRT on 15. April
'The whole thing starts on the 15th April.'

Space precludes going into all the details here, but the analysis treats apparent multiple frontings parallel to partial verb phrase frontings. A lexical rule is used for multiple frontings which is a special case of the head-movement rule that was discussed in Section 5.1. So, apparent multiple frontings are analyzed with means that are available to the grammar anyway. This analysis allows us to keep the insight that German is a V2 language and it also gets the same-clause constraint and the linearization of elements right. As for (44b): *los damit* 'off there.with' forms a verbal constituent placed in the *Vorfeld* and within this verbal domain, we have the topological fields that are needed: the right sentence bracket for the verbal particle and the verbal trace and the *Nachfeld* for *damit* 'there.with'. See Müller (2005a,b; 2017) for details.

This paper so far has discussed the tools that have been suggested in HPSG to account for constituent order: flat vs. binary branching structures, linearization domains, head-movement via DSL. I showed that analyses of German relying on discontinuous constituents and constituent order domains are not without problems and that head-movement approaches with binary branching and continuous constituents can account for the data. I also demonstrated in Section 6.2 that languages like Warlpiri that allow for much freer constituent order than German can be accounted for in models allowing for discontinuous constituents. The following section discusses a proposal by Bender (2008) that shows that even languages like the Australian free constituent order languages can be handled without discontinuous constituents.

7 Free constituent order languages without order domains

Bender (2008) discusses the Australian language Wambaya and shows how phenomena parallel to those treated by Donohue & Sag (1999) can be handled without discontinuous constituents. Bender assumes that all arguments of a head are projected to higher nodes even when they are combined with the head; that is, arguments are not canceled off from valence lists. See also Meurers (1999); Przepiórkowski (1999) and Müller (2008) for earlier non-cancellation approaches.[16] The example (33) from Section 6.2 can be recast with continuous constituents as shown in Figure 13. The figure shows that arguments are not removed from the valence representation after combination with the head. Rather they are marked as satisfied: $\boxed{\cancel{1}}$. Since they are still in the representation, schemata may refer to them. Bender suggests a schema that identifies the MOD value of an element that could function as an adjunct in a normal head-adjunct structure with an element in the valence representation. In Figure 13, the MOD value of the second ergative nominal *wita-jarra-rlu* 'small' is identified with an argument of the auxiliary verb ($\boxed{1}$). The adjunct hence has access to the referential index of the argument and it is therefore guaranteed that both parts of the noun phrase refer to the same discourse referent. The NP for *kurdu-jarra-rlu* is combined with the projection

[16] Higginbotham (1985) and Winkler (1997) make similar suggestions with regard to the representation of theta roles.

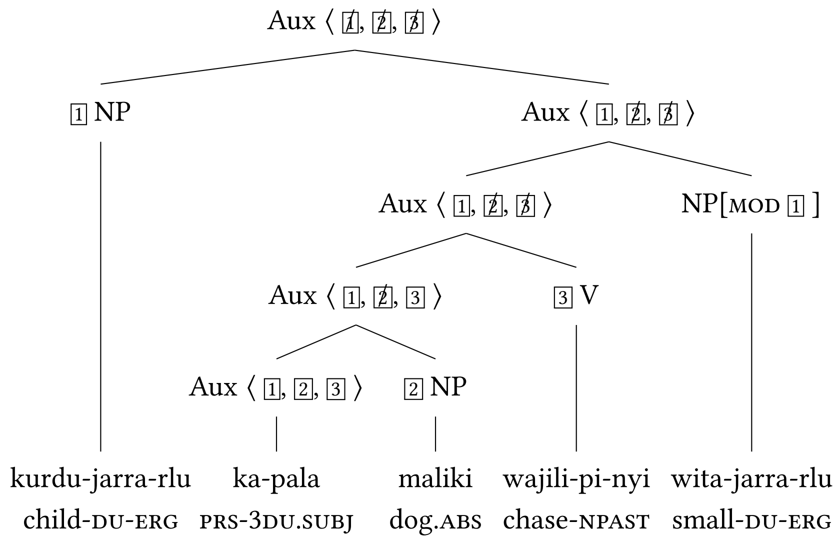


Figure 13: Analysis of free constituent order in Warlpiri using non-cancellation

of the auxiliary to yield a complete sentence. Since $\boxed{1}$ does not only contain the semantic index and hence information about number (the dual) but also case information, it is ensured that distributed noun phrases have to bear the same case. Since information about all arguments are projected along the head path, $\boxed{2}$ would also be available for an adjunct referring to it. So in the place of *wita-jarra-rlu* 'small-DU-ERG' we could also have another adjunct referring to *maliki* 'dog.ABS'. This shows that even languages with constituent order as free as the Australian languages can be handled within HPSG without assuming discontinuous constituents.

8 Summary

A major feature of constraint-based analyses is that when no constraints are stated, there is freedom. The paper discussed the order of head and adjunct: if the order of head and adjunct is not constrained, both orders are admitted.

This paper explored general approaches to constituent order in HPSG. On the one hand, there are approaches to constituent order assuming flat constituent structure, allowing permutation of daughters as long as no LP constraint is violated. On the other hand, there are approaches assuming binary branching structures. Approaches that assume flat structures can serialize the head to the left or to the right or somewhere between other daughters in the structure. Approaches assuming binary branching have to use other means. One possibility is "head movement", which is analyzed as a series of local dependencies by passing information about the missing head up along the head path. The alternative to head movement is linearization of elements in special linearization domains, allowing for discontinuous constituents. I showed that there are reasons for assuming head-movement for German and how even languages with extremely free constituent order can be analyzed without assuming discontinuous constituents.

There is also the work on heaviness and lightness, to account for the (reduced) mobility of various categories: bare nominals in various languages, certains pronouns (A&G 1999a), certains adverbs (**AG2004a-u-French**; Abeillé & Godard 1997), negation (Abeillé & Godard 2001), and attributive adjectives (**AG99b-u-French**). In various papers, Abeillé and Godard propose a three valued WEIGHT feature to account for light, middle-weight and heavy constituents (**AG2004a-u-French**; Abeillé & Godard 2000). The LP rules ususally apply to middle weight ones. The light/heavy distinction is also relevant for the Complex predicate chapter, but it should definitely be mentioned, maybe as a subsection or word order/constituent order since it is part of what LFG people call X° syntax.

Acknowledgements

I thank Anne Abeillé, Bob Borsley, and Doug Ball for very detailed and very useful comments. I thank Elizabeth Pankratz for comments and for proofreading.

References

Abeillé, Anne. 2019. Control and raising. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 351–352. Berlin: Language Science Press. DOI:??

Abeillé, Anne & Rui P. Chaves. 2019. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 469–516. Berlin: Language Science Press. DOI:??

Abeillé, Anne & Danièle Godard. 1997. The syntax of French negative adverbs. In Danielle Forget, Paul Hirschbühler, France Martineau & María Luisa Rivero (eds.), *Negation and polarity syntax and semantics: Selected papers from the colloquium Negation: Syntax and Semantics. Ottawa, 11–13 May 1995* (Current Issues in Linguistic Theory 155), 1–28. Amsterdam: John Benjamins Publishing Co. DOI:10.1075/cilt.155.02abe

Abeillé, Anne & Danièle Godard. 2000. French word order and lexical weight. In Robert D. Borsley (ed.), *The nature and function of syntactic categories*, vol. 32 (Syntax and Semantics), 325–360. San Diego: Academic Press.

Abeillé, Anne & Danièle Godard. 2001. A class of "lite" adverbs in French. In Joaquim Camps & Caroline R. Wiltshire (eds.), *Romance syntax, semantics and L2 acquisition: Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000* (Current Issues in Linguistic Theory 216), 9–25. Amsterdam: John Benjamins Publishing Co. DOI:10.1075/cilt.216.04abe

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1. 1–27.

Bach, Emmon. 1962. The order of elements in a Transformational Grammar of German. *Language* 8(3). 263–269.

Bender, Emily M. 2008. Radical non-configurationality without shuffle operators: An analysis of Wambaya. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 6–24. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/, accessed 2018-2-25.

Beneš, Eduard. 1971. Die Besetzung der ersten Position im deutschen Aussagesatz. In Hugo Moser (ed.), *Fragen der strukturellen Syntax und der kontrastiven Grammatik* (Sprache der Gegenwart – Schriften des IdS Mannheim 17), 160–182. Düsseldorf: Pädagogischer Verlag Schwann.

Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs* (studia grammatica 2). Berlin: Akademie Verlag.

Bildhauer, Felix. 2011. Mehrfache Vorfeldbesetzung und Informationsstruktur: Eine Bestandsaufnahme. *Deutsche Sprache* 39(4). 362–379.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topic-hood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Bjerre, Tavs. 2006. *Object positions in a topological sentence model for Danish: A linearization-based HPSG approach*. http://www.hum.au.dk/engelsk/engsv/objectpositions/workshop/Bjerre.pdf, accessed 2018-2-25. Presentation at Ph.D.-Course at Sandbjerg, Denmark.

Blevins, James P. 1994. Derived constituent order in unbounded dependency constructions. *Journal of Linguistics* 30(2). 349–409.

Bonami, Olivier & Danièle Godard. 2007. Integrating linguistic dimensions: The scope of adverbs. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 25–45. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2007/bonami-godard.pdf, accessed 2018-2-25.

Bonami, Olivier, Danièle Godard & Jean-Marie Marandin. 1999. Constituency and word order in French subject inversion. In Gosse Bouma, Erhard W. Hinrichs, Geert-Jan M. Kruijff & Richard T. Oehrle (eds.), *Constraints and resources in natural language syntax and semantics* (Studies in Constraint-Based Lexicalism 3), 21–40. Stanford, CA: CSLI Publications.

Borsley, Bob & Anne Abeillé. 2019. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 3–5. Berlin: Language Science Press. DOI:??

Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 1987. *Subjects and complements in HPSG*. Report No. CSLI-87-107. Stanford, CA: Center for the Study of Language & Information.

Borsley, Robert D. 1989. Phrase-Structure Grammar and the Barriers conception of clause structure. *Linguistics* 27(5). 843–863.

Borsley, Robert D. 2006. *Syntactic and lexical approaches to unbounded dependencies*. Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.

Borsley, Robert D. & Stefan Müller. 2019a. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 937–1012. Berlin: Language Science Press. DOI:??

Borsley, Robert D. & Stefan Müller. 2019b. Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. To appear.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press.

Branco, António. 2019. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 613–665. Berlin: Language Science Press. DOI:??

Chaves, Rui P. 2009. Construction-based cumulation and adjunct extraction. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 47–67. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2009/, accessed 2018-2-25.

Chomsky, Noam. 2010. *Restricting stipulations: consequences and challenges*. Talk given in Stuttgart.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York, Oxford: Oxford University Press.

Crysmann, Berthold. 2013. On the locality of complement clause and relative clause extraposition. In Gert Webelhuth, Manfred Sailer & Heike Walker (eds.), *Rightward movement in a comparative perspective* (Linguistik Aktuell/Linguistics Today 200), 369–396. Amsterdam: John Benjamins Publishing Co.

Donohue, Cathryn & Ivan A. Sag. 1999. Domains in Warlpiri. In *Sixth International Conference on HPSG–Abstracts. 04–06 August 1999*, 101–106. Edinburgh.

Dowty, David R. 1996. Toward a Minimalist theory of syntactic structure. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 11–62. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

Eisenberg, Peter. 1994. German. In Ekkehard König & Johan van der Auwera (eds.), *The Germanic languages* (Routledge Language Family Descriptions), 349–387. London: Routledge.

Engelkamp, Judith, Gregor Erbach & Hans Uszkoreit. 1992. Handling linear precedence constraints by unification. In Henry S. Thomson (ed.), *30th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 201–208. Newark, DE: Association for Computational Linguistics. auch als CLAUS-Report, Nr. 19, Universität des Saarlandes erschienen.

Eynde, Frank Van. 2019. Nominal structures. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 185–223. Berlin: Language Science Press. DOI:??

Fanselow, Gisbert. 1993. Die Rückkehr der Basisgenerierer. *Groninger Arbeiten zur Germanistischen Linguistik* 36. 1–74.

Feldhaus, Anke. 1997. *Eine HPSG-Analyse ausgewählter Phänomene des deutschen* w*-Fragesatzes*. Working Papers of the Institute for Logic and Linguistics 27. IBM Scientific Center Heidelberg: Institute for Logic & Linguistics.

Fillmore, Charles J. 1999. Inversion and Constructional inheritance. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 113–128. Stanford, CA: CSLI Publications.

Frank, Anette. 1994. *Verb second by lexical rule or by underspecification.* Arbeitspapiere des SFB 340 No. 43. Heidelberg: IBM Deutschland GmbH. ftp://ftp.ims.uni-stuttgart.de/pub/papers/anette/v2-usp.ps.gz, accessed 2018-3-20.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar.* Cambridge, MA: Harvard University Press.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, 1–21. Stanford, CA: CSLI Publications.

Haider, Hubert. 2019. VO/OV-base ordering. In B. Richard Page & Michael T. Putnam (eds.), *Cambridge handbook of Germanic linguistics* (Cambridge Handbooks in Language and Linguistics). Cambridge, UK: Cambridge University Press.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16(4). 547–593.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989. Subcategorization and VP structure in German. In *Aspects of German VP structure* (SfS-Report-01-93). Tübingen: Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–38. Stanford, CA: CSLI Publications.

Hoberg, Ursula. 1997. Die Linearstruktur des Satzes. In Hans-Werner Eroms, Gerhard Stickel & Gisela Zifonun (eds.), *Grammatik der deutschen Sprache*, vol. 7.2 (Schriften des Instituts für deutsche Sprache), 1495–1680. Berlin: Walter de Gruyter.

Höhle, Tilman N. 1982. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 75–153. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag. Republished as Höhle (2018).

Höhle, Tilman N. 2018. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 107–191. Berlin: Language Science Press. DOI:10.5281/zenodo.1145680

Jacobson, Pauline. 1987. Phrase structure, grammatical relations, and discontinuous constituents. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Discontinuous constituency* (Syntax and Semantics 20), 27–69. New York: Academic Press.

Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 39–70. Stanford, CA: CSLI Publications.

Kathol, Andreas. 1998. Constituency and linearization of verbal complexes. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 221–270. San Diego: Academic Press.

Kathol, Andreas. 2000. *Linear syntax*. New York, Oxford: Oxford University Press.

Kathol, Andreas. 2001. Positional effects in a monostratal grammar of German. *Journal of Linguistics* 37(1). 35–66.

Kathol, Andreas & Carl J. Pollard. 1995. Extraposition via complex domain formation. In Hans Uszkoreit (ed.), *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 174–180. Cambridge, MA: Association for Computational Linguistics.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

Keller, Frank. 1995. Towards an account of extraposition in HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, 301–306. Dublin: Association for Computational Linguistics.

Kim, Jong-Bok & Ivan A. Sag. 2002. Negation without head-movement. *Natural Language and Linguistic Theory* 20(2). 339–412. DOI:10.1023/A:1015045225019

King, Paul. 1999. Towards truth in Head-Driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitsberichte des SFB 340 No. 132), 301–352. Tübingen: Universität Tübingen. http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/132/132abs.html, accessed 2018-2-25.

Kiss, Tibor. 1993. *Infinite Komplementation – Neue Studien zum deutschen Verbum infinitum*. Arbeiten des SFB 282 No. 42. Bergische Universität Gesamthochschule Wuppertal.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.

Kiss, Tibor. 2005. Semantic constraints on relative clause extraposition. *Natural Language and Linguistic Theory* 23(2). 281–334. DOI:10.1007/s11049-003-1838-7

Kiss, Tibor & Birgit Wesche. 1991. Verb order and head movement. In Otthein Herzog & Claus-Rainer Rollinger (eds.), *Text understanding in LILOG* (Lecture Notes in Artificial Intelligence 546), 216–242. Berlin: Springer Verlag.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions* (CSLI Lecture Notes 166). Stanford, CA: CSLI Publications.

Link, Godehard. 1984. Hydras: On the logic of relative constructions with multiple heads. In Fred Landmann & Frank Veltman (eds.), *Varieties of formal semantics*, 245–257. Dordrecht: Foris Publications.

Meurers, Walt Detmar. 1999. Raising spirits (and assigning them case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43. 173–226. http://www.sfs.uni-tuebingen.de/~dm/papers/gagl99.html, accessed 2018-2-25.

Meurers, Walt Detmar. 2000. *Lexical generalizations in the syntax of German non-finite constructions*. Arbeitspapiere des SFB 340 No. 145. Tübingen: Eberhard-Karls-Universität. http://www.sfs.uni-tuebingen.de/~dm/papers/diss.html, accessed 2018-2-25.

Müller, Stefan. 1995. Scrambling in German – Extraction into the *Mittelfeld*. In Benjamin K. T'sou & Tom Bong Yeung Lai (eds.), *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*, 79–83. City University of Hong Kong.

Müller, Stefan. 1996. The Babel-System: An HPSG fragment for German, a parser, and a dialogue component. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, 263–277. London.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2003a. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

Müller, Stefan. 2003b. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI:10.1017/S0022226703002032

Müller, Stefan. 2004. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation* 2(2). 209–257.

Müller, Stefan. 2005a. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201. 3–39.

Müller, Stefan. 2005b. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330.

Müller, Stefan. 2008. Depictive secondary predicates in German and English. In Christoph Schroeder, Gerd Hentschel & Winfried Boeder (eds.), *Secondary predicates in Eastern European languages and beyond* (Studia Slavica Oldenburgensia 16), 255–273. Oldenburg: BIS-Verlag.

Müller, Stefan. 2013. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950. DOI:10.1353/lan.2013.0061

Müller, Stefan. 2015a. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), 937–973. Berlin: Walter de Gruyter. DOI:10.1515/9783110363708-004

Müller, Stefan. 2015b. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.91

Müller, Stefan. 2016. Satztypen: Lexikalisch oder/und phrasal. In Rita Finkbeiner & Jörg Meibauer (eds.), *Satztypen und Konstruktionen im Deutschen* (Linguistik – Impulse & Tendenzen 65), 72–105. Berlin: de Gruyter.

Müller, Stefan. 2017. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Revise and resubmit.

Müller, Stefan. 2018. *A lexicalist account of argument structure: Template-based phrasal LFG approaches and a lexical HPSG alternative* (Conceptual Foundations of Language Science 2). Berlin: Language Science Press. DOI:10.5281/zenodo.1441351

Müller, Stefan. 2019. *Germanic syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin. https://hpsg.hu-berlin.de/~stefan/Pub/germanic.html, accessed 2019-6-30.

Müller, Stefan & Bjarne Ørsnes. 2013. Towards an HPSG analysis of object shift in Danish. In Glyn Morrill & Mark-Jan Nederhof (eds.), *Formal Grammar: 17th and 18th International Conferences, FG 2012, Opole, Poland, August 2012, revised* *selected papers, FG 2013, Düsseldorf, Germany, August 2013: Proceedings* (Lecture Notes in Computer Science 8036), 69–89. Berlin: Springer Verlag.

Müller, Stefan & Bjarne Ørsnes. 2015. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In Preparation.

Netter, Klaus. 1992. On non-head non-movement: An HPSG treatment of finite verb position in German. In Günther Görz (ed.), *Konvens 92. 1. Konferenz „Verarbeitung natürlicher Sprache". Nürnberg 7.–9. Oktober 1992* (Informatik aktuell), 218–227. Berlin: Springer Verlag.

Nykiel, Joanna & Jong-Bok Kim. 2019. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 585–612. Berlin: Language Science Press. DOI:??

Oliva, Karel. 1992. *Word order constraints in binary branching syntactic structures*. CLAUS-Report 20. Saarbrücken: Universität des Saarlandes.

Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Przepiórkowski, Adam. 1999. On case assignment and "adjuncts as complements". In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 231–245. Stanford, CA: CSLI Publications.

Przepiórkowski, Adam. 2019. Case. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 155–183. Berlin: Language Science Press. DOI:??

Pullum, Geoffrey K. 1977. Word order universals and grammatical relations. In Peter Cole & Jerrold M. Sadock (eds.), *Grammatical relations* (Syntax and Semantics 8), 249–277. New York, San Francisco, London: Academic Press.

Reape, Mike. 1991. *Word order variation in Germanic and parsing*. DYANA Report Deliverable R1.1.C. University of Edinburgh.

Reape, Mike. 1992. *A formal theory of word order: A case study in West Germanic*. University of Edinburgh dissertation.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Reis, Marga. 1974. Syntaktische Hauptsatzprivilegien und das Problem der deutschen Wortstellung. *Zeitschrift für Germanistische Linguistik* 2(3). 299–327.

Reis, Marga. 1980. On justifying topological frames: 'Positional field' and the order of nonverbal constituents in German. *Documentation et Recherche en Linguistique Allemande Contemporaine. Revue de Linguistique* 22/23. 59–85.

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 45–46. Berlin: Language Science Press. DOI:??

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A., Rui Chaves, Anne Abeillé, Bruno Estigarribia, Frank Van Eynde, Dan Flickinger, Paul Kay, Laura A. Michaelis-Cummings, Stefan Müller, Geoffrey K. Pullum & Tom Wasow. 2020. Lessons from the English auxiliary system. *Journal of Linguistics* 56. Ahead of print (online). DOI:10.1017/S002222671800052X

Steedman, Mark J. 2000. *The syntactic process* (Language, Speech, and Communication). Cambridge, MA: MIT Press.

Thiersch, Craig L. 1978. *Topics in German syntax*. M.I.T. Dissertation. http://hdl.handle.net/1721.1/16327, accessed 2018-2-25.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

van Noord, Gertjan & Gosse Bouma. 1994. The scope of adjuncts and the processing of lexical rules. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 250–256. Kyoto, Japan: Association for Computational Linguistics. http://www.aclweb.org/anthology/C94-1039, accessed 2018-7-5.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Wells, Rulon S. 1947. Immediate constituents. *Language* 23(2). 81–117.

Wetta, Andrew C. 2011. A Construction-based cross-linguistic analysis of V2 word order. In Stefan Müller (ed.), *Proceedings of the 18th International Con-*

ference on Head-Driven Phrase Structure Grammar, University of Washington, 248–268. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2011/, accessed 2018-2-25.

Wetta, Andrew Charles. 2014. *Construction-based approaches to flexible word order*. Buffalo: State University of New York at Buffalo dissertation. http://gradworks.umi.com/36/83/3683106.html, accessed 2016-3-11.

Winkler, Susanne. 1997. *Focus and secondary predication* (Studies in Generative Grammar 43). Berlin, New York: Mouton de Gruyter.

Chapter 11

Clitics

Berthold Crysmann
Université Paris Diderot

Gerald Penn
University of Toronto

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla

Berthold Crysmann & Gerald Penn. 2019. Clitics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 313–314. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 12

Complex predicates

Danièle Godard
Université Paris Diderot

Pollet Samvelian
Université Sorbonne Nouvelle

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 What are Complex Predicates?

1.1 Definition

The term *complex predicate* does not have a universally accepted definition. In the HPSG tradition, a complex predicate (CP) is composed of two or more words, which are predicates, of which one is a verb, head of the construction, and the other belongs to diverse categories: non-finite verb, noun, adjective, preposition. Both predicates merge their argument structures: the head inherits the arguments of its complement.

To take an example, tense auxiliaries and the participle in Romance languages are two different words, since they can be separated by adverbs, as in (1), but

Danièle Godard & Pollet Samvelian. 2019. Complex predicates. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 315–350. Berlin: Language Science Press. DOI:??

the two verbs belong to the same clause, and, more precisely, the syntactic arguments belong to one argument structure. We admit that the property of monoclausality can manifest itself differently in different languages (Butt (2010) , *pace* Haspelmath (2016)). In the case of Romance auxiliary constructions, the first verb (the auxiliary) hosts the clitics which pronominalize the arguments of the main predicate: corresponding to the NP complement in (1a) the pronominal clitic *l(e)* is hosted by the auxiliary *a* (1b), (1c). This contrasts with the construction of a control verb such as *vouloir* ('to want') in French, where the clitic corresponding to the argument of the infinitive is hosted by the infinitive:

(1) a. Paul a rapidement lu son livre. (French)
'Paul has quickly read his book.'
b. Paul l'a rapidement lu.
Paul it-has quickly read
c. * Paul a rapidement le-lu.

(2) a. Paul veut lire son livre.
'Paul wants to read his book.'
b. Paul veut le lire.
Paul wants to it-read
c. * Paul le veut lire.[1]

Complex predicates are sometimes given a semantic definition: the two elements constitute one predicate which describes one situation. However, a semantic definition is not sufficient. It is true that the head verb tends to add tense, aspectual or modal information while the other element describes a situation type. Thus, in (1), the two verbs jointly describe one situation, the auxiliary adding tense and aspect information. But the semantics of a complex predicate is not always different from that of ordinary verbal complements. Thus, there is no evident semantic distinction depending on whether the Italian restructuring verb *volere* ('want') is the head of a complex predicate (3a) or not (3b).

(3) a. Anna lo vuole comprare. (Italian, Monachesi 1998)
Anna it wants to-buy
'Anna wants to-buy it.'

[1]Possible in an earlier stage of French.

b. Anna vuole comprarlo.
Anna wants to-buy-it
'Anna wants to buy it.'

In other cases, the term is used to describe the complex content of a word. For instance, the verb *dance* has been analyzed as incorporating the noun *dance* and considered a complex predicate (Hale & Keyser 1997). In the sense adopted here, complex predicates involve at least two words.

It is more difficult to distinguish CPs from *Serial verb constructions* (SVC) which are composed of verbs which belong to the same clause (4). However, following Haspelmath (2016), there are a number of differences: (i) SVC are sequences of verbs, and cannot involve various categories, contrary to CP. (ii) These verbs are independent of each other: they have the same form and meaning when they occur outside the SVC, and they do not involve predicate argument relationship between them, in particular infinitival complements and causative constructions are excluded (see 3a). (iii) Finally, they tend to each describe a subevent, being frequently directionals or resultatives (Prince 2017), while a tense auxiliary as in the French example in (1) is not associated with a subevent.

(4) Òzó sàán rrá ógbà.
Ozo jump cross fence
'Ozo jumped over the fence.' (edo, Haspelmath 2016)

1.2 Constructions based on complex predicates

Complex predicates enter into a number of constructions across languages. The following have been particularly studied in HPSG:

- Tense auxiliaries, the copulas and predicative verbs, restructuring verbs, and certain causative and perception verbs in Romance languages (Abeillé & Godard 2002; 2005; 2010; Abeillé et al. 1998; Monachesi 1998);
- Coherent constructions in German and Dutch, headed by tense auxiliaries, copulas, certain raising and control verbs, certain verbs with predicative complements, and particle verbs (De Kuthy & Meurers 2001; Hinrichs & Nakazawa 1994; 1998; Kathol 1998; Kiss 1994; Meurers 2002; Müller 2002; Bouma & van Noord 1998; Rentier 1994);
- Korean auxiliaries, control verbs and causative verb *ha* (Chung 1998; Kim 2016; Sells 1991; Yoo 2003);

- Light verb constructions (combination of a semantically light verb with a predicate belonging to diverse categories) in Persian (Bonami & Samvelian 2010; Müller 2010; Samvelian 2012; Samvelian & Faghiri 2016) and Korean (Choi & Wechsler 2002; Kim 2016; Lee 2001; Ryu 1993).

In this chapter, we examine some of these constructions which illustrate different aspects of the phenomenon of complex predicates.

2 The basic mechanism in HPSG: Argument inheritance

In HPSG, complex predicates are analyzed in the following way: one of the predicates is the head of the construction, and it inherits the (syntactic) arguments of the other predicate, that is, its complements and, possibly, its subject. We illustrate this phenomenon with tense auxiliaries in French (Abeillé & Godard 2002). The phenomenon is called *argument inheritance*, *attraction*, *composition* or *sharing*.

In French, auxiliary constructions consist of a tense auxiliary (*avoir* or *être*) followed by a past participle and its complements as illustrated in (1). The auxiliary is the head: it bears inflectional affixes (for tense and person), like any other verb; in (1), it has the form of a present indicative, 3rd person; it is in the indicative, as expected in a declarative sentence. It hosts pronominal clitics, like verbal heads in a general way (1b), (1c). Moreover, it can be gapped alone (5d), while the participle can only be gapped with the auxiliary (5e), (5f); this is expected if the auxiliary is the head, since it behaves like *pense* in (5a) and (5b), while the participle behaves like the infinitive.

(5) a. Lola pense acheter des pommes, et Alice (pense) cueillir des pêches.
   ‘Lola is thinking of buying apples, and Alice (is thinking of) picking peaches.’

   b. Lola pense acheter des pommes, et Alice (pense acheter) des pêches.
   ‘Lola is thinking of buying apples, and Alice (is thinking of picking) peaches.’

   c. * Lola pense cueillir des pommes et Alice pense des pêches.
   ‘Lola is thinking of picking apples and Alice is thinking of (picking) peaches.’

   d. Lola a acheté des pommes, et Alice (a) acheté des pêches.
   ‘Lola has bought apples, and Alice (has) bought picked peaches.’

e. Lola a a acheté des pommes, et Alice (a acheté) des pêches.
   'Lola has bought apples, and Alice (has bought) peaches.'

f. * Lola a acheté des pommes, et Alice a des pêches.
   Lola has bought apples, and Alice has peaches

The auxiliary construction in French is a CP: The clitic corresponding to a complement of the participle is hosted by the auxiliary (it is said to "climb"), as in (1); moreover, it occurs in bounded dependencies such as the infinitival complement of adjectives like *facile* 'easy', whose nominal complement is unexpressed (6a). This unexpressed complement can be that of a participle (6c), but not that of an embedded infinitive (6b).

(6) a. Cette technique est impossible à maîtriser en un jour. (French)
   'This technique is impossible to master in one day.'

b. * Cette technique est impossible à promettre de maîtriser en un jour.
   This technique is impossible to promise to master in one day

c. Cette technique est impossible à avoir maîtrisé en un jour.
   This technique is impossible to have mastered in one day

The syntactic description of the auxiliary *avoir* is given in (7):[2]

(7) Lexeme of the French tense auxiliary *avoir*

$$\begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{T-AUX} & \boxed{3} \\ \text{AUX} & + \end{bmatrix} \\ \text{ARG-ST} & \boxed{1} \bigcirc \left\langle \text{V}|\text{CAT} \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{basic-vb} & \\ \text{VFORM} & \textit{past-part} \\ \text{T-AUX} & \boxed{3} \end{bmatrix} \\ \text{ARG-ST} & \boxed{1} \\ \text{LIGHT} & + \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The auxiliary takes the participle as its complement, from which it inherits its syntactic arguments. The value for ARG-ST uses the shuffle operator ○, which allows the elements of the lists to be reordered (Reape 1994). Thus, it is the auxiliary itself, as head of the construction, which is responsible for argument inheritance:

[2] The feature [AUX ±] indicates whether the infinitival can occur before the negative adverb (*pas*, *jamais*) or not. The feature T-AUX ensures that the appropriate auxiliary (*avoir* or *être*) combines with a participle: it has the same value for the auxiliary and the participle, and is defined lexically.

the phenomenon is lexically driven. These arguments are realized as indicated in (8) (see Ginzburg & Sag 2000: for the Argument Realization Principle). The argument structure contains canonical and non-canonical elements, but the valence only contains canonical ones. The valents in turn are realized are the subject and the complements (and, possibly, the specifier for nouns).

(8) Argument Realization (Bonami & Samvelian 2015)

$$\textit{word} \Rightarrow \left[\textsc{ss}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{arg-st} \ \boxed{1} \circ \textit{list}\left(\textit{non-canon}\right) \\ \textsc{val} \ \boxed{1} \ \textit{list}\left(\textit{canon}\right) \end{bmatrix} \right]$$

The arguments are *synsems*, which can have different subtypes as in (9). Usually, they are not specified on the lexeme description, but they are on words. Thus, (8) says that if an argument is non-canonical, it stays on the argument list, but is not realized as a valent.

(9) *synsem* hierarchy:

- *synsem*
  - *non-canon*
    - *aff*
    - *gap*
    - *null-pro*
  - *canon*

In (1a), the participle *lu* selects the argument *son livre*, which is inherited by the auxiliary *a*. Accordingly, it is realized as the complement of the auxiliary *a* in (1a). The structure is as in Figure 1.

Add example with glosses and translation, add glosses in tree

Romance clitics are analyzed as affixes (Miller & Sag 1997). As arguments, they are inherited by the auxiliary. A morphological function may apply to verbs containing affixes in their argument list, including auxiliaries, which induces a certain morphology: a value of FORM (final form) different from that of I-FORM (inflected form). Hence the description of *l'a* in *l'a lu* differs from that of *a* in *a lu* (10) (Abeillé et al. 1998).

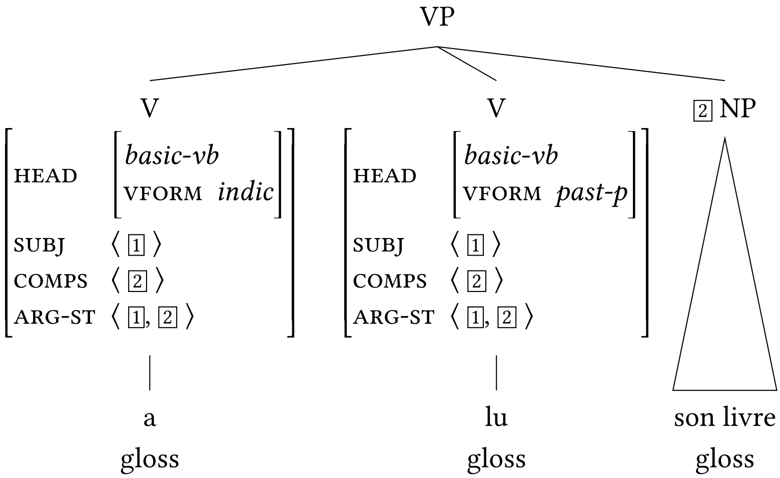


Figure 1: Add caption

(10) Description of *l'a* in *l'a lu*:

$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{FORM} & \textit{l'a} \\ \text{I-FORM} & \textit{a} \\ \text{STEM} & \textit{av-} \end{bmatrix} \\ \text{SS}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{reduced-verb} & \\ \text{VFORM} & \textit{indic} \\ \text{T-AUX} & \boxed{5} \end{bmatrix} \\ \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \oplus \boxed{4} \\ \text{ARG-ST} & \left\langle \boxed{1}, \boxed{2}\text{V}\begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{LIGHT} & + \\ \text{HEAD} & \begin{bmatrix} \textit{basic-verb} & \\ \text{FORM} & \textit{past-part} \\ \text{T-AUX} & \boxed{5} \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}, \boxed{3}[\textit{aff}] \rangle \oplus \boxed{4} \end{bmatrix} \end{bmatrix}, \boxed{3} \right\rangle \oplus \boxed{4} \end{bmatrix} \end{bmatrix}$$

In (10) , the complements of the auxiliary are the participle $\boxed{2}$, and the list $\boxed{4}$; the argument list contains in addition $\boxed{3}$, an affix, 3rd person, masculine singular. In the sequence *l'a lu*, as in (1b), [4] is the empty list; in the sequence *l'a lu à son père* 'has read it to his father', $\boxed{4}$ is the PP complement.

This description is sufficient to ensure that the clitic is always realized on the auxiliary in French, and not on the participle: *Paul l'a lu* contrasts with **Paul a le-lu*. In French, past participles never host clitics, which we assume to be a morphological property. But, in Italian, past participles may host clitics, although never when they combine with the auxiliary. To account for this, we distinguish between basic verbs, with no cliticized argument, and reduced verbs, with at least one cliticized argument which belongs to the argument list, but not to the COMPS list. The verb complement of *avoir* is specified as *basic-verb* (see (7), (10)). We distinguish between *basic-verbs* and *reduced-verbs*. Basic verbs can have arguments typed as affixes, but not cliticized, while reduced verbs have at least one cliticized argument: present in the argument list, but not in the complement list. The complement of *avoir* is specified as *basic-verb* (see (7), (10)).

3 Different structures for complex predicates: Restructuring verbs in Romance languages

In addition to tense auxiliaries, Romance languages have Restructuring verbs as heads of complex predicates. A comparison of their properties illustrates an important aspect of the phenomenon: argument inheritance is compatible with different syntactic structures, in this case, a flat structure, and a verbal complex (Abeillé & Godard 2010).

3.1 Romance Restructuring verbs

Certain verbs in Romance Languages, called *Restructuring verbs*, exhibit two behaviors: either as ordinary verbs taking a VP complement or as complex predicates (Aissen & Perlmutter 1983; Rizzi 1982). They are: modal verbs, aspectuals, or movement verbs. However, it must be kept in mind that this behavior is lexical: verbs which are close semantically may not be heads of CPs.

Several properties show that they head CPs. First, clitic climbing, which is optional (while it is obligatory with tense auxiliaries):

gloss and translate everything

(11) a. Giovanni *le* vuole mangiare.
Giovanni vuole mangiar*le*. (Italian)

b. Juan *las* quiere comer.
Juan quiere comer*las*. (Spanish)

c. O Joãn quere-*as* comer.
O Joãn quer comê-*las*. (Portuguese)

d. En Joan *les* vol menjar.
En Joan vol menjar-*les*. (Catalan)

'John want to eat them.' (Abeillé & Godard 2010)

Second, the medio-passive or middle *si* construction, where the verb hosts the reflexive clitic *si* or *se* and the subject corresponds to the object of the active construction (12a), with an interpretation close to that of passive (12b). The construction is possible with restructuring verbs (12c), but not with ordinary verbs taking an infinitival complement (12d).

(12) a. Giovanni stira queste camicie facilmente. (Italian)
'Giovanni irons these shirts easily.'

b. Queste camicie si stirano facilmente.
'These shirts iron easily.'

c. Queste camicie si possono stirare facilmente.
'These shirts can be ironed easily.'

d. * Queste camicie si paiono stirare facilmente.
These shirts appear to be ironed easily.'

Third, their acceptability in bounded dependencies, as illustrated in (6) for tense auxiliaries, and (13) for Restructuring verbs. (13b) relies on *cominciare* being a Restructuring verb, while *promettere* is not (13c).

(13) a. Questa canzone è facile da apprendere.
'This song is easy to learn.'

b. Questa canzone è facile da cominciare a apprendere.
'This song is easy to begin to learn.'

c. * Questa canzone è facile da promettere a apprendere
'This song is easy to promise to learn.'

Finally, the possibility of preposing the verbal complement of a Restructuring verb disappears when there is evidence of a complex predicate. For the sake of simplification, we now concentrate on Italian and Spanish. The data in (14), with a VP, contrast with those in (15), where the head verb inherits a cliticized complement of the infinitival. Preposing of the verbal complement is associated with pronominalization (*lo*) in Italian (14a) not in Spanish (14b), and is more natural in contrastive contexts in Spanish. The infinitive cannot be preposed if its argument is cliticized onto the head verb (15).

(14) Does he want to talk to Mary?

a. – Parlare a Maria, certamente lo vuole.

b. – Hablarle a María, seguramente quiere (pero no a su madre).
‘Talk to Maria, certainly he wants to (but not to her mother).’

(15) a. *Parlare, certamente glielo vuole.
b. *Hablar, le quiere (pero no mucho tempo).

We assume that Restructuring verbs have two possible descriptions: as ordinary verbs taking an infinitival VP complement, or as heads of complex predicates. They are related by the Argument Composition Lexical Rule (16) (Monachesi 1998).

(16) Argument Composition Lexical Rule for Romance restructuring verbs:

$$\begin{bmatrix} \textsc{head}\ \textit{verb} \\ \textsc{Light}\ + \\ \textsc{subj}\ \left\langle \textsc{np}\begin{bmatrix}\textsc{ind} & i\end{bmatrix} \right\rangle \\ \textsc{comps} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textit{verb} \\ \textsc{vform} & \textit{inf} \end{bmatrix} \\ \textsc{subj} \left\langle \textsc{np}\begin{bmatrix}\textsc{ind} & i\end{bmatrix} \right\rangle \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \end{bmatrix} \mapsto \begin{bmatrix} \textsc{comps} \left\langle \begin{bmatrix} \textit{basic-verb} \\ \textsc{comps}\ \boxed{2} \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

The left hand side represents a verb taking a VP complement. The identity of indices for the expected subjects of the verb and its infinitival complement indicates that they are control or raising verbs; the verbal complement is saturated for its complements. The right hand side represents the head of a complex predicate: it inherits the expected complement(s) of the infinitival complement (list $\boxed{2}$). The categorization of the infinitive as *basic verb* ensures that it does not host cliticized complements, and the feature [LIGHT +], shared with the input, that it has not combined with complements (see below).

The complement of adjectives such as ‘easy’ in RL is a bounded dependency: they take an unsaturated infinitival complement which expects a complement coindexed with its subject. Complex predicates can occur in this construction because they inherit the complement of their complement. In (13a), *apprendere* is expecting an object, co-indexed with *questa conzone*. In (13b), *cominciare* inherits the expected complement of *apprendere*, a configuration which is similar to that of (13a).

The medio-passive verb is also the result of a Lexical Rule, which takes a transitive verb like *stirare* (12a) to give a verb whose subject corresponds to the expected object of the input verb, and acquires a clitic SE (realized *si* or *se* in the different languages) as in (12b) (Monachesi 1998). While a verb taking a VP complement like Italian *potere* cannot be the input to this rule, since it lacks an NP object, the corresponding verb which is the output of LR (16) can if it inherits a nominal complement from the infinitive, as in (12c). The verb *potere* as in *Giovanni può stirare queste camicie*, which inherits *queste camicie* from *stirare*, is the input to Rule (16), giving the verb which occurs in (12c).

3.2 The structure of CP constructions: Verbal complex and flat structure

Argument composition is compatible with different structures. In Romance Languages, complex predicates enter either a flat structure, or a verbal complex. We contrast Italian and Spanish.[3]

The impossibility of Preposing illustrated in (15) shows that the sequence of the complement verb and its complements does not form a constituent (a VP) when there is a complex Predicate, a point made by Rizzi (1982) for Italian, on the basis of a series of constructions (pied-piping, clefting, Right Node Raising, complex NP shift). However, regarding other properties, the two languages do not behave in the same way. Note that in Spanish, there is variation among speakers: we illustrate here Spanish-1 variety. The presence of a clitic on the head verb indicates that there is a complex predicate.

First, adverbs occur between the Restructuring verb and the infinitive in Italian (17a), but not in Spanish-1 in a general way (17b) (a few adverbs, such as *casi* 'nearly', *ya* 'already', *apenas* 'barely' are possible).

(17) a. Giovanni *lo* vuole spesso leggere. (Italian)
 b. * Juan *lo* quiere a menudo leer. (Spanish-1)
 c. Juan quiere a menudo leer*lo*.
 'John wants to often read it.'

Second, an inverted subject NP can occur between the two verbs of a complex predicate in Italian, but not in Spanish-1. The subject can occur postverbally in interrogative sentences. In Italian, it can occur between the two verbs with a special prosody, and with speaker's variation (Salvi 2010) (18a). In Spanish-1, it is not possible (except for the pronominal subject) (Suñer 2010) (18b), while it can occur between the verb and an infinitival VP complement (18c).

[3] In Portuguese complex predicate constructions are also a flat structure, but with different ordering constraints from Italian; in Spanish-2, they are similar to Portuguese. Complex predicates with verbs which have only one structure also distribute between these two structures: tense auxiliaries in French, Italian, Portuguese as well as Romanian modal a *putea* 'can' are the head of a flat structure, while Spanish-1 and Romanian tense auxiliaries enter a verbal complex (Abeillé & Godard 2010).

(18) a. % Lo comincia MARIA a capire, il problema, oppure no?
'Maria, she's beginning to understand it, the problem, yes or no.'

b. *? Lo comienza Juan a comprender?

c. Comienza Juan a comprenderlo?
'Is John beginning to understand it?'

Finally, Italian heads of complex predicates can have scope over the coordination of infinitives with their complements (19a), while this is not the case in Spanish-1 (19b) (from Abeillé & Godard 2010). Again, the presence of the clitic on the head verb (*lo vuole*, *le volvió*) shows that this is a complex predicate construction.

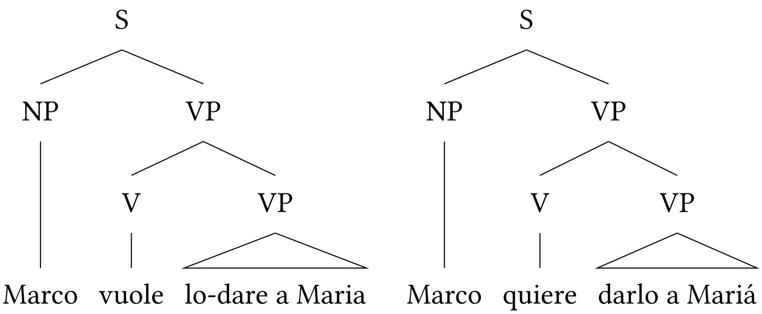
(19) a. Giovanni lo vuole comprare subito e dare a Maria
Giovanni it wants buy immediately and give to Maria
'Giovanni wants to buy it immediately and give it to Maria.'

b. * Le volvió a pedir un autografo y a hacer
to-him started again to ask an autograph and to make
proposiciones
propositions
'He started again to ask him for an autograph and to make propositions to him'

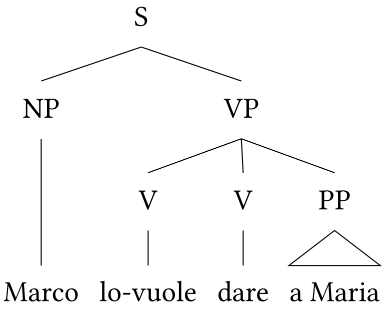
Constituency tests such as preposing (15) show that the verbal complement is not a VP. The verbal complex, in which the two verbs form a constituent without the complements is well-suited to account for the absence of adverbs (in a general way) and of subject NPs, if such combinations exclude non-verbal elements and adverbs (in general). This constraint can be captured by the feature [LIGHT +],[4] which has been used in Romance languages for other phenomena as well (Abeillé & Godard 2005; 2010). Hence, complex predicate constructions in Spanish-1 contain a verbal complex, while they represent a flat structure in Italian.

[4] The feature [LIGHT ±] is used rather than the feature WEIGHT as in Abeillé and Godard; it substitutes for [LEX ±] used in HPSG formalizations of German (Bonami & Webelhuth 2013). The adverbs admissible in the Spanish-1 verbal complex are light.

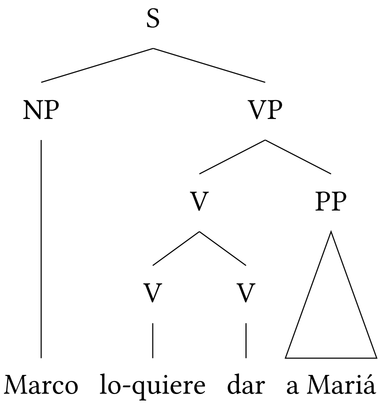
(20) a. VP complement:



b. Flat structure:



c. Verbal complex:



The possibility of the coordination in (19a) has been viewed as an argument in favor of a VP complement. However, if such sequences can be analyzed as coordinations of VP, they can also be Non Constituent Coordinations (NCC) (*John gives a book to Maria and discs to her brother*). So, the question is: why is (19b) not an acceptable NCC in Spanish? We propose that coordinations are subject to a general constraint in Romance languages: the parallel elements of the coordination must be at the same syntactic level, otherwise the acceptability is degraded. An example is the contrast between (21a) and (21b) in Spanish. In (21b), the constituant [*de Camus*] in the second conjunct is at the same level as [*el libro de Proust*], but not at the same level as [*de Proust*], with which it is parallel.

Add gloss and translation

(21) a. Juan da [el libro de Proust] [a Mariá] y [el (libro) de Camus] [a Pablo]. (Spanish)

b. ?? Juan da [el libro de Proust] [a Mariá] y [de Camus] [a Pablo].

3.3 Analysis of Romance CP constructions in HPSG

The complex predicate in Italian is a flat structure (20b), while it is a verbal complex costruction in Spanish (20c). Abeillé & Godard (2005) do not propose a special phrase for dealing with verbal complexes, but treat them as a subtype of head-complements-phrases. The difference between flat structure and verbal complex is attributed to the feature [LIGHT ±] and linearization constraints. The head-complements-phrase is as follows:

(22) a. *Head-Complements-phrase* ⇒

$$\left[\begin{array}{ll} \text{MOTHER}|\text{SS}|\text{CAT} & \left[\begin{array}{l}\text{HEAD } \boxed{1} \\ \text{COMPS } /\langle\rangle \end{array}\right] \\ \text{HEAD-DTR}|\text{SS}|\text{CAT} & \left[\begin{array}{l} \text{HEAD } \boxed{1}\\ \text{LIGHT } + \\ \text{COMPS } \boxed{2} \circ \boxed{3}\end{array}\right] \\ \text{NON-HEAD-DTRS } \boxed{2} \textit{ non-empty list} & \end{array}\right]$$

b. $\left[\textit{phrase}\left[\text{LIGHT } +\right]\right] \Rightarrow \left[\text{DTRS}\left[\text{LIGHT } +\right]\right]$

The head-complements-phrase is usually saturated for the expected complements, but this constraint can be violated, notably in the case of a verbal complex. In this case , the infinitive combines with a head verb without being saturated, and the head verb is not saturated for the complements it inherits from the infinitive. The requirement is passed up because of the Generalized Head Feature Principle (Ginzburg & Sag 2000), which says that the *synsem* of the phrase is identical to that of the head daughter by default (each phrase description says in what way they differ).

The LIGHT feature is appropriate both for words and phrases; it has ordering as well as structural consequences. Words can be light or non-light; lexical verbs (finite verbs, participles or infinitives without complements) are light. Most phrases are non-light, but phrases made up of light elements can be light or non light; if the phrase is light, the daughters are light (22b). Accordingly, a verbal complex, made up of two lexical verbs can be light: this allows it to be the head of the VP (see (22a)).

In order to enforce the verbal complex in Spanish-1, we assume an additional constraint on the non-light head-complements-phrase in this language, which does not exist in Italian (or French).

(23) *Head-complements-phrase* [LIGHT −] ⇒ [NON-HEAD-DTRS list([LIGHT −]]) (Spanish)

Constraint (23) says that all the complements in a non-light phrase must be non-light. This precludes a light complement such as a (bare) infinitive from occurring at the same level as an NP or a PP, for instance, which are non-light, contrary to what happens in the flat structure in (20b).

In order to prevent a verbal complex in Italian (and French), we assume a different additional constraint (which differs from that in Abeillé & Godard (2005)).

(24) *Head-complements-phrase* [LIGHT +] ⇒ [[NON-HEAD-DTRS list([HEAD *non-verbal*]] (Italian)

The Verbal Complex has to be light, since it is the head of the VP (see (22b)). Since it is made of verbal constituents, it is excluded by constraint (24) in Italian.[5]

Romance languages follow the general constraints on ordering in head-complements-phrases in non-final languages. They are given here informally. Constraint (26) implies that the infinitival complement in complex predicates (or the participle for the auxiliary constructions), precedes the complements it subcategorizes for.[6]

(25) The head DTR precedes the complements (Romance languages)

(26) [V [COMPS < $\boxed{1}$[ADV −]>]] precedes $\boxed{1}$. (Romance languages)

[5] Note that head-only phrases are non-light.

[6] The feature [ADV −] excludes adverbial complements from the ordering rule: they may precede the past participle as in *Paul s'est bien comporté* (Paul SE-is well behaved, 'Paul has behaved well').

4 Complex predicates and word order

In certain languages, a complex predicate construction signals itself by properties of word order. This is the case for instance in German (De Kuthy & Meurers 2001; Hinrichs & Nakazawa 1994; 1998; Kathol 1998; Kiss 1994; Meurers 2002; Müller 2002) and Dutch (Bouma & van Noord 1998; Rentier 1994), as well as Korean (Chung 1998; Kim 2016; Sells 1991; Yoo 2003).

4.1 Coherent and incoherent constructions in German

Verbs with an infinitival complement in German enter into so-called incoherent or coherent constructions: in the first case, they combine with a VP (saturated for its complements), but not in the second case (Bech 1955). We speak of constructions rather than verbs, because, although the constructions are triggered by lexical properties of verbs, many verbs can be constructed either way. Verbs entering coherent constructions, obligatorily or optionally, belong to different classes: tense auxiliaries, modals, subject and object raising, object control verbs, the copulas and predicative constructions (see Müller 2002).

Incoherent constructions are constructions where a verb takes a (saturated) VP complement. They are illustrated by the combination of *überreden* with the infinitival complement in (27). They allow for VP fronting as in (27a), VP extraposition as in (27b), and VP relatives such as (27c), where the infinitival complement is"pied-piped" with the relative pronoun (examples from Hinrichs & Nakazawa 1998).[7]

(27) a. [Das auto zu kaufen] wird Peter Maria *überreden.*
to buy the car will Peter Maria persuade
'Peter will persuade Maria to buy the car.'

b. … dass Peter Maria *überredet*, [das Auto zu kaufen].
that Peter persuade Marie, the car to buy
'… that Peter persuade Maria to buy the car.'

c. Das ist das Auto, das zu kaufen er Peter *überreden* wird.
that is the car, which to buy he Peter persuade will
'That is the car, which he will persuade Peter to buy.'

On the other hand, coherent constructions, of which the combination of the auxiliary *wird* or the raising verb *scheinen* with an infinitival complement are

[7] The verb taking the VP complement is underlined.

typical examples, do not allow for VP extraposition (28b), (28d) (from Müller 2002), or VP relatives (28e), (28f).[8]

(28) a. … dass Karl das Buch lesen *wird*
that Karl the book read will
'… that Karl will read the book'

b. * … dass Karl *wird* das Buch lesen

c. … weil Karl das Buch zu lesen *scheint*
because Karl the book to read seems
'… because Karl seems to read the book'

d. * … weil Karl *scheint* das Buch zu lesen

e. * Das ist das Buch das lesen Karl *wird*.

f. * Das ist das Buch das zu lesen Karl *scheint*.
'This is the book that Karl will read.'

Similarly, scrambling of the complements is possible in a coherent construction, not in an incoherent one. In (29a) and (29c), the complements of *sehen* or *überreden*, respectively, do not mix. In (29b), *Peter*, the complement of *sehen*, occurs between *das Auto*, which is the complement of *kaufen*, and *kaufen*. Such a word order is possible with *sehen*, not *überreden* (29d) (examples from Hinrichs & Nakazawa 1998).

(29) a. … dass er [Peter] [das Auto kaufen] sehen wird
that he Peter the car buy see will
'that he will see Peter buy the car'

b. … dass er das Auto Peter kaufen sehen wird

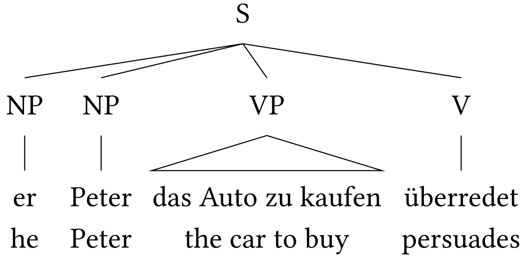
c. … dass er [Peter] [das Auto zu kaufen] überreden wird
that he Peter the car to buy persuade will
'that he will persuade Peter to buy the car'

d. * … dass er das Auto Peter zu kaufen überreden wird

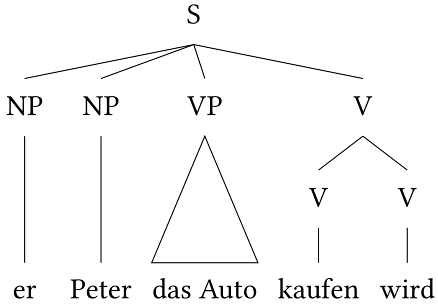
These data point to the following analysis: incoherent constructions involve a saturated VP complement (30a), while coherent constructions do not: they form a complex predicate, with a verb inheriting the complements of its complement. We assume here a verbal complex, a structure largely accepted for German (30b).

[8] The head verb of the coherent construction is underlined. We leave aside fronting of a verbal complement in a coherent construction, because the data are less straightforward (De Kuthy & Meurers 2001; Müller 2002).

(30) a. Incoherent construction (embedded clause)



b. Coherent construction (embedded clause)



4.2 Coherent constructions in HPSG

One might wonder whether it is possible to analyze the data in terms of word order instead of structure: a verb governing a coherent construction would trigger a modification of the ordering domain. More precisely, it would induce domain union of the two ordering domains associated with the two verbal projections (Reape 1994), thus allowing the order in (29b), for instance, while the structure would remain the same (as in (30a)). The existence of the remote (or long) passive goes against such an analysis (Kathol 1998; Müller 2002). A complex predicate construction can be passivized in such a way that the subject (in the nominative case) of the passive auxiliary corresponds to the object of the active infinitive complement. An (impersonal) passive construction like (31a) with an infinitival VP containing an accusative object (*den Wagen*) alternates with a coherent construction such as (31b), with a corresponding nominative (*der Wagen*).

(31) a. … weil oft versucht wurde, [den Wagen zu reparieren]
because often tried was the car to repair
'… because many attempts were made to repair the car'

b. … weil der Wagen oft zu reparieren versucht wurde

In (31b), there is no infinitival VP, as shown by the occurrence of the adverb *oft*, which modifies *versucht*, between *der Wagen* and *zu reparieren*.

The German verbal complex, like the Spanish-1 one, instantiates a light head-complements-phrase. Thus, the two elements are light (22b). The constraint in (32b) renders explicit that in a verbal complex, the head inherits the complements of its complement. This complement can itself be a verbal complex (as in (29b)).[9]

Except for the verbal complex, we adopt here a flat structure for German (Uszkoreit 1987; Pollard 1996).[10] German differs from RL in not distinguishing structurally between the subject and the complements: they occur at the same level, can be interspersed, and are introduced by the same phrasal constraint. For this reason, it is useful to use the feature VAL (for valence, that is, the subject and the complements) in the description of lexemes and words, distinguishing the subject via the feature XARG on the lexemes, and the complements by COMPS only when it is necessary (Sag 2012). Thus, coherent constructions obey two phrasal constraints: the Head-Valents-phrase and the verbal-complex-phrase.

(32) a. *Head-Valents-phrase* ⇒

$$\begin{bmatrix} \textsc{mother} \mid \textsc{ss} \mid \textsc{cat} \begin{bmatrix} \textsc{val} \ /\langle\rangle \\ \textsc{light} \ - \end{bmatrix} \\ \textsc{head-dtr}|\textsc{ss}|\textsc{cat} \begin{bmatrix} \textsc{light} \ + \\ \textsc{val} \ \textit{nelist} \ \boxed{1} \end{bmatrix} \\ \textsc{non-hd-dtrs} \ \boxed{1} \end{bmatrix}$$

b. *Verbal-Complex-phrase* (German) ⇒

$$\begin{bmatrix} \textsc{mother} \mid \textsc{ss} \mid \textsc{cat} \begin{bmatrix} \textsc{val} \ \boxed{1} \oplus \boxed{2} \\ \textsc{light} \ + \end{bmatrix} \\ \textsc{hd-dtr}|\textsc{ss}|\textsc{cat}|\textsc{val} \ \boxed{1} \oplus \langle \boxed{3} \rangle \oplus \boxed{2} \\ \textsc{non-hd-dtrs} \ \boxed{3} \begin{bmatrix} \textit{verb} \\ \textsc{light} \ + \\ \textsc{comps} \ \boxed{2} \end{bmatrix} \end{bmatrix}$$

The two phrases differ by the value of LIGHT, and also in that (32a) allows for a number of daughters while (32b) allows just for one non-head daughter, a verbal complex as illustrated in (29a) corresponding to a multi-level structure. The

[9] The representations of a verbal complex across RL, German and Korean, are homogenized. There is no need for XCOMP (Müller 2002) or GOV (Chung 1998) to distinguish this complement from the others in addition to specifying inheritance as in (32b) (Bonami & Webelhuth 2013).

[10] Alternatively, the structure can be represented as a hierarchy of binary branching structures, with a flat linearization domain (Müller 2002).

syntactic descriptions of the passive auxiliary *werden* and of *sehen* (in a coherent construction) are as in (33). *Sehen* is associated with two different descriptions: it has a incoherent version (no inheritance of valents) and a coherent one, which can be related via a lexical rule in the manner of Romance restructuring verbs as in (16).

(33) a. *werden* (passive auxiliary):

$$\begin{bmatrix}\textsc{xarg}\ \langle \boxed{1} \rangle \\ \textsc{val}\ \left\langle \boxed{1}\ ,\ \mathrm{V}\left[\textit{passive-part},\ \textsc{light}\ +\ ,\ \textsc{val}\ \langle \boxed{1} \rangle \oplus \boxed{2}\right] \oplus \boxed{2} \right\rangle\end{bmatrix}$$

b. *sehen* (coherent version):

$$\begin{bmatrix}\textsc{x-arg}\ \langle \mathrm{NP}_i \rangle \\ \textsc{val}\ \left\langle \mathrm{NP}_i,\ \mathrm{NP}_k,\ \mathrm{V}\left[\textit{inf},\ \textsc{val}\ \langle \mathrm{NP}_k \rangle \oplus \boxed{2}\right] \right\rangle \oplus \boxed{2}\end{bmatrix}$$

We leave aside the placement of the head verb in the German clause, which is a question in itself. In a verbal complex, the head is ordered after the complement. The structure of (29b) is illustrated in Figure 2.

Please use feature geometry of Sag (1997). People working on German usually assume binary branching structures. See Chapter on order.

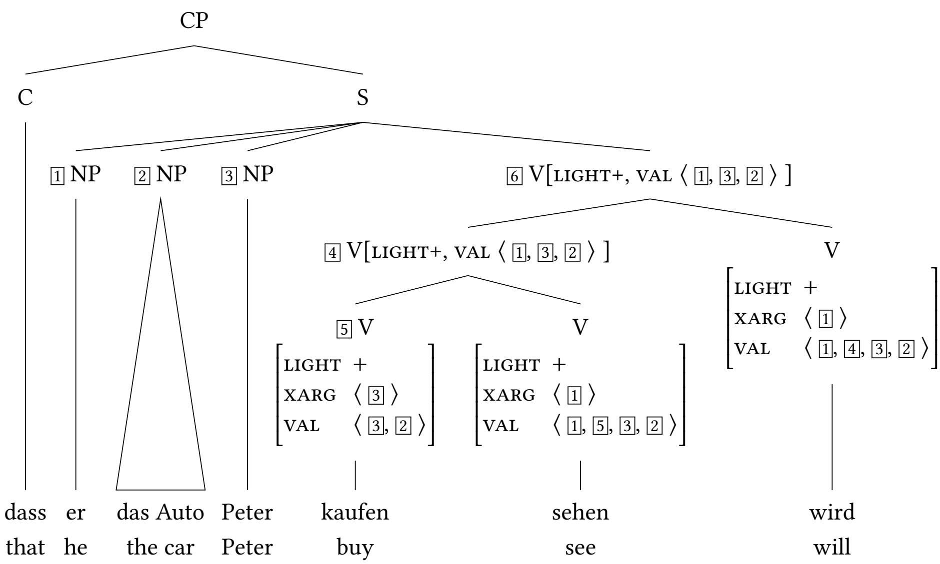


Figure 2: Analysis of *dass er das Auto Peter kaufen sehen wird*

In Figure 2, the verb $\boxed{6}$ is a verbal complex, whose head is *wird*, which inherits the complements of its complement $\boxed{4}$, also a verbal complex; the head of $\boxed{4}$ is *sehen*, which identifies its subject (or xarg) with that of *wird* (see (33a)), and inherits the complement of its complement *kaufen*, that is $\boxed{2}$, in addition to its other complement $\boxed{3}$ (see (33b)); the verbal complex $\boxed{4}$ inherits its valence from its head, minus the complement which is saturated ($\boxed{5}$).

4.3 Argument inheritance in Korean

4.3.1 Scrambling in Korean

Korean resembles German in that a complex predicate signals itself mainly by word order properties. Complex predicates are formed from auxiliaries, control verbs, and causative *ha*. We follow in particular Chung (1998), from which the examples are borrowed.

Korean auxiliaries semantically resemble aspectual or modal verbs rather than tense auxiliaries. They include such verbs as *iss* 'to be in the process/state of', *chiwu* 'to do resolutely', *siph* 'want', but also the verb of negation *anh* 'not'. Scrambling with auxiliaries is illustrated in (34). In (34b) the subject of the auxiliary *issta* (*Mary-ka*) occurs between the embedded verb (*ilke*) and its complement (*chayk-ul*).

(34) a. Mary-ka ku chayk-ul ilke poko issta.
Mary-NOM the book-ACC read try be.in.the.process.of-DECL
'Mary is giving the book a trial reading.'

b. Ku chayk-ul Mary-ka ilke poko issta.

Control verbs such as *seltuk* 'persuade' (35a), (35b) , *cisi* 'order' (object control), *yaksok* 'promise', *sito* 'try' (subject control) (35c), (35d), as well as causative *ha*, also allow valents of the embedded verb to be interleaved with their own valents, as in (35b) and (35d).

(35) a. Mary-ka John-hantley [ku chayk-ul ilkulako] seltukhayssta.
Mary-NOM John-DAT the book-ACC read persuade-past-DECL
'Mary persuaded John to read the book.'

b. Ku chayk-ul Mary-ka John-hantley ilkulako seltukhayssta.
the book-ACC Mary-NOM John-DAT read-will persuade-PAST-DECL
'Mary persuaded John to read the book.'

c. Mary-ka John-hantley [ku chayk-ul pilyecwu-kesstako]
Marie-NOM John-DAT the book-ACC lend-will
yaksokhayssta.
promise-PAST-DECL
'Mary promised John to lend the book.'

d. Ku chayk-ul John-hantley Mary-ka pilyecwu-kesstako
the book-ACC John-DAT Marie-NOM lend-will
yaksokhayssta.
promise-PAST-DECL
'Mary promised John to lend the book.'

These data can be explained in two ways: either the auxiliary or the control verb always takes a VP complement, and scrambling is due to linearization (the domains of the two verbs being unioned), or there is a complex predicate: the complement of the embedded verb (the book) is inherited by the auxiliary or the control verb, and, assuming a flat structure for the Korean sentence (with a head-valents-phrase as in (32a)), it occurs at the same level as the subject. As in German, long passivization is possible, which cannot be accounted for by appeal to linearization and domain union.

Certain auxiliary and control verbs can be passivized, so that the expected complement of the embedded verb becomes the subject of the construction: *malssengmanhun so-ka* is subject of the passive auxiliary *ci-esta* in (36b) and *ku cengchayk-i* is subject of the passive of *order* in (37b).

(36) a. Ku mongpwu-ka malssengmanhun so-lul phala
The farmer-NOM troublesome cow-ACC sell
chiw-essta.
do.resultely-PAST-DECL
'The farmer resolutely sold the troublesome cow.'

b. Malssengmanhun so-ka (ku nongpwu-eyuyhay) phala
troublesome cow-NOM the farmer-by sell
chiwe ci-essta.
do.resolutely passive-PAST-DECL
'The troublesome cow was resolutely sold (by the farmer).'

(37) a. Nay-ka Mary-hantley ku cengchayk-ul sihaynghalako
I-NOM Mary-DAT the policy-ACC carry.out
cisihayssta.
order-PAST-DECL
'I ordered Mary to carry out the policy.'

b. Ku cengchayk-i naey-uyhayse Mary-hantley sihaynghalako
the policy-NOM I-by Mary-DAT carry.out

cisi-toy-essta.
order-PASS-PAST-DECL
(lit. the policy was ordered by me for Mary to carry out)

Since passivization only affects the complement of the verb which is itself passivized, it follows that *malssengmanhun so-lul* is the complement of the auxiliary in (36a), and *ku cengchayk-ul* the complement of the control verb in (37a).[11]

4.3.2 The structures of CP in Korean

In Korean, the structure of a complex predicate formed on an auxiliary contrasts with that formed on a control verb. Auxiliaries are the head of a verbal complex (as in German complex predicates), while control verbs can either take a saturated VP complement, or be the head of a flat structure (like Restructuring verbs in Italian).

A characteristic property of Korean auxiliary constructions is that nothing can separate the two verbs. While an NP may occur after the verbs in a construction called afterthought (38a), this is not possible for the embedded verb alone (38b) or with its complement (38c), and no parenthetical expression can intervene as shown by *hayekan* in (39b) (examples from Chung 1998). This behavior follows if the two verbs form a verbal complex (see section 3.2).

(38) a. Mary-ka mekko issta, sakwa-lul.
Mary-NOM eat be.in.the.process.of-DECL apple-ACC
'Mary is in the process of eating an apple.'

b. * Mary-ka sakwa-lul issta, mekko

c. * Mary-ka issta, sakwa-lul mekko

(39) a. Mary-ka hayekan sakwa-lul mekko issta.
Mary-NOM anyway apple-ACC eat be.in.the.process.of-DECL
'Anyway, Mary is eating an apple.'

b. * Mary-ka sakwa-lul mekko hayekan issta

Korean has a Verbal-Complex-phrase similar to (32b). In Korean, this structure characterizes the class of auxiliaries. This is captured by the head feature AUX: auxiliaries are [AUX +], while the other verbs are [AUX −]. Moreover, unlike the

[11]More complicated data, showing that the case of the inherited complement depends on the agentivity of the combination of auxiliary *siph* with its verbal complement, can be explained if the auxiliary and the verbal complement form a verbal complex (Yoo 2003).

head of a verbal complex in German, which can be a control verb (like *sehen*), they inherit the whole argument structure of their complement. Note the use of 'o', the shuffle operator, which allows to reorder the elements of the two lists.

(40) *Verbal-Complex-phrase* (Korean) ⇒

$$\begin{bmatrix} \text{MOTHER}|\text{SS}|\text{CAT} & \begin{bmatrix}\text{AUX} & + \\ \text{LIGHT} & + \\ \text{VAL} & \boxed{1}\end{bmatrix} \\ \text{HEAD-DTR}|\text{SS}|\text{CAT} & \begin{bmatrix}\text{AUX} & + \\ \text{LIGHT} & + \\ \text{VAL} & \boxed{1} \circ \left\langle \boxed{2} \right\rangle\end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \boxed{2} \begin{bmatrix} \textit{verb} \\ \text{LIGHT} \ + \\ \text{VAL} \ \boxed{1}\end{bmatrix}\right\rangle \end{bmatrix}$$

An auxiliary such as *iss* 'to be in the process of' has the following syntactic description:

(41)
$$\begin{bmatrix}\text{HEAD} \left[\text{AUX} \ +\right] \\ \text{LIGHT} \ + \\ \text{VAL} \left\langle \text{V}\left[\text{VFORM } \textit{ko}, \text{ LIGHT} \ +, \text{ VAL } \boxed{1}\right]\right\rangle \bigcirc \ \boxed{1}\end{bmatrix}$$

The auxiliary *iss* takes as its complement a light verb which is constrained to end in −*ko* (different auxiliaries put different syntactic and semantic restrictions on their verbal complement, (Yoo 2003)), from which it inherits its valence. The sentence in Korean is generally assumed to have a flat structure, which corresponds to the Head-Valents-phrase in (32a). The structure of (34a), with a series of two auxiliaries, is similar to (2). *Issta* takes as its complement *ilke poko*, whose head is *poko*, and which is also a verbal complex. *Poko*, being an auxiliary like *iss* (41), takes as its complement the verb *ilke*, from which it inherits its valence, which is transmitted to the verbal complex *ilke poko*; *ilke poko* saturates the verbal complement expected by *issta*, and transmits its valence to the head auxiliary. The structure is schematized in Figure 3.

Contrary to auxiliaries, control verbs such as *seltuk* (35) or *cisi* (37) can be separated from their verbal complement, for instance by the adverb in (42a). They also allow for the infinitive and its complement to form a VP constituent as in (42b), where they constitute an afterthought. Thus, control verbs are analyzed in the same way as Italian Restructuring verbs: they can take a saturated VP complement, or be the head of a complex predicate. They contrast with Korean Raising verbs which only take a VP complement.

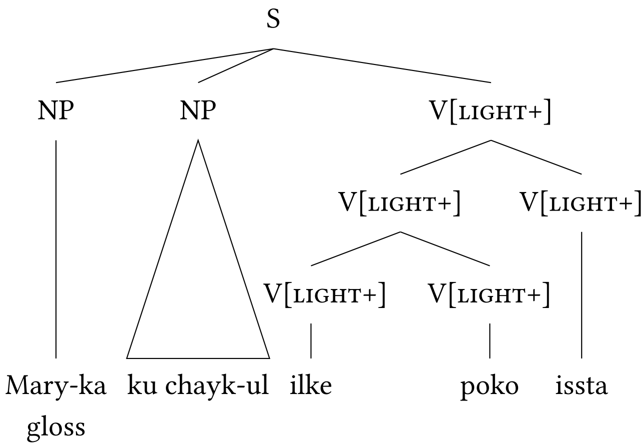


Figure 3: Add caption

(42) a. Mary-ka ku mwuncey-lul phwulye-ko (kkuncilkikey)
Mary-NOM the problem-ACC solve ceaselessly
sitohayssta
try-PAST-DECL
'Mary tried (ceaselessly) to solve the problem.'

b. Mary-ka sitohayssta, [ku mwuncey-lul phwulye-ko].

c. Ku chayk-ul Mary-ka ilkulako sitohayssta.
The book-ACC Mary-NOM read try-PAST-DECL
'Mary tried to read the book.'

d. Ku chayk-ul Mary-ka ilkulako John-hantley seltukhayssta.
The book-ACC Mary-NOM read John-DAT persuade-PAST-DECL
'Mary persuaded John to read the book.'

The scrambling data, together with the possibility of long passivization, show that there is a complex predicate (section 4.3.1). For instance, the subject of the head verb (*Mary-ka*) comes between the complement of the infinitive and the infinitive in (42c), (42d). More precisely, there is no verbal complex in this case: the two verbs do not have to be contiguous, but can be separated, for instance by a complement as in (42d) (*John-hantley*). Thus, they are the head of a flat structure as in Figure 4 corresponding to (42d).

As stated in the description of verbal complexes in Korean (40), the head of the structure is [AUX +]. Since control verbs are [AUX −], they cannot enter

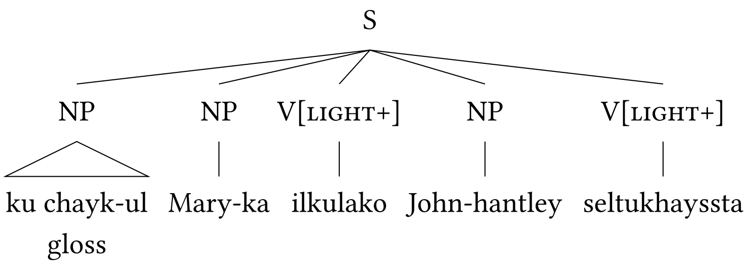


Figure 4: Add caption

verbal complexes, and the structure for complex predicates is a flat Head-Valents-phrase, as described in (32a).

The head comes last in Korean, except in the afterthought construction (42b), which requires an additional mechanism. This is true in the verbal complex, in the embedded VP and in the flat structure, as well. The alternative ordering for (42a) **Mary-ka [phwulye-ko] [ku mwuncey-lul] sitohayssta*, where the verb *phwulye-ko* precedes its complement *ku mwuncey-lul* is thus impossible. In (42d), the NP *John-hantley* can follow the complement verb *ilkulako* because it is the complement of the head verb *seltukhayssta*, not of *ilkulako*. We can state the constraint informally as in (43).

(43) X precedes the HEAD

Chung (1998) extends the possibility of inheritance to adjuncts, as well as to constructions with a S complement, which are somewhat marginal. The behavior of adjuncts is easily accounted for, if adjuncts can be treated as complements (Bouma et al. 2001) by a verb. For the second case, Chung proposes a flat structure, in which the valents and (lexical) verbs are all sisters. In this analysis, the definition of a complex predicate, which relies on a syntactic relation between words can still be maintained (but see Lee 2001: for a different proposal based on linearization).

5 Light verb constructions in Persian: Syntax and morphology, syntax and semantics

Light verb constructions, where the head is a verb and the second predicate belongs to different categories, depending on the language and the construction, allow us to focus on other aspects of complex predicates. Such constructions are frequent in different languages (see Choi & Wechsler 2002; Lee 2001: for noun-verb combinations in Korean). We focus on Persian (Bonami & Samvelian 2010; Müller 2010; Samvelian 2012), examining two aspects: morphosyntactic and semantic.

Persian complex predicates are multiword expressions. More precisely, they are lexemes which are realized by several words (Bonami & Samvelian 2010). To solve the morphosyntactic and semantic problems, we rely crucially on the same property of HPSG as in the preceding syntactic cases, that is, the view of heads as sharing information with their expected complements.

5.1 Complex predicates and derivational processes

Although complex predicates have properties expected of the combination of words (they can be separated by the future auxiliary, for instance), they give rise to agent nominalization. Examples (44) show that although no agent noun can be derived from the verb *kon* 'do', an agent noun can be derived from the complex predicate formed with the verb *kon* and the adjective *bāz* 'open'.

(44) a. kon 'do', *kon-ande 'do-er'

b. bāz kon-ande 'opener'

We follow Müller (2010), who exploits the possibility for a derived word to keep track of its formation via a daughter. Thus, the verb part of *bāz kon-ande* is the target of a nominalization rule, and is combined with an affix (-*ande*), but this is possible only if it is waiting for the adjective complement. The agent nominalization rule takes the verb as input and gives a noun which refers to the expected subject of the verb, and, syntactically takes as its complement (inherited from the verb) the adjectival predicate. This is represented in (5) (adapted from Müller 2010).

The content of the verb *kon* in the complex predicate *bāz kon* is a cause relation, with the state expressed by the adjective (noted as $\boxed{5}$) as a result. The derived N is thus the opener.

5.2 The Semantics of Light verb constructions

Although Persian complex predicates can be viewed as lexemes, the nominal element and the verb in noun-verb combinations behave nevertheless as two independent syntactic units ((Karimi-Doostan 1997; Megerdoomian 2002; Samvelian 2012). All inflection is prefixed or suffixed on the verb, as is the negation in (45),

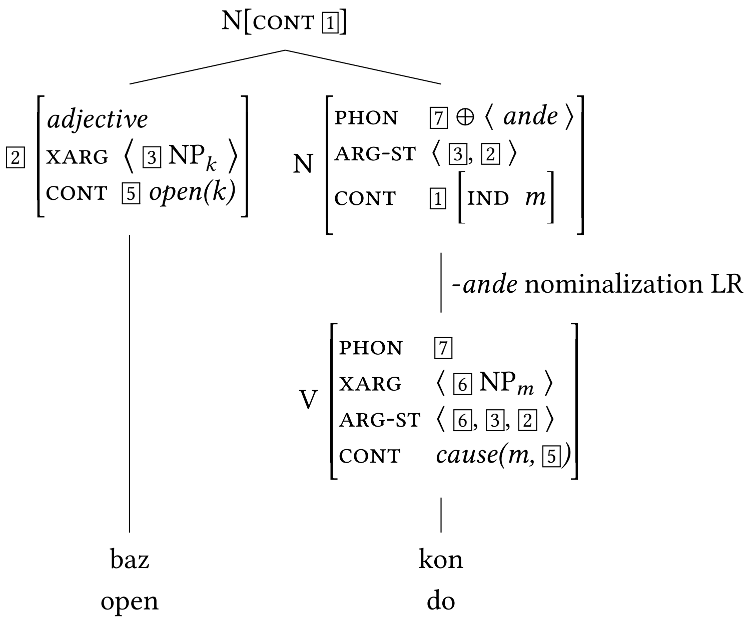


Figure 5: Insert caption

and never on the noun. The two elements can be separated by pronominal clitics, the future auxiliary, or even by clearly syntactic constituents, like the complement in (45). Both the noun and the verb can be coordinated, as shown in (46) and (47) respectively. The noun can be extracted, as in the topicalization in (48). CPs can be passivized. In this case, the nominal element of the CP (*tohmat* in (49a)) becomes the subject of the passive construction (49b), as does the object of a transitive construction. Finally, the noun can be modified (50). The nominal part of the CP is underlined in the examples.

(45) *Dast* be gol-hā na-zan.
hand to flower-PL NEG-strike.PRS
'Don't touch the flowers.'

(46) Mu-hā=yaš=rā *boros* yā *šāne* zad.
hair-PL=3SG=RA brush or comb strike.PAST.3SG
'(S)he brushed or combed her hair.'

(47) Omid *sili* zad va xord.
Omid slap hit.PAST.3SG and strike.PAST.3SG
'Omid gave and received slaps.'

(48) *Dast* goft-am [be gol-hā —– na-zan].
hand said-1SG to flower-PL —– NEG-strike.PRS
'I told you not to touch the flowers.'

(49) a. Maryam be Omid *tohmat* zad.
Maryam to Omid slander strike.PAST.3SG
'Maryam slandered Omid.'

b. be Omid *tohmat* zade šod.
to Omid slander hit.PP become.PAST.3SG
'Omid was slandered.'

(50) [In *xabar*=e mohem]=rā be mā dād
this news=EZ important=RA to us give.PAST.3SG
'(S)he gave us this important piece of news.'

Nevertheless, there is evidence that the verb and the noun may share one argument structure. For instance, in 51b the verb *dādan* and the noun *āb* take together a direct object.

(51) a. Maryam be bāqče āb dād
Maryam to garden water give.PAST.3SG
'Maryam watered the gadern.'

b. Maryam bāqče=rā āb dād
Maryam garden-DOM water give.PAST.3SG
'Maryam watered the gadern.'

Persian complex predicates do not have a homogeneous semantics. The general idea is that the verb serves to turn a noun into a verb (Bonami & Samvelian 2010), but there is a spectrum, going from a semantically compositional combination, with the head taking as its argument the content of its complement, like in Romance languages CP (see section §3) or "causative" predicates formed with in adjective, like *bāz kardan* 'to open' (open do) (see section 5.1), to idioms whose semantics is not predictable from the components. Such is the case of *dast zadan (be)* 'to begin' (hand strike):[12]

(52) kārgar-ān be e'tesāb dast zad-and
worker-PL to strike hand strike.PAST-3PL
'The workers went on strike.'

We begin with idioms illustrated by *dast zadan*, following analyses in Müller (2010), which draw on Goldberg (1996), Nunberg et al. (1994), and Sag (2007): the head of the idiom is an item in the lexicon, with special selection properties and the appropriate meaning for the idiom. More precisely, the verb is associated with an idiomatic reading, an i-relation which arises in the special combination; it selects its complement via the lexical identifier or LID feature, which is associated with lexemes in the lexicon, and contains morpho-syntactic as well as semantic information (Sag 2007; 2012). The lexical entry for the verb *zadan* 'to hit' in the idiomatic combination *dast zadan* 'to start' is given in (53):

(53) Lexical entry of *zadan* 'strike' in *dast zadan* 'to start'

$$\begin{bmatrix} \textit{zadan1-lxm} \\ \text{CAT}\ \textit{verb} \\ \text{CONT} \begin{bmatrix} \textit{soa} \\ \text{NUCLEUS} \begin{bmatrix} \text{REL} & \textit{i-start-rel} \\ \text{AGENT} & k \\ \text{THEME} & m \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} \left\langle \text{NP}_k, \left(\textit{be}\right) \text{NP}_m, \text{N} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{LIGHT} & + \\ \text{LID} & \textit{dast} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The same mechanisms can be used to describe the intermediate cases. We illustrate three cases, drawing on the detailed study of the verb *zadan* 'to strike' in Samvelian (2012). They show different ways in which the noun takes preeminence over the verb. The meaning of the verb can be close to a basic meaning that it has outside the light verb construction; the noun contributes to a more specific interpretation, and the meaning of the verb is absorbed into that of the noun. This is the case for the light verb construction involving *zadan* 'to strike' and a noun describing a type of blow, for instance, *lagad* 'kick', *sili* 'slap'. So *lagad zadan* simply means 'to kick'.

[12] Note that the combination *dast zadan* has several meanings. When used with a PP argument introduced by *be* 'to', it either means 'to touch' or 'to start'. In its intransitive use, it means 'to clap'.

(54) olāq be Omid lagad zad
donkey to Omid kick strike.PAST.3SG
'The donkey kicked Omid.'

The description of *zadan* (here, *zadan-2*) when it combines with a noun depicting a type of blow is as in (55).

(55)
$$\begin{bmatrix} \textit{zadan2-lxm} \\ \textsc{cat}\ \textit{verb} \\ \textsc{cont} \begin{bmatrix} \textit{soa} \\ \textsc{nucleus} \begin{bmatrix} \textsc{rel} & \textit{i-inflict-}\boxed{1}\textit{-rel} \\ \textsc{agent} & k \\ \textsc{patient} & m \end{bmatrix} \end{bmatrix} \\ \textsc{arg-st} \left\langle \textsc{np}_k, \left(\textit{be}\right) \textsc{np}_m, \textsc{n} \begin{bmatrix} \textsc{cat} \begin{bmatrix} \textsc{pred} & + \\ \textsc{lid} & \boxed{1}\ \textit{blow-rel} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

In other cases, the meaning of the noun takes over: the construction is similar to a derivation rule, and converts the noun into a verb with the same meaning, 57. When *zadan* combines with a noun describing a kind of fraud or deception, the complex predicate means 'deceive' in the way described more precisely by the noun.

(56) Hālā mā kalak be=hešun mi-zan-im
Now we deceit to=them IPFFV-strike.PRS-3PL
'Now, it is our turn to deceive them.'

(57)
$$\begin{bmatrix} \textit{zadan3-lxm} \\ \textsc{cat}\ \textit{verb} \\ \textsc{cont} \begin{bmatrix} \textit{soa} \\ \textsc{nucleus} \begin{bmatrix} \textsc{rel} & \textit{i-deceit-}\boxed{1}\textit{-rel} \\ \textsc{agent} & k \\ \textsc{patient} & m \end{bmatrix} \end{bmatrix} \\ \textsc{arg-st} \left\langle \textsc{np}_k, \left(\textit{be}\right) \textsc{np}_m, \textsc{n} \begin{bmatrix} \textsc{cat} \begin{bmatrix} \textsc{pred} & + \\ \textsc{lid} & \boxed{1}\ \textit{deceit-rel} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Finally, there are also cases where the meaning of the noun is preserved, but a new meaning emerges for the verb, in a non predictable manner. When *xod=rā zadan* (lit. self strike) combines with a subset of nouns describing illnesses, handicaps or problematic states (like stupidity, ignorance), it means 'to pretend, feign' the illness or state in question.

(58) Maryam xod=rā be divanegi zad
Maryam self=RA to madness strike.PAST.3SG
'Maryam feigned madness.'

(59)
$$\begin{bmatrix} \textit{zadan4-lxm} \\ \text{CAT } \textit{VERB} \\ \text{CONT} \begin{bmatrix} \text{SOA} \\ \text{NUCLEUS} \begin{bmatrix} \text{REL} & \textit{i-pretend}\boxed{1}\textit{-rel} \\ \text{AGENT} & k \\ \text{THEME} & e \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} \left\langle \text{NP}_k, \textit{pro}_k, \text{N} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{INDEX} & e \\ \text{LID} & \boxed{1}\textit{i-illness-rel} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Not all nouns for illnesses and problematic states can occur in this complex predicate. To account for this, we indicate that the noun, too, has an idiomatic meaning, an *i-rel*, in addition to its usual meaning, and only the nouns which have such an *i-rel* can participate in the complex Predicate with *xod=rā zadan*. Here, the case is close to that of an idiom, but the noun does not have a special meaning: it is special in that not all nouns with a comparable semantics can form such a complex Predicate with *xod=rā zadan*.

Acknowledgements

References

Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.

Abeillé, Anne & Danièle Godard. 2005. The diversity of complex predicates in Romance languages: argument sharing and syntactic structure. Manuscript, University of Université Paris-Diderot.

Abeillé, Anne & Danièle Godard. 2010. Complex predicates in the Romance languages. In Danièle Godard (ed.), *Fundamental issues in the romance languages*, 107–170. Stanford: CSLI Publications.

Abeillé, Anne, Danièle Godard & Ivan A. Sag. 1998. Two kinds of composition in French complex predicates. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 1–41. San Diego: Academic Press.

Aissen, Judith & David Perlmutter. 1983. Clause reduction in spanish. In David Perlmutter (ed.), *Studies in relational grammar 1*, 360–403. Chicago: The University of Chicago Press.

Bech, Gunnar. 1955. *Studien über das deutsche Verbum infinitum* (Linguistische Arbeiten 139). Tübingen: Max Niemeyer Verlag. 2nd unchanged edition 1983.

Bonami, Olivier & Pollet Samvelian. 2010. *Persian complex predicates: lexeme formation by itself*. Talk presented at Les Décembrettes 7. Colloque international de morphologie, 2-3 December, 2-3, 2010.

Bonami, Olivier & Pollet Samvelian. 2015. The diversity of inflectional periphrasis in Persian. *Journal of Linguistics* 51(1). 327–382. DOI:10.1017/S0022226714000243

Bonami, Olivier & Gert Webelhuth. 2013. The phrase-structural diversity of periphrasis : a lexicalist account. In Marina Chumakina et Greville G. Corbett (ed.), *Periphrasis : the role of syntax and morphology in paradigms*, 141–167. Oxford: Oxford University Press.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press.

Butt, Miriam. 2010. The light verb jungle: Still hacking away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex predicates: cross-linguistic perspectives on event structure*, 48–78. Cambridge University Press.

Choi, Incheol & Stephen Wechsler. 2002. Mixed categories and argument transfer in the Korean light verb construction. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *The proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 103–120. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2/, accessed 2018-3-20.

Chung, Chan. 1998. Argument composition and long-distance scrambling in Korean. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 159–220. San Diego: Academic Press.

De Kuthy, Kordula & Walt Detmar Meurers. 2001. On partial constituent fronting in German. *Journal of Comparative Germanic Linguistics* 3(3). 143–205.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. *Berkeley Linguistic Society* 22. 132–146. http://roa.rutgers.edu/files/415-0900/roa-415-goldberg-2.pdf, accessed 2018-3-20.

Hale, Kenneth & Samuel Jay Keyser. 1997. On the complex nature of simple predicators. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates* (CSLI Lecture Notes 64), 29–65. Stanford, CA: CSLI Publications.

Haspelmath, Martin. 2016. The serial verb construction: comparative concept and cross linguistic generalizations. *Language and Linguistics* 3(17). 291–319.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–38. Stanford, CA: CSLI Publications.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1998. Third Construction and VP extraposition in German: An HPSG analysis. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 115–157. San Diego: Academic Press.

Karimi-Doostan, Gholamhossein. 1997. *Light verb constructions in Persian*. Department of Language & Linguistics, University of Essex dissertation.

Kathol, Andreas. 1998. Constituency and linearization of verbal complexes. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 221–270. San Diego: Academic Press.

Kim, Jong-Bok. 2016. *The syntactic structures of Korean: A Construction Grammar perspective*. Cambridge, UK: Cambridge University Press.

Kiss, Tibor. 1994. Obligatory coherence: The structure of German modal verb constructions. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 71–108. Stanford, CA: CSLI Publications.

Lee, Sun-Hee. 2001. Argument composition and linearization: Korean complex predicates and scrambling, *OSUWPL* 56. 53–78.

Megerdoomian, Karine. 2002. *Beyond words and phrases: A unified theory of predicate composition*. University of Southern California dissertation. http://www.zoorna.org/publications.html, accessed 2015-12-23.

Meurers, Detmar. 2002. To flip or not to flip: on the nature of irregularities in the german verbal complex. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *The proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 235–246. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2/, accessed 2018-3-20.

Miller, Philip H. & Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3). 573–639.

Monachesi, Paola. 1998. Italian restructuring verbs: A lexical analysis. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 313–368. San Diego: Academic Press.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. DOI:10.1017/S0022226709990284

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.

Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Rentier, Gerrit. 1994. Dutch cross serial dependencies in HPSG. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 818–822. Kyoto, Japan: Association for Computational Linguistics. http://www.aclweb.org/anthology/C94-2130, accessed 2018-2-25.

Rizzi, Luigi. 1982. *Issues in Italian syntax* (Studies in Generative Grammar 11). Dordrecht: Foris.

Ryu, Byong-Rae. 1993. *Structure sharing and argument transfer: An HPSG approach to verbal noun constructions*. SfS Report 04-93. Eberhard-Karls-Universität Tübingen: Seminar für Sprachwissenschaft.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/, accessed 2018-2-25.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.

Salvi, Giampaolo. 2010. Gli ausiliari essere e avere in italiano. *Acta linguistica Academiae scientiarum Hungaricae* 30(1-2). 137–162.

Samvelian, Pollet. 2012. *Grammaire des prédicats complexes : les constructions nom-verbe*. Paris: Lavoisier.

Samvelian, Pollet & Pegah Faghiri. 2016. In *Proceedings of the 39th Annual Meeting of the Berkeley Linguistics Society*, 212–227. Berkeley: Berkeley Linguistics Society.

Sells, Peter. 1991. Complex verbs and argument structures in Korean. In Susumo Kuno (ed.), *Harvard studies in korean linguistics 4*, 395–406. Seoul: Hanshin.

Suñer, Margarita. 2010. Two properties of clitics in clitic-doubled constructions. In C.-T. James Huang & Robert May (eds.), *Logical structure and linguistic structure*, vol. 40 (Studies in Linguistics and Philosophy), 233–252. Dordrecht: Kluwer.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

Yoo, Eun-Jung. 2003. Case marking in Korean auxiliary verb constructions. In Jong-Bok Kim & Stephen Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 413–438. Stanford, CA: CSLI Publications.

Chapter 13

Control and raising

Anne Abeillé
Université Paris Diderot

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse

Anne Abeillé. 2019. Control and raising. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 351–352. Berlin: Language Science Press. DOI:??

rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 14

Unbounded dependencies

Bob Borsley
University of Essex

Berthold Crysmann
Université Paris Diderot

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

finality

Abbreviations

Acknowledgements

Bob Borsley & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Chapter 15

Relative Clauses in HPSG

Doug Arnold
University of Essex

Danièle Godard
University de Paris Diderot

We provide an extended discussion of analyses of relative clauses (prototypically clauses with a noun modifying function) and related constructions that have appeared in the HPSG literature. The basic theoretical approaches are presented (specifically, the lexical "head-driven" approach associated with earlier work in HPSG, and the more recent constructional approach), followed by descriptions of analyses of different kinds of relative clause across a range of typologically diverse languages (notably Arabic, English, French, German, Japanese, and Korean). Phenomena discussed include *wh*-relatives, relatives headed by complementisers, "bare" relatives, non-restrictive relatives, extraposition of relative clauses, relative clause-like constructions that function as complements, various kinds of "dependent noun" and "pseudo" relative clause, and free (headless) relatives.

1 Introduction

The goal of this paper is to give an overview of HPSG analyses of relative clauses. Relative clauses are, typically, sentential constructions that function as nominal modifiers, like the italicised part of (1), for example.

(1) The person *to whom Kim spoke yesterday* claimed to know nothing.

Relative clauses have been an important topic in HPSG: not only as the focus on a considerable amount of descriptive and theoretical work across a range of languages, but also in terms of the theoretical development of the framework.

Doug Arnold & Danièle Godard. 2019. Relative Clauses in HPSG. in Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 355–414. Berlin: Language Science Press. DOI:??

Notably, Sag's (1997) analysis of English relative clauses was the first fully developed realisation of the constructional approach involving cross-classifying phrase types that has dominated work in HPSG in the last two decades, and was thus the first step towards the development of Sign-based Construction Grammar (cf. Müller (2019b), Chapter P of this volume).

The basic organisation of the discussion is as follows. Section 2 introduces basic ideas and overviews the main analytic techniques that have been used, focusing on one kind of relative clause. Section 3 looks at other kinds of relative clause in a variety of languages. Section 4 looks at a variety of constructions which have some similarity with relative clauses, but which are in some way untypical (e.g. clauses that resemble relative clauses, but which are not nominal modifiers, or which are not adjoined to the nominals they modify). Section 5 provides a conclusion.

2 Basic ideas and approaches

This section introduces basic ideas and intuitions about relative clauses, viewed from an HPSG perspective (Section 2.1), then introduces the two main approaches that have been taken in HPSG: the lexical approach of Pollard & Sag (1994) which makes use of phonologically empty elements (Section 2.2), and the constructional approach of Sag (1997), which makes phonologically empty elements unnecessary (Section 2.3). Section 2.4 presents some interim conclusions, and provides some discussion of some brief discussion alternative approaches.

2.1 Basic ideas and intuitions

Relative clauses are, prototypically, sentential constructions which modify a nominal. (2) is an example of one kind of English relative clause, which we will call a "*wh*-relative". In (3) it is used as a modifier of the nominal *person* (the *antecedent* of the relative clause).

(2) to whom Kim spoke yesterday

(3) The person to whom Kim spoke yesterday claimed to know nothing.

Syntactically, this kind of relative clause consists of a preposed *wh*-phrase (*to whom*), i.e. a phrase containing a relative pronoun (*whom*), and a clause with a missing constituent – a gap (the complement of *speak*: *Kim spoke __ yesterday*). This is often called the *relativised constituent*. Semantically, in (2) the interpretation of the relative clause is *intersective*: (2) denotes the intersection of the set of people and the set of entities that Kim spoke to. Getting this interpretation involves combining the descriptive content of the antecedent nominal and the propositional content of the relative clause, and equating the referential indices of the nominal and the relative pronoun, to produce something along the lines of "the set of x where x is a person and Kim spoke to x".

Not all relative clauses have these properties, but they provide a good starting point. In the remainder of this section, we will show, in broad terms, how these properties can be accounted for.

As regards their function and distribution, relative clauses are subordinate clauses, which can be captured by assuming they have a HEAD feature like [MC −], "MAIN-CLAUSE *minus*". They are naturally assumed to be adjuncts: their distribution as nominal adjuncts can be dealt with by assuming that (like other adjuncts) they indicate the sort of head they can modify via a feature like MOD or SELECT (cf. **chapters/ChXXX** Chapter **??** of this volume on the treatment of adjuncts). That is, relative clauses such as (2) will be specified as in (4a), whereas adjunct clauses headed by a subordinator like *because* (as in *We're late because it's raining*) will be specified as (4b), and normal, non-adjunct, clauses will typically be specified as (4c):

(4) a. $\left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}|\text{MOD}\ \left[\text{LOC}\ |\ \text{CAT}\ |\ \text{HEAD}\ \ \textit{noun}\right]\right]$

b. $\left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}|\text{MOD}\ \left[\text{LOC}\ |\ \text{CAT}\ |\ \text{HEAD}\ \ \textit{verb}\right]\right]$

c. $\left[\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}|\text{MOD}\ \ \textit{none}\right]$

With this in hand, we will look in more detail at the internal structure of this kind of relative clause (Section 2.1.1), and at the relation between the relative clause and its antecedent (Section 2.1.2).

2.1.1 The internal structure of the relative clause

As regards internal structure, it is characteristic of *wh*-relatives that they consist of a preposed *wh*-phrase, and a clause containing a gap. The dependency between the *wh*-phrase and the associated gap is potentially unbounded, as can be seen from examples like (5).

(5) the person to whom [Sam said [Kim intended [to speak ___ yesterday]]]

As regards the *wh*-phrase, it is notable that it must be preposed — English does not allow examples like (6a) without a relative phrase, or (6b) where the relative phrase is *in situ*.

(6) a. *a person Kim spoke to her yesterday

b. *a person Kim spoke to whom yesterday

Despite being forbidden *in situ*, the preposed *wh*-phrase behaves in some respects as though it occupied the gap. For example, in the examples above *to whom* satisfies the subcategorisation requirements of *speak*, and makes a semantic contribution in the gapped clause. Assuming some kind of co-indexation relation between the antecedent and the *wh*-phrase, the same behaviour can be seen with subject-verb agreement, as in (7a), and binding, as in (7a):

(7) a. a person who [everyone thinks [__ is/*are weird]]
b. a person who [everyone thinks [__ hates herself/*her]]

In fact, this dependency between the *wh*-phrase and the gap appears to be a typical filler-gap dependency, with the *wh*-phrase as the filler, which can be handled by standard SLASH inheritance techniques (see Borsley & Crysmann (2019), Chapter 14 of this volume), so that these properties are accounted for.

In examples like (2) the *wh*-phrase must contain a relative pronoun. Here we have another apparently unbounded dependency, because the relative pronoun can be embedded arbitrarily deeply inside the *wh*-phrase (example (8d) is due to Ross 1967):

(8) a. the person [to [whose friends]] Kim spoke __
b. the person [to [[whose children's] friends]] Kim spoke __
c. the person [to [the children [of [whose friends]]] Kim spoke __
d. books [the height [of [the letters [on [the covers [of which]]]]] the government regulates __

This dependency between a relative pronoun and the phrase that contains it is often called "*wh*-percolation", "relative percolation", or, following Ross (1967), "pied-piping". We will talk about *relative inheritance*.

Notice that as well as being unbounded, relative inheritance resembles SLASH inheritance in that the "bottom" of the inheritance path (i.e. the actual relative pronoun, or the gap in a filler-gap a dependency) is typically not a head (e.g. *whom* is not the head of *to whom*). Moreover, though examples involving multiple independent relative pronouns are rather rare in English (i.e. there are few, if any, relative clauses parallel to interrogatives like *Who gave what to whom?*) they exist in other languages, so it is reasonable to assume that relative inheritance involves a set of some kind.[1] This motivates the introduction of a REL feature

[1]Examples of languages which allow multiple relative pronouns include Hindi (e.g. Srivastav 1991) and Marathi (e.g. Dhongde & Wali 2009: Ch7). See Pollard & Sag (1994: 227–232) for

which is subject to the same kind of formal mechanisms as SLASH.[2]

The idea is that a relative pronoun will register its presence by introducing a non-empty REL value, which will be inherited upwards until it reaches the preposed *wh*-phrase at the top of the relative clause (equivalently: a relative clause introduces a non-empty REL value on its *wh*-phrase daughter that is inherited downwards till it is realised as a relative pronoun). Within the *wh*-phrase, REL inheritance can be handled by the same sort of formal apparatus as is used for handling SLASH inheritance. Blocking REL inheritance from carrying a REL element upwards beyond the top of relative clause can be achieved with the same formal apparatus as is used to block SLASH inheritance from carrying information about a gap higher than the level at which the associated filler appears.[3]

Co-indexation of the antecedent nominal and the relative pronoun can be achieved simply if the REL value contains an index which is shared by both the antecedent and the relative pronoun. As regards the relative pronoun, at the "bottom" of the REL dependency, this can be a matter of lexical stipulation: relative pronouns can be lexically specified as having a REL value that contains their INDEX value, roughly as in (9a), which we abbreviate to (9b).[4]

(9) a. $\left[\text{SYNSEM}\ \left[\begin{array}{ll}\text{LOC} & \left[\begin{array}{ll}\text{CAT} & \left[\text{HEAD}\ \textit{noun}\right]\\ \text{CONT} & \left[\text{INDEX}\ \boxed{1}\right]\end{array}\right]\\ \text{NON-LOC} & \left[\text{INHER}|\text{REL}\ \{\boxed{1}\}\right]\end{array}\right]\right]$

HPSG analyses. In English, multiple relative pronouns occur in cases of co-ordination (e.g. *the person with whom or for whom you work*), but they are not independent (they relate to the same entity). Kayne (2017) gives some English examples that appear to involve multiple relative pronouns, but they are rather marginal.

[2] The assumption that relative inheritance should be treated as involving an unbounded dependency (i.e. handled with a NON-LOCAL feature, like SLASH), has been challenged in Van Eynde (2004) (Van Eynde argues it should be treated as local dependency).

[3] In case it is not obvious why further upwards inheritance of a REL value would be problematic, notice that while a relative clause can *contain* a *wh*-phrase, it cannot *be* a *wh*-phrase, e.g. it cannot function as the filler in a relative clause. Suppose, counter-factually, the REL value of *who* could be inherited beyond the relative clause *to whom Kim spoke*, so that e.g. *a person to whom Kim spoke* was marked as [REL $\{\boxed{1}\}$]. This phrase would be able to function as the *wh*-phrase in a relative clause like **[a person to whom Kim spoke] Sam recognised __* , which would be able to combine with a noun specified as [INDEX $\boxed{1}$] to produce something like **a person [[a person to whom Kim spoke] Sam recognised __].*

[4] Here, and below, we will abbreviate attribute paths where no confusion arises, and use a number of other standard abbreviations, in particular, we write INDEX values as subscripts on nouns and NPs. We use $\overline{\text{N}}$ to indicate a noun with an empty COMPS list, i.e. one which has combined with its complements, if any, and NP for a $\overline{\text{N}}$ with an empty SPEC (SPECIFIER) list (e.g. a combination of determiner and a $\overline{\text{N}}$). Similarly, we use PP to abbreviate a phase consisting of a preposition and its complement, VP for a phrase consisting of verb with its complements, and S for a phrase consisting of a subject and a VP.

b. $\overline{\mathrm{N}}_{\boxed{1}}$ $\left[\textsc{rel}\ \{\boxed{1}\}\right]$

This index can then be inherited upwards via the REL value to the level of the *wh*-phrase. At the top, the index of the antecedent can be accessed via the MOD value of the relative clause: this is simply a matter of replacing the specification of the MOD value in (4a) with that in (10a), abbreviated as in (10b), where $\boxed{1}$ is the index that appears in the REL value of the associated *wh*-phrase.[5]

(10) a. $\left[\text{SYNSEM|LOC|CAT|HEAD|MOD}\ \left[\text{LOC}\ \begin{bmatrix}\text{CAT} & [\text{HEAD}\ \textit{noun}] \\ \text{CONT} & [\text{INDEX}\ \boxed{1}]\end{bmatrix}\right]\right]$

b. S $\left[\text{MOD}\ \overline{\mathrm{N}}_{\boxed{1}}\right]$

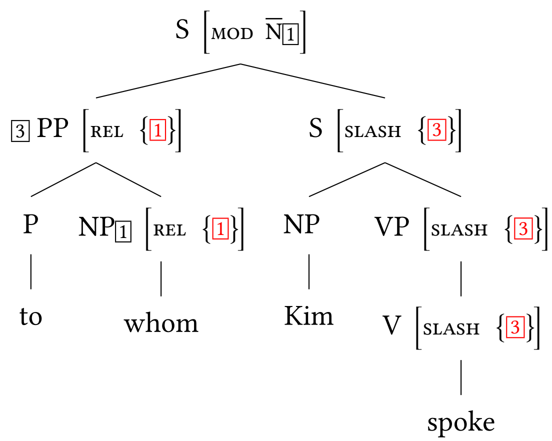


Figure 1: Representation of *to whom Kim spoke.*

Schematically, then, *wh*-relatives should have structures along the lines of Figure 1. The top structure here is a head-filler structure. Notice how SLASH inheritance ensures the relevant properties of the PP are shared by lower nodes so that the subcategorization requirements of the verb can be satisfied, with the PP being interpreted as a complement of the verb (equivalently: SLASH inheritance ensures that the gap caused by the missing complement of *speak* is registered on higher nodes until it is filled by the PP). Similarly, REL inheritance means that the

[5]We assume, for simplicity, that the value of REL is a set of indices. This is consistent with e.g. Pollard & Sag (1994) and Sag (1997), but not with Ginzburg & Sag (2000: 188), who assume it is a set of *parameters*, that is, indices with restrictions (a kind of *scope-object*), like the QUE attribute which is used for *wh*-inheritance in interrogatives. It is not clear that anything important hangs on this.

INDEX of the relative pronoun appears on higher nodes so that it can be identified with the INDEX of the antecedent noun, via the MOD value of the highest S (equivalently: the index of the antecedent nominal appears on lower nodes down to the relative pronoun, so that the nominal and the relative pronoun are co-indexed).

As regards (CONTENT, the effect of this will be to give the relative clause *to whom*$_i$ *Kim spoke* an interpretation along the lines of *Kim spoke to whom*$_i$, where *i* is the index of its antecedent. In terms of standard HPSG semantics, this "internal" content (i.e. the content associated with a verbal head with its complements and modifiers) is a *state-of-affairs* (*soa*), and can be represented as in (11a), abbreviated to (11b):[6]

(11) a. $$\begin{bmatrix} \textit{soa} & \\ \textsc{nuc} & \begin{bmatrix} \textit{speak\_to} & \\ \textsc{speaker} & \textit{Kim} \\ \textsc{addressee} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

b. *speak_to(Kim,*$\boxed{1}$*)*

There are restrictions on what can occur as the preposed *wh*-phrase in a relative clause, as can be seen in (12). To a first approximation, NPs and PPs are fine in English, but VPs and Ss are not.[7]

(12) a. the person [$_{\text{NP}}$ who] we want Kim to speak to ___

b. the person [$_{\text{PP}}$ to whom] we want Kim to speak ___

c. * the person [$_{\text{VP}}$ speak to whom] we want Kim to ___

d. * the person [$_{\text{VP}}$ to speak to whom] we want Kim ___

e. * the person [$_{\text{S}}$ Kim to speak to whom] we want ___

It is rather natural to interpret this as indicating restrictions on relative inheritance (i.e. pied-piping in relative clauses) – e.g. as indicating that while relative inheritance from NP to PP (and through an upward chain of NPs and PPs, as in (8d)), is permitted, it is blocked by VP and S nodes. The can be achieved by requiring that the REL value on VP and S nodes to be empty (cf. the Clausal REL Prohibition of Pollard & Sag 1994: 220). While important, these restrictions are

[6] In fact (17) is already somewhat abbreviated: [SPEAKER *Kim*] is an abbreviation for a structure including an index, and a BACKGROUND restriction on that index indicating that it stands in the *naming* relation to the name *Kim*.

[7] This is a considerable simplification: e.g. English allows VPs like that in (12d) so long as they function as subjects, so *person to speak to whom is a privilege* is allowed. German allows REL inheritance to VP more freely than English, and an analogue of (12d) is grammatical in German. See De Kuthy (1999), Hinrichs & Nakazawa (1999) and Müller (1999b) for discussion and HPSG analyses.

poorly understood and we will have nothing to say about them here, except to emphasise that they are different from the restrictions on QUE inheritance (i.e. pied-piping in interrogatives). For example, QUE-inheritance is not possible from the complement of a noun, but REL inheritance is fine, so *some pictures of whom* is not possible as the focus of a question, as in (13), but is fine as the initial phrase of a relative clause, as in (14):

(13) *I wonder [some pictures of whom] they were admiring.

(14) the children [some pictures of whom] they were admiring

Notice that REL and QUE also differ in other ways: e.g. as Sag (2010: 490–493) emphasises, there are *wh*-words that can function as interrogative pronouns, but not as relative pronouns (i.e. which have non-empty QUE values, but empty REL values), and *vice versa*. For example, *how* and (in standard English) *what* function as interrogative pronouns, but not relative pronouns, as the following show (as Sag 2010: 493 puts it, there is "no morphological or syntactic unity underlying the concept of an English *wh*-expression):[8]

(15) a. I wonder how she did it. (interrogative)
b. *the way how she did it (relative)

(16) a. I wonder what (things) she bought. (interrogative)
b. *the book what (things) she bought (relative)

With this overview of the internal structure of a relative clause in place, we now turn to relation between the relative clause and the nominal it modifies (its antecedent).

2.1.2 The relative clause and its antecedent

The combination of a relative clauses and the nominal it modifies is traditionally regarded as a head-adjunct structure, where the nominal is the head and the relative clause is the adjunct, as in Figure 2.

[8]See also Müller (1999a: 81–85) on differences between interrogative and relative pronouns in German. Several non-standard English dialects allow NP *what* as a relative pronoun like *which* (cf. non-standard *%the book what she bought*, vs standard *the book which she bought*), no dialect allows determiner *what* as a relative pronoun (though it is fine as an interrogative, as can be seen in (16a)). Sag (2010: 491, note 10) suggests that NP *which* is only ever a relative pronoun (an apparent counter-example like *Which did you buy?* involves determiner *which* with an elliptical noun.)

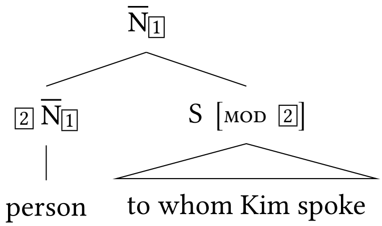


Figure 2: A relative clause and its antecedent.

The content we want for a modified nominal such as *person to whom Kim spoke*, as for an unmodified nominal such as *person*, is a *restricted index*, i.e. in HPSG terms a *scope-object* — an INDEX and a RESTR (RESTRICTION) set (a set of objects of type *fact*).[9] For *person*, this is as in (17), abbreviated as in (18), for *person to whom Kim spoke* it is as in (19), abbreviated as in (20).

(17) $$\begin{bmatrix} \textit{scope-obj} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \left\{ \begin{bmatrix} \textit{fact} & \\ \text{PROP}|\text{SOA} & \begin{bmatrix} \textit{soa} & \\ \text{NUC} & \begin{bmatrix} \textit{person} & \\ \text{INSTANCE} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\} \end{bmatrix}$$

(18) $\boxed{1} : \{person(\boxed{1})\}$

(19) $$\begin{bmatrix} \textit{scope-obj} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \left\{ \begin{array}{l} \begin{bmatrix} \textit{fact} & \\ \text{PROP}|\text{SOA} & \begin{bmatrix} \textit{soa} & \\ \text{NUC} & \begin{bmatrix} \textit{person} & \\ \text{INSTANCE} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \\ \begin{bmatrix} \textit{fact} & \\ \text{PROP}|\text{SOA} & \begin{bmatrix} \textit{soa} & \\ \text{NUC} & \begin{bmatrix} \textit{speak\_to} & \\ \text{SPEAKER} & \textit{Kim} \\ \text{ADDRESSEE} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{array} \right\} \end{bmatrix}$$

(20) $\boxed{1} : \{person(\boxed{1}), speak\_to(Kim, \boxed{1})\}$

[9] In Pollard & Sag (1994), *scope-objects* were called *nom-objects*, and restrictions were sets of *parameterized states of affairs* (*psoas*), rather than *facts*. The difference reflects the more comprehensive semantics of Ginzburg & Sag (2000), which involves different kinds of *message* (e.g. *proposition*, *outcome*, and *question*, as well as *fact*). For our purposes, this is just a minor change in feature geometry: *facts* contain Pollard & Sag style *state-of-affairs* content as the value of the PROP | SOA path, as can be seen in (17).

To get the content of *person to whom Kim spoke* from the content of *person* is a matter of producing a *scope-object* whose index is the index of *person* (and the relative pronoun), and whose restrictions are the union of the restrictions of *person* with a set containing a *fact* corresponding to the *state-of-affairs* that is the content of the relative clause. Unioning the restrictions gives the intersective interpretation.

Conceptually, this is straightforward, but there is technical difficulty: the structure in Figure 2 is a head-adjunct structure, and in such structures the content should come from the adjunct daughter, the relative clause. That is, for "external" semantic purposes (purposes of semantic composition) relative clauses should have *scope-object* content, but as we have seen their "internal" content is a *soa*. So some special apparatus will be required, as will appear in the following discussion.[10]

This should give the reader an idea of the general shape of an approach to relative clauses like (2) using HPSG apparatus. In the following sections we will make this more precise by outlining the two main approaches that have been taken to the analysis of relative clauses in HPSG: the lexical approach of Pollard & Sag (1994: Chapter 5), which makes use of phonologically empty elements, and the constructional approach of Sag (1997), which does not.[11]

2.2 The lexical approach of Pollard & Sag (1994)

The idea that relative clauses have a lexical head is appealing for some kinds of relative clause in many languages (see below, e.g. Section 3.2, Section 3.3), but it is

[10]Though the details are HPSG specific, this is a general problem, regardless of semantic theory. For example, in a setting using standard logical types, relative clauses *qua* clauses (saturated predications) might be assigned type t, but in order to act as nominal modifiers this predicative semantics must be converted into "attributive" (noun-modifying) semantics, i.e. logical type $\langle et, et \rangle$. See, e.g. Sag (2010: 521–524) where an HPSG syntax is combined with a conventional predicate-logic based semantics for relative clauses.

[11]Müller (1999a) presents what might be considered a third approach, which resembles Sag (1997) in avoiding empty elements, but uses a rule schema for German relative clauses rather than constructional apparatus of phrasal types (see Müller 1999a: 95 for details). The overview of HPSG in Müller & Machicao y Priemer (2019) also presents a rule schema for relative clauses (*loc cit* Section 6.1). Rule schemas were a crucial piece of apparatus in the framework of Pollard & Sag (1994), but they have fallen out of favour with the rise of construction based analyses since Sag (1997). A rule schema is essentially just a phrasal type — that is, a type describing constraints on a mother and daughters — with the difference that unlike phrasal types, rule schemas do not stand in inheritance relations, so it is not possible to factor out generalisations in the way of construction based analyses. This is not an issue for Müller, who claims that a description of restrictive relatives in German requires only a single schema (Müller 1999a: 74).

problematic for relative clauses like (2) — there is no obvious candidate to serve as the head. This is clearly problematic for a lexical, "head-driven" approach, such as HPSG. Building on an approach originally proposed by Borsley (1989), the analysis proposed in Pollard & Sag (1994: Chapter 5), overcomes this problem by assuming that relative clauses involve a phonologically empty head, which Pollard & Sag call R ("relativiser"), and which projects an RP (that is, a relative clause).

R is lexically specified to be a nominal modifier (i.e. [MOD *noun*]) which takes two arguments. The first is an XP, the *wh*-phrase, with a REL value which contains the index of the antecedent nominal. The second is sentential, and constrained to have a SLASH value that includes the XP. With some simplifications and some minor modifications to fit the framework we assume here, this is along the lines of (21) (cf. Pollard & Sag 1994: 216). Here XP $\boxed{4}$ is intended to mean an XP whose LOCAL value is $\boxed{4}$, and S:$\boxed{3}$ means a clause (a saturated *verb* – i.e. one with empty SUBJ and COMPS specifications) whose CONTENT is $\boxed{3}$. The $\boxed{2}$ that appears in the value of RESTR is intended to be the RESTR set of the antecedent nominal (this should be specified as part of the MOD value, but we have not done this, in the interests of readability).

(21)
$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \overline{\text{N}}\boxed{1} \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \textit{scope-obj} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \boxed{2} \cup \left\{ \begin{bmatrix} \textit{fact} & \\ \text{PROP}|\text{SOA} & \boxed{3} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{matrix} \text{XP}\,\boxed{4} \\ \begin{bmatrix} \text{REL} & \{\boxed{1}\} \end{bmatrix} \end{matrix}, \begin{matrix} \text{S:}\boxed{3} \\ \begin{bmatrix} \text{SLASH} & \{\boxed{4}\} \end{bmatrix} \end{matrix} \right\rangle \end{bmatrix}$$

Standard schemas for combining heads with arguments will produce structures like the RP in Figure 3, which (since MOD is a head feature) will inherit the MOD feature from R, and hence combine with a nominal like *person* in a head-adjunct phrase to produce the structure in Figure 3.[12]

This captures the properties described above, and resolves the issues mentioned in the following way.

The first argument of R is specified as [REL $\{\boxed{1}\}$]. Thus, it must contain a relative pronoun. Moreover, (21) specifies that the first argument must correspond to a gap in the second argument. Hence cases like (6) where there is no *wh*-phrase, or where the *wh*-phrase is *in situ*, are excluded.

[12]Here again we have used PP $\boxed{4}$ to indicate a PP whose LOCAL value is $\boxed{4}$.

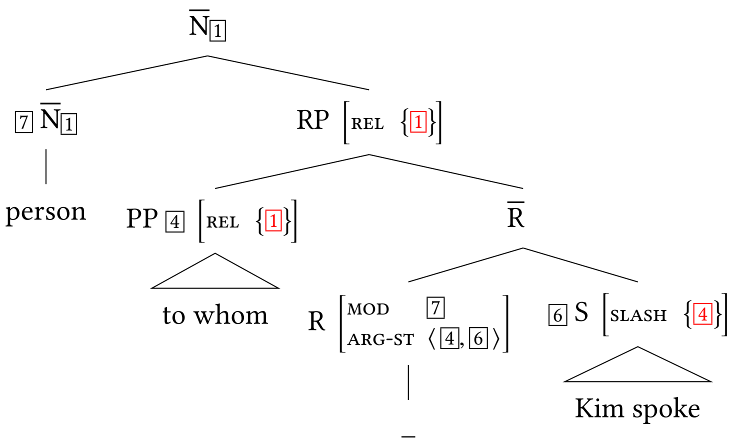


Figure 3: A Pollard & Sag (1994) style structure involving a finite *wh*-relative clause.

Since R, not the slashed S, is the head of RP, there is no problem of mismatch between the content of the S and the relative clause: R is lexically specified as having *fact* (i.e. *scope-object*) content incorporating the "internal" content of its complement clause (tagged $\boxed{3}$) in the appropriate way. This *fact* content will be projected to RP by normal principles of semantic composition relating to heads, complements, and subjects, and RP will produce the right content by unioning the restrictions that come from the head nominal with this *fact* content.

This leaves the question of how upwards inheritance of the REL and SLASH values can be prevented. The same method is used for both. The idea is that for features like REL and SLASH (non-local features) the value on the mother is the union of the values on the daughters, less any indicated as being discharged ("bound-off") on the head daughter (the values that are bound-off in this way are specified as elements of the value of a TO-BIND attribute). Thus, R can be specified so as to discharge the SLASH value on its S sister (so that $\overline{\mathrm{R}}$ is [SLASH {}]), and we can ensure that the topmost $\overline{\mathrm{N}}$ is [REL {}], so long as its head $\overline{\mathrm{N}}$ daughter is specified as binding-off the REL value on RP. This specification can be imposed by stipulation in the MOD value of R. See Pollard & Sag (1994: 164) for details.

The approach can be extended to deal with other kinds of relative clause by positing alternative forms of empty relativiser (see below and Pollard & Sag 1994: Chapter 5).

The great attraction of the approach is that, apart from R, it requires no special apparatus of any kind. On the other hand, it requires the introduction of a novel part of speech (R), and the need to posit phonologically empty elements for which there is no independent evidence. Reservations about this lead Sag to develop the constructional approach presented in Sag (1997).[13]

2.3 The constructional approach of Sag (1997)

The analysis of English relative clauses in Sag (1997) is constructional and completely dispenses with phonologically empty elements.[14] It involves three main constructions: one for combining relative clauses and nominals, and two for relative clauses themselves. One of these is the standard construction for head-filler phrases. The other involves a number of sub-constructions specific to relative clauses, which are treated as a subtype of *clause* (alongside e.g. *declaratives* and *imperatives*). These are outlined (with some simplifications and minor adjustments) in Figure 4.[15]

The *rel-cl* clause type is associated with the constraints in (22), which simply state that relative clauses are subordinate clauses ([MC −]) that modify nouns and have *propositional* content, and that they do not permit subject-aux inversion ([INV −]).[16]

[13] One detail we ignore here concerns the analysis of "subject" relatives: relative clauses where the relative phrase is a grammatical subject inside the relative clause, as in (i):

(i) person who spoke to Kim

Pollard & Sag (1994) treat such examples specially (cf. Pollard & Sag 1994: 218–219), using the "Subject Extraction Lexical Rule" (SELR) which in essence permits a VP to replace an S in an ARG-ST in the presence of a gap, Pollard & Sag (1994: 174), so that R combines with a VP rather than an S. But this is not an essential part of the analysis of relative clauses: it is motivated by quite independent theoretical considerations (specifically, the assumption that gaps are associated only with non-initial members of ARG-ST lists — cf. the "Trace-Principle", Pollard & Sag (1994: 172)). Hence we ignore it here.

[14] See Müller (2019b), Chapter P of this volume, for broader discussion of the constructional approach to HPSG.

[15] See Kim & Sells (2008: Chapter 11) for an introductory overview of English relative clauses on similar lines to Sag (1997). Sag (2010: 521–524) outlines an approach which is stated using the Sign-based Construction Grammar style notation (Boas & Sag 2012). Apart from the semantics (which is formulated using conventional λ-calculus apparatus), it is generally compatible with the earlier analysis described here. One simplification we make here is that we follow the more recent work (e.g. Sag 2010: 523) and do not distinguish subject and non-subject finite relative clauses: Sag (1997) follows Pollard & Sag (1994) in treating them differently (cf. footnote 13; and see Sag 1997: 452–454), but it is not clear how important this is in the framework of Sag (1997).

[16] Giving relative clauses *propositional* content puts them on a par with other kinds of clause, and is not very different from Pollard & Sag's assumption that clauses have *state-of-affairs* content (since *propositions* are simply semantic objects which contain a SOA).

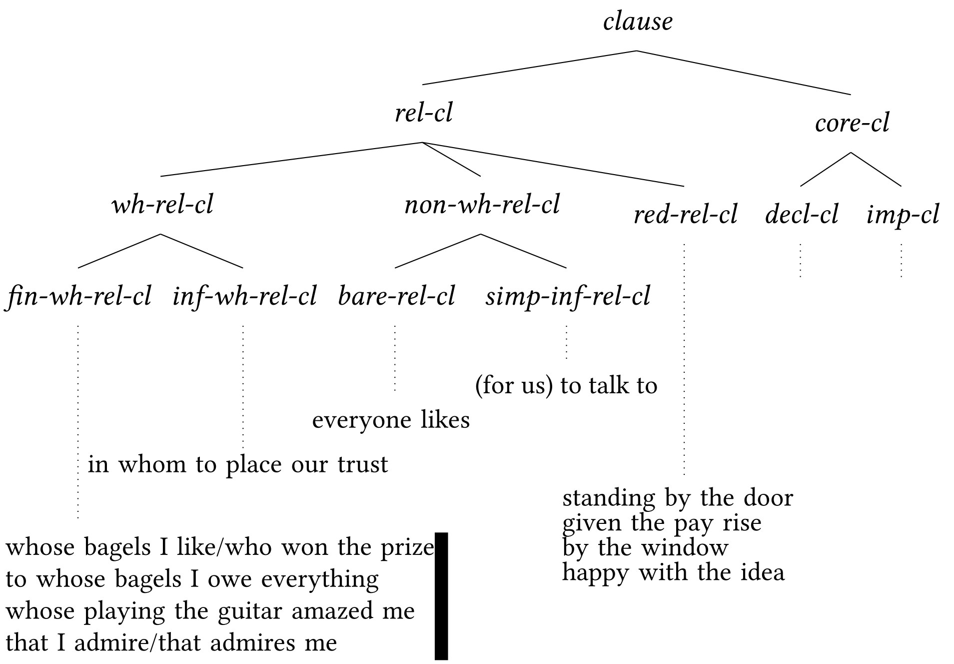


Figure 4: Type hierarch for *clause*, based on Sag (1997).

(22) $$\textit{rel-cl} \Rightarrow \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{mc} & - \\ \textsc{inv} & - \\ \textsc{mod} & \begin{bmatrix} \textsc{head} & \textit{noun} \end{bmatrix} \end{bmatrix} \\ \textsc{cont} & \textit{proposition} \end{bmatrix}$$

Relative clauses such as that in (2) are what Sag calls *fin-wh-rel-cl*, a sub-type of *wh-rel-cl*. This is associated with the constraints in (23). In words: *wh*-relatives are a subtype of relative clause (as stated in the type hierarchy in Figure 4), where the non-head daughter is required to have a REL value which contains the INDEX of the antecedent.[17]

[17]For simplicity and to avoid distractions, we have presented *wh*-relatives as $\overline{\text{N}}$ modifiers in (23). This is a conventional assumption, because standard methods of semantic composition ensure that the content of the relative clause is included in the restrictions of a quantificational determiner (as in *every person to whom Kim spoke*), but it is not Sag's analysis. Instead he takes *wh*-relatives to be NP modifiers, which allows him to account for facts about the ordering of *wh*-relatives and bare relatives, see Sag (1997: 465–469). Kiss (2005: 293–294) gives a number of arguments in favour of this view, for example, the existence of what Link (1984) called "hydras", like (i), where the relative clause must be interpreted as modifying the coordinate structure consisting of the conjoined NPs.

(23) $\textit{wh-rel-cl} \Rightarrow \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{mod} & \overline{\mathrm{N}}\boxed{1} \end{bmatrix} \\ \textsc{non-hd-dtrs} & \left\langle \begin{bmatrix} \textsc{rel} & \{\boxed{1}\} \end{bmatrix} \right\rangle \end{bmatrix}$

The framework assumed in Sag (1997) allows multiple inheritance of constraints from different dimensions (cf. **chapters/ChXXXX** Chapter **??** of this volume). As well as inheriting properties in the clausal dimension, expressions of type *fin-wh-rel-cl* are also classified in the phrasal dimension as belonging to a sub-type of head-filler phrase (*hd-fill-ph*), thus inheriting constraints as in (24).[18]

(24) $\textit{hd-fill-ph} \Rightarrow \begin{bmatrix} \textsc{slash} & \boxed{2} \\ \textsc{hd-dtr} & \begin{bmatrix} \textsc{head} & \textit{verbal} \\ \textsc{slash} & \{\boxed{1}\} \uplus \boxed{2} \end{bmatrix} \\ \textsc{non-hd-dtrs} & \left\langle \begin{bmatrix} \textsc{local} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix}$

In words: they are *verbal* — e.g. clausal — phrases where the SLASH value of the head daughter is the SLASH value of the mother plus the LOCAL value of the non-head daughter (equivalently, the SLASH value of the mother is the SLASH value of the head daughter less the LOCAL value of the non-head daughter). Head-filler phrases are a sub-type of another phrase type (*head-nexus-phrase*) which specifies identity of content between mother and head daughter.

Putting these together with a constraint that requires clauses to have empty REL values will license local trees like that in Figure 5 for a finite relative clause (*fin-wh-rel-cl*) like (2) (simplifying, and disregarding most empty and irrelevant attributes).[19] The REL specification on the non-head daughter (the PP) in (23) ensures the presence of a *wh*-phrase, and the fact that this is a head-filler phrase ensures that the *wh*-phrase cannot be *in situ* (cf. (6), above); the [REL {}] on the daughter S excludes the possibility of additional relative pronouns inside the S (i.e. the possibility of multiple relative pronouns, cf. **(the person) to whom Kim*

(i) The boy$_i$ and the girl$_j$ who$_{i+j}$ dated each other are Kim's friends.

Sag's analysis requires a different approach to semantic composition to that assumed here, e.g. one using Minimal Recursion Semantics (MRS, Copestake et al. 2005) or Lexical Resource Semantics (LRS, Richter & Sailer 2004) — see, in particular Walker (2017), where such an approach is worked out in detail using LRS.

[18] The ⊎ symbol here signifies *disjoint union*. This is like normal set union, except that it is undefined for pairs of sets that share common elements. Here, the intention is that restrictions are to distributed between the noun and the clause, so the restrictions associated with the noun do not include the restrictions associated with the clause, and *vice versa*.

[19] This assumption about REL values is one of many minor technical differences between Sag (1997) and Pollard & Sag (1994), where the non-empty REL value is inherited upwards to RP, and is discharged there. This means that for Pollard & Sag, but not for Sag (1997), a *wh*-relative clause is a REL marked clause.

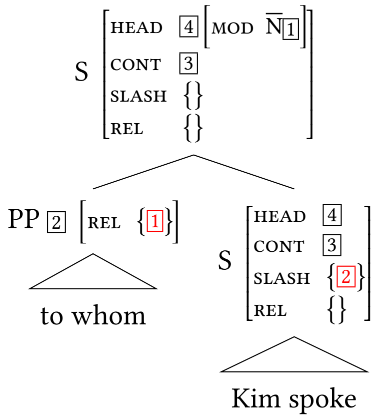


Figure 5: A Sag (1997) style structure for a finite *wh*-relative clause.

spoke about whom). REL inheritance will carry the index of the antecedent down into the PP, guaranteeing the presence of a relative pronoun co-indexed with any nominal that this relative clause is used to modify. Further upwards inheritance of this REL value is prevented by a requirement that all clauses (including relative clauses) have empty REL values.[20] The SLASH specification on the head S daughter will ensure that the LOCAL value of the PP is inherited lower down inside the S, so that the subcategorisation requirements of *speak* can be satisfied, and the right content produced for this S (and passed to the mother S, because this is a head-filler phrase).

The task of combining a nominal and a relative clause (in particular, identifying indices and unioning restrictions) involves a further phrase type *head-relative-phrase*, as in (25).

[20] Sag's account of the propagation of REL values is a special case of the apparatus that is now standardly assumed for propagation of all non-local features, SLASH, WH (i.e. QUE), and BACKGROUND (Ginzburg & Sag 2000: Ch5). Upwards inheritance is handled by a constraint on *words* that says that (by default) the REL value of a word is the union of the REL values of its arguments. In the absence of a lexical head with arguments (e.g. in *of whom* and *of whose friends* if *of* is treated simply as a marker) the REL value on a phrase is that of its head daughter (the "Wh-Inheritance Principle", WHIP), see Sag (1997: 449). Since these are only default principles, they can be overridden, e.g. by the requirement that clauses have empty REL values.

(25) $\textit{head-relative-phrase} \Rightarrow \left[\begin{array}{ll} \text{HEAD} & \textit{noun} \\ \text{CONT} & \left[\begin{array}{ll} \text{INDEX} & \boxed{2} \\ \text{RESTR} & \boxed{3} \uplus \left\{ \left[\begin{array}{ll} \textit{fact} & \\ \text{PROP} & \boxed{4} \end{array}\right] \right\} \end{array}\right] \\ \text{HD-DTR} & \left[\begin{array}{ll} \text{INDEX} & \boxed{2} \\ \text{RESTR} & \boxed{3} \end{array}\right] \\ \text{NON-HD-DTR} & \left[\text{CONT}\ \boxed{4}\right] \end{array}\right]$

In words, this specifies a nominal construction (i.e. one whose head is a noun), whose CONTENT is the same as that of its head daughter, except that the content of the non-head-daughter (the relative clause) has been added to its restriction set. (Thus, it is this construction that takes care of the mismatch between the "internal", propositional, CONTENT of the relative clause itself, and its "external" contribution of restrictions on the nominal it modifies). Since *head-relative-phrase*s are a subtype of *head-adjunct-phrase*, which requires the MOD value of the non-head to be identical to the SYNSEM value of the head (Sag 1997: 475), this will give rise to structures like that in Figure 6.[21]

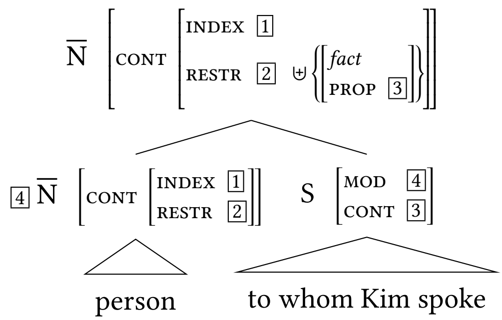


Figure 6: Sag's (1997) analysis of a relative clause plus its antecedent.

From a purely formal point of view the *head-relative-phrase* construction is not strictly necessary. It would be possible to build its semantic effects into the *rel-cl* construction, so that the structure in Figure 6 would be an entirely normal head-adjunct phrase where the content comes from the adjunct daughter. There

[21] This is not the normal semantics associated with head-adjunct phrases (where the content is simply the content of the adjunct daughter). This could be dealt with by introducing a separate sub-type of *head-adjunct-ph* which deals with content in this way: *head-adjunct-ph* itself would impose no constraints on content. Notice that we again follow Ginzburg & Sag (2000) in taking restrictions to be sets of *facts* (Sag 1997 assumes they are sets of *propositions*). Nothing hangs on this.

are two arguments against this. One is that it would require the relative clause to have nominal (i.e. *scope-object*) content, which is somewhat at odds with its status as a clause. The other is that it would push the semantic mismatch into the relative clause itself. That is, semantically, relative clauses like *to whom Kim spoke* would no longer be normal head-filler phrases where CONTENT is shared between head and mother. Perhaps neither argument is compelling — in fact, the discussion of relative clauses in Sag (2010) suggests essentially this approach (Sag 2010: 522).

2.4 Interim Conclusions

The discussion so far has focused on one kind of relative clause, sketched the basic ideas and intuitions behind the HPSG approach, and outlined the two main approaches: that of Pollard & Sag (1994) and that of Sag (1997). At some levels they seem very different (e.g. in the use of phonologically empty lexical heads vs. the use of constructions), and there are differences in terms of low level technical details (e.g. precisely which phrases are specified as having empty REL values, and in the precise way inheritance of SLASH and REL values is terminated). But in other respects they are very similar: for the most part the same features are used in ways that are not radically different.

More significantly, the approaches involve a common view of the relation between relative clause and antecedent: the view that the relative clause is adjoined to the antecedent, with the relation between the antecedent and the relativised constituent within the relative clause being one of co-indexation (a more or less anaphoric relation): a view that can be traced back to Chomsky (1977).

Outside HPSG this style of analysis stands in contrast to two others: the *raising* analysis (see *inter alia* Schachter 1973; Vergnaud 1974; Kayne 1994), and the *matching* analysis (see *inter alia* Chomsky 1965; Lees 1961; Sauerland 1998). Under the raising analysis, the relative clause contains a DP of the form *which*+noun, which is preposed to the beginning of the clause; then the noun is moved out of the relative clause ("raised") to combine with a determiner, which selects both the noun and the relative clause. According to the matching analysis, the relative clause is adjoined to the antecedent, as in the adjunction analysis, but, as in the raising analysis, the relative clause contains a DP *which*+noun, which is preposed to the beginning of the clause; the noun is not raised, but the noun is deleted under identity with the antecedent nominal.

Neither analysis has any appeal from an HPSG perspective: as normally understood, both are fundamentally derivational in nature, presupposing at least two levels of syntactic structure. Moreover, many of the motivations usually cited are absent given standard HPSG assumptions (e.g. arguments from binding theory which taken be taken as indicating the presence of *wh*-phrase inside the relative clause fall out naturally without this assumption given the argument-structure based account of binding theory which is standard in HPSG, see Wechsler, Koenig & Davis (2019), Chapter 9 of this volume). More important, as discussed in Webelhuth et al. (2018), both face numerous empirical difficulties and miss important generalisations which are unproblematic for the style of analysis described here.[22]

3 Varieties of relative clause

In this section we will look at how the approaches introduced above have been adapted and extended to deal with other kinds of relative clause in a variety of languages.[23] Section 3.1 looks at other kinds of relative clauses which involve a relative pronoun, notably ones which do not involve a finite verb. Section 3.2 and Section 3.3 look at relative clauses which do not involve relative pronouns: Section 3.2 looks at relative clauses which can be analysed as involving a complementiser; Section 3.3 looks at "bare" relatives, which involve neither relative pronouns nor complementisers. Section 3.4 looks at non-restrictive relative clauses, which lack the intersective semantics associated with prototypical relative clauses.

One dimension of variation among relative clause constructions which we will discuss only in passing relates to whether, in the case of relative clauses that involve a filler-gap construction, the gap is genuinely absent phonologically (as in the examples we have looked at so far), or whether it is realised as a full pronoun (a so-called *resumptive pronoun*) as in (26) from Alqurashi & Borsley (2012: 28), or the English example in (27) — the resumptive pronouns are indicated in bold.

(26) wajadtu l-kitab-a [llaði tuħib-**hu** Hind-un] (Arabic)
found.1.SG DEF-book-ACC that.M.SG like.3.F.SG-3.M.SG Hind-NOM
'I found the book that Hind likes.'

[22]For example, for example, both analyses treat *wh*-words like *who*, *what* and *which* and their equivalents as determiners, whereas in fact they behave like pronouns. Case assignment appears to pose a fundamental problem for the raising analysis, since it seems to predict that the case properties of the antecedent NP should be assigned "downstairs" inside the relative clause. But they never are. (see Webelhuth et al. 2018).

[23]In addition to the phenomena and languages we discuss, the HPSG literature includes more or less detailed treatments of relative clauses in Bulgarian (Avgustinova 1996), German (Müller 1999a,b; Müller & Machicao y Priemer 2019), Hausa (Crysmann 2016), Polish (Mykowiecka et al. 2003; Bolc 2005), and Turkish (Güngördü 1996).

(27) This is the road which I don't know where **it** goes.

The analysis of resumptive pronouns is discussed elsewhere in this volume (**chapters/ChYYYY** Chapter **??** of this volume), and while they are an important feature of relative clause constructions in many languages (see e.g. Vaillette 2001a;Vaillette 2001b;Taghvaipour 2005; Abeillé & Godard 2007; Alotaibi & Borsley 2013), the issues seem to be similar in all constructions involving unbounded dependencies, and not specific to relative clauses.

3.1 *Wh*-relatives

Finite *wh*-relatives in English have been discussed above (Section 2). English also allows *wh*-relatives which are headed by non-finite verbs, such as (28); (29) is a similar example from French.

(28) a person [on whom to place the blame]

(29) un paon [dans les plumes duquel] mettre le courrier (French)
a peacock in the feathers of.which to.place the mail
'a peacock in whose feathers to place the mail'

Non-finite relatives were not discussed by Pollard & Sag (1994), but Sag's (1997) constructional approach provides a straightforward account. It involves distinguishing two sub-types of *hd-fill-ph*: a finite subtype which has an empty SUBJ list, and a non-finite subtype whose SUBJ list is required to contain just a PRO (that is, a pronominal that is not syntactically expressed as a syntactic daughter). This requirement reflects the fact that non-finite *wh*-relatives do not allow overt subjects:

(30) *a person [on whom (for) Sam to place the blame]

The relative clause in (28) receives a structure like that in Figure 7. Apart from the finite specification, this differs from the finite *wh*-relative in (5) above only in the presence of the PRO on the SUBJ list.[24]

The exclusion of overt subjects is not peculiar to non-finite relatives (it is shared by non-finite interrogatives, cf. *I wonder on whom (*for Sam) to put the*

[24] The use of S_{inf} in Figure 7 is an approximation. First, S is standardly an abbreviation for something of type *verb* with empty SUBJ and COMPS values, and here there is a non-emtpy SUBJ. Second, Sag would have CP instead of S here, reflecting his analysis of *to* as a complementiser rather than an auxiliary verb, as is often assumed in HPSG analyses (e.g. Ginzburg & Sag 2000: 51–52). S and CP are not very different (both *verb* and *comp* are subtypes of *verbal*), but Sag is careful to treat *to* as a *comp* and non-finite *wh*-relatives as CPs because this gives a principled basis for excluding overt subjects.

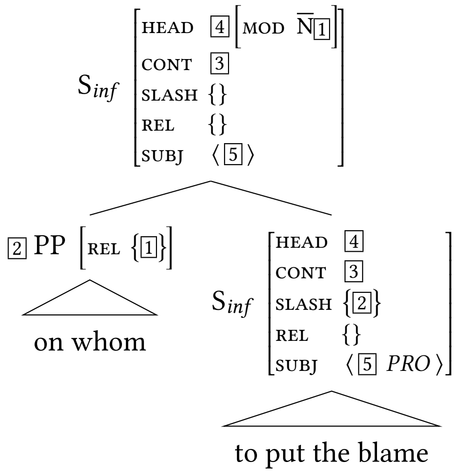


Figure 7: Sag's (1997) analysis of a non-finite *wh*-relative clause (*inf-wh-rel-cl*).

blame), but non-finite *wh*-relatives are subject to the apparently idiosyncratic restriction that the *wh*-phrase must be a PP:

(31) a. * a person who(m) to place the blame on (relative)

b. I wonder who(m) to place the blame on (interrogative)

The relevant constraints can be stated directly — roughly as in (32) (disregarding constraints that are inherited from elsewhere). In words these constraints say that a non-finite head-filler phrase must have an unexpressed subject, and a non-finite *wh*-relative clause is a non-finite head-filler phrase whose non-head daughter is a PP.

(32) a. $\textit{inf-hd-fill-ph} \Rightarrow \left[\textsc{hd-dtr} \begin{bmatrix} \textsc{head} & \left[\textsc{vform} \;\; \textit{non-finite}\right] \\ \textsc{subj} & \langle \textit{PRO} \rangle \end{bmatrix}\right]$

b. $\textit{inf-hd-fill-rel-cl} \Rightarrow \textit{inf-hd-fill-ph} \;\&\; \left[\textsc{non-hd-dtrs} \;\; \langle \textit{PP} \rangle\right]$

3.2 Complementizer relatives

As well as *wh*-relatives, which involve relative pronouns, there are cases of relative clauses which appear to be headed by what is plausibly analysed as a complementiser. In this section we look first at Arabic, where a complementiser analysis has been proposed, then at English, where such an analysis seems possible for some cases, but where it is controversial, and an interesting construction in French.[25]

3.2.1 Arabic

Alqurashi & Borsley (2012) argue that in Arabic finite relatives the word *ʔallaði* ‘that’ (transliterated as *llaði* in (33), from Alqurashi & Borsley 2012: 27) and its inflectional variants should be analysed as a complementiser, with a SYNSEM value roughly as in (34).[26]

(33) jaaʔa l-walad-u llaði qaabala l-malik-a. (Arabic)
came.3.M.SG DEF-boy-NOM that.M.SG met.3.M.SG DEF-king-ACC
‘The boy who met the king came.’

(34)
$$
\begin{bmatrix}
\textit{synsem} & \\
\text{LOC} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix}
\text{HEAD} & \begin{bmatrix} \textit{c} & \\ \text{MOD} & \text{NP}_{\textit{def}}\text{:} \begin{bmatrix}\text{INDEX} & \boxed{1}\\ \text{RESTR} & \boxed{2}\end{bmatrix}\end{bmatrix}\\
\text{COMPS} & \left\langle \begin{bmatrix}\text{LOC} & \begin{bmatrix}\text{CAT} & \text{S}_{\textit{fin}}\\ \text{CONT} & \boxed{3}\end{bmatrix}\\ \text{NON-LOC} & \begin{bmatrix}\text{SLASH} & \{\text{NP}_{\boxed{1}}\}\end{bmatrix}\end{bmatrix}\right\rangle
\end{bmatrix}\\
\text{CONT} & \begin{bmatrix}\text{INDEX} & \boxed{1}\\ \text{RESTR} & \boxed{2} \uplus \{\boxed{3}\}\end{bmatrix}
\end{bmatrix}\\
\text{NON-LOC} & \begin{bmatrix}\text{SLASH} & \{\}\end{bmatrix}
\end{bmatrix}
$$

According to this, *ʔallaði* will combine with a slashed finite sentential complement, to produce a phrase which will modify a definite NP. When it combines with that NP, its content will have the same INDEX as the NP, and the restrictions

[25] There are also cases which involve a relative pronoun *and* a complementiser, as in the following from Hinrichs & Nakazawa’s (2002) discussion of Bavarian German:

(i) der Mantl (den) wo i kaffd hob (Bavarian German)
the coat which that I bought have
‘the coat which I bought’

Hinrichs & Nakazawa (2002) analyse these as *wh*-relatives, even when the relative pronoun is omitted, as it can be under certain circumstances.

[26] Here $\text{S}_{\textit{fin}}$ means a finite clause (a *verb* which is COMPS and SUBJ saturated) $\text{NP}_{\textit{def}}$ in the MOD means a fully saturated definite nominal whose CONTENT is given after the colon. According to (34) the content of the $\text{S}_{\textit{fin}}$ is merged with the restrictions of this modified NP. This is imprecise: as discussed above, what should be merged is a *fact* constructed from the content of the $\text{S}_{\textit{fin}}$.

of the NP combined with the propositional content of the sentential complement. The SLASH value on the sentential complement means that it will contain a gap (or a resumptive pronoun) which also bears the same index.

Notice that there no role for a REL feature here (obviously, since there is no relative pronoun). The presence of the SLASH value indicates that Alqurashi & Borsley assume that Arabic relatives involve an unbounded dependency (i.e. that the gap or resumptive pronoun may be embedded arbitrarily deeply within the relative clause). In *wh*-relatives, as described above, the unbounded dependency is what Pollard & Sag (1994: 155) call a "strong" unbounded dependency, i.e. one that is terminated by at the top by a filler (the *wh*-phrase), in a head-filler phrase. This is not the case here — here there is no filler, and upward inheritance of the gap is halted by the head *ʔallaði* itself (cf. its own empty SLASH specification). That is, Arabic relatives (and complementiser relatives generally) are normal head-complement structures, involving what Pollard & Sag (*loc cit*) call a "weak" unbounded dependency construction (like English purpose clauses and *tough*-constructions).[27]

Since *ʔallaði* shows inflections agreeing with the antecedent NP for NUMBER, GENDER, and CASE, different forms will impose additional restrictions on the modified NP (e.g. the form transliterated as *llaði* in (33) will add to (34) the additional requirement that the NP which is modified must be masculine singular).

Notice that Alqurashi & Borsley's account is entirely lexical: no constructional apparatus is used at all. Hahn (2012) argues for a constructional alternative.[28]

3.2.2 English

A similar analysis could be proposed for English *that*-relatives as in (35). However, this is controversial: Pollard & Sag (1994) treat some uses of *that* as simply a marker (i.e. the realisation of a MARKING feature whose value is *that*, as opposed to *unmarked*), and others as a relative pronoun, see Pollard & Sag (1994: 221–222). Sag (1997: 462–464) prefers to treat *that* as a relative pronoun.[29]

[27] Alqurashi & Borsley (2012: 42) assume that the SLASH inheritance is governed by a default principle, so the empty SLASH specification on *ʔallaði* prevents upwards inheritance. The same effect could be achieved with an appropriate TO-BIND specification.

[28] Arabic also has finite relatives that do not have an overt relativiser (and which occur with indefinite antecedents). Alqurashi & Borsley analyse these as involving a phonetically null complementiser. In addition, Arabic also has non-finite and free relatives, which have received some attention. See Melnik (2006); Haddar et al. (2009); Zalila & Haddar (2011); Hahn (2012); Crysmann & Reintges (2014) for further discussion.

[29] Pollard & Sag (1994) treat instances of *that* in relative clauses involving relativisation of a top level subject, like (35a), as a relative pronoun. In other relative clauses, in particular those

(35) a. person that ___ admires Kim
b. person that everyone thinks ___ admires Kim

On Pollard & Sag's (1994) analysis, some support for a relative pronoun analysis comes from coordination. It is possible to coordinate *that* relatives with normal *wh*-relatives quite freely, as in (36). This is a natural consequence if the REL value of the coordinate structure is shared by both conjuncts (implying that both conjuncts contain relative pronouns, of course).[30]

(36) a book [that/which you own or that/which you can borrow]

Potential evidence against this, and in favour of a complementiser (or perhaps marker) style analysis would be that *that* differs from normal relative pronouns in not allowing pied-piping, cf. (37b).

(37) a. the person that I spoke to ___
b. * the person to that I spoke ___

Sag (1997: 464) and Pollard & Sag (1994: 220) argue that this restriction is compatible with a relative pronoun analysis on the assumption that *that* has nominative case, so that it cannot occur as e.g. the complement of a preposition. Notice also that *who* (which is generally regarded as a relative pronoun) follows the same pattern:

(38) a. the person who I spoke to ___
b. * the person to who I spoke ___

However, this response is not very convincing. What (37) and (38) show is that *that* and *who* cannot appear as complement of a preposition, but can be associated with a gap that is complement of a preposition. But this is inconsistent with them being fillers in a head-filler phrase, where SLASH inheritance ensures

involving relativisation of embedded subjects, like (35b), or non-subjects, *that* is treated as a marker, meaning that such clauses are treated as instances of bare relatives. It is hard to find clear empirical evidence against this, but an analysis which provides a uniform treatment of English *that*-relatives is clearly more appealing.

[30] The same argument can be made given Sag's (1997) assumptions, but it is less direct. Recall that, on Sag's (1997) analysis, relative clauses have empty REL values, so a coordination of relative clauses will have an empty REL value too (cf. above Section 2.3, especially Footnote 19). For Sag (1997) the argument relies on the assumption that all and only *wh*-relatives are NP modifiers, rather than $\overline{\mathrm{N}}$ modifiers as we have presented them here (cf. Footnote 17). Since coordination involves identity of MOD values, data like (36) show that *that*-relatives must be NP modifiers, and consequently must be *wh*-relatives, i.e. must contain a relative pronoun (namely, *that*).XXXXX

identity between the LOCAL values of filler and gap (including, of course CASE): if *that* and *who* are nominative, then they should not be compatible with non-nominative gaps, such as we see in (37a) and (38a). But if they are not fillers, then they must be heads (or markers). Developing an analysis along these lines is beyond the scope of this paper, but it is worth pointing out that it would not involve a radical change to the analyses described above (for example, modifying Sag's (1997) analysis might involve creating a new subtype of *rel-cl* for *that* and *who* relatives, separate from *wh-rel-cl*, and new lexical entries for *that* and *who*, but could otherwise use the same apparatus, and produce the same distribution of properties).

3.2.3 French

Besides *wh*-relatives, French has relatives introduced by complementisers: *que* 'that' and *dont* 'of-which'. *Dont*-relatives present something of a challenge, which is addressed in Abeillé & Godard (2007). *Dont* is generally analysed as a complementiser introducing finite relatives (Godard 1992). It can introduce a relative with a PP_{de} gap (i.e. a gap that could be occupied by a PP marked with the preposition *de* 'of'). The contrast between the grammatical (39a) and the ungrammatical (39b) arises because whereas *parler* 'talk' in (39a) takes a PP_{de} complement, *comprendre* 'understand' in (39b) takes an NP complement, and so cannot cannot contain a gap licensed by *dont*, as can be seen in (40a) and (40b).

(39) a. un problème dont on a parlé (French)
a problem of-which one has talked
'a problem that we have talked about'

b. * un problème dont on résoudra (French)
a problem of-which one will.resolve
'intended: a problem that we will resolve'

(40) a. On a parlé d' un problème. (French)
One has talked of a problem
'We have talked about a problem.'

b. * On résoudra d' un problème. (French)
One will.resolve of a problem

Abeillé & Godard suggest a lexical entry for *dont* with a SYNSEM value along the lines of (41).

(41)
$$\begin{bmatrix} \textit{synsem} \\ \textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} c \\ \textsc{mod} & \overline{\mathrm{N}}\boxed{1} \end{bmatrix} \\ \textsc{comps} & \left\langle \begin{bmatrix} \textsc{loc} & \begin{bmatrix}\textsc{cat} & \mathrm{S}\textit{fin} \\ \textsc{cont} & \boxed{2}\end{bmatrix} \\ \textsc{non-loc} & \begin{bmatrix} \textsc{slash} & \left\{ \boxed{3} \begin{bmatrix} \textit{nprl} \\ \textsc{cat} & \mathrm{PP}_{\textit{de}}\boxed{1} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{cont} & \boxed{2} \end{bmatrix} \\ \textsc{to-bind} & \{\boxed{3}\} \end{bmatrix}$$

In words: *dont* is a complementiser that takes a finite S complement, and heads a phrase that can act as an $\overline{\mathrm{N}}$ modifier. *Dont* itself has no inherent semantic content (its CONTENT is just that of its complement S). The complement S is associated with a SLASH value that contains a PP_{de} which is co-indexed with the antecedent nominal, as specified in the MOD value. The TO-BIND value simply prevents this SLASH element being inherited upwards beyond the phrase headed by *dont*. This SLASH element is non-pronominal (*nprl*) — that is, a genuine gap, rather than a resumptive pronoun.[31]

Given this, one might expect that it is generally impossible for a *dont*-relative to have an NP as the relativised constituent, but this is not the case. It is in fact possible providing the relativised constituent is realised by an overt pronoun (i.e. a resumptive pronoun) and is somewhere inside the complement of (some) propositional attitude and communication predicates. For example, in (42) the pronoun *le* represents the relativised constituent, which appears in the complement of *être certain* 'be sure'.[32]

[31] Abeillé & Godard (2007) assume that gaps and resumptive pronouns are associated with distinct subtypes of *local* value: *plr* (pronominal) for pronouns and *nprl* (non-pronominal) for genuine gaps. The relevance of this will appear directly.

[32] One might consider an alternative analysis where *dont* is associated with a PP_{de} gap dependent of *certain*, and the resumptive pronoun is a normal anaphoric pronoun — this would correspond to a main clause along the lines of *Paul is sure, of this problem, that we will resolve it*. One problem with this alternative is that this sort of PP_{de} dependent is not very good with *certain*, see (i). Another is that it would not explain the fact that the personal pronoun is obligatory — (ii), with no personal pronoun, is ungrammatical, though semantically coherent:

(i) ?? Paul est certain de ce problème qu' on le résoudra. (French)
Paul is sure of this problem that one it will.solve.

(ii) * un problème dont [Paul est certain que tout va se résoudre]
a problem of-which Paul is sure that everything goes itself to.solve
(French)

(42) un problème dont [Paul est certain [qu' on **le** résoudra]] (French)
a problem of-which Paul is sure that one it will.solve
'a problem that Paul is sure that we will solve'

Unsurprisingly, the presence of a resumptive pronoun is associated with immunity to island constraints. So, for example, in (43) we have a relative where the relativised constituent is within a relative clause inside an embedded NP, which is impossible for a genuine gap.

(43) un problème dont [Paul est certain [qu' il y a [quelqu'un qui **le** résoudra]]] (French)
a problem of-which Paul is sure that there is someone that it will.solve
'a problem such that Paul is sure that there is someone who will solve it'

What is surprising, however, is that the path between *dont* and the predicate that licenses the resumptive *is* sensitive to island constraints. To see this, compare the grammatical (42) and (43) with the ungrammatical (44). All involve a *dont* relative containing a resumptive pronoun licensed by *être certain*, but in (44) *être certain* is separated from *dont* by an island boundary (*être certain* is inside a relative clause).

(44) * un problème dont il y a [quelqu'un qui est certain qu' on le résoudra] (French)
a problem of-which there is someone who is sure that one it will.solve

In short, though the dependency between the licensing predicate and the resumptive pronoun can cross island boundaries, the dependency between the licensing predicate and *dont* cannot. Abeillé & Godard's (2007) account of this is that while the dependency between the licensing predicate and the relativised constituent involves inheritance of a resumptive element, that between the licensing predicate and *dont* involves inheritance of a gap. They suggest that this should be dealt with by a lexical rule along the lines of (45):

(45) Lexical Rule for Propositional attitude predicates in French

$$\left[\textsc{comps}\left\langle \overset{\text{CP}}{\left[\textsc{slash}\left\{\boxed{1}\begin{bmatrix}\textit{prl}\\ \textsc{cont|index}\ \boxed{2}\end{bmatrix}\right\}\right]} \right\rangle \oplus \ldots\right] \mapsto \begin{bmatrix}\textsc{slash} & \left\{\begin{bmatrix}\textit{nprl}\\ \textsc{cat}\ \ \text{PP}_{de\boxed{2}}\end{bmatrix}\right\}\\ \textsc{to-bind} & \{\boxed{1}\}\end{bmatrix}$$

In words, the left-hand side of this describes a lexeme that takes a CP complement with a SLASH value containing pronominal (*plr*) elements (that is, a CP that can contain resumptive pronouns). The effect of the rule is to provide a lexical entry that binds-off the resumptive pronoun, and introduces an PP_{de} gap co-indexed to the resumptive pronoun. That is, the sort of gap that can legitimately be associated with *dont*. Thinking from the top down, this rule produces a predicate that can appear in a context with an inherited requirement for a PP_{de} gap (e.g. a relative clause headed by *dont*), and convert this into a requirement for a resumptive pronoun further down. Thinking from the bottom up, the predicate can bind-off a resumptive pronoun, and replace it with a gap dependency.[33]

3.3 Bare relatives

Not all languages realise relative clauses using relative pronouns or complementisers. In this section we will discuss HPSG analyses of what we will call *bare relatives* in Japanese and Korean (Section 3.3.1) and in English, where they are often called "*that*-less" relatives (Section 3.3.2). The absence of relative pronouns means there is no question of pied-piping, hence no role for a REL feature in these constructions.

3.3.1 Bare relatives in Japanese and Korean

Japanese relative clauses corresponding to (2) contain a gap, but are otherwise similar to normal clauses, cf. (46) (from Sirai & Gunji 1998: 18); in Korean they are distinguished by special marking on the topmost verb — cf. the *-nun* affix on *sayngkakha* "think" in (47) (from Kim 2016b: 285).

[33] As Abeillé & Godard (2007) point out, the facts are not quite as simple as this. In particular there is an interesting complication involving coordination. It is possible for a *dont*-clause containing a predicate like *être certain* to involve a coordinate structure, where one conjunct contains a PP_{de} gap and the other contains a pronoun, as in (i) (the second conjunct here contains the pronominal *y* 'to-it'; the English translation is intended to make it clear that the second conjunct is in the scope of *être certain*).

(i) un problème dont Paul est certain [que nous avons parlé ___] [et que
a problem of-which Paul is sure that we have spoken and that
nous y reviendrons plus tard] (French)
we to-it will.come.back more late
Lit: 'a problem of which Paul is sure that we have spoken and that he is sure that we will come back to it later'

Dealing with this involves a formal complication that we leave aside here. See Abeillé & Godard (2007).

(46) Naomi-ga ___$_i$ yon-da hon$_i$ (Japanese)
Naomi-NOM read-PAST book
'the book (that) Naomi read'

(47) [motwu-ka [Kim-i ___$_i$ ilk-ess-ta-ko] sayngkakha-nun] chayk$_i$ (Korean)
everyone-NOM Kim-NOM read-PST-DECL-COMP think-PRES.MOD book
'the book (that) everyone thinks Kim read'

Evidence for a gap in these examples is that it is not possible to put an overt NP in place of the gap (e.g. putting *sore-wo* 'it-ACC' in (46), or *sosel-u* 'novel-ACC' in (47) renders them ungrammatical).[34]

Sirai & Gunji (1998) provide a non-constructional account of Japanese bare relatives like (46). They show how an account that uses SLASH inheritance could work, but their actual proposal is SLASH-less. They assume that the tense affixes are heads of verbal predications, and operate via "predicate composition" — by inheriting the subcategorisation requirements of the associated verb. The adnominal tense affixes are special in that a) they are specified as nominal modifiers, and b) they inherit the subcategorisation requirements of the associated verb, less an NP that is co-indexed with the modified nominal. (A lexical equivalent of this could be implemented with a lexical rule which removes an element from a verb's ARG-ST and introduces a MOD value containing a nominal with the corresponding index). Of course, a SLASH-less account like this will only deal with cases of local relativization — where the relativised NP is an argument of the highest verb. Sirai & Gunji argue that cases of non-local relativization, like (48), should be treated as involving null-pronominals (which are a common feature of Japanese). They suggest that the requirement that the modified noun and the pronoun be co-indexed should be captured via a pragmatic condition that requires the relative clause be "about" the modified noun.

(48) [Ken-ga [Eiko-ga ___$_i$ yon-da] to sinzitei-ru] hon$_i$ (Japanese)
Ken-NOM Eiko-NOM read.PAST COMP believe-PRES book
'the book that Ken believes Eiko read'

Kim (2016b) provides a constructional analysis for Korean which resembles Sag's (1997) analysis of English — see also Kim (1998a); Kim & Yang (2003). He

[34] As well as these "standard" relatives, Korean and Japanese both have other kinds of relative construction, notably what are sometimes called *internally headed* relatives, and so-called *pseudo-relatives*, which are briefly discussed below. See Section 4.2.2.

suggests that Korean allows verb lexemes to be realised as "modifier verbs" (*v-mod*) subject to a constraint along the lines of (49) — these are verbs that can head a subordinate clause ([MC −]) which modifies a nominal (N).[35]

(49) $\left[\text{HEAD}\ \begin{bmatrix} \textit{verb} & \\ \text{MC} & - \\ \text{MOD} & \textit{noun} \end{bmatrix}\right]$

He also proposes a construction (the *head-relative-mod* construction, see Kim 2016b: 290) to combine a structure headed by such a modifier verb with a head nominal, along the lines of (50).[36]

(50) $\textit{hd-relative-mod-phrase} \Rightarrow \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SLASH} & \{\} \\ \text{HD-DTR} & \boxed{2}\ \text{N}\boxed{1} \\ \text{NON-HD-DTRS} & \left\langle \text{S}\begin{bmatrix} \text{HEAD}|\text{MOD} & \boxed{2} \\ \text{SLASH} & \{\text{NP}\boxed{1}\} \end{bmatrix} \right\rangle \end{bmatrix}$

In words: a nominal structure can consist of a head noun, and a clause headed by a modifier verb containing an NP gap which is co-indexed with the head noun. The empty SLASH value on the mother is necessary to prevent the gap being inherited upwards. The SLASH value on the S daughter ensures the presence of an appropriate gap, the MOD value on the S daughter ensures that it is headed by a verb with the right morphology. It will license structures like that in Figure 8. Kim does not discuss the semantics, but it would be straightforward to add constraints to this construction along the lines of those presented above.

3.3.2 Bare relatives in English

English also has bare relative clauses, both finite, as in (51a), and non-finite as in (51b):

(51) a. the cakes Kim bought ___
b. some cakes (for Sam) to eat ___

In English, there is no obvious motivation for suggesting a special sub-type of "relative clause heading" verb, so an alternative way of licensing noun-modifying

[35] Different sub-types of *v-mod* are associated with different tense affixes. (49) differs from Kim's formulation, e.g. Kim's formulation involves a POS (part-of-speech) feature and he assumes that MOD is list valued (see Kim 2016b: 285). This is not important here.

[36] Again, our formulation is slightly different from Kim's for the sake of consistency with the rest of our presentation.

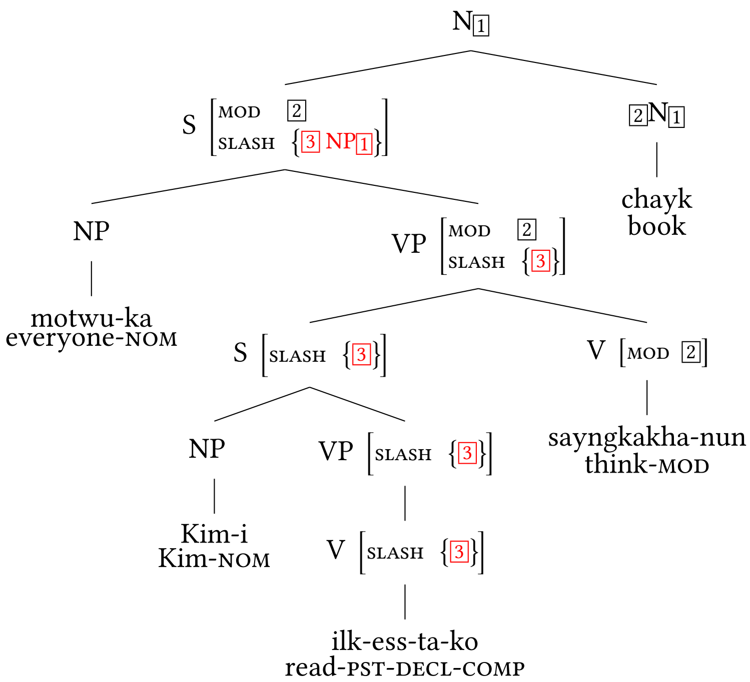


Figure 8: A Korean relative clause, based on Kim (2016b: 295).

clauses with appropriate SLASH values is required. In Pollard & Sag (1994) this was the role of an empty relativiser similar to that described above, differing only in taking a single argument — a slashed clause (see Pollard & Sag 1994: 222; recall that the relativiser discussed above takes two arguments: a *wh*-phrase, and a slashed clause). This gives structures like that in Figure 9.[37]

In Sag (1997) the task of licensing such bare relatives is carried out by a construction (an immediate subtype of *rel-cl*) as in (52). In words: a relative clause can be a noun-modifying clause whose head daughter contains an NP gap that is co-indexed with the modified nominal.

[37] According to Pollard & Sag (1994: 222), the clausal argument of this single argument version of R can either be bare, as here, or marked by *that*. Thus, terminological accuracy demands the observation that for Pollard & Sag some instances of *that*-relatives are actually “bare” in the sense of containing neither a relative pronoun nor a complementiser (though others, in particular those involving relativisation of a top level subject, are analysed as containing a version of *that* which is actually a relative pronoun). See above Footnote 29.

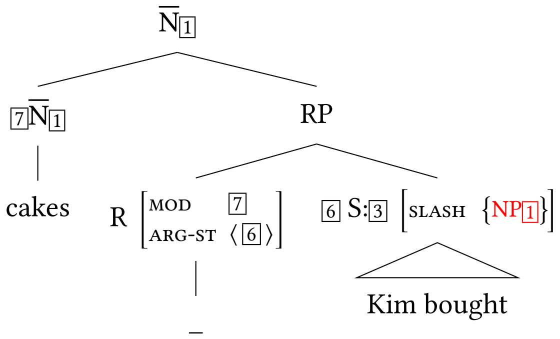


Figure 9: A Pollard & Sag (1994) style structure for an English bare relative.

(52) $$\textit{non-wh-rel-cl} \Rightarrow \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \begin{bmatrix} \text{HEAD} & \overline{\text{N}}\boxed{1} \end{bmatrix} \end{bmatrix} \\ \text{SLASH} & \{\} \\ \text{HD-DTR} & \begin{bmatrix} \text{SLASH} & \{\text{NP}\boxed{1}\} \end{bmatrix} \end{bmatrix}$$

This licenses structures like that in Figure 10.[38]

This differs from Kim's proposal for Korean in where the SLASH value is bound-off: in particular, where Kim's analysis involves a nominal and a slashed S, Sag's involves a nominal and an *un*slashed S — the clause is [SLASH {}], it is the VP which is [SLASH {NP}]. This reflects the fact that in English the gap in the relative clause cannot be the subject, accounting for the contrast in (53).[39]

(53) a. * person spoke to Sam
b. person who spoke to Sam

[38] Sag also proposes a subtype of (52) to deal with non-finite bare relatives, like (i), which he calls *simple infinitival relatives*, cf. *simp-inf-rel-cl* in Figure 4. See Sag (1997: 469). Abeillé et al. (1997) includes discussion of a similar construction in French — 'infinitival *à*-relatives', like (ii):

(i) book (for Sam) to read

(ii) un livre à lire (French)
a book to read

Neither discussion addresses the special modal semantics associated with non-finites, e.g. (i) means something like "books that Sam can (or should) read".

[39] Examples like (53a) are acceptable in some non-standard dialects of English. Sag suggests this is not problematic, since they could be analysed as reduced relatives (see Sag 1997: 471), but see immediately below where we cast doubt on this. If we are right, then the non-standard dialects would have something like (50) instead of (52).

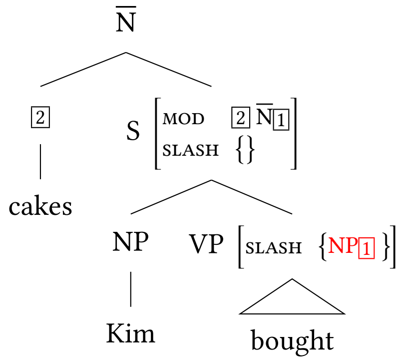


Figure 10: A Sag (1997) style structure for an English bare relative.

The issue of where upwards termination of SLASH inheritance should occur highlights the impossibility of having an entirely lexical and non-constructional account of bare relatives that does not employ empty elements. At first glance, a purely lexical approach might seem simple: since all we need is to create clauses specified as [MOD $\overline{\text{N}}$] which contain a co-indexed gap, all we seem to need is verbs specified as in (54).

(54) $\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} \\ \text{MOD} \ \ \overline{\text{N}}\boxed{1} \end{bmatrix} \\ \text{COMPS} & \langle \text{PP} \rangle \\ \text{SLASH} & \{ \ \text{NP}\boxed{1} \ \} \end{bmatrix}$

In the absence of special constructions or empty elements, this would license structures like that in Figure 10, except that the upwards inheritance of the SLASH value will not be terminated, allowing an additional spurious filler for the gap, as in (55):[40]

(55) *That book$_i$, I enjoyed [the book$_i$ Kim read ___ $_i$]

There is one class of exceptions to this — that is, phrases which might be analysed as relative clauses for which a purely lexical account *is* possible. Examples involving participal phrases and a variety of other post-nominal modifiers, notably APs and PPs, are often called *reduced relatives*, and analysed as a type of

[40]The SLASH based analysis of Japanese relatives outlined in Sirai & Gunji (1998) manages to avoid this problem, without either special constructions or empty elements, but it is not fully lexical, because it assumes tense affixes combine with the associated lexical verb in the syntax (hence the affix is able to block higher inheritance of the gap introduced by the lexical verb).

relative clause. Sag (1997: 471) follows this tradition (*red-rel-cl* in Figure 4). What this comes down to is the assumption that such examples involve clauses containing predicative phrases with PRO subjects, co-indexed with the nominals they modify.

(56) a. a person standing by the door (VP-*pres-part*)
b. a train recently arrived at platform four (VP-*past-part*)
c. a person given a pay rise (VP-*passive-part*)
d. a person in the doorway (PP)
e. a person fond of children (AP)

It is not obvious to us what is gained by treating these as relative clauses introduced by a special construction. A lexical account seems at least as appealing, where the relevant properties of the phrases (e.g. noun modifying semantics) are projected directly from lexical entries for the head words. The reason such a non-constructional approach is possible is that such examples involve neither relative pronouns nor genuine gaps, so there are neither REL nor SLASH dependencies to terminate.[41] This approach seems particularly appealing in the cases like (56e), which would be analysed as just involving an attributive adjective (*fond*) which happens to take a complement, along the lines of (57), where {...} stands for the restrictions the adjective itself imposes. But we think a similar account of verbal participles and prepositions is equally plausible.[42]

(57)
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{MOD} & \begin{bmatrix} \textit{noun} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{RESTR} & \boxed{2} \uplus \{\ldots\} \end{bmatrix} \end{bmatrix}$$

Notice that in (57) we omit mention of the SUBJ. If we assume the noun-modifying entry is derived from a predicative entry, there are two obvious alternatives: a) that the predicative subject is suppressed; or b) that it is constrained to be unexpressed (i.e. PRO). In the latter case, the two approaches are very similar, the only difference being whether examples like those in (56) are classified as clausal. It is not clear whether this has empirical consequences.

[41] This argument does not necessarily carry over to languages which allow relativisation of non-subjects in reduced relatives, such as Arabic. See Melnik (2006: 241).

[42] For example, Müller (2002: 159–164) deals with adjectival passive participles in this way.

3.4 Non-restrictive (supplemental) relatives

The examples of relative clauses considered so far have been *restrictive relatives* (RRCs); they are interpreted as restricting the denotation of their antecedent to a subset of what it would be without the relative clause. So-called *supplemental*, *supplementary*, *appositive*, or *non-restrictive* relatives (NRCs) are different. They do not affect the interpretation of any associated nominal, and are generally interpreted with wide scope, much like independent utterances. For example, if *who understand logic* is read as an NRC as in (58a) it will be interpreted outside the scope of *Kim thinks*.

(58) a. Kim thinks linguists, who understand logic, are clever. (NRC)

b. Kim thinks linguists who understand logic are clever. (RRC)

NRCs are often set off intonationally, and are subject to a number of surface morphosyntactic restrictions in English. In particular, they must be finite and contain a *wh*-pronoun, witness the ungrammaticality of (59a) and (59b).[43]

(59) a. *Kim, for Sandy to speak to, will arrive later.

b. *Kim, (that) Sandy spoke to, will arrive later.

The analysis of non-restrictives has attracted some attention in the HPSG literature.[44]

Where RRCs are typically nominal modifiers, NRCs are compatible with a wide range of antecedents. Holler (2003) provides an analysis of German non-restrictives which are adjoined to S, as in (60). Her account uses a version of the empty relativiser from Pollard & Sag (1994) whose MOD value specifies a clausal (rather than nominal) target for modification, and looks for an appropriate antecedent for its first argument (the *wh*-phrase) among the discourse referents

---

[43] More extensive discussion of differences between NRCs and RRCs can be found in Arnold (2007).

[44] Bîlbîie & Laurens (2009) discuss what they call *verbless relative adjuncts*, such as (i), in French and Romanian:

(i) Trois personnes, [parmi lesquelles Jean], sont venues. (French)
three people(FEM) among which.FEM John AUX come
'Three people, among which John, have come.'

These have non-restrictive semantics, and some similarities with relative clauses, but Bîlbîie & Laurens point out significant differences, and argue for an analysis that treats them rather differently, as a distinct construction.

contributed by the modification target (for example, the discourse referent corresponding to the proposition expressed by the main clause in (60)). The relative pronoun is thus treated rather like a normal pronoun.

(60) Anna gewann die Schachpartie, was Peter ärgerte. (German)
Anna won the game of chess which Peter annoyed
'Anna won the game of chess, which annoyed Peter.'

Arnold (2004) provides an analysis of English non-restrictives of all kinds. This analysis also takes the relative pronouns involved in NRCs to be much like normal pronouns, but accounts for the syntactic restrictions by making minor modifications to constructions given in Sag's (1997) analysis of restrictives. It assumes a uniform syntax for restrictives and NRCs, but provides a way for relative clauses to combine with their heads in two semantically distinct ways, either restrictively (in the normal way) or non-restrictively (making their semantic contribution at the same level as the root clause, accounting for the wide-scope interpretation). The fact that supplementary relatives are required to be finite and contain a *wh*-pronoun can then be simply stated (e.g. non-restrictive semantics entails a non-head daughter which is a *fin-wh-rel-cl*).[45] Likewise, the wider range of antecedents available to NRCs can be captured by relaxing the [MOD *noun*] constraint associated with *rel-cl* (so in principle all kinds of relative clause are compatible with any antecedent), and adding it as requirement associated with restrictive semantics.

The approach to NRCs developed in Arnold (2004) is *syntactically integrated* – NRCs are treated as normal parts of the syntactic structure on a par with restrictive relatives. On the face of it, examples like (61b) are problematic for such an approach:

(61) a. What did Jo think?
b. You should say nothing, which is regrettable.

When uttered in the context provided by (61a), the interpretation of (61b) is that it is regrettable that *Jo thinks* you should say nothing. This has been taken as an indication that the interpretation of NRCs requires antecedents that are not syntactically realised and only available at a level of conceptual structure (see Blakemore 2006). However, Arnold & Borsley (2008) show that this is incorrect,

[45] As stated, given Sag's (1997) assumption that *that*-relatives are a variety of *wh*-relative, this wrongly predicts that supplemental *that*-relatives should normally be allowed. One way round this is to adopt an different analysis of *that*, but Arnold (2004) also considers an analysis whereby *that* has a different kind of REL value from "real" relative pronouns.

and in fact a syntactically integrated account combined with the approach to ellipsis and fragmentary utterances of Ginzburg & Sag (2000) makes precisely the right predictions in this case and in a range of others.

Arnold & Borsley (2010) look at NRCs where the antecedent is a VP, and where the gap is the complement of an auxiliary, as in (62).

(62) Kim has ridden a camel, which Sam never would ___.

This is unexpected, because such examples seem to involve an NP filler (*which*) being associated with a gap in a position where an NP is generally impossible, cf. **Sam never would that activity*. Arnold & Borsley consider a number of analyses, including an analysis which treats *which* as a potential VP, and an analysis which introduces a special relative clause construction. However, they argue that the best analysis is one which relates examples like (62) to cases of VP ellipsis (as in *Kim has ridden a camel but Sam never would*), which involve the VP argument of an auxiliary verb being omitted from its COMPS list. The idea is that auxiliary verbs allow such an elided VP argument to have (optionally) a SLASH value that contains an appropriately co-indexed NP. If such a SLASH value is present, normal SLASH amalgamation and inheritance will yield (62) as a normal relative clause, without further stipulation.

NRCs normally follow their antecedents. However, as Lee-Goldman (2012) observes, there are some special cases where the NRC precedes the antecedent. Such cases involve the relative pronouns *which* and *what* with antecedents that have clausal interpretations, i.e. either actual clauses, as in (63a) and (63b), or other expressions interpreted elliptically as with *later* in (63c).

(63) a. It may happen now, or – *which would be worse* – it may happen later.
b. It may happen now, or – *which would be worse* – later.
c. It may happen now. *What is worse*, it may happen later.

Lee-Goldman provides a constructional account. It makes use of a feature RELZR, introduced by Sag (2010), which is shared between a relative clause and its filler daughter, and whose value reflects the identity of the relative pronoun (so possible values include *which*, *what*, etc.). Cases like (63c) are dealt with simply by means of a special construction which combines a *what*-relative clause with its antecedent in the desired order. The account of cases like (63a) and (63c) makes use of the idea of constituent order domains for linearisation originally proposed by Reape (e.g. Reape (1994), and Müller (2019a), Chapter 10 of this volume). The relevant construction combines a phrase whose RELZR value is *which* (e.g. *which would be worse*) with a clause whose constituent order DOMAIN has a coordinator as its first element (e.g. the DOMAIN associated with *or it may happen later*) and produces a phrase where the DOMAIN value of the *which* phrase appears after the coordinator and before the remainder of the clause, giving the desired result.[46]

4 Other functions, other issues

For reasons of space, we have so far restricted the notion *relative clause* to the typical case: clauses which are nominal modifiers, adjoined to nominals. This ignores a number of relevant phenomena, notably the fact that relative clauses are not necessarily nominal modifiers, and the possibility that even when they function as nominal modifiers they need not be adjoined to nominals. In this section we will provide some discussion of these issues. Section 4.1 will briefly review HPSG analyses of cases where relative clauses are not adjoined to nominals. Section 4.2 will overview HPSG approaches on cases where clauses resembling relative clauses are not nominal modifiers.[47]

4.1 Extraposition

As noted above, relative clauses are typically nominal modifiers, and typically adjoined to the nominals they modify. However, this is not invariably the case: under certain circumstances relative clauses can be *extraposed*, as in (64), where the relative clauses (emphasised) have been extraposed from the subject NP to the end of the clause.

(64) a. Someone might win *who does not deserve it.*
b. Something happened then *(that) I can't really talk about here.*
c. Something may arise *for us to talk about.*

[46] Lee-Goldman handles the wide scope interpretation of NRCs by implementing a multidimensional notion of CONTENT inspired by Potts (2005). He also extends the analysis described here to deal with cases of *as*-parentheticals (e.g. *As most of you are aware, we have been under severe stress lately*), arguing that *as* should be analysed as a relativiser, and that such clauses should be analysed as relative clauses.

[47] Among the other phenomena we have neglected, one should mention *amount* relatives (e.g. Grosu & Landman 2017), that is, relative clauses where what is modified semantically is not a nominal, but an *amount* related to the nominal, as for example in (i) where the relative clause clause gives information about the *amount* of wine, rather than the wine itself.

(i) It would take me a year to drink the wine [that Kim drinks on a normal night].

Several different approaches to extraposition have been proposed in the HPSG literature.

One approach uses the idea of constituent order domains, mentioned briefly in Section 3.4 above (and see Müller (2019a), Chapter 10 of this volume). The idea is that an extraposed relative clause is composed with its antecedent nominal in the normal way as regards syntax and semantics, but that rather than being *compacted* into a single DOMAIN element, the nominal and the relative clause remain as separate DOMAIN elements, with the effect that that relative clause can be *liberated* away from the nominal, so that its phonology is contributed discontinuously from the phonology of the nominal, as in the examples in (64). See e.g. Nerbonne (1994), and Kathol & Pollard (1995) for details.

A second approach treats extraposition as involving a non-local dependency, introducing a non-local feature, typically called something like EXTRA, which functions much like other non-local features (e.g. SLASH). The idea is that a relative clause can make its semantic contribution as a nominal modifier "downstairs", but rather than being realised as a syntactic DAUGHTER (sister to the nominal), the relevant properties (e.g. the LOCAL features) are added to the EXTRA list of the head, and inherited up the tree until they are discharged from the EXTRA list by the appearance of an appropriate phrase-final daughter constituent, which contributes its phonology in the normal way, but makes no semantic contribution. Thinking from the top downwards, this is equivalent having a construction which allows a relative clause to appear e.g. as sister to a VP (as in (64a)) without affecting the VP's syntax or semantics, so long as it is pushed onto the EXTRA list of the VP, from where it will be inherited downwards until a nominal occurs which it can be interpreted as modifying (the apparatus needed to deal with the "bottom" of the dependency might be a family of lexical items derived by lexical rule, or a non-branching construction). See e.g. Keller (1995); Bouma (1996); Müller (1999a); Müller (2004); Crysmann (2005); Crysmann (2013).

A third approach is suggested in Kiss (2005), and adopted in Crysmann (2004) and Walker (2017). This approach exploits the more flexible approach to semantic composition provided by Minimal Recursion Semantics (MRS, Copestake et al. 2005), in the case of Kiss (2005), and Lexical Resource Semantics (LRS, Richter & Sailer 2004) in Walker (2017). The idea is that an extraposed relative clause appears as a normal syntactic daughter in its surface position, but the notion of semantic modification is generalised so that rather than the index of a modifying phrase being identified with that of a sister constituent (as standardly assumed), it may be identified with that of any suitable constituent *within* the sister. That is, adjuncts can be interpreted as modifying not just their sisters, but anything *contained in* their sisters — words and phrase to which they have no direct syntactic connection. This is implemented by means of a set valued ANCHORS feature, which is inherited upwards in the manner of a non-local feature, and which allows access to the indices of constituents from lower down. The flexibility of semantic composition afforded by MRS and LRS means that the right interpretations can be obtained.

A number of authors have argued for the superiority of an approach using EXTRA-style apparatus (e.g. Müller (2004) and Crysmann (2013)), but in terms of theoretical costs and benefits there seems to be little to choose between these alternatives — the first and third approaches rely on particular approaches to constituent order and semantic composition, while EXTRA-style analyses involve only on the more commonplace apparatus of non-local features (though with the added cost of special constructions or lexical operations to introduce and remove elements from EXTRA lists). Empirically, there are several issues that all approaches deal with more or less successfully (for example, the Right Roof Constraint from Ross 1967 that prevents extraposition beyond the clause, cf. (65b)). However, a more significant factor may be how well different accounts integrate with analyses of extraposition involving other kinds of adjunct and complement (e.g. complement clauses, as in (66)), capturing similarities and differences (see e.g. Crysmann 2013).

(65) a. [That someone might win *who does not deserve it*] is irrelevant.

b. * [That someone might win] is irrelevant *who does not deserve it.*

(66) The question then arises *whether we should continue in this way.*

4.2 Other functions

In this section we will briefly discuss phenomena involving clauses whose internal structures resemble relative clauses but which do not function as nominal modifers.[48]

[48] One omission here is discussion of *relative-corelative* constructions, which can be found in Hindi and Marathi, *inter alia*, and which were given an analysis in Pollard & Sag (1994: 227–232). These involve the paratactic combination of a clause that contains one or more relative pronouns, and what looks like a main clause containing coreferential pronouns, something like 'which boy$_i$ saw which girl$_j$, he$_i$ proposed to her$_j$' (meaning *the boy who saw the girl proposed to her*). Pollard & Sag's analysis involves associating a set of indices in the REL value of the first clause, which are realised by relative pronouns in the normal way, and an identical set of indices as encoded as the value of a CORRELATIVE feature in the main clause, which are realised by normal pronouns.

4.2.1 Complement clauses

Perhaps the most obvious cases of this kind involve clauses with the internal structure of a relative clause which occur as complements, rather than adjuncts. The following are some examples.[49]

(67) a. This story is the *(most) interesting *that we have heard.*

b. diejenige Frau *(*die dort steht*) (German)
the.that woman who there stands
'the very woman who is standing there'

c. It was Kim *that solved the problem.*

d. It was from Kim *that we got the news.*

e. On l' a vu *qui s'enfuyait* (French)
We him have seen who run.away.IMPERF
'We saw him running away'

In (67a) we have what looks like a *that* relative which is plausibly analysed as the complement of the superlative (notice that omitting the superlative makes (67a) ungrammatical).

The German example in (67b) exemplifies the *diejenigen* class of determiners, which require a complement that looks like a relative clause (and is analysed as such in Walker (2017)).

In (67c) we have a so-called *it*-cleft, a construction which features a clause resembling a relative clause, but rather than adding information about an associated nominal (as it would if it were a normal relative clause), the clause is interpreted as providing a presupposition ("someone/something solved the problem"), for an associated focus phrase (here the nominal *Kim*, so the interpretation is roughly "… and that person/thing was Kim"). Notice that the focus phrase need not be nominal (e.g. in (67d) it is a PP *from Kim*), again this is unlike normal (restrictive) relatives clauses (which are nominal modifiers).[50] In HPSG, following Pollard & Sag (1994: 260–262), *it*-clefts have typically been analysed as involving a lexical entry for *be* that takes an *it* subject, and two complements: an XP and an S which is marked as containing an XP gap. This makes *it*-clefts look rather different from relative clauses (the only real similarity being the existence of an

[49] Another case where a relative clause should be analysed as a complement is discussed in Arnold & Lucas (2016).

[50] Notice also that *that*-relatives are usually incompatible with proper name antecedents, but proper names are perfectly acceptable as the focus of an *it*-cleft with a *that*-clause, as in (67c) (Huddleston & Pullum 2002: 1416-1417).

unbounded dependency). One problem is that is not clear how this approach can be extended to examples like (68), where we seem to have an NP focus (*Sam*) which is not directly associated with an XP gap — we have instead a PP gap that seems to be associated with a normal relative phrase filler (*on whom*), i.e. where the similarity of the clefted clause to a relative clause is quite strong. It is not obvious how this problem should be dealt with.

(68) It was Sam [on whom she particularly focused her attention ___].

The French example in (67e) contains a so-called *predicative relative clause* (PRC).[51] Such clauses have the superficial form of a finite relative clause, but differ from them syntactically, semantically, and pragmatically. Koenig & Lambrecht (1999) analyse them as a form of *secondary predicate* (cf. *running away* in English *We saw them running away*). Syntactically, they are restricted to post-verbal positions, and are only permitted with certain kinds of verb (notably verbs of perception, like *voir* 'see', and discovery, like *trouver* 'find'), and the relative pronoun must be a top level subject. Semantically, they are subject to constraints on tense, modality and negation (there must be temporal overlap between the perception/discover event and the event reported in the relative clause, and the relative clause content cannot be either modal or negative). Pragmatically, their content must be asserted (rather than presupposed). Koenig & Lambrecht provide an analysis which treats PRCs as REL marked clauses with both an internal and an external subject (instances of *head-subject-ph* which have a non-empty SUBJ value), and which can consequently function as secondary predicates.

4.2.2 Dependent noun and pseudo-relative constructions

The following exemplifies a Korean structure that contains what looks superficially like a relative clause:

(69) Kim-u [[sakwa-ka cayngpan-wi-ey iss-nun] kes]-ul mek-ess-ta.
Kim-TOP apple-NOM tray-TOP-LOC exist-MOD KES-ACC eat-PST-DECL
(Korean)

'Kim ate an apple which was on the tray.'

Here what is traditionally called a *dependent noun* (*kes*) is preceded by a clause whose verb bears the morphological marking that is characteristic of relative clauses (the *-nun* affix).[52]

[51] The French term is *proposition relative dépendante attribut* (Sandfeld 1965).

[52] Japanese has a similar construction, involving the nominalising particle *no*, which has received some attention in the HPSG literature (e.g. Kikuta 1998; 2001; 2002). A difference is that there

However, unlike a normal relative clause, this "dependent" clause does not contain a gap, instead it contains what might be regarded as the semantic head of the construction (in this case, *sakwa-ka* 'apple'), notice that the clause+*kes* constituent satisfies the selection restriction of the verb *mek-ess-ta* 'ate'; this is what motivates the translation and explains why such clauses are often regarded as "internally headed" relatives. Kim (2016b: 303–317) notes a number of differences between *kes*-clauses and normal relatives (e.g. *kes*-clauses do not allow the full range of relative affixes to appear), and suggests these clauses are better analysed as complements of *kes*. See also Kim (1996), Chan & Kim (2002), Kim (2016a), and references there.[53]

Another Korean structure that has some similarity with relative clauses is the so-called *pseudo-relative* construction, exemplified in (70).[54]

(70) [komwu-ka tha-nun] naymsay (Korean)
rubber-NOM burn-MOD smell
'the smell that characterises the burning of rubber'

There is again no gap in the relative clause; again, only one kind of marking is allowed on the verb (only past tense *-un*); and only a limited range of nouns allow this kind of relative clause; this makes them rather like complement clauses. However, it is less plausible to think of a noun like *naymsay* 'smell' taking a com-

is no special morphology on the clause in Japanese, as noted above, in Section 3.3.1.

[53] Pollard & Sag (1994: 232–236) discuss a number of cases of what appear to be more plausible instances of *internally headed* relatives from a number of languages (Lakhota, Dogon, and Quechua); the following is from Dogon:

(i) [ya indɛ mi wɛ gɔ] yimaa boli. (Dogon)
yesterday person 1sg see.PN.∅ DEF die.PSP go.PN.3sg
'The person I saw yesterday is dead.'

Here we have a determiner *gɔ* preceded by a clause containing what would be the external head of a standard relative clause (in this case *indɛ* 'person'). The key difference between this and the Korean case is the absence here of any obvious clause external nominal like *kes* which can be treated as the head which takes the relative clause as a complement. Pollard & Sag (1994: 234) suggest (following Culy 1990) that NPs like that in (i) involve an exocentric construction, but no empty elements (neither an empty nominal, nor an empty relativiser). The NP consists of a determiner and a nominal, where the nominal consists of just a clause whose REL value contains the index of the nominal. This REL value is inherited downwards into the clause where it is identified with the index of one of the NPs, here the index of *indɛ* 'person': the effect of this is that the index of *indɛ* 'person' becomes the index of the whole NP (This ignores a number of technical and empirical issues to do with the inheritance and binding-off of REL values).

[54] A similar construction can be found in Japanese, cf. Kikuta (1998; 2001; 2002); Chan & Kim (2002).

plement (unlike *kes*), and these clauses are like prototypical relative clauses in not allowing topic marking. Kim suggests this is a special construction where the relation of head noun and relative clause is that the noun describes the perceptive result of the situation described by the clause (e.g. the smell is the perceptive result of the rubber burning. See Kim (1998b); Yoon (1993); Chan & Kim (2002); Cha (2005); Kim (2016b).

4.2.3 Free relatives

Perhaps the most significant case of a clause type that resembles a relative clause but which does not function as a nominal modifier consists of the so-called *free* (*headless*, or *fused*) *relatives*, exemplified in (71). These have received considerable attention in the HPSG literature.

(71) a. She ate *what I suggested.*
b. She ate *whatever I suggested.*
c. She put it *where I suggested.*

As these examples suggest, free relatives can be interpreted as involving either definite descriptions, as in (71a) "the thing that I suggested", or universal quantification, as in (71b) "everything that I suggested". They can also have adverbial or prepositional interpretations, as in (71c) "in the place that I suggested". The interpretation is related to the choice of *wh*-phrase. There are some special restrictions. For example, in English free relatives must be finite, as can be seen from (72a), and there are restrictions on what *wh*-words are allowed (e.g. *what* is permitted, as in (71a), but *which* is not, witness (72b)).

(72) a. *She ate *what to cook.*
b. *She ate *which I suggested.*

Free relatives resemble prototypical *wh*-relatives (and interrogative clauses) in containing a gap, and an initial *wh*-phrase in which is interpreted as filling the gap. They differ from interrogatives in having the external distribution of NPs or PPs rather than clauses (for example in (71a) *what I suggested* is the complement of *eat*, and in (71c) *where suggested* is a complement of *put*, neither of which allow clausal complements). They differ from prototypical relative clauses in not being associated with a nominal antecedent. They can contain relative pronouns which are not permitted in normal *wh*-relatives, notably the *-ever* pronouns, *whatever*, *whoever*, etc, and *what*, witness the ungrammaticality of the following:[55]

[55] *What* is not a relative pronoun in standard English, but it is in some other varieties, and (73b) is grammatical in those.

(73) a. *She ate the thing(s) *whatever I suggested.*
b. *She ate the things(s) *what I suggested.*

In general the possibilities of relative inheritance (pied-piping) in free relatives are dramatically reduced compared to prototypical relatives and interrogatives. For example in English, relative inheritance is not possible from the complement of a preposition, as can be seen from (74b):

(74) a. Try to describe *what you talked about.*
b. *Try to describe *about what you talked.*

In fact, in English relative inheritance only seems to be possible from *wh*-phrases in in pre-nominal position (determiners and genitive NPs), as in (75), and (77a) below.[56]

(75) They will steal *what(ever) things they can carry.*

As with prototypical relatives, the initial *wh*-phrase in a free relative has to satisfy restrictions imposed "downstairs" in the relative clause (i.e. restrictions that follow from the location of the gap). In addition, however, it seems that with free relatives the *wh*-phrase is also sensitive to restrictions imposed from the outside the relative clause — the *wh*-phrase of a free relative has to be of the appropriate category for the position where the free relative appears. For example, as a first approximation, a free relative with *what* is only possible where an NP is possible, and a free relative with *where* is only possible where a locative PP is possible. This is the so-called *matching effect* in free relatives.[57]

One interesting instance of this involves case marking. Consider, for example, the German data in (76). These show a free relative in a position which requires nominative case marking, containing a relative pronoun whose role within the relative clause requires nominative marking. Since *wer* 'who' is nominative, all is well. By contrast, in (76b) while the nominative *wer* satisfies the requirements within the relative clause, there is a case conflict because the free relative as a whole is the complement of a verb *vertrauen* 'trust' that requires a dative complement. The result is ungrammatical. Examples like (76c) show a complication. Here again there is a case conflict: within the relative clause, the relative pronoun is required to be accusative (complement of *empfehlen* 'recommend'), and

[56]Other languages are less restrictive, e.g. Müller (1999a: 57) gives German examples analogous to (74b). See Footnote 59.

[57]In fact, things are more complicated. For example, in *He walked to [where his horse was waiting].* we have a free relative with *where* in an NP position (object of a preposition) rather than a PP position. See e.g. Kim (2017) for discussion.

the free relative as a whole is in a nominative position. However, the result is grammatical, presumably because the morphological form of the neuter relative pronoun *was* 'what' can realise either nominative or accusative case (unlike the masculine *wer*).

(76) a. Wer schwach ist, muss klug sein. (German)
who.NOM weak is must clever be
'Whoever is weak must be clever.'

b. * Wer klug ist, vertraue ich immer. (German)
who.NOM clever is trust I ever
intended: 'I trust whoever is clever.'

c. Was du mir empfiehlst, macht einen guten Eindruck.
what.NOM/ACC you me recommend makes a good impression (German)
'What you recommend me makes a good impression.'

The agreement properties of free relatives are somewhat surprising, and reveal a potential complication in the matching effect. Notice that in (77a) the *wh*-phrase, *whoever's dogs*, is plural, and triggers plural agreement on the verb in relative clause.

(77) a. [[Whoever's$_{sg}$ dogs]$_{pl}$ are running around]$_{sg}$ is in trouble.

b. Whoever is/*are running around (is in trouble).

This is not surprising since *whoever's dogs* is headed by a plural noun (*dogs*). However, the free relative as a whole triggers singular agreement, consistent with the agreement properties coming from the relative pronoun — *whoever* is singular, as can be seen from (77b). This is also consistent with the semantics: the free relative in (77a) denotes the person whose dogs are running around, not the dogs (in this it resembles an NP like *anyone whose dogs are running around*, which involves a normal relative clause construction).[58] This shows a complication of the matching effect: it seems that within clause requirements are reflected on the initial *wh*-phrase (*whoever's dogs* is the subject of the relative), but the external distribution reflects the properties of the relative *word* (*whoever*). Of course, the fact that relative inheritance is so limited in free relatives means that usually the *wh*-phrase consists of just the *wh*-word, so that is very difficult to tease these

[58]This is not a universal property: Borsley (2008) notes that examples in Welsh resembling (77a) are interpreted as meaning that the dogs are in big trouble, not the owner.

things apart.[59]

Following Müller (1999a) on German, free relatives have received considerable attention in the HPSG literature, with analyses dealing with a variety of languages, including: Arabic (Alqurashi 2012; Hahn 2012), Danish (Bjerre 2012; 2014), English (Kim & Park 1996; Kim 2001; Wright & Kathol 2003; Francis 2007; Yoo 2008; Kim 2017), German (Hinrichs & Nakazawa 2002; Kubota 2003), Persian (Taghvaipour 2005), and Welsh (Borsley 2008).

The central analytic problem is this: leaving aside the complication arising from case syncretism and relative inheritance just mentioned, the existence of matching effects has suggested to some (e.g. Kubota (2003)) that the *wh*-phrase should be the head of the free relative, because the distribution of free relatives depends on the properties of the *wh*-phrase. So, for example, the NP *what* would be the head of *what I suggested*. But this is inconsistent with *what* being the filler of the gap in *what I suggested* (i.e. the missing object of *suggested*), because in a normal filler-gap construction the filler is *non*-head. If, instead, we assume that *what* is primarily the filler of the gap in the free relative, then we should assume that the clause *I suggested* ___ is the head of the free relative – and the distributional properties of the free relative are unexplained.

4.2.4 Pseudo-clefts and transparent free relatives

Two constructions that show some similarity with free relatives, and have received some attention in the HPSG literature, are *specificational pseudo-clefts*, exemplified in (78), and so-called *transparent free relatives* (TFRs), exemplified in (79).

(78) a. A new coat is [what Kim will be wearing].
b. [What Kim will be wearing] is a new coat.
c. [What she did] was cut her hair.

[59] Müller (1999a: 90) discusses the following German example of a free relative with an initial PP containing the nominal relative word *wem* 'whom' (i.e. showing relative inheritance to PP):

(i) Ihr könnt beginnen, [mit wem ihr (beginnen) wollt]. (German)
you can start with whom you start want
'You can start with whoever you like.'

He observes that the free relative functions as a PP, just like *mit wem*, and in the variant where the parenthesised instance of *beginnen* is present, the within clause role is also that of a PP. However, *mit* here is a non-predicative preposition, so the index associated with the PP is just that of the *wh*-word that it contains, so it is still not possible to fully distinguish the properties of the *wh*-phase and the properties of the *wh*-word it contains.

d. [What she did not bring] was any wine.

(79) a. She replied in [what anyone would consider ___ a belligerent tone].

b. Her reply was [what anyone would consider ___ belligerent].

Specificational pseudo-clefts typically consist of a *wh*-clause, *be*, and a *focal phrase* (e.g. *any wine* in (78d)). The focal phrase corresponds to a gap in the *wh*-clause (e.g. in (78d) *any wine* is interpreted as the missing object of *bring*). They raise a number of issues that are not typical of relative clauses, notably the existence of *connectivity effects* whereby the focal phrase behaves as though it was part of the *wh*-clause (e.g. in (78d) the negative polarity item *any* is licensed by the negation in the *wh*-clause). Beyond this, it is not obvious whether the *wh*-clauses should be analysed as related to interrogatives, as in Yoo (2003), or as related to free relatives, as in Gerbl (2007).[60]

In TFRs the relative appears to function somewhat like a parenthetical modifier of a *nucleus* (e.g. *a belligerent tone* in (79a)), which seems to provide the head properties of the phrase as a whole — so for example the TFR in (79a) has the characteristics of an NP, that in (79b) has those of an AP (it is a natural starting point to assume the nucleus is internal to the relative clause, since otherwise one has the puzzle of a relative clause which is both incomplete and occurs before the head it modifies). TFRs are in some ways even more restricted than other kinds of relative (only *what* is allowed as the relative expression), but in others less restricted (e.g. free relatives have the external distribution of NPs, but the TFR in (79b) has the distribution of an AP, like its nucleus *belligerent*). Some approaches to TFRs employ novel kinds of structure (e.g. *grafts*, cf. van Riemskijk 2006), but Yoo (2008) and Kim (2011) provide HPSG analyses which capture the relevant properties using existing apparatus with only minor adjustments.

5 Conclusion

The analysis of relative clauses has been important in the theoretical evolution of HPSG, notably in the development of constructional approach involving inheritance from cross classifying dimensions of description. Empirically, relative

[60] It can be difficult to distinguish this kind of pseudo-cleft from cases involving a normal free relative. An example like *What she is wearing is a mess* is superficially similar to (78b), but it involves a free relative. Notice, for example, it can be paraphrased with a normal NP plus relative clause (as “The thing that she is wearing is a mess”) and *what* can be replaced with *whatever*. It does not have a paraphrase with an *it*-cleft or a simple proposition — it cannot be paraphrased as “It is mess that she is wearing” or “She is wearing a mess”.

clauses have been the focus of a significant amount of descriptive work in a variety of typologically diverse languages. Our goal in this paper has been exposition and survey rather than argumentation towards particular conclusions, but, perhaps paradoxically given what we have just said, we think one conclusion that clearly emerges is that, from an HPSG perspective at least, *relative clauses are not a natural kind*. There is *nothing* one can say that will be true of everything that has been described as a "relative clause" in the literature. As regards internal structure, some are *head-filler* structures (*wh*-relatives), while others are *head-complement* structures (complementiser relatives, some kinds of bare relative); correspondingly, some involve relative pronouns (hence a REL feature), some do not. It is true that most involve some kind of SLASH dependency, but this is hardly unique to relative clauses, and even this does not hold of the dependent noun and pseudo-relatives mentioned in Section 4.2.2. There is no semantic unity — while restrictive relatives are noun-modifiers, non-restrictives function more like independent clauses, and free relatives have nominal or adverbial semantics. Similarly, as regards external distribution: prototypical relatives are noun modifiers, and appear in *head-adjunct-ph* structures, but expressions with similar internal structure occur as complements (e.g. free relatives, clefts, and complements of superlative adjectives).

We do not think it is a bad thing that this conclusion should emerge from a discussion of HPSG approaches. Rather, it suggests to us that an approach that tries to impose unity will end up being procrustean. In fact, discussion of relative clauses seems to us to show some of the best features of HPSG — the analyses we have summarised are generally well formalised, carefully constructed (detailed, precise, and coherent), and both empirically satisfying and insightful, with relatively few *ad hoc* assumptions or special stipulations. The discussion shows how the expressivity and flexibility of the descriptive machinery of the framework are compatible with a wide range of phenomena across a range of languages.

Abbreviations

- AP, XP, PP, NP, VP, CP, S, DP, PRO (standard linguistic abbreviations)
- RP – a phrase headed by the empty relativiser R – put in the index?
- SELR – Subject Extraction Lexical Rule
- MRS – Minimal Recursion Semantics
- LRS – Lexical Resource Semantics

- WHIP – Wh-Inheritance Principle
- NRC – non-restrictive relative clause
- RRC – restrictive relative clause
- PRC – predicative relative clause
- TFR – transparent free relative

Acknowledgements

We are grateful to the editors of this volume and two anonymous referees for their careful and insightful comments on earlier versions of this contribution. Remaining flaws are our sole responsibility, of course.

References

Abeillé, Ann & Danièle Godard. 2007. Les relatives sans pronom relatif. In Michaël Abécassis, Laure Ayosso & Elodie Vialleton (eds.), *Le francais parlé au 21ème siècle: normes et variations dans le discours et en interaction*, vol. 2, 37–60. Paris: L'Harmattan.

Abeillé, Anne, Danièle Godard, Philip H. Miller & Ivan A. Sag. 1997. French bounded dependencies. In Sergio Balari & Luca Dini (eds.), *Romance in HPSG* (CSLI Lecture Notes 75), 1–54. Stanford, CA: CSLI Publications.

Alotaibi, Mansour & Robert D. Borsley. 2013. Gaps and resumptive pronouns in Modern Standard Arabic. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 6–26. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2013/alotaibi-borsley.pdf, accessed 2018-2-25.

Alqurashi, Abdulrahman. 2012. An HPSG approach to free relatives in Arabic. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University, Daejeon*, 6–25. http://cslipublications.stanford.edu/HPSG/2012/alqurashi.pdf, accessed 2019-6-5.

Alqurashi, Abdulrahman & Robert D. Borsley. 2012. Arabic relative clauses in HPSG. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University, Daejeon*, 26–44. Stanford, CA. http://cslipublications.stanford.edu/HPSG/2012/alqurashi-borsley.pdf, accessed 2019-6-5.

Arnold, Doug. 2004. Non-Restrictive relative clauses in construction based HPSG. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 27–47. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/arnold.pdf, accessed 2019-6-5.

Arnold, Doug. 2007. Non-Restrictive relatives are not orphans. *Journal of Linguistics* 43(2). 272–309. DOI:10.1017/S0022226707004586

Arnold, Doug & Robert D. Borsley. 2008. Non-restrictive relative clauses, ellipsis and anaphora. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar, National Institute of Information and Communications Technology, Keihanna*, 325–345. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/9/arnold-borsley.pdf, accessed 2019-6-5.

Arnold, Doug & Robert D. Borsley. 2010. Auxiliary-stranding relative clauses. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot, Paris 7*, 47–67. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2010/arnold-borsley.pdf, accessed 2019-6-5.

Arnold, Doug & Christopher Lucas. 2016. One of those constructions that really needs a proper analysis. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on head-driven phrase structure grammar and lexical functional grammar, polish academy of sciences, warsaw, poland*, 43–63. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2016/headlex2016-arnold-lucas.pdf, accessed 2019-9-10.

Avgustinova, Tania. 1996. *Relative clause constructions in Bulgarian HPSG*. CLAUS-Report 71. Saarbrücken: University of the Saarland.

Bîlbîie, Gabriela & Frédéric Laurens. 2009. A construction-based analysis of verbless relative adjuncts in French and Romanian. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen*, 5–25. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2009/bilbiie-laurens.pdf, accessed 2019-6-5.

Bjerre, Anne. 2012. An analysis of Danish free relatives. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University, Daejeon*, 45–63. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2012/bjerre.pdf, accessed 2019-5-20.

Bjerre, Anne. 2014. Danish non-specific free relatives. *Nordic Journal of Linguistics* 37(1). 5–27. DOI:10.1017/S0332586514000043

Blakemore, Diane. 2006. Divisions of labour: the analysis of parentheticals. *Lingua* 116. 1670–1687.

Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based Construction Grammar* (CSLI Lecture Notes 193). Stanford, CA: CSLI Publications.

Bolc, Leonard. 2005. Formal description of natural languages: an HPSG grammar of Polish. In Dieter Hutter & Werner Stephan (eds.), *Mechanizing mathematical reasoning: essays in honor of jörg h. siekmann on the occasion of his 60th birthday* (Lecture Notes in Artificial Intelligence 2605), 297–320. Berlin/Heidelberg: Springer Verlag.

Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 1989. Phrase-structure grammar and the Barriers conception of clause structure. *Linguistics* 27(5). 843–864.

Borsley, Robert D. 2008. *On some Welsh unbounded dependency constructions*. Essex Research Reports in Linguistics 57.4. Dept. of Language & Linguistics, University of Essex. http://repository.essex.ac.uk/id/eprint/98, accessed 2019-5-15.

Bouma, Gosse. 1996. Extraposition as a nonlocal dependency. In Geert-Jan Kruijff, Glynn V. Morrill & Dick Oehrle (eds.), *Proceedings of Formal Grammar 96*, 1–14. Prag. http://www.let.rug.nl/gosse/papers.html, accessed 2018-2-25.

Cha, Jong-Yul. 2005. *Constraints on clausal complex noun phrases in Korean with focus on the gapless relative clause construction*. The University of Illinois at Urbana-Champaign dissertation. https://pdfs.semanticscholar.org/b61d/878a9ec8283741a9f21cea4c32f14048856a.pdf, accessed 2019-6-7.

Chan, Chung & Jong-Bok Kim. 2002. Differences between externally and internally headed relative clause constructions. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 43–65. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/3/chungkim.pdf, accessed 2019-6-7.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1977. On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Crysmann, Berthold. 2004. Underspecification of intersective modifier attachment: Some arguments from German. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 378–392. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

Crysmann, Berthold. 2005. Relative clause extraposition in German: An efficient and portable implementation. *Research on Language and Computation* 1(3). 61–82.

Crysmann, Berthold. 2013. On the locality of complement clause and relative clause extraposition. In Gert Webelhuth, Manfred Sailer & Heike Walker (eds.), *Rightward movement in a comparative perspective* (Linguistik Aktuell/Linguistics Today 200), 369–396. Amsterdam: John Benjamins Publishing Co.

Crysmann, Berthold. 2016. An underspecification approach to hausa resumption. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on head-driven phrase structure grammar and lexical functional grammar, polish academy of sciences, warsaw, poland*, 194–214. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2016/headlex2016-crysmann.pdf, accessed 2019-9-10.

Crysmann, Berthold & Chris H. Reintges. 2014. The polyfunctionality of Coptic Egyptian relative complementisers. In Stefan Müller (ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo*, 63–82. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2014/crysmann-reintges.pdf, accessed 2019-6-7.

Culy, Chris. 1990. *The syntax and semantics of internally headed relative clauses*. Stanford University dissertation.

De Kuthy, Kordula. 1999. Linearization versus movement: Evidence from German pied-piped infinitives. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 97–112. Stanford, CA: CSLI Publications.

Dhongde, Ramesh Vaman & Kashi Wali. 2009. *Marathi* (London Oriental and African Library 13). John Benjamins Publishing.

Francis, Elaine J. 2007. The role of default constructions in the processing of mismatch: The case of possessive free relatives. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Department of Linguistics, Stanford, CA*, 345–363. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/8/francis.pdf, accessed 2019-6-7.

Gerbl, Niko. 2007. *An analysis of pseudoclefts and specificational clauses in Head-Driven Phrase Structure Grammar*. Philosophische Fakultät der Georg-August-Universität Göttingen dissertation.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Godard, Danièle. 1992. Extraction out of NP in French. *Natural Language and Linguistic Theory* 10(2). 233–277.

Grosu, Alexander & Fred Landman. 2017. Amount relatives. In Martin Everaert, Henk C. van Riemsdijk, Rob Goedemans & Bart Hollebrandse (eds.), *The Wiley Blackwell Companion to Syntax* (Wiley Blackwell Companions to Linguistics 3). Oxford: John Wiley & Sons, Inc.

Güngördü, Zelal. 1996. An HPSG analysis of Turkish relative clauses. In Claire Grover & Enric Vallduv'i (eds.), *Edinburgh working papers in cognitive science, vol. 12: studies in HPSG*, chap. 3, 71–119. Scotland: Centre for Cognitive Science, University of Edinburgh. https://www.upf.edu/documents/2983731/3019795/1996-ccs-evcg-hpsg-wp12.pdf, accessed 2019-9-10.

Haddar, Kais, Ines Zalila & Sirine Boukedi. 2009. A parser generation with the LKB for the Arabic relatives. *International Journal of Computing and Information Sciences* 7(1). 51–60. http://www.ijcis.info/Vol7N1/Vol7P5N1-PP-51-60.pdf, accessed 2015-3-1.

Hahn, Michael. 2012. Arabic relativization patterns: A unified HPSG analysis. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University, Daejeon*, 144–164. http://cslipublications.stanford.edu/HPSG/2012/hahn.pdf, accessed 2019-7-6.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1999. VP relatives in German. In Gosse Bouma, Erhard W. Hinrichs, Geert-Jan M. Kruijff & Richard T. Oehrle (eds.), *Constraints and resources in natural language syntax and semantics* (Studies in Constraint-Based Lexicalism 3), 63–81. Stanford, CA: CSLI Publications.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 2002. Case matching in Bavarian relative clauses: A declarative, non-derivational account. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the 8th International*

Conference on Head-Driven Phrase Structure Grammar, Norwegian University of Science and Technology, Trondheim, 180–188. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2/HinrichsNakasawa2-pn.pdf, accessed 2019-6-7.

Holler, Anke. 2003. An HPSG analysis of the non-integrated *Wh*-relative clauses in German. In Stefan Müller (ed.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 163–180. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/4/holler.pdf, accessed 2018-6-7.

Huddleston, Rodney & Geoffrey K. Pullum (eds.). 2002. *The Cambridge grammar of the English language*. Cambridge, UK: Cambridge University Press.

Kathol, Andreas & Carl J. Pollard. 1995. Extraposition via complex domain formation. In *Proceedings of the Thirty-Third Annual Meeting of the Association for Computational Linguistics*, 174–180. Cambridge, MA. https://www.aclweb.org/anthology/P95-1024, accessed 2019-6-6.

Kayne, Richard. 2017. A note on some even more unusual relative clauses. In Laura Bailey & Michelle Sheehan (eds.), *Order and structure in syntax: Festschrift for Anders Holmberg*, 363–371. Berlin: Language Science Press.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Keller, Frank. 1995. Towards an account of extraposition in HPSG. In *Proceedings of the 7th Conference of the European Association for Computational Linguistics (Student Session)*, 301–306. Dublin. https://www.aclweb.org/anthology/E95-1045, accessed 2019-6-6.

Kikuta, Chiharu Uda. 1998. A multiple inheritance analysis of the internally-headed relative clause in Japanese. In Jin Guo, Kim Teng Lua & Jie Xu (eds.), *Proceedings of the 12th Pacific Asia Conference on Language, Information and Computation*, 82–93. Singapore. https://www.aclweb.org/anthology/papers/Y/Y98/Y98-1008/, accessed 2015-10-6.

Kikuta, Chiharu Uda. 2001. The Japanese internally-headed relative clause as a marked head-complement structure. In Benjamin K. T'sou, Olivia O.Y. Kwong & Tom B.Y. Lai (eds.), *Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation*, 317–324. Hong Kong, China. https://www.aclweb.org/anthology/papers/Y/Y01/Y01-1029/, accessed 2019-6-6.

Kikuta, Chiharu Uda. 2002. Clausal complement or adverbial clause? Toward an integral account of the Japanese internally-headed relative clause. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *The proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 202–

220. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2/, accessed 2006-9-24.

Kim, Jong-Bok. 1998a. A head-driven and constraint-based analysis of Korean relative clause constructions. *Language Research* 34(4). 1–41.

Kim, Jong-Bok. 1998b. On the mixed properties of pseudo relative clause constructions. In *Proceedings of the 11th International Conference on Korean*, 83–93. Seoul, Korea: International Circle of Korean Linguistics.

Kim, Jong-Bok. 2001. Constructional constraints in English free relative constructions. *Language and Information* 5(1). 35–53.

Kim, Jong-Bok. 2011. English transparent free relatives: Interactions between the lexicon and constructions. *English Language and Linguistics* 17(2). 153–181.

Kim, Jong-Bok. 2016a. Copular constructions and asymmetries in the specificational pseudocleft constructions in Korean. *Language and Linguistics* 17(1). 89–112.

Kim, Jong-Bok. 2016b. *The syntactic structures of Korean*. Cambridge: Cambridge University Press.

Kim, Jong-Bok. 2017. Mixed properties and matching effects in English free relatives: A construction-based perspective. *Linguistic Research* 34(3). 361–385.

Kim, Jong-Bok & Byung-Soo Park. 1996. English free relative clause constructions: From a constraint-based perspective. In *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation*, 31–41.

Kim, Jong-Bok & Peter Sells. 2008. *English syntax: Ang introduction*. Stanford, CA: CSLI Publications.

Kim, Jong-Bok & Jaehyung Yang. 2003. Korean phrase structure grammar and its implementations into the LKB system. In *Proceedings of the 17th Pacific Asia Conference on Language, Information and Computation*, 88–97.

Kim, Yong-Beom. 1996. Internally headed relative clause constructions in Korean. In *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation*, 403–413.

Kiss, Tibor. 2005. Semantic constraints on relative clause extraposition. *Natural Language and Linguistic Theory* 23(2). 281–334. DOI:10.1007/s11049-003-1838-7

Koenig, Jean-Pierre & Knud Lambrecht. 1999. French relative clauses as secondary predicates: A case study in Construction Theory. In Francis Corblin, Carmen Dobrovie-Sorin & Jean-Marie Marandin (eds.), *Empirical issues in formal syntax and semantics 2, selected papers from the Colloque de Syntaxe et Sémantique à Paris (CSSP 1997)*, 191–214. The Hague: Thesus.

Kubota, Yusuke. 2003. Yet another HPSG-analysis for free relative clauses in German. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 147–167. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/3/kubota.pdf, accessed 2019-6-6.

Lee-Goldman, Russell. 2012. Supplemental relative clauses: internal and external syntax. *Journal of Linguistics* 48(3). 573–608. DOI:10.1017/S0022226712000047

Lees, Robert B. 1961. The constituent structure of noun phrases. *American Speech* 36. 159–168.

Link, Godehard. 1984. Hydras: On the logic of relative constructions with multiple heads. In Fred Landmann & Frank Veltman (eds.), *Varieties of formal semantics*, 245–257. Dordrecht: Foris Publications.

Melnik, Nurit. 2006. Hybrid agreement as a conflict resolution strategy. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Bulgarian Academy of Sciences, Sofia*, 228–246. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/7/melnik.pdf, accessed 2019-6-6.

Müller, Stefan. 1999a. An HPSG-analysis for free relative clauses in German. *Grammars* 2(1). 53–105.

Müller, Stefan. 1999b. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2004. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation* 2(2). 209–257.

Müller, Stefan. 2019a. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

Müller, Stefan. 2019b. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

Müller, Stefan & Antonio Machicao y Priemer. 2019. Head-Driven Phrase Structure Grammar. In András Kertész, Edith Moravcsik & Csilla Rákosi (eds.), *Cur-*

rent approaches to syntax: A comparative handbook (Comparative Handbooks of Linguistics 3). Berlin: Mouton de Gruyter. DOI:10.1515/9783110540253-012

Mykowiecka, Agnieszka, Małgorzata Marciniak, Adam Przepiórkowski & Anna Kupść. 2003. An implementation of a Generative Grammar of Polish. In Peter Kosta, Joanna Błaszczak, Jens Frasek, Ljudmila Geist & Marzena Żygis (eds.), *Investigations into formal slavic linguistics: Contributions of the Fourth European Conference on Formal Description of Slavic Languages – FDSL IV held at Potsdam University, November 28–30, 2001*, 271–285. Frankfurt am Main: Peter Lang.

Nerbonne, John. 1994. Partial verb phrases and spurious ambiguities. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 109–150. Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago & Stanford: University of Chicago Press & CSLI Publications.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In Arnold Beckmann & Norbert Preining (eds.), *European Summer School in Logic, Language and Information 2003, Course Material I* (Collegium Logicum 5), 87–143. Wien: Publication Series of the Kurt Gödel Society.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation. Appeared as Ross (1986): *Infinite Syntax*. Ablex Publishing Corporation,Norwood, New Jersey.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484. DOI:10.1017/S002222679700652X

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545. http://www.jstor.org/stable/40961690, accessed 2019-6-6.

Sandfeld, Kristian. 1965. *Syntaxe du Français contemporain, les propositions subordonnées*. Genève: Droz.

Sauerland, Uli. 1998. *The meaning of chains*. Cambridge, MA: MIT dissertation.

Schachter, Paul. 1973. Focus and relativization. *Language* 49(1). 19–46.

Sirai, Hidetoshi & Takao Gunji. 1998. Relative clauses and adnominal clauses. In Takao Gunji & Koiti Hashida (eds.), *Topics in constraint-based grammar of Japanese*, 17–38. Dordrecht: Kluwer.

Srivastav, Veneeta. 1991. The syntax and semantics of correlatives. *Natural Language and Linguistic Theory* 9(4). 637–686. DOI:10.1007/BF00134752

Taghvaipour, Mehran A. 2005. Persian free relatives. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 364–374. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/taghvaipour.pdf, accessed 2019-6-6.

Vaillette, Nathan. 2001a. Hebrew relative clauses in HPSG. In Daniel P. Flickinger & Andreas Kathol (eds.), *Proceedings of the HPSG-2000 Conference, University of California, Berkeley*, 305–324. CSLI Publications. http://csli-publications.stanford.edu/HPSG/1/, accessed 2018-2-25.

Vaillette, Nathan. 2001b. Irish gaps and resumptive pronouns in HPSG. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 Conference, Norwegian University of Science and Technology, Trondheim*, 284–299. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2001/, accessed 2015-12-23.

Van Eynde, Frank. 2004. Pied piping is a local dependency. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 313–334. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/vaneynde.pdf, accessed 2007-4-3.

van Riemskijk, Henk. 2006. Grafts follow from Merge. In Mara Frascarelli (ed.), *Phases of interpretation* (Studies in Generative Grammar 91), 17–44. Berlin: Mouton de Gruyter.

Vergnaud, J.-R. 1974. *French relative clauses*. Cambridge, MA: MIT dissertation.

Walker, Heike. 2017. *The syntax and semantics of relative clause attachment*. Johann Wolfgang Goethe-Universität dissertation. http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/42757, accessed 2017-6-6.

Webelhuth, Gert, Sascha Bargmann & Christopher Götze. 2018. Idioms as evidence for the proper analysis of relative clauses. In Manfred Krifka & Matthias Schenner (eds.), *Reconstruction effects in relative clauses* (Studia Grammatica 75). Berlin: Akademie Verlag.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Wright, Abby & Andreas Kathol. 2003. When a head is not a head: A Constructional approach to exocentricity in English. In Jongbok Kim & Stephen Mark

Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 373–389. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/3/, accessed 2018-9-25.

Yoo, Eun Jung. 2003. Specificational pseudoclefts in English. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-2-25.

Yoo, Eun-Jung. 2008. Transparent free relatives in English. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar, National Institute of Information and Communications Technology, Keihanna*, 274–293. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2008/yoo.pdf, accessed 2019-7-2.

Yoon, Jae-Hak. 1993. Different semantics for different syntax: relative clauses in Korean. In Andreas Kathol & Carl J. Pollard (eds.), *Papers in syntax* (OSU Working Papers in Linguistics 42), 199–226. Ohio State University: Department of Linguistics.

Zalila, Ines & Kais Haddar. 2011. Construction of an HPSG grammar for the Arabic relative sentences. In Galia Angelova, Kalina Bontcheva, Ruslan Mitkov & Nicolas Nicolov (eds.), *Proceedings of Recent Advances in Natural Language Processing, RANLP 2011*, 574–579. Hissar, Bulgaria. http://www.aclweb.org/anthology/R11-1081, accessed 2019-4-9.

Chapter 16

Island phenomena and related matters

Rui Chaves
University at Buffalo, SUNY

Extraction constraints on long-distance dependencies – so-called *islands* – have been the subject of intense linguistic and psycholinguistic research for the last half century. Despite of their importance in syntactic theory, the heterogeneity of island constraints has posed many difficult challenges to linguistic theory, across all frameworks. The HPSG perspective of island phenomena is that they are unlikely to be due to a unitary syntactic constraint given the fact that virtually all such island constraints have known exceptions. Rather, it is more plausible that island constraints result from a combination of independently motivated syntactic, semantic, pragmatic and processing phenomena. The present chapter is somewhat different from others in this volume in that its focus in not on HPSG analyses of some phenomena, but rather on the nature of the phenomena itself. This is because there is evidence that most of the phenomena are not purely grammatical, and to that extent independent from HPSG or indeed any theory of grammar. One may call this view of island phenomena 'minimalist' in the sense that much of it does not involve formal grammar.

1 Introduction

This chapter provides an overview of various island effects that have received attention from members of the HPSG community. I begin with the extraction constraints peculiar to coordinate structures, because they not only have a special status in the history of HPSG, but also because they illustrate well the non-unitary nature of island constraints. I then argue that, at a deeper level, some of these constraints are in fact present in many other island types, though not necessarily all. For example, I take it as relatively clear that *factive islands* are purely pragmatic in nature (Oshima 2007), as are *negative islands* (Kroch 1989;

Rui Chaves. 2019. Island phenomena and related matters. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 415–468. Berlin: Language Science Press. DOI:??

Szabolcsi & Zwarts 1993; Abrusán 2011; Fox & Hackl 2006; Abrusán & Spector 2011), although one can quibble about the particular technical details of how such accounts are best articulated. Similarly, the *NP Constraint* in the sense of Horn (1972) is likely to be semantic-pragmatic in nature (Kuno 1987; Godard 1988; Davies & Dubinsky 2009). Conversely, I take it as relatively uncontroversial that the *Clause Non-Final Incomplete Constituent Constraint* is due to processing difficulty (Hukari & Levine 1991; Fodor 1992). See also Kothari (2008) for evidence that 'bridge' verb effects in filler-gap dependencies are partly due to lack of contextualization.

In the present chapter I focus on islands that have garnered more attention from members of the HPSG community, and that have caused more controversy cross-theoretically. My goal is to provide an overview of the range of explanations that have been proposed to account for the complex array of facts surrounding islands, and to show that no single unified account is likely.

2 Background

As already detailed in Borsley & Crysmann (2019), Chapter 14 of this volume, HPSG encodes filler-gap dependencies in terms of a set-valued feature SLASH. Because the theory consists of a feature-based declarative system of constraints, virtually all that goes on in the grammar involves constraints stating which value a given feature takes. By allowing SLASH sets to be unified (or unioned), it follows that constructions in which multiple gaps are linked to the same filler are trivially obtained, as in (1).

(1)
  a. Which celebrity did [the article insult _ more than it praised _]?
  b. Which celebrity did you expect [[the pictures of _] to bother _ the most]?
  c. Which celebrity did you [inform _ [that the police was coming to arrest _]]?
  d. Which celebrity did you [compare [the memoir of _] [with a movie about _]?
  e. Which celebrity did you [hire _ [without auditioning _ first]]?
  f. Which celebrity did you [[meet _ at a party] and [date _ for a few months]]?

But another advantage of encoding the presence of filler-gap dependencies as a feature is that certain lexical items and constructions can easily impose idiosyncratic constraints on SLASH values. For example, to account for languages that do not allow preposition stranding, it suffices to state that prepositions are necessarily specified as [SLASH {}]. Thus, their complements cannot appear in SLASH instead of COMPS. The converse also occurs. Certain uses of the verb *assure*, for example, are lexically required to have one complement in SLASH rather than in COMPS. Thus, extraction is obligatory as (2) shows.

(2) a. * I can assure you him to be the most competent.
 b. Who$_i$ can you assure me _ $_i$ to be the most competent? (Kayne 1980).

As we shall see, it would be rather trivial to impose the classic island constraints in the standard syntactic environments in which they arise.[1] The problem is that island effects are riddled with exceptions which defy purely syntactic accounts of the phenomena. Hence, HPSG has generally refrained from assuming that islands are syntactic, in contrast to mainstream linguistic theory.

3 The Coordinate Structure Constraint

Ross (1967) first observed that coordinate structures impose various constraints on long-distance dependencies, shown in (3), collectively dubbed the *Coordinate Structure Constraint*. For perspicuity, I follow Grosu (1973a) in referring to (i) as the *Conjunct Constraint* and to (ii) as the *Element Constraint*.

(3) COORDINATE STRUCTURE CONSTRAINT (CSC):
 In a coordinate structure, (i) no conjunct may be moved, (ii) nor may any element contained in a conjunct be moved out of that conjunct … unless each conjunct properly contains a gap paired with the same filler.

The *Conjunct Constraint* (CC) is illustrated by the unacceptability of the extractions in (3.1). No such constraint is active in other constructions like those in (5) and (6), for example.

(4) a. * Which celebrity did you see [Priscilla and _]?
 (cf. with 'Did you see Priscilla and Elvis?')
 b. * Which celebrity did you see [_ and Priscilla]?
 (cf. with 'Did you see Elvis and Priscilla?')

[1] In such a view, island effects could perhaps result from grammaticized constraints, induced by parsing and performance considerations (Pritchett 1991; Fodor 1978; 1983a).

c. * Which celebrity did you see [_ or/and a picture of _]?
(cf. with 'Did you see Elvis or/and a picture of Elvis?')

(5) a. Which celebrity did you see Priscilla with _ ?
(cf. with 'Did you see Priscilla with Elvis?')
b. Which celebrity did you see _ with Elvis?
(cf. with 'Did you see Priscilla with Elvis?')

(6) a. Which celebrity is Kim as tall as _ ?
b. Which celebrity did you say Robin arrived earlier than _ ?

In HPSG accounts of extraction that assume the existence of traces (Pollard & Sag 1994; Levine & Hukari 2006) the CC must be stipulated at the level of the coordination construction, by stating that conjuncts cannot be empty elements.[2] On the other hand, the CC follows immediately in a traceless account of filler-gap dependencies (Sag & Fodor 1994; Bouma et al. 2001; Ginzburg & Sag 2000; Sag 2010) since there is simply nothing to conjoin in (3.1), and thus nothing else needs to be said about conjunct extraction; see Sag (2000) for more criticism of traces.

HPSG's traceless account of the CC is semantic in nature, in a sense. Coordinators like *and*, *or*, *but* and so on are not regarded as heads that select arguments, and therefore have empty ARG-ST and valence specifications. And given that HPSG assumes that the signs that can appear in a given lexical head SLASH values are valents, then it follows that the signs that coordinators combine with cannot instead be registered in the coordinator's SLASH feature. Hence, words like *and* have no valents, no arguments and therefore no conjunct extraction. Incidentally, adnominal adjectives cannot be extracted either, for exactly the same reason, as they are not selected by any head, and therefore are not listed in any ARG-ST list.

In order to allow certain adverbials to be extractable, Ginzburg & Sag (2000) assume that those adverbials are members of ARG-ST. See Levine & Hukari (2006) for more on adverbial extraction, and see Borsley & Crysmann (2019), Chapter 14 of this volume for further discussion.[3]

[2] See however Levine (n.d.: 317–318) for the claim that each conjunct must contain at least one stressed syllable. Given that traces are phonologically silent, nothing is there to bear stress and the CC is obtained. This raises the question of why no such stress constraint exists in P-stranding, for example, or indeed in any kind of extraction.

[3] The empirical facts are less clear when it comes to adnominal PPs, however. Even PPs that are usually regarded as modifiers can sometimes be extracted, as in *From which shelf am I not*

Let us now turn to the *Element Constraint*, illustrated in (7). As before, the constraint appears to be restricted to coordination structures, as no oddness arises in the comitative counterparts like (8), or in comparatives like (9).[4]

(7) a. * Which celebrity did you see [Priscilla and a picture of _]?
(cf. with 'Did you see Priscilla and a picture of Elvis?')
b. * Which celebrity did you see [a picture of _ and Priscilla]?
(cf. with 'Did you see a picture of Elvis and Priscilla?')

(8) a. Which celebrity did you see [the brother of _ with Priscilla]?
b. Which celebrity did you see [Priscilla with the brother of _]?

(9) a. Which celebrity did [[you enjoy the memoir of _ more] than [any other non-fiction book]]?
b. Which celebrity did you say that [[the sooner we take a picture of _], [the quicker we can go home]]?

The ATB exception to the CSC is illustrated by the acceptability of (10), where each conjunct hosts a gap, linked to the same filler. As already noted above in (1), the fact that multiple gaps can be linked to the same filler is not unique to coordination.

(10) a. Which celebrity did you buy [[a picture of _ and a book about _]]?
b. Which celebrity did you [[meet _ at a party] and [date _ for a few months]]?

---

supposed to read any books? In many such extractions the PP can alternatively be parsed as VP modifier, which complicates judgements.

[4] Although Winter (2001: 83) and others claim that coordination imposes semantic scope islands, Chaves (2007: §3.6) shows that this is not the case, as illustrated in examples like those below.

(i) a. The White House is very careful about this. An official representative [[will personally read each document] and [reply to every letter]].
(∀ doc-letter >∃ representative / ∃ representative >∀ doc-letter)
b. We had to do this ourselves. By the end of the year, some student [[had proof-read every document] and [corrected each theorem]].
(∀ doc-theorem >∃ student / ∃ student >∀ doc-theorem)
c. Your task is to document the social interaction between [[each female] and [an adult male]].
(∀ female >∃ adult male / ∃ adult male >∀ female)

Gazdar (1981) and Gazdar et al. (1985) assumed that the coordination rule requires SLASH values to be structure-shared across conjuncts and the mother node, thus predicting both the Element Constraint and the ATB exceptions. The failure of movement-based grammar to predict multiple gap extraction facts was also seen as a major empirical advantage of GPSG/HPSG. A similar constraint is assumed in Pollard & Sag (1994: 202) and Beavers & Sag (2004: 60), among others, illustrated in (11). See Abeillé & Chaves (2019), Chapter 17 of this volume for more discussion about coordination.

(11) Coordination Construction (abbreviated)

$$\textit{coordinate-phr} \rightarrow \begin{bmatrix} \text{SYNSEM} \mid \text{NONLOCAL} \mid \text{SLASH} \ \boxed{1} \\ \text{DTRS} \left\langle \begin{bmatrix} \text{SYNSEM} \mid \text{NONLOCAL} \mid \text{SLASH} \ \boxed{1} \end{bmatrix}, \begin{bmatrix} \text{SYNSEM} \mid \text{NONLOCAL} \mid \text{SLASH} \ \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix}$$

Because the SLASH value $\boxed{1}$ is structure-shared between the mother and the daughters in (11), all three nodes must bear the same SLASH value. This predicts the CSC and the ATB exceptions straightforwardly. The failure of mainstream Chomskyan grammar to predict these and related multiple gap extraction facts in a precise way is regarded as one of the major empirical advantages of HPSG over movement-based accounts.

But the facts about extraction in coordination structures are more complex than originally assumed, and than (11) allows for. A crucial difference between the Conjunct Constraint and the Element Constraint is that the latter is only in effect if the coordination has a symmetric interpretation (Ross 1967; Goldsmith 1985; Lakoff 1986; Levin & Prince 1986), as in (12).[5]

(12)
a. Here's the whiskey which I [[went to the store] and [bought _]].
b. Who did Lizzie Borden [[take an ax] and [whack _ to death]]?
c. How much can you [[drink _] and [still stay sober]]?

The coordinate status of (12) has been questioned since Ross (1967). After all, if these are subordinate structures rather than coordinate structures, then the possibility for non-ATB long-distance dependencies ceases to be exceptional. But as Schmerling (1972), Lakoff (1986), Levine (2001) and Kehler (2002) point out,

[5] In asymmetric coordination, the order of the conjuncts has a major effect on the interpretation. Thus, *Robin jumped on a horse and rode into the sunset* does not mean the same as *Robin rode into the sunset and jumped on a horse*. Conversely, in symmetric coordination the order of the conjuncts has no interpretational differences, as illustrated by the paraphrases *Robin drank a beer and Sue ate a burger* and *Sue ate a burger and Robin drank a beer*.

there is no empirical reason to assume that the examples in (12) are anything other than coordination structures.

Another reason to reject the idea that the SLASH values of the daughters and the mother node are are simply equated in ATB extraction is the fact that sometimes multiple gaps are 'cumulatively' combined into a 'pluralic gap'.[6] As an example, consider the extractions in (13). There are two possible interpretations for such extractions: one in which the *ex situ* signs (i.e. the gap signs) and the filler phrase are co-indexed, and therefore co-referential, and a second reading in which the two *ex situ* phrases are not co-indexed even though they are linked to the same filler phrase. Rather, the filler phrase refers to a plural referent composed by the referents of the *ex situ* signs, as indicated by the subscripts in (13). For different speakers, the preferred reading is the former, and in other cases, the latter, often depending on the example.

(13) a. [What]$_{\{i,j\}}$ did Kim eat _$_i$ and drink _$_j$ at the party?
(answer: 'Kim at pizza and drank beer')

b. [Which city]$_{\{i,j\}}$ did Jack travel to _$_i$ and Sally decide to live in _$_j$?
(answer: 'Jack traveled to London and Sally decided to live in Rome')

c. [Who]$_{\{i,j\}}$ did the pictures of _$_i$ impress _$_j$ the most?
(answer: 'Robin's pictures impressed Sam the most')

d. [Who]$_{\{i,j\}}$ did the rivals of _$_i$ shoot _$_j$?
(answer: 'Robin's rivals shot Sam')

e. [Who]$_{\{i,j\}}$ did you send nude photos of _$_i$ to _$_j$?
(answer: 'I sent photos of Sam to Robin')

In conclusion, the non-ATB exceptions in (12) suggest that the coordination rule should not constrain SLASH at all, as argued for in Chaves (2003). Rather, the Element Constraint, its ATB exceptions in (10a,b) and the asymmetric non-ATB exceptions in (12) are more likely to be the consequence of an independent semantic-pragmatic constraint that requires the filler phrase to be 'topical' relative to the clause (Lakoff 1986; Kuno 1987; Kehler 2002; Kubota & Lee 2015). Thus, if the coordination is symmetric, then the topicality requirement distributes over each conjunct, to require that the filler phrase be topical in each conjunct. Consequently, extraction must be ATB in symmetric coordination. No distribution needs to take place in asymmetric coordination, and thus both ATB and non-ATB

[6] See for example Munn (1998; 1999), Postal (1998: 136, 160), Kehler (2002: 125), Gawron & Kehler (2003), Zhang (2007), Chaves (2012a), and Vicente (2016).

extraction is licit in asymmetric coordination. For an attempt to transfer some of Kuno's and Kehlers' insights into HPSG see Chaves (2003). In the latter proposal, the coordination rule is like most other rules in the grammar in that it says nothing about the SLASH values of the mother and the daughters, along the lines of Levine & Hukari (2006: 354). In other words, the constraints on SLASH in (11) are unnecessary. Rather, pragmatics is the driving force behind how long-distance dependencies propagate one or more conjuncts, depending on the coordination being interpreted symmetrically or not. See Abeillé & Chaves (2019), Chapter 17 of this volume for more discussion.

Let us take stock. The CSC does not receive a unitary account in modern HPSG, given that the Conjunct Constraint and the Element Constraint are of a very different nature. Whereas the former does not admit ATB extraction, and is predicted by a traceless analysis, the latter allows ATB extraction as seen by the contrast between (3.1c) and (10). Upon closer inspection, the Element Constraint and the ATB exceptions are semantic-pragmatic in nature. As we shall see, a similar conclusion is plausible for various other island phenomena.

4 Complex NP Constraint

The Complex NP Constraint concerns the difficulty in extracting out of complex NPs formed with either relative clauses (14) or complement phrases (15).

(14) a. * [What]$_i$ does Robin know [someone who has _$_i$]?
(cf. with 'Does Robin know someone who has a drum kit?')

b. * [Which language]$_i$ did they hire [someone [who speaks _$_i$]]?
(cf. with 'Did they hire someone who speaks Arabic?')

(15) a. * [Which book]$_i$ do you believe the claim [that Robin plagiarized _$_i$]?
(cf. with 'Do you believe the claim that Robin plagiarized *this book*?')

b. * What$_i$ did you believe [the rumor [that Ed disclosed _$_i$]]?
(cf. with 'Did you believe [the rumor [that Ed disclosed *that*]]?')

It is tempting to prevent extractions out of adnominal clauses by simply stipulating that the SLASH value of the modifier must be empty, as (4) illustrates. Perhaps, along the lines of Fodor (1978; 1983b), Berwick & Weinberg (1984), and Hawkins (1999; 2004), processing difficulties lead to the grammaticization of such a constraint, effectively blocking any modified head from hosting any gaps.

(16) HEAD-MODIFIER CONSTRUCTION (abbreviated)

$$\textit{head-mod-phr} \rightarrow \begin{bmatrix} \text{HEAD-DTR} \ \boxed{1} \\ \text{DTRS} \left\langle \boxed{1}, \begin{bmatrix} \text{SYNSEM} \begin{bmatrix} \text{LOC} \mid \text{MOD} \ \boxed{1} \\ \text{NONLOC} \mid \text{SLASH} \ \{\} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

However, the robustness of the CNPC has been challenged by various counterexamples over the years (Ross 1967; Pollard & Sag 1994; Kluender 1998; Postal 1998; Sag et al. 2007). The sample in (17) involves acceptable extractions from NP-embedded complement CPs (some of which are definite), and (18) involves acceptable extractions from NP-embedded relative clauses.[7]

(17) a. The money which I am making [the claim [that the company squandered _]] amounts to $400,000.
(Pollard & Sag 1994: 206, 207)

b. Which rebel leader did you hear [rumors [that the CIA assassinated _]]?

c. Which company did Simon spread [the rumor [that he had started _]]?

d. What did you get [the impression [that the problem really was _]]?
(Kluender 1998)

(18) a. This is the kind of weather$_i$ that there are [many people [who like _$_i$]].
(Erteschik-Shir & Lappin 1979)

b. Violence is something$_i$ that there are [many Americans [who condone _$_i$]].
(McCawley 1981: 108)

c. There were several old rock songs$_i$ that she and I were [the only two [who knew _$_i$]].
(Chung & McCloskey 1983)

d. This is the chapter$_i$ that we really need to find [someone [who understands _$_i$]].
(Kluender 1992: 238)

[7] Counterexamples to the CNPC can be found in a number of languages, including Japanese and Korean (Kuno 1973; Nishigauchi 1999), Ahan (Saah & Goodluck 1995), Danish (Erteschik-Shir 1973: Chapter 2), Swedish (Allwood 1976; Engdahl 1982) Norwegian (Taraldsen 1982) and Romance languages (Cinque 2010). In some languages that support verb constructions, the CNPC is not active, which goes with the analysis as complex predicates Abeillé & Vivès (2019).

e. Which diamond ring did you say there was [nobody in the world [who could buy _ $_i$]]? (Pollard & Sag 1994: 206)

f. John is the sort of guy that I don't know [a lot of people [who think well of _ $_i$]]. (Culicover 1999: 230)

In the above counterexamples, the relative clauses contribute to the main assertion of the utterance, rather than expressing background information. For example, (18a) asserts 'There are many people who like this kind of weather' rather than 'This the kind of weather', and so on. Some authors have argued that it is precisely because such relatives express new information that the extraction can escape the embedded clause (Erteschik-Shir & Lappin 1979; Kuno 1987; Deane 1992; Goldberg 2013). If this is correct, then the proper account of CNPC effects is not unlike that of the CSC. In both cases, the information structural status of the clause that contains the gap is crucial to the acceptability of the overall long-distance dependencies.[8]

In addition to pragmatic constraints, Kluender (1992; 1998) proposed that processing factors also influence the acceptability of CNPC violations. Consider for example the acceptability hierarchy in (19); more specific filler phrases increase acceptability, whereas the presence of more specific phrases between the filler and the gap seem to cause increased processing difficulty, and therefore lower the acceptability of the sentence. The symbol '<' reads as 'is less acceptable than'.

(19) a. What do you need to find the expert who can translate _ ? <

b. What do you need to find an expert who can translate _ ? <

c. What do you need to find someone who can translate _ ? <

d. Which document do you need to find an expert who can translate _ ?

[8] Although it is sometimes claimed that such island effects are also active in logical form and semantic scope (May 1985; Ruys 1993; Fox 2000; Sabbagh 2007; Bachrach & Katzir 2008), there much reason to be skeptical. For example, the universally quantified noun phrase in (i) is embedded in a relative clause can have wide scope over the indefinite *someone*, constituting a semantic CNPC violation. Note that these relatives are not presentational, and therefore are not specially permeable to extraction.

(i) a. We were able to find someone who was an expert on each of the castles we planned to visit. (Copestake et al. 2005: 304)

b. John was able to find someone who is willing to learn every language that we intend to study. (Chaves 2014)

There is on-line sentence processing evidence that CNPC violations with more informative fillers are more acceptable and are processed faster at the gap site than violations with less informative fillers (Hofmeister & Sag 2010), as in (20).

(20) a. ? Who did you say that nobody in the world could ever depose _ ?
b. Which military dictator did you say that nobody in the world could ever depose _ ?

The same difference in reading times is found in sentences without CPNP violations, in fact. For example, (21b) was found to be read faster at *encouraged* than (21a). Crucially, that critical region of the sentence is not in the path of any filler-gap dependency.

(21) a. The diplomat contacted the dictator who the activist looking for more contributions encouraged to preserve natural habitats and resources.
b. The diplomat contacted the ruthless military dictator who the activist looking for more contributions encouraged to preserve natural habitats and resources.

Given that finite tensed verbs can be regarded as definite, and infinitival verbs as indefinite (Partee 1984), and given that finiteness can create processing difficulty (Kluender 1992; Gibson 2000), then acceptability clines like (22) are to be expected. See Levine & Hukari (2006: Chapter 5) and Levine (n.d.: 308) for more discussion.

(22) a. Who did you wonder what Mary said to _ ? <
b. Who did you wonder what to say to _ ? <
c. Which of the people at the party did you wonder what to say to _ ?

4.1 On D-Linking

The amelioration caused by more specific (definite) *wh*-phrases as in (19d), (20b) and (22c) has been called a 'D-Linking' effect (Pesetsky 1987; 2000). It purportedly arises if the set of possible answers is pre-established or otherwise salient. But there are several problems with the D-Linking story. First, there is currently no non-circular definition of D-Linking; see Pesetsky (2000: 16), Ginzburg & Sag (2000: 247–250), Chung (1994: 33, 39) and Levine & Hukari (2006: 242, 268–271). Second, the counterexamples above are given out of the blue, and therefore cannot evoke any preexisting set of referents, as D-Linking requires. Furthermore, nothing should prevent D-Linking with a bare *wh*-item, as Pesetsky himself acknowledges, but on the other hand there is no experimental evidence that context can lead to D-linking of a bare *wh*-phrase (Sprouse 2007; Villata et al. 2016).[9]

Kluender & Kutas (1993), Sag et al. (2007), Hofmeister (2007b), Hofmeister (2007a) and Hofmeister & Sag (2010) argue that more definite *wh*-phrases improve the acceptability of extractions because they resist memory decay better than indefinites, and are compatible with fewer potential gap sites. In addition, Kroch (1989) and Levine & Hukari (2006: 270) point out that D-Linking amelioration effects may simply result from the plausibility of background assumptions associated with the proposition.

4.2 On memory limitations

Sprouse et al. (2012a) use *n*-back and serial recall tasks to argue that there is no evidence that working memory limitations correlate with island acceptability, and therefore that the 'processing-based' account of islands put forth by Kluender (1992; 1998), Kluender & Kutas (1993), Hofmeister & Sag (2010) and others is unfounded. To be sure, it cannot be stressed enough that the accounts in Kluender (1992) and Hofmeister & Sag (2010) are not strictly based on performance, and involve other factors as well, most notably plausibility and pragmatic factors. See in particular Hofmeister et al. (2013: 49), where it is argued that at least some extraction constraints may be due to a combination of syntactic, semantic, pragmatic, and performance factors. Basically, if the correct location of a gap is syntactically, semantically, or pragmatically highly unlikely in that particular utterance, then it is less likely for the sentence to be acceptable. Indeed, there is independent experimental evidence that speakers attend to probabilistic information about the syntactic distribution of filler-gap dependencies (van Schijndel et al. 2014), and that gap predictability is crucial for on-line processing of islands (Michel 2014).[10] But as Sprouse et al. (2012b) point out, there is no reason to believe that *n*-back and serial recall tasks are strongly correlated to working memory capacity to begin with. Second, one of the main points of Hofmeister

[9]For more detailed criticism of D-Linking see Hofmeister et al. (2007).

[10]More broadly, there is good evidence that speakers deploy probabilistic information when processing a variety of linguistic input, including words (Altmann & Kamide 1999; Arai & Keller 2013; Creel et al. 2008; DeLong et al. 2005; Kutas & Hillyard 1984), lexical categories (Gibson 2006; Levy & Keller 2013; Tabor et al. 1997), syntactic structures (Levy et al. 2012; Lau et al. 2006; Levy 2008; Staub & Clifton 2006), semantics (Altmann & Kamide 1999; Federmeier & Kutas 1999; Kamide et al. 2003), and pragmatics (Ni et al. 1996; Mak et al. 2008; Roland et al. 2012).

& Sag (2010) is that the literature on experimental island research has not systematically controlled for multiple factors that can impact the processing and comprehension of complex sentences. If the experimental items are excessively complex, then readers are more likely to give up understanding the utterances and subtler effects will not be measurable. Phillips (2013b), however, regard such concerns as irrelevant. Although it is unclear to what extent expectations and processing constraints contribute to island effects, it is likely that they play some role in CNPC effects, as well as other island types discussed below.

5 Right Roof Constraint

Rightward movement is traditionally regarded as being clause bounded. Such *Right Roof Constraint* (Ross 1967) effects are illustrated in (23), in which a phrase appears *ex situ* in a position to the right of its *in situ* counterpart; see Akmajian (1975), Baltin (1978), and Stowell (1981), among others.

(23) a. * I [met a man [who knows _$_i$] yesterday] [all of your songs]$_i$.
b. * [[That a review _$_i$ came out yesterday] is catastrophic] [of this article]$_i$.
c. * It was believed _$_x$ that [there walked into the room] _$_y$ [by everyone]$_x$ [a man with long blond hair]$_y$.
(Rochemont 1992)

When treated as a form of extraction, rightward movement has been predominantly accounted for via a feature EXTRA(POSED) (Keller 1995; van Noord & Bouma 1996; Van Eynde 1996; Keller 1994; Müller 1999; Kim & Sag 2005), rather than by SLASH. Thus, RRC island effects can be easily modeled by stipulating that the EXTRA value of an S node must be empty. One way to do so is to state that any S dependent (valent or adjunct) must be [EXTRA {}]. Thus, no extraposed element may escape its clause. However, the oddness of (23) may not be due any such syntactic stipulation, given the acceptability of counterexamples like (24). Note that the adverbial interveners in such examples do not require parenthetical prosody. Conversely, even strong parenthetical prosody on the adverbs in (23) fails to improve those data.

(24) a. I've [been requesting [that you pay back _] [ever since May]] [the money I lent to you a year ago].
(Kayne 2000: 251)

b. I've [been wanting to [meet someone who KNOWS _] [ever since I was little]] [exactly what happened to Amelia Earhart].

c. I've been wondering if it is possible _] [for many years now] [for anyone to memorize the Bible word for word].
(Chaves 2014: 861)

The durative semantics of *I've been wanting/requesting/wondering* raises an expectation about the realization of a durative adverbial expression like *ever since* or *for many years* that provides information about the durative semantics of the main predicate. Hence, the adverb is cued by the main predication, in some sense, and coheres much better in a high attachment than with a lower one.

The fact that the RRC is prone to exceptions has been noted by multiple authors as the sample in (25) illustrates. In all such cases, a phrase is right-extracted from an embedded clause, which should be flat out impossible if extraposition is clause-bounded. Again, the adverbial interveners in (25) do not require any special prosody, which means that these data cannot be easily discarded as parenthetical insertions.

(25) a. I have [wanted [to know _] for many years] [exactly what happened to Rosa Luxemburg].
(attributed to Witten 1972 in Postal 1974: 92n)

b. I have [wanted [to meet _] for many years] [the man who spent so much money planning the assassination of Kennedy].
(attributed to Janet Fodor (p.c.) in Gazdar 1981: 177)

c. Sue [kept [regretting _] for years] [that she had not turned him down].
(Van Eynde 1996)

d. She has been [requesting that he [return _] [ever since last Tuesday]] [the book that John borrowed from her last year].
(Kayne 2000: 251)

e. Mary [wanted [to go _] until yesterday] [to the public lecture].
(Howard Lasnik 2007 course handout[11])

Further evidence against a syntactic account of RRC comes from corpora (Müller 2004; 2007) and experimental findings (Strunk & Snider 2008; 2013), which confirm that extraposition does not always obey island constraints. The counterexamples in (26a–c) are adapted from Strunk & Snider (2008) and Strunk & Snider (2013), and those in (26d–f) are from Chaves (2014: 863).

[11] http://ling.umd.edu/~lasnik/LING819%202007/Multiple%20Sluicing%20819%20.pdf; Retr. 2009.

(26) a. [In [what noble capacity _]] can I serve him [that would glorify him and magnify his name]?
b. We drafted [a list of basic demands _] last night [that have to be unconditionally met or we will go on strike].
c. For example, we understand that Ariva buses have won [a number of contracts for routes in London _] recently, [which will not be run by low floor accessible buses].
d. Robin bought [a copy of a book _] yesterday [about ancient Egyptian culture].
e. I'm reading [a book written by a famous physicist _] right now, [who was involved in the Manhattan Project].
f. I saw [your ad in a magazine _] yesterday [on the table at the dentist office].

Grosu (1973b), Gazdar (1981) and Stucky (1987) argued that the RRC is the result of performance factors such as syntactic and semantic parsing expectations and memory resource limitations, not grammar proper. Indeed, we now know that there is a general well-known tendency for the language processor to prefer attaching new material to the more recent constituents (Frazier & Clifton 1996; Gibson et al. 1996; Traxler et al. 1998; Fodor 2002; Fernández 2003). Indeed, eye-tracking studies like Staub et al. (2006) indicate that the parser is reluctant to adopt extraposition parses. This explains why extraposition in written texts is less common in proportion to length of the intervening material (Uszkoreit et al. 1998): the longer the structure, the bigger the processing burden. Crucially, however, the preference for the closest attachment can be weakened by many factors (Fernández 2003; Desmet et al. 2006; De Vicenzi & Job 1993; Carreiras 1992). For example, Levy et al. (2012) show that relative clause extraposition creates significant processing difficulty when compared with non-extraposed counterparts of the same sentences, but that a preceding context that sets up a strong expectation for a relative clause modifying a given noun can facilitate comprehension of an extraposed relative clause modifying that noun. In other words, in spite of a larger processing burden, some extrapositions can be made easier to process by parsing expectations.

A detailed account of extraposition island phenomena does not exist in any framework, as far as I am aware. But the line of inquiry first proposed by Grosu (1973b), Gazdar (1981) and Stucky (1987), and later experimentally supported by Levy et al. (2012), Strunk & Snider (2008), and Strunk & Snider (2013) seems to be on the right track. If so, then there is no syntactic constraint on EXTRA.

Rather, RCC effects are to a large extent the result of difficulty in integrating the extraposed phrase in its *in situ* position.

5.1 Freezing

A related island phenomenon also involving rightward displacement, first noted in Ross (1967: 305), is *Freezing*: leftward extraction (27a) and extraposition (27b) cause low acceptability when they interact, as seen in (28). In (28a) there is extraction from an extraposed PP, in (28b) there is extraction from an extraposed NP, and in (28c) an extraction from a PP crossed with direct object extraposition.

(27) a. Who$_j$ did you [give [a picture of _$_j$] [to Robin]]?
b. Did you [give _$_i$ [to Robin] [a picture of my brother]$_i$]?

(28) a. * Who$_j$ did you [give a picture _$_i$] [to Robin] [of _$_j$]$_i$?
b. * Who$_j$ did you [give _$_i$ [to Robin] [a picture of _$_j$]$_i$]?
c. * Who$_j$ did you [give _$_i$ [to _$_j$] [a picture of my brother]$_i$]?

Fodor (1978: 457) notes that (28c) has a syntactically highly probable temporary alternative parse in which *to* combines with the NP *a picture of my brother*. The existence of this local ambiguity likely disrupts parsing, especially as it occurs in a portion of the sentence that contains two gaps in close succession. Indeed, constructions with two independent gaps in close proximity are licit, but not trivial to process, as seen in (29), specially if the extraction paths cross (Fodor 1978), as in (29b).

(29) a. * This is a problem which$_i$ John$_j$ is difficult to talk to _$_j$ about _$_i$.
b. * Who$_j$ can't you remember which papers$_i$ you sent copies of _$_i$ to _$_j$?

A similar analysis is offered by Hofmeister et al. (2015: 477), who note that constructions like (28c) must cause increased processing effort since the point of retrieval and integration coincides with the point of reanalysis. The existence of a preferential alternative parse that is locally licit but globally illicit can in turn lead to a "digging-in" effect (Ferreira & Henderson 1991; 1993; Tabor & Hutchins 2004), in which the more committed the parser becomes to a syntactic parse, the harder it is to backtrack and reanalyze the input. The net effect of these factors is that the correct parse of (28c) is less probable and therefore harder to identify than that of (28b), which suffers from none of these problems, and is regarded to be more acceptable than (28c) by Fodor (1978: 453) and others. See Chaves (2018) for experimental evidence that speakers can adapt and to some extent overcome some of these parsing biases.

Finally, prosodic and pragmatic factors are likely also at play in (28), as in the RRC. Huck & Na (1990) show that when an unstressed stranded preposition is separated from its selecting head by another phrase, oddness ensues for prosodic reasons. Finally, Huck & Na (1990) and Bolinger (1992) also argue that freezing effects are also in part due to a pragmatic conflict created by extraposition and extraction: *wh*-movement has extracted a phrase leftward, focusing interest on that expression, while at the same time extraposition has moved a constituent rightward, focusing interest on that constituent as well. Objects tend to be extraposed when they are discourse new, and even more so when they are heavy (Wasow 2002: 71). Therefore, the theme phrase *a picture of John* in (28c) is strongly biased to be discourse new, but this clashes with the fact that an entirely different entity, the recipient, is leftward extracted, and therefore is the *de facto* new information that the open proposition is about. No such mismatch exists in (28a) or (28b), in contrast, where the extraposed theme is more directly linked to the entity targeted by leftward extraction.

6 Subject islands

Extraction out of subject phrases like (30) is broadly regarded to be impossible in several languages, including English (Ross 1967; Chomsky 1973), an effect referred to as a *Subject Island* (SI). This constraint is much less severe in languages like Japanese , German, and Spanish, among others (Stepanov 2007; Jurka et al. 2011; Goodall 2011; Sprouse et al. 2015; Fukuda et al. 2018; Polinsky et al. 2013).

(30) a. * Who did stories about terrify John?
(Chomsky 1977: 106)

b. * Who was a picture of laying there?
(Kayne 1981: 114)

c. * Who do you think pictures of would please John?
(Huang 1982: 497)

d. * Who does the claim that Mary likes upset Bill?
(Lasnik & Saito 1992: 42)

e. * Which candidate were posters of all over the town?
(Lasnik & Park 2003)

However, English exceptions were noticed early on, and have since accumulated in the literature. In fact, for Ross (1967), English extractions like (31a) are not illicit, and more recently Chomsky (2008: 147) has added more such counterexamples. Other authors noted that certain extractions from subject phrases are naturally attested, as in (31b,c). Indeed, Abeillé et al. (2018) shows that extractions like those in (31c) are in fact acceptable to native speakers, and that no such island effect exists in French either.[12]

(31) a. [Of which cars]$_i$ were [the hoods _$_i$] damaged by the explosion? (Ross 1967: 4.252)

b. They have eight children [of whom]$_i$ [five _$_i$] are still living at home.
(Huddleston et al. 2002: 1093)

c. Already Agassiz had become interested in the rich stores of the extinct fishes of Europe, especially those of Glarus in Switzerland and of Monte Bolca near Verona, [of which]$_i$, at that time, [only a few _$_i$] had been critically studied.
(Santorini 2007)

English exceptions to the SI constraint are not restricted to PP extractions, however. Although Ross (1967) claimed NP extractions from NP subjects like (32) are illicit, it was arguably premature to generalize from such a small sample.

(32) a. The hoods of these cars were damaged by the explosion.

b. *Which cars were the hoods of damaged by the explosion?
(Ross 1967)

Indeed, a number of authors have noted that some NP extractions from subject NPs are either passable or fairly acceptable, as illustrated in (33). See also Pollard & Sag (1994: 195, ft. 32), Postal (1998), Sauerland & Elbourne (2002: 304), Culicover (1999: 230), Levine & Hukari (2006: 265), Chaves (2012b: 470, 471), and Chaves & Dery (2014).

(33) a. [What]$_i$ were [pictures of _$_i$] seen around the globe?
(Kluender 1998: 268)

b. It's [the kind of policy statement]$_i$ that [jokes about _$_i$] are a dime a dozen.
(Levine et al. 2001: 204)

[12]For completeness, other authors argue that PP extractions from NP subjects are illicit, such as Lasnik & Park (2003: 653), among many others.

c. There are [certain topics]$_i$ that [jokes about _$_i$] are completely unacceptable.
(Levine & Sag 2003: 252, ft. 6)

d. [Which car]$_i$ did [some pictures of _$_i$] cause a scandal?
(Jiménez–Fernández 2009: 111)

e. [What]$_i$ did [the attempt to find _$_i$] end in failure?
(Hofmeister & Sag 2010: 370)

f. [Which president]$_i$ would [the impeachment of _$_i$] cause outrage?
(Chaves 2012b)

g. I have a question$_i$ that [the probability of you knowing the *answer to* _$_i$] is zero.
(Chaves 2013)

Whereas SI violations involving subject CPs are not attested, those involving infinitival VP subjects like (34) are. See Chaves (2012b: 471) for more natural occurrences.

(34) a. The eight dancers and their caller, Laurie Schmidt, make up the Farmall Promenade of nearby Nemaha, a town$_i$ that [[to describe _$_i$ as tiny] would be to overstate its size].
(Huddleston et al. 2002: 1094, ft.27)

b. In his bedroom, [which]$_i$ [to describe _$_i$ as small] would be a gross understatement, he has an audio studio setup.

c. They amounted to near twenty thousand pounds, [which]$_i$ [to pay _$_i$] would have ruined me. (Benjamin Franklin, William Temple Franklin and William Duane. 1834. Memoirs of Benjamin Franklin, vol 1. p.58)

Incidentally, subject phrases are not extraposition islands either, as in (35). See also Guéron & May (1984). Odd examples like *[*Pictures* _] *frighten people* [*of John*] from Drummond (2009), are more likely due to a digging-in effect, caused by speakers assuming that the subject is syntactically and semantically complete by the end of the verb phrase.

(35) a. [The circulation of a rumor _$_i$] has started [that Obama will not seek re-election]$_i$.

b. [A copy of a new book _$_i$] arrived yesterday [about ancient Egyptian culture]$_i$.

c. [Concerns about the deaths _ $_i$] were raised [of two diplomatic envoys recently abducted in Somalia]$_i$.

6.1 Clausal Subject Constraint

Let us now consider SI effects involving more complex subjects. Infinitival subject clauses seem to impose no SI constraint, an observation going back to Kuno & Takami (1993), but noted elsewhere a few times:

(36) a. This is something [which]$_i$ – for you to try to understand _ $_i$ – would be futile.
(Kuno & Takami 1993: 49)

b. I just met Terry's eager-beaver research assistant [who]$_i$ – for us to talk to _ $_i$ about any subject other than linguistics – would be absolutely pointless.
(Levine & Hukari 2006: 265)

c. There are [people in this world]$_i$ that – for me to describe _ $_i$ as despicable – would be an understatement.
(Chaves 2012b: 471).

Infinitival subjects contrast dramatically with finite subjects. The latter are renowned for being particularly hard to extract from, as in (37). Ross (1967) dubbed this extreme kind of SI the *Sentential Subject Constraint* (SSC). See also Chomsky (1973), Huang (1982), Chomsky (1986), and Freidin (1992).[13]

(37) a. * [Who]$_i$ did [that Maria Sharapova beat _ $_i$] surprise everyone?
(cf. with 'That Maria Sharapova beat Serena Williams surprised everyone')

b. * [Who]$_i$ did [that Robin married _ $_i$] surprise you?
(cf. with 'Did that Robin married Sam surprise you?')

There are some functional reasons for why clausal SI violations may be so strong. First, subject clauses are notorious for being particularly difficult to process, independent of extraction. Clausal subjects are often stylistically marked

[13]That said, Chaves (2013) reports that some native speakers find SSC violations like (i) to be fairly acceptable, again raising some doubt about the robustness of English SI effects:

(i) [Which actress]$_i$ does [whether Tom Cruise marries _ $_i$] make any difference to you?

and difficult to process, as (38a) illustrates. Thus, it is extremely hard to embed a clausal subject within another clausal subject, even though such constructions ought to be perfectly grammatical, like (38b, c). In addition, it is known that tense can induce greater processing costs (Kluender 1992; Gibson 2000).

(38) a. That the food that John ordered tasted good pleased him. (Cowper 1976; Gibson 1991)
b. * That that Jill left bothered Sarah surprised Max. (Kimball 1973)
c. * That that the world is round is obvious is dubious. (Kuno 1974)

Interestingly, clausal subjects become more acceptable if extraposed as shown in (39). The explanation offered by Fodor et al. (1974: 356–357) is that speakers tend to take the initial clause in the sentence to be the main clause. Thus, *that* is taken to be the subject, but the remainder of the structure does not fit this pattern. Thus, a sentence like (39a) causes increased processing load because it has a different structure than the parser expects. This processing problem does not arise in the counterpart in (39b).[14]

(39) a. ? That [it is obvious that [the world is round]] is dubious.
b. It is dubious that [it is obvious that [the world is round]]. (Kuno 1974)

Indeed, Fodor & Garrett (1967), Bever (1970), and Frazier & Rayner (1988) also show that extraposed clausal subject sentences like (40a) are easier to process than their *in-situ* counterparts like (40b). Not surprisingly, the former are much more frequent than the latter, which explains why the parser would expect the former more than the latter.

(40) a. It surprised Max that Mary was happy.
b. That Mary was happy surprised Max.

[14] See Gibson (2006) for online evidence that the word *that* is preferentially interpreted as a determiner even in syntactic contexts where it cannot be a determiner. The use of 'determiner' corresponds to the traditional term, referring to a certain category of prenominal constituent rather than to the whole nominal phrase including the noun and all its dependents. Gibson's evidence suggests that both top-down (syntactic) expectations are independent from bottom-up (lexical) frequency-based expectations in sentence processing. Thus, clausal subject phrase starting with the complementizer *that* is likely to be misparsed as a matrix clause with sentence-initial pronominal or determiner *that*.

Note that embedded conditionals like (41a) are also exceedingly difficult to parse, although perfectly syntactically well-formed. However, such structures can become more acceptable with fewer embeddings to strain working memory, and with judicious use of prosodic cues that indicate the relevant structure, as illustrated by the acceptability of (41b). This is evidence that the oddness of (41) is at least in part due to performance.

(41) If – if every time it snowed there was no school – then there wouldn't be classes all winter.

If we add a filler-gap dependency to a sentence that already is complex by virtue of having a clausal subject, the resulting structure may be too difficult to parse. This point is illustrated by the contrast in (42).

(42) a. ?* What does that he will come prove?
b. What does his coming prove?
(Lewis 1993)

As argued by Davies & Dubinsky (2009), the low acceptability of extraction in subject-auxiliary inversion sentences with clausal subjects is more likely to be the result of extragrammatical factors than of grammatical conditions. For example, not all extractions like (43b) are unacceptable, as Delahunty (1983: 382–387) and Davies & Dubinsky (2009: 115) point out.

(43) a. That the food that John ordered tasted good pleased me.
b. * Who did that the food that John ordered tasted good please _ ?

The evidence discussed so far suggests that sentences involving extraction and clausal subjects are odd at least in part due to the likely cumulative effect of various sources of processing complexity. Sentences with sentential subjects are unusual structures, which can mislead the parser into the wrong analysis. A breakdown in comprehension can occur because the parser must hold complex incomplete phrases in memory while processing the remainder of the sentence. The presence of a filler-gap dependency will likely only make the sentence harder to process. It is independently known that the more committed the parser becomes to a syntactic parse, the harder it is to reanalyze the string (Ferreira & Henderson 1991; 1993; Tabor & Hutchins 2004). For example, unless prosodic or contextual cues are employed to boost the activation of the correct parse, (6.1) will be preferentially misanalysed as having the structure [NP [V [NP]]].

(44) Fat people eat accumulates.

The garden-path effect that the digging-in causes in example (6.1) serves as an analogy for what may be happening in particularly difficult subject island violations. In both cases, the sentences have exactly one grammatical analysis, but that parse is preempted by a highly preferential alternative which ultimately cannot yield a complete analysis of the sentence. Thus, without prosodic cues indicating the extraction site, sentences like (45) induce a significant digging-in effect as well.

(45) a. * Which problem will a solution to be found by you?
b. * Which disease will a cure for be found by you?

This also explains why SI violations like (46) are relatively acceptable: the processing cost of an NP-embedded causal is smaller than the processing cost incurred by processing a clausal subject.[15] Clausal subjects are unusual structures, inconsistent with the parser expectations (Fodor et al. 1974), and the presence of filler-gap dependency in an NP-embedded clausal subject is less likely to cause difficulty for the parse to go awry than a filler-gap dependency in a clausal subject.[16]

(46) a. [Which puzzle]$_i$ did the fact that nobody could solve _$_i$ astonish you the most?
b. [Which crime]$_i$ did the fact that nobody was accused of _$_i$ astonish you the most?
c. [Which question]$_i$ did the fact that none of us could answer _$_i$ surprise you the most?
d. [Which joke]$_i$ did the fact that nobody laughed at _$_i$ surprise you the most?

[15] For claims that NP-embedded clausal SI violations are illicit see Lasnik & Saito (1992: 42), Phillips (2006: 796), and Phillips (2013a: 67).

[16] Clausen (2010; 2011) provide experimental evidence that complex subjects cause a measurable increase in processing load, with and without extraction. Moreover, it is known that elderly adults have far more difficulty repeating sentences with complex subjects than sentences with complex objects (Kemper 1986). Similar difficulty is found in timed reading comprehension tasks (Kynette & Kemper 1986), and in disfluencies in non-elderly adults (Clark & Wasow 1998). Speech initiation times for sentences with complex subjects are also known to be longer than for sentences with simple subjects (Ferreira 1991; Tsiamtsiouris & Cairns 2009), and sentences with center-embedding in subjects are harder to process than sentences with center-embedding in objects (Amy & Noziet 1978; Eady & Fodor 1981). Finally, Garnsey (1985), Kutas et al. (1988), and Petten & Kutas (1991) show that the processing of open-class words, particularly at the beginning of sentences, require greater processing effort than closed-class words.

6.2 Accounts of SI effects

This complex array of effects suggests that the SI constraint is not due to a single factor (Chomsky 2008; Chaves 2013; Jiménez–Fernández 2009), be it grammatical or otherwise. One possibility is that SIs are partly due to pragmatic and processing constraints, perhaps not too different from those that appear to be active in the island effects discussed so far. As Kluender (2004: 495) notes: "Subject Island effects seem to be weaker when the *wh*-phrase maintains a pragmatic association not only with the gap, but also with the main clause predicate, such that the filler-gap dependency into the subject position is construed as of some relevance to the main assertion of the sentence". Indeed, many authors (Erteschik-Shir 1981; Van Valin 1986; Kuno 1987; Takami 1992; Deane 1992; Goldberg 2013) have argued that extraction is in general restricted to the informational focus of the proposition, and that SIs (among others) are predicted as a consequence. In a nutshell, since subjects are typically reserved for topic continuity, subject-embedded referents are unlikely to be the informational focus of the utterance. Although it is not easy to construct sentences where a dependent of the subject can be easily deemed as the informational focus, it is by no means impossible. For instance, (47a) is particularly acceptable because whether or not an impeachment causes outrage crucially depends on who is impeached (cf. with *Would the impeachment of Donald Trump cause outrage?*). Similarly, in (47b) whether or not an attempt failed or succeeded crucially depends on what was attempted (cf. with *The attempt to find the culprit ended in failure*).

(47) a. Which President would [the impeachment of _] cause outrage? (Chaves 2012b)

b. What did [the attempt to find _] end in failure? (Hofmeister & Sag 2010: 370)

Although experimental research has confirmed that sentences with SI violations tend to be less acceptable than grammatical controls (Sprouse 2009; Goodall 2011; Crawford 2011; Clausen 2011; Sprouse et al. 2015), and that their acceptability remains consistently low during repeated exposition (Sprouse 2009; Crawford 2011), other research has found that the acceptability of SI violations is not consistently low, and can be made to increase significantly (Hiramatsu 2000; Clausen 2011; Chaves & Dery 2014; Chaves 2012a). This mixed evidence is consistent with the idea that SI effects are very sensitive to the particular syntax, semantics, and pragmatics of the utterance in which they occur. If the items are too complex, or stylistically awkward, or presuppose unusual contexts, then SI effects are strong. For example, if the extraction is difficult to process because the sentence gives rise to local garden-path and digging-in effects, and is pragmatically infelicitous in the sense that the extracted element is not particularly relevant for the proposition (i.e. unlikely to be what the proposition is about) or comes from the presupposition rather than the assertion, then we obtain a very strong SI effect. Otherwise, the SI effect is weaker, and in some cases nearly non-existent like (47), (34), or the pied-piping examples studied by Abeillé et al. (2018). The latter involve relative clauses, in which subjects are not strongly required to be topics, in contrast to the subjects or main clauses.

This approach also explains why subject-embedded gaps often become more acceptable in the presence of a second non-island gap: since the two gaps are co-indexed, then the fronted referent is trivially relevant for the main assertion, as it is predicated by the main verb. For example, the low acceptability of (48a) is arguably caused by the lack of plausibility of the described proposition: without further contextual information, it is unclear how the attempt to repair an unspecified thing *x* is connected to the attempt causing damage to a car.

(48) a. * What did [the attempt to repair _] ultimately damage the car?
b. What did [the attempt to repair _] ultimately damage _ ? (Phillips 2006)

The example in (48a) becomes more acceptable if it is contextually established that *x* is a component of the car. In contrast, (48b) is felicitous even out-of-the-blue because it conveys a proposition that is readily recognized as being plausible according to world knowledge: attempting to fix *x* can cause damage to *x*. If Subject Island effects are indeed contingent on how relevant the extracted subject-embedded referent is for the assertion expressed by the proposition, then a wide range of acceptable patterns is to be expected, parasitic or otherwise. This includes cases like (49), where both gaps are in SI environments. As Levine & Sag (2003), Levine & Hukari (2006: 256) and Culicover (2013: 161) note, cases like (49) should be completely unacceptable, contrary to fact.

(49) This is a man who [friends of _] think that [enemies of _] are everywhere.

The conclusion that SI effects are contingent on the particular proposition expressed by the utterance and its pragmatics thus seems unavoidable (Chaves & Dery 2019). In order to test this hypothesis, Chaves & Dery (2019) examine the acceptability of sentences like (50), which crucially express nearly-identical truth conditions and have equally acceptable declarative counterparts. This way, any source of acceptability contrast must come from the extraction itself, not from the felicity of the proposition.

(50) a. Which country does the King of Spain resemble [the President of _]?
b. Which country does [the President of _] resemble the King of Spain?

The results indicate that although the acceptability of the SI counterpart in (50b) is initially significantly lower than (50a), it gradually improves. After eight exposures, the acceptability of near-truth-conditionally-equivalent sentences like (50) becomes non-statistically different. What this suggests is that SI effects are at least in part probabilistic: the semantic, syntactic and pragmatic likelihood of a subject-embedded gap likely matters for how acceptable such extractions are. This is most consistent with the claim that – in general – extracted phrases must correspond to the informational focus of the utterance (Erteschik-Shir 1981; Van Valin 1986; Kuno 1987; Takami 1992; Deane 1992; Goldberg 2013), and in particular with the intuition that SI violations are weaker when the extracted referent is relevant for the main predication (Kluender 2004: 495).

7 Adjunct islands

Ross (1967) and Cattell (1976) noted that adjunct phrases often resist extraction, as illustrated in (51), a phenomenon usually referred to as *The Adjunct Island Constraint* (AIC).

(51) a. * What$_i$ did John die [whistling _$_i$]?
b. * What$_i$ did John build [whistling _$_i$]?
c. * Which club$_i$ did John meet a lot of girls [without going to _$_i$]? (Cattell 1976: 38)
d. * Who$_i$ did Mary cry [after John hit _$_i$]? (Huang 1982: 503)

Although a constraint on SLASH could effectively ban all extraction from adjuncts, the problem is that the AIC has a long history of exceptions, noted as early as Cattell (1976: 38), and by many others since, including Chomsky (1982: 72), Engdahl (1983), Hegarty (1990: 103), Cinque (1990), Pollard & Sag (1994), Culicover (1997: 253), and Borgonovo & Neeleman (2000). A sample of representative counterexamples is provided in (52).

(52) a. Who did he buy a book [for _]?

b. Who would you rather [sing with _]?
c. What temperature should I wash my jeans [at _]?
d. That's the symphony that Schubert [died without finishing _].
e. Which report did Kim [go to lunch without reading _]?
f. A problem this important, I could never [go home without solving _ first].
g. What did he [fall asleep complaining about _]?
h. What did John [drive Mary crazy trying to fix _]?
i. Who did you [go to Girona in order to meet _]?
j. Who did you go to Harvard [in order to work with _]?

Exceptions to the AIC include tensed adjuncts, as noted by Grosu (1981: 88), Deane (1991: 29), Kluender (1998), Levine & Hukari (2006: 287), Goldberg (2006: 144), Chaves (2012b: 471), Truswell (2011: 175, ft.1) and others. A sample is provided in (53).[17]

(53) a. These are the pills that Mary died [before she could take _].
b. This is the house that Mary died [before she could sell _].
c. The person who I would kill myself [if I couldn't marry _] is Jane.
d. Which book will Kim understand linguistics better [if she reads _]?
e. This is the watch that I got upset [when I lost _].
f. Robin, Pat and Terry were the people who I lounged around at home all day [without realizing were _ coming for dinner].
g. Which email account would you be in trouble [if someone broke into _]?
h. Which celebrity did you say that [[the sooner we take a picture of _], [the quicker we can go home]]?

To be sure, some of these sentences are complex and difficult to process, which in turn can lead speakers to prefer the insertion of an "intrusive" resumptive

[17] Truswell (2011) argues that the AIC and its exceptions are best characterized in terms of event-semantic constraints, such that the adjunct must occupy an event position in the argument structure of the main clause verb. However, recent experimental research has been unable to validate Truswell's acceptability predictions (Kohrt et al. 2018), and moreover, such an account incorrectly predicts that extractions from tensed adjuncts is impossible (Truswell 2011: 175, ft. 1).

pronoun at the gap site, but they are certainly more acceptable than the classic tensed AIC violations examples like Huang's (51d). Acceptable tensed AIC violations are more frequent in languages like Japanese, Korean, and Malayalam.

Like Subject Islands, AIC violations sometimes improve "parasitically" in the presence of a second gap as in (54). First of all, note that these sentences express radically different propositions, and so there is no reason to assume that all of these are equally felicitous. Second, note that (54a, c) describe plausible states of affairs in which it is clear what the extracted referent has to do with the main predication and assertion, simply because of the fact that *document* is predicated by *read*. In contrast, (54b) describes an unusual state of affairs in that it is unclear what the extracted referent has to do with the main predication *read the email*, out of the blue. Basically, what does reading emails have to do with filing documents?

(54) a. Which document did John read _ before filing _ ?
b. * Which document did John read the email email before filing _ ?
c. Which document did John read _ before filing a complaint?

If AIC violations were truly only salvageable parasitically, then counterexamples like (55a) should not exist. As Levine & Sag (2003) and Levine & Hukari (2006: 256) note, both gaps reside in island environments and should be completely out and less acceptable than (55b, c), contrary to fact.

(55) a. What kinds of books do [the authors of _] argue about royalties [after writing _]?
b. * What kinds of books do [authors of _] argue about royalties after writing malicious pamphlets?
c. * What kinds of books do authors of malicious pamphlets argue about royalties [after writing _]?

In (55a), there is no sense in which the gap inside the subject is parasitic on the gap inside the adjunct, or vice-versa – under the assumption that neither gaps is supposed to be licit without the presence of a gap outside an island environment. In conclusion, the notion of parasitic gap is rather dubious. See Levine & Hukari (2006) for a more in-depth discussion of parasitism and empirical criticism of null resumptive pronoun accounts.

As in the case of other island phenomena discussed so far, it is doubtful that any purely syntactic account can describe all the empirical facts. Rather, extractions out of adjuncts are licit to the degree that the extracted referent can be interpreted as being relevant for the assertion.

8 Superiority effects

Contrasts like those below have traditionally been taken to be due to a constraint that prevents a given phrase from being extracted if another phrase in a higher position can be extracted instead (Chomsky 1973; 1980). Thus, the highest *wh*-phrase is extractable, but the lowest is not.

(56) a. Who _ saw what?
b. * What did who see _ ?

(57) a. Who did you persuade _ to buy what?
b. * What did you persuade who to buy _ ?

Several different kinds of exceptions to this *Superiority Constraint* (SC) have been noted in the literature. First, it is generally recognized that *which*-phrases are immune to the SC:

(58) a. I wonder which book which of our students read _ over the summer?
b. Which book did which professor buy _ ?

Pesetsky (1987) proposed to explain the lack of SC effects in (58) by stipulating that *which*-phrases are interpreted as indefinites which do not undergo LF movement. Rather, they require "D-linking" and obtain wide scope via an entirely different semantic mechanism called unselective binding. In order for a phrase to be D-linked, it must be associated with a salient set of referents. But as Ginzburg & Sag (2000: 248ff.) note, there is no independent evidence for saliency interpretational differences between *which* and other *wh*-phrases like *what* and *who*. For example, it is implausible that speakers have a specific referent in mind for the *which*-phrases in examples like (59).

(59) a. I don't know anything about cars. Do you have any suggestions about which car – if any – I should buy when I get a raise?
b. I don't know anything about cars. Do you have any suggestions about what – if anything – I should buy when I get a raise? (Ginzburg & Sag 2000: 248)

Furthermore, there are acceptable SC violations involving multiple *wh*-questions such as those in (60). See Bolinger (1978), Kayne (1983) and Pesetsky (1987:

109) for more such examples and discussion.[18]

(60) a. Who wondered what WHO was doing _ ?
b. What did WHO take _ WHERE?
c. Where did WHO take WHAT _ ?

Finally, there are also SC violations that involve echo questions like (61) and reference questions like (62). See Ginzburg & Sag (2000: Chapter 7) for a detailed argumentation that echo questions are not fundamentally different, syntactically or semantically, from other uses of interrogatives.

(61) a. What did Agamemnon break?
b. What did WHO break _ ?

(62) a. What did he break?
b. What did WHO break _ ?

There are two different, yet mutually consistent, possible explanations for SC effects in HPSG circles. One potential factor concerns processing difficulty (Arnon et al. 2007). Basically, long-distance dependencies where a *which*-phrase is fronted are generally more acceptable and faster to process than those where *what* or *who* if fronted, presumably because the latter are semantically less informative, and thus decay from memory faster, and are compatible with more potential gap sites before the actual gap. The second potential factor is prosodic in nature. Drawing from insights by Ladd (1996: 170–172) about the English interrogative intonation, Ginzburg & Sag (2000: 251) propose that in a multiple *wh*-interrogative construction, all *wh*-phrases must be in focus except the first. Crucially, focus is typically – but not always – associated with clearly discernible pitch accent. Thus, (56) and (57) are odd because the second *wh*-word is unaccented. In this account, a word like *who* has two possible lexical descriptions, shown in (63).

[18] Fedorenko & Gibson (2010) and others have found no evidence that the presence of a third *wh*-phrase improves the acceptability of a multiple interrogative, even with supporting contexts. However, the examples in (60) require peculiar intonation phrasings and contours, which may be difficult to elicit with written stimuli.

(63) a. *Ex situ* interrogative *who*:

$$\begin{bmatrix} \text{PHON} & \langle \text{who/WHO} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & i \\ \text{RESTR} & \{\} \end{bmatrix} \\ \text{STORE} & \left\{ \boxed{1} \begin{bmatrix} \text{IND} & i \\ \text{RESTR} & \{\textit{person}(i)\} \end{bmatrix} \right\} \end{bmatrix} \\ \text{NONLOC} & \begin{bmatrix} \text{WH} & \{\boxed{1}\} \\ \text{REL} & \{\} \\ \text{SLASH} & \{\} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \langle\rangle \end{bmatrix}$$

b. Optionally *ex situ* interrogative *who*:

$$\begin{bmatrix} \text{PHON} & \langle \text{WHO} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & i \\ \text{RESTR} & \{\} \end{bmatrix} \\ \text{STORE} & \left\{ \begin{bmatrix} \text{IND} & i \\ \text{RESTR} & \{\textit{person}(i)\} \end{bmatrix} \right\} \end{bmatrix} \\ \text{NONLOC} & \begin{bmatrix} \text{WH} & \{\} \\ \text{REL} & \{\} \\ \text{SLASH} & \{\} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \langle\rangle \end{bmatrix}$$

Since only the (optionally accented) lexical entry in (63a) is specified with a non-empty WH value, the theory of extraction proposed in Ginzburg & Sag (2000) predicts that (63a) must appear *ex situ*. In contrast, the accented lexical entry in (63b) is can appear *in situ*. For more discussion see Levine & Hukari (2006: 261).

A related range of island phenomena concern extraction from *whether*-clauses, which is traditionally assumed to be forbidden, as (64) illustrates.

(64) a. * Which movie did John wonder whether Bill liked _ ?

b. * Which movie did John ask why Mary liked _ ?

But again, the oddness of (64) is unlikely to be due to syntactic constraints, given the existence of passable counterexamples like (65).

(65) a. He told me about a book which I can't figure out whether to buy _ or not.
(Ross 1967)

b. Which glass of wine do you wonder whether I poisoned _ ?
(Cresti 1995: 81)

c. Who is John wondering whether or not he should fire _ ?

d. Which shoes are you wondering whether you should buy _ ?
(Chaves 2012b)

As noted by Kroch (1989), the reduced acceptability of an example like (66a) is better explained simply by noting the difficulty of accommodating its presupposition in (66b).

(66) a. How much money was John wondering whether to pay?

b. There was a sum of money about which John was wondering whether to pay it.

9 The Left Branch Condition

Ross (1967) discovered that the leftmost constituent of an NP cannot be extracted, as in (67), a constraint he dubbed as the *Left Branch Condition* (LBC).[19]

(67) a. * Whose$_i$ did you meet [_$_i$ friend]?
(cf. with '*You met whose friend?*')

b. * Which$_i$ did you buy [_$_i$ book]?
(cf. with '*You bought which book?*')

c. * How much$_i$ did you find [_$_i$ money]?
(cf. with '*You found how much money?*')

These facts are accounted for in versions of HPSG like Sag (2012) where Determiner Phrases (DPs) are not valents of the nominal head. If the DP is not listed in the argument structure of the nominal head, then there is no way for the DP to appear in SLASH. Rather, the DP selects the nominal head:

[19] As in other island environments discussed above, the LBC is not operative in constraining semantic scope, as illustrated below.

(i) Someone took a picture of each student's bicycle.
(Copestake et al. 2005: 303)

(68)
$$\begin{bmatrix} \text{PHON} & \langle \textit{the} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{determiner} & \\ \text{SELECT} & N' \begin{bmatrix} \text{IND} & i \\ \text{RESTR} & \{\boxed{1}\} \end{bmatrix} \end{bmatrix} \\ \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & i \\ \text{RESTR} & \{\} \end{bmatrix} \\ \text{STORE} & \left\{ \begin{bmatrix} \textit{the-rel} & \\ \text{VAR} & i \\ \text{ARG} & \boxed{1} \end{bmatrix} \right\} \end{bmatrix} \\ \text{NON-LOC} & \begin{bmatrix} \text{WH} & \{\} \\ \text{REL} & \{\} \\ \text{SLASH} & \{\} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \langle\rangle \end{bmatrix}$$

Analogously, in Sag (2012: 133), genitive DPs combine with nominal heads and bind their X-ARG index via a dedicated construction, not as valents. For example, in nominalizations like *Kim's description of the problem* the DP *Kim's* is not a valent of *description*, and therefore the genitive DP cannot appear in SLASH. Rather, gentive DPs are instead constructionally co-indexed with the agent role of the noun *description* via X-ARG. Moreover, the clitic *s* in *Kim's* must lean phonologically on the NP it selects, and therefore cannot be stranded for independently motivated phonological reasons, predicting the oddness of **It was Kim who I read 's description of the problem.*

There are various languages which do not permit extraction of left branches from noun phrases, but have a particular PP construction that appears to allow LBC violations. This is illustrated below in (9), with French data.

(69) a. Combien$_i$ a-t-il vendu [_$_i$ de livres]?
how-many has-he sold of books
'How many books did he sell?'

b. Quels$_i$ avez-vous acheté [_$_i$ livres]?
how-many have-you bought books
'How many books have you bought'

But the LBC violation is only apparent. The *de livres* is in fact a post-verbal *de*-N′ nominal, and thus no LBC violation occurs in (9a). See Abeillé et al. (2004) for details. Finally, Ross (1967) also noted that some languages do not obey the LBC at all. A small sample is given in (70). However, the languages in question lack determiners, and therefore it is possible that the extracted phrase is has a similar independent status to the French *de-N′* phrase in (9).

(70) a. Jaką$_i$ kupiłeś [_$_i$ książkę]
what you-bought book
'Which book did you buy?' (Polish)

b. Cju$_i$ citajes [_$_i$ knigu]?
whose you-are-reading book
'whose book are you reading?' (Russian)

c. Ki-nek$_i$ akarod, hogy halljam [_$_i$ a hang-já-t]?
who-DAT you-want that I-hear the voice-POSS.3SG-ACC
'Whose voice do you want me to hear?' (Hungarian)

10 The Complementizer Constraint

Perlmutter (1968) noted that subject phrases have different extraction properties than that of object phrases, as illustrated in (71). The presence of the complementizer hampers extraction of the subject, but not of the complement.[20]

(71) a. * [Who]$_i$ did Tom say (?that) _$_i$ had bought the tickets?

b. * [Who]$_i$ do you believe (?that) _$_i$ got you fired?

c. [The things]$_i$ that they believe (?that) _$_i$ will happen are disturbing to contemplate.

d. * [Who]$_i$ did you ask if _$_i$ bought the tickets?

e. * [Who]$_i$ do you expect for _$_i$ to fire you?

Bresnan (1977) and others also noted that Complementizer Constraint effects can be reduced in the presence of an adverbial intervening between the complementizer and the gap:

[20] There is no evidence that the Complementizer Constraint applies at the semantic level, however. The subject phrase of the embedded clause that can outscope the subject phrase of the matrix:

(i) a. Some teacher claimed that each student had cheated.

b. Every teacher claimed that a student had cheated.

(72) a. [Who]$_i$ did Tom say that – as far he could remember – _$_i$ had bought the tickets?

b. [Who]$_i$ do you believe that – for all intents and purposes – _$_i$ got you fired?

c. [Who]$_i$ do you think that after years and years of cheating death _$_i$ finally died?

In Bouma et al. (2001) and Ginzburg & Sag (2000), extracted arguments are typed as *gap-ss* rather than *canon-ss*. Only the latter are allowed to correspond to *in situ* signs and to reside in valence lists. However, subject extraction is different. If a subject phrase is extracted, then the list SUBJ contains the respective *gap-ss* sign. If one assumes that the lexical entry for the complementizer *that* requires S complements specified as [SUBJ ⟨ ⟩] then the oddness of (71) follows. For Bouma et al. (2001) and Ginzburg & Sag (2000), the adverbial circumvention effect in (72) is the result of assuming that the construction which allows the adverb to combine with the clause forces the mother node to be [SLASH { }], a rather *ad-hoc* account.[21]

A simpler account of the Complementizer Constraint has emerged recently, however, in principle compatible with any theory of grammar. For Kandybowicz (2006; 2009) and others, the Complementizer Constraint is prosodic in nature. Complementizers must cliticize to the following phonological unit, but if a pause is made at the gap site then the complementizer cannot do so. Accordingly, if the pronunciation of *that* is produced with a reduced vowel [ðət] rather than [ðæt] then the Complementizer Constraint violations in (71) improve in acceptability. Though promising, Ritchart et al. (2016) found no experimental evidence for amelioration of the Complementizer Constraint effects either with phonological reduction of the complementizer or with contrastive focus. Further research is needed to determine the true nature of Complementizer Constraint effects.

11 Island circumvention via ellipsis

Ellipsis somehow renders island constraints inactive, as in (73). A deletion-based analysis of such phenomena such as Merchant (2001) relies on moving the *wh*-phrase before deletion takes place, but since movement is assumed to be sensitive to syntactic island constraints, the prediction is that (73) should be illicit, contrary to fact.

[21]Except that in French, when the subject of the complement CP is extracted, the complementizer is *qui* instead of *que*, which could easily be captured by such an account.

(73) a. Terry wrote an article about Lee and a book about someone else from East Texas, but we don't know who$_i$ (*Terry wrote an article about _$_i$ Lee and a book about _$_i$).
[CSC violation]

b. Bo talked to the person who discovered something, but I still don't know what$_i$ (*Bo talked to the person who discovered _$_i$).
[CNPC violation]

c. That he'll hire someone is possible, but I won't divulge who$_i$ (*that he'll hire _$_i$ is possible).
[SSC violation]

d. She bought a rather expensive car, but I can't remember how expensive (*she bought a _ car).
[LBC violation]

The account adopted in HPSG is one in which remnants are assigned an interpretation based on the surrounding discourse context (Ginzburg & Sag 2000; Culicover & Jackendoff 2005; Jacobson 2008; Sag & Nykiel 2011). See Nykiel & Kim (2019), Chapter 20 of this volume for more detailed discussion. In a nutshell, the *wh*-phrases in (73) are "coerced" into a proposition-denoting clause via a unary branching construction that taps into contextual information. This straightforwardly explains not only why the antecedent for the elided phrase need not correspond to overt discourse – e.g. sluices like *What floor?* or *What else?* – but also why the examples in (73) are immune to island constraints: there simply is no island environment to begin with, and thus, no extraction to violate it.

12 Conclusion

HPSG remains relatively agnostic about many island types, given the existence of robust exceptions. It is however clear that many island effects are not purely due to syntactic constraints, and are more likely the result of multiple factors, including pragmatics, semantics and processing difficulty. To be sure, it is yet unclear how these factors can be brought together are articulate an explicit and testable account of island effects. In particular, it is unclear how to combine probabilistic information with syntactic, semantic and pragmatic representations, although one fruitful avenue of approach to this problem may be via *Data-Oriented Parsing* (Neumann & Flickinger 2003; 2015; Arnold & Linardaki 2007; Bod et al. 2003;

Bod 2009).

From its inception, HPSG has been meant to be compatible with models of language comprehension and production (Sag 1992; Sag & Wasow 2010; 2015), but not much work has been dedicated to bridging these worlds; see Wasow (2019), Chapter G of this volume. The challenge that island effects posit to any theory of grammar is central to linguistic theory and cognitive science: how to integrate theoretical linguistics and psycholinguistic models of on-line language processing so that fine-grained predictions about variability in acceptability judgements across nearly isomorphic clauses can be explained.

Acknowledgements

References

Abeillé, Anne, Olivier Bonami, Danièle Godard & Jesse Tseng. 2004. The syntax of french n' phrases. In Stefan Müller (ed.), *Proceedings of the hpsg-2004 conference, center for computational linguistics, katholieke universiteit leuven*, 6–26. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.

Abeillé, Anne & Rui P. Chaves. 2019. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 469–516. Berlin: Language Science Press. DOI:??

Abeillé, Anne, Barbara Hemforth, Elodie Winckel & Edward Gibson. 2018. A construction-conflict explanation of the subject-island constraint. 31th Annual CUNY Conference on Human Sentence Processing.

Abeillé, Anne & Robert Vivès. 2019. Les constructions à verbe support. In Anne Abeillé Danièle Godard & Antoine Gautier (eds.), *La Grande Grammaire du français, Actes Sud*. (In preparation).

Abrusán, Márta. 2011. Presuppositional and negative islands: a semantic account. *Natural Language Semantics* 19. 257–321.

Abrusán, Márta & Benjamin Spector. 2011. A semantics for degree questions based on intervals: negative islands and their obviation. *Journal of Semantics* 28(1). 107–147.

Akmajian, Adrian. 1975. More evidence for an np cycle. *Linguistic Inquiry* 1(6). 115–129.

Allwood, Jens. 1976. The complex np constraint as a non-universal rule and some semantic factors influencing the acceptability of swedish sentences which violate the cnpc. In Justine Stillings (ed.), *University of massachusetts occasional pa-*

pers in linguistics, volume ii, 1–20. Amherst, MA: University of Massachusetts at Amherst.

Altmann, Gerry T. M. & Yuki Kamide. 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition* 73. 247–264.

Amy, Gérard & Georges Noziet. 1978. Memory requirements and local ambiguities of parsing strategies. *Journal of Psycholinguistic Research* 20(3). 233–250.

Arai, Manabu & Frank Keller. 2013. The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes* 4(28). 525–560.

Arnold, Doug & Evita Linardaki. 2007. A data-oriented parsing model for HPSG. In Anders Søgaard & Petter Haugereid (eds.), *2nd international workshop on typed feature structure grammars (TFSG'07)*, 1–9. Tartu, Estonia: Center for Sprogteknologi, Kobenhavens Universitet, Working Papers, Report No. 8, ISN1600-339X. http://cst.dk/anders/publ/tfsg/proceedings07.pdf.

Arnon, Inbal, Neal Snider, Philip Hofmeister, T. Florian Jaeger & Ivan Ivan Andrew Sag. 2007. Cross-linguistic variation in a processing account: the case of multiple wh-questions. In *Proceedings of bls 32*.

Bachrach, Asaf & Roni Katzir. 2008. Right-node raising and delayed spell-out. In Kleanthes K. Grohmann (ed.), *Interphases: phase-theoretic investigations of linguistic interfaces*, 284–316. Oxford University Press.

Baltin, Mark. 1978. PP as a bounding node. In Mark Stein (ed.), *Proceedings of the eighth regional meeting of the northeast linguistic society*. Amherst, Massachusetts, Graduate Linguistics Students Association, University of Massachusetts, Amherst.

Beavers, John & Ivan A. Sag. 2004. Ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th international conference on head-driven phrase structure grammar, katholieke universiteit leuven*, 48–69. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.

Berwick, Robert & Amy S. Weinberg. 1984. *The grammatical basis of linguistic performance*. Cambridge. MAL MIT Press.

Bever, Thomas G. 1970. The influence of speech performance on linguistic structure. In G. B. Flores d'Arcais & W. J. M. Levelt (eds.), *Advances in psycholinguistics*, 65–88. Hayes.

Bod, Rens. 2009. From exemplar to grammar: Integrating analogy and probability in language learning. *Cognitive Science* 33(4). 752–793. DOI:10.1111/j.1551-6709.2009.01031.x

Bod, Rens, Remko Scha & Khalil Sima'an (eds.). 2003. *Data-oriented parsing* (CSLI Studies in Computational Linguistics). CSLI Publications.

Bolinger, Dwight. 1978. Asking more than one thing at a time. In Henry Hiz (ed.), *Questions*, 107–150. Dordrecht, Reidel.

Bolinger, Dwight. 1992. The role of accent in extraposition and focus. *Studies in Language* 16. 265–324.

Borgonovo, Claudia & Ad Neeleman. 2000. Transparent adjuncts. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 45. 199–224.

Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Bouma, Gosse, Rob Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bresnan, Joan. 1977. Variables in the theory of transformations. In Adrian Akmajian, Peter Culicover & Thomas Wasow (eds.), *Formal syntax*. New York: Academic Press.

Carreiras, Manuel. 1992. Estrategias de análisis sintáctico en el procesamiento de frases: cierre temprano versus cierre tardío. *Cognitiva* 4(1). 3–27.

Cattell, Ray. 1976. Constraints on movement rules. *Language* 52. 18–59.

Chaves, Rui P. 2003. Non-Redundant Scope Disambiguation in Underspecified Semantics. In Balder ten Cate (ed.), *Proceedings of the 8th ESSLLI student session*, 47–58.

Chaves, Rui P. 2007. *Coordinate structures – constraint-based syntax-semantics processing*. University of Lisbon Doctoral Dissertation.

Chaves, Rui P. 2012a. Conjunction, cumulation and *respectively* readings. *Journal of Linguistics* 48(2). 297–344.

Chaves, Rui P. 2012b. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30(2). 465–512.

Chaves, Rui P. 2013. An expectation-based account of subject islands and parasitism. *Journal of Linguistics* 2(49). 285–327.

Chaves, Rui P. 2014. On the disunity of right node raising phenomena: extraposition, ellipsis, and deletion. 4(90). 834–886.

Chaves, Rui P. 2018. Freezing as a probabilistic phenomenon. In Jutta Hartmann, Marion Knecht, Andreas Konietzko & Susanne Winkler (eds.), *Freezing - theoretical approaches and empirical domains*. De Gruyter.

Chaves, Rui P. & Jeruen E. Dery. 2014. Which subject islands will the acceptability of improve with repeated exposure? In R. E. Santana-LaBarge (ed.), *31st west coast conference on formal linguistics*, 96–106. Somerville, MA: Cascadilla Proceedings Project.

Chaves, Rui P. & Jeruen E. Dery. 2019. Frequency effects in subject islands. *Journal of Linguistics* 55. 47.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A festschrift for morris halle*, 232–286. New York: Holt, Reinhart & Winston.

Chomsky, Noam. 1977. *Essays on form and interpretation*. New York: North-Holland.

Chomsky, Noam. 1980. *Rules and representations*. Columbia University Press, New York.

Chomsky, Noam. 1982. *Concepts and consequences of the theory of government and binding*. Cambridge: MIT Press.

Chomsky, Noam. 1986. *Barriers* (Linguistic Inquiry Monographs 13). Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, David Michaels, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: essays in honor of Jean-Roger Vergnaud*, 133–165. Cambridge, MA, USA: MIT Press.

Chung, Sandra. 1994. Wh-agreement and 'referentiality' in chamorro. *Linguistic Inquiry* 25. 1–44.

Chung, Sandra & James McCloskey. 1983. On the interpretation of certain island facts in GPSG. *Linguistic Inquiry* 14. 703–714.

Cinque, Guglielmo. 1990. *Types of $\overline{A}$-dependencies*. Cambridge, Mass.: MIT Press.

Cinque, Guglielmo. 2010. Structure preserved: studies in syntax for jan koster. In C. Jan-Wouter Zwart & Mark de Vries (eds.), 81–90. John Benjamins Publishing Company.

Clark, Herbert H. & Thomas Wasow. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37. 201–242.

Clausen, David R. 2010. Processing factors influencing acceptability in extractions from complex subjects. http://www.davidclausen.org/papers/Clausen10SubjectIslands.pdf. Unpublished manuscript. Stanford University.

Clausen, David R. 2011. Informativity and acceptability of complex subject islands. Poster presented at the 24th Annual CUNY Sentence Processing Conference, Stanford, CA.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 4(3). 281–332.

Cowper, Elizabeth A. 1976. *Constraints on sentence complexity: a model for syntactic processing*. University of California, Irvine, CA PhD dissertation.

Crawford, Jean. 2011. Using syntactic satiation effects to investigate subject islands. In *WCCFL 29 proceedings*. University of Arizona.

Creel, Sarah C., Richard N. Aslin & Michael K. Tanenhaus. 2008. Heeding the voice of experience: the role of talker variation in lexical access. *Cognition* 106. 633–664.

Cresti, Diana. 1995. Extraction and reconstruction. *Natural Language Semantics* 3(1). 79–122.

Culicover, Peter. 1997. *Principles and parameters: an introduction to syntactic theory*. New York: Oxford University Press.

Culicover, Peter W. 1999. *Syntactic nuts: hard cases in syntax* (Volume One of Foundations of Syntax). Oxford University Press.

Culicover, Peter W. 2013. *Grammar & complexity*. Oxford University Press.

Culicover, Peter & Ray S. Jackendoff. 2005. *Simpler syntax*. Oxford University Press.

Davies, William D. & Stanley Dubinsky. 2009. On the existence (and distribution) of sentential subjects. In Donna B. Gerdts, John C. Moore & Maria Polinsky (eds.), *Hypothesis a/hypothesis b. linguistic explorations in honor of david m. perlmutter*, 111–128. Cambridge, MA: MIT Press.

De Vicenzi, Marica & Remo Job. 1993. Some observations on the universality of the late closure strategy. *Journal of Psycholinguistic Research* 22(2). 189–206.

Deane, Paul. 1991. Limits to attention: a cognitive theory of island phenomena. *Cognitive Linguistics* 2(1). 1–63.

Deane, Paul D. 1992. *Grammar in mind and brain*. Mouton de Gruyter.

Delahunty, Gerald P. 1983. But sentential subjects do exist. *Linguistic Analysis* 12. 379–398.

DeLong, Katherine A., Thomas P. Urbach & Marta Kutas. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience* 8(8). 1117–1121.

Desmet, Timothy, Constantijn De Baecke, Denis Drieghe, Mark Brysbaert & Wietske Vonk. 2006. Relative clause attachment in dutch: on-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes* 21(4). 453–485.

Drummond, Alex. 2009. The unity of extraposition and the a/a distinction. In *Proceedings of cls 45*. University of Chicago.

Eady, Stephen J. & Janet D. Fodor. 1981. Is center embedding a source of processing difficulty? Presented at the Linguistics Society of America Annual Meeting.

Engdahl, Elisabet. 1982. Restrictions on unbounded dependencies in swedish. In Elisabet Engdahl & Eva Ejerhed (eds.), *Readings on unbounded dependencies in scandinavian languages*, 151–174. Stockholm: Almqvist & Wiksell International.

Engdahl, Elisabet. 1983. Parasitic gaps. *Linguistics and Philosophy* 6. 3–34.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MIT Doctoral Dissertation.

Erteschik-Shir, Nomi. 1981. More on extractability from quasi-nps. *Linguistic Inquiry* 12. 665–670.

Erteschik-Shir, Nomi & Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6. 41–86.

Federmeier, Kara D. & Marta Kutas. 1999. A rose by any other name: long-term memory structure and sentence processing. *Journal of Memory and Language* 41. 469–495.

Fedorenko, Evelina & Edward Gibson. 2010. Adding a third wh-phrase does not increase the acceptability of object-initial multiple-wh-questions. *Syntax* 3(13). 183–195.

Fernández, Eva M. 2003. *Bilingual sentence processing: relative clause attachment in english and spanish*. Amsterdam, The Netherlands: John Benjamins.

Ferreira, Fernanda. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30(2). 210–233.

Ferreira, Fernanda & John M. Henderson. 1991. Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language* 31. 725–745.

Ferreira, Fernanda & John M. Henderson. 1993. Basic reading processes during syntactic analysis and reanalysis. *Canadian Journal of Psychology: Special Issue on Reading and Language Processing* 47. 247–275.

Fodor, Janet A. & M. F. Garrett. 1967. Some syntactic determinants of sentential complexity. *Perception and Pshychophisics* 2. 282–296.

Fodor, Janet Dean. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9. 427–473.

Fodor, Janet Dean. 1983a. Phrase structure parsing and the island constraints. *Linguistics and Philosophy* 6. 163–223.

Fodor, Janet Dean. 1983b. Phrase structure parsing and the island constraints. *Linguisitcs and Philosophy* 6. 163–223.

Fodor, Janet Dean. 1992. Islands, learnability and the lexicon. In Helen Good luck & Michael Rochemont (eds.), *Island constraints: theory, acquisition and processing*, 109–180. Dordrecht: Kluwer.

Fodor, Janet Dean. 2002. Prosodic disambiguation in silent reading. In M. Hirotani (ed.), *Nels 32*, 113–132. Amherst, MA: GLSA Publications.

Fodor, Janet Dean, Thomas G. Bever & Merrill F. Garrett. 1974. *The psychology of language*. New York: McGraw Hill.

Fox, Danny. 2000. *Economy and semantic interpretation*. MIT: Cambridge, Massachusetts.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29. 537–586.

Frazier, Lyn & Charles E. Jr. Clifton. 1996. *Construal*. Cambridge, MA: MIT Press.

Frazier, Lynn & K. Rayner. 1988. Parameterizing the language processing system: left versus right branching within and across languages. In J. A. Hawkins (ed.), *Explaining language universals*, 247–279. Oxford: Blackwell.

Freidin, Robert. 1992. *Foundations of generative syntax*. MIT Press, Cambridge, Mass.

Fukuda, Shin, Chizuru Nakao, Akira Omaki & Maria Polinsky. 2018. Revisiting subject-object asymmetry: subextraction in Japanese. In *Proceedings of the 10th Workshop on Altaic Formal Linguistics*. MIT Working Papers in Linguistics. to appear.

Garnsey, Susan. 1985. *Function words and content words: reaction time and evoked potential measures of word recognition*. Cognitive Science Technical Report URCS-29. University of Rochester.

Gawron, Jean Mark & Andrew Kehler. 2003. Respective answers to coordinated questions. In *Semantics and linguistic theory 13 (salt 13)*.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12(2). 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized phrase structure grammar*. Oxford: Blackwell, & Cambridge, Ma: Harvard University Press.

Gibson, Edward. 1991. *A computational theory of human linguistic processing: memory limitations and processing breakdown*. Carnegie Mellon University PhD dissertation.

Gibson, Edward. 2000. The dependency locality theory: a distance-based theory of linguistic complexity. In Yasuhi Miyashita, Alec P. Marantz & Wayne O'Neil (eds.), *Image, language, brain*, 95–126. Cambridge, MA: MIT Press.

Gibson, Edward. 2006. The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. *Journal of Memory and Language* 54. 363–388.

Gibson, Edward, Neal Pearlmutter, Enriqueta Canseco-Gonzalez & Gregory Hickok. 1996. Recency preference in the human sentence processing mechanism. *Cognition* 59(1). 23–59.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative Investigations: the form, meaning and use of English interrogative constructions*. Stanford: CSLI Publications.

Godard, Danièle. 1988. *La syntaxe des relatives en Français*.

Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, Adele E. 2013. Backgrounded constituents cannot be extracted. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and island effects*, 221–238. Cambridge University Press.

Goldsmith, John. 1985. A Principled Exception to the Coordinate Structure Constraint. In *Papers from the twenty-first annual regional meeting of the chicago linguistic society*, 133–143. Chicago: Chicago Linguistic Society.

Goodall, Grant. 2011. Syntactic satiation and the inversion effect in English and Spanish *wh-* questions. *Syntax* 14. 29–47.

Grosu, Alexander. 1973a. On the nonunitary nature of the coordinate structure constraint. *Linguistic Inquiry* 4. 88–92.

Grosu, Alexander. 1973b. On the status of the so-called right roof constraint. *Language* 49. 294–311.

Grosu, Alexander. 1981. *Approaches to island phenomena*. North-Holland.

Guéron, Jacqueline & Robert May. 1984. Extraposition and logical form. *Linguistic Inquiry* 15(1). 1–31.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2). 244–249.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford.

Hegarty, Michael. 1990. On adjunct extraction from complements. In Lisa L.S. Cheng & Hamida Demirdash (eds.), *Mit working papers in linguistics 13*, 101–124.

Hiramatsu, Kazuko. 2000. *Accessing linguistic competence: evidence from children's and adults' acceptability judgments*. University of Connecticut Ph.D. Thesis.

Hofmeister, Philip. 2007a. *Facilitating memory retrieval in natural language comprehension*. Stanford University. Ph.D. Thesis.

Hofmeister, Philip. 2007b. Retrievability and gradience in filler-gap dependencies. In M. Elliott, J. Kirby, O. Sawada, E. Staraki & S. Yoon (eds.), *Proceedings from the annual meeting of the chicago linguistics society*, vol. 43, 1, 109–123.

Hofmeister, Philip, Laura Staum Casasanto & Ivan A. Sag. 2013. Islands in the grammar? Standards of evidence. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and islands effects*, 42–63. Cambridge: Cambridge University Press.

Hofmeister, Philip, Peter Culicover & Susanne Winkler. 2015. Effects of processing on the acceptability of 'frozen' extraposed constituents. *Syntax*, 18(4). 464–483.

Hofmeister, Philip, Florian T. Jaeger, Ivan A. Sag, Inbal Arnon & Neal Snider. 2007. Locality and accessibility in wh-questions. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: linguistics in search of its evidential base*, 185–206. Berlin: Mouton de Gruyter.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86(2). 366–415.

Horn, Laurence. 1972. *On the semantic properties of logical operators in english*. University of California, LA PhD. Dissertation.

Huang, Cheng-Teh James. 1982. *Logical relations in chinese and the theory of grammar*. MIT, Cambridge, Massachusetts Ph.D Dissertation.

Huck, Geoffrey J. & Younghee Na. 1990. Extraposition and focus. *Language* 66. 51–77.

Huddleston, Rodney D., Geoffrey Pullum & Peter Peterson. 2002. The Cambridge Grammar of the English Language. In, chap. 12: Relative clause constructions and unbounded dependencies, 1033–1096. Cambridge University Press.

Hukari, Thomas & Robert Levine. 1991. On the disunity of unbounded dependency constructions. *Natural Language and Linguistic Theory* 9. 97–144.

Jacobson, Pauline. 2008. Direct compositionality and variable free semantics: the case of antecedent contained ellipsis. In Kyle Johnson (ed.), *Topics in ellipsis*, 30–68. Cambridge University Press.

Jiménez–Fernández, Ángel. 2009. On the composite nature of subject islands: a phase-based approach. *SKY Journal of Linguistics* 22. 91–138.

Jurka, Johannes, Chizuru Nakao & Akira Omaki. 2011. It's not the end of the ced as we know it: revisiting german and japanese subject islands. In Mary Byram Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer & Barbara Tomaszewicz (eds.), *Proceedings of the 28th west coast conference on formal linguistics*, 124–132. Somerville, MA: Cascadilla Proceedings Project.

Kamide, Yuki, Christoph Scheepers & Gerry T. M. Altmann. 2003. Integration of syntactic and semantic information in predictive processing: crosslinguistic evidence from German and English. *Journal of Psycholinguistic Research* 32(1). 37–55.

Kandybowicz, Jason. 2006. Comp-trace effects explained away. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th west coast conference on formal linguistics*, 220–228. Somerville, MA: Cascadilla Press.

Kandybowicz, Jason. 2009. Comp-trace effects and the syntax–phonology interface. In Jacqueline Bunting, Sapna Desai, Robert Peachey, Christopher Straughn & Zuzana Tomková (eds.), *Proceedings of cls 42 (volume 1: the main session)*, 103–115. Chicago: Chicago Linguistics Society.

Kayne, Richard. 1980. Extensions of binding and case-marking. *Linguistic Inquiry* 11. 75–96.

Kayne, Richard. 1981. ECP Extensions. *Linguistic Inquiry* 12. 93–133.

Kayne, Richard. 1983. Connectedness. *Linguistic Inquiry* 14. 223–249.

Kayne, Richard. 2000. Overt vs. covert movement. In *Parameters and universals*, 261–276. Oxford University Press.

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Keller, Frank. 1994. *Extraposition in hpsg*. Verbmobil-Report 30. IBM Germany, Institute for Logic & Linguistics, Heidelberg.

Keller, Frank. 1995. Towards an account of extraposition in HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, 301–306. Dublin: Association for Computational Linguistics.

Kemper, Susan. 1986. Imitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics* 7. 277–288.

Kim, Jong-Bok & Ivan A. Sag. 2005. English object extraposition: a constraint-based approach. In Stefan Müller (ed.), *Proceedings of the 12th international conference on head-driven phrase structure grammar, department of informatics, university of lisbon*, 192–212. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Kimball, John. 1973. Seven principles of surface structure parsing in natural language. *Cognition* 2. 15–57.

Kluender, Robert. 1992. Deriving island constraints from principles of predication. In Helen Goodluck & Michael Rochemont (eds.), *Island constraints: theory, acquisition and processing*, 223–258. Dordrecht: Kluwer.

Kluender, Robert. 1998. On the distinction between strong islands and weak islands: a processing perspective. In Peter W. Culicover & Louise McNally (eds.), *Syntax and semantics 29: the limits of syntax*, 241–279. New York, NY: Academic Press.

Kluender, Robert. 2004. Are subject islands subject to a processing account? In B. Schmeiser, V. Chand, A. Kelleher & A. Rodriguez (eds.), *In proceedings of the WCCFL 23*, 101–125. Somerville, MA: Cascadilla Press.

Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8. 573–633.

Kohrt, Annika, Trey Sorensen & Dustin A. Chacón. 2018. The real-time status of semantic exceptions to the adjunct island constraint. In *Proceedings of wecol 2018: western conference on linguistics.*

Kothari, Anubha. 2008. Frequency-based expectations and context influence bridge quality. In M. Grosvald & D. Soares (eds.), *Proceedings of wecol 2008: western conference on linguistics.*

Kroch, Anthony. 1989. Amount quantification, referentiality, and long *wh*-movement. Ms., Department of Linguistics, University of Pennsylvania.

Kubota, Yusuke & Jungmee Lee. 2015. The coordinate structure constraint as a discourse-oriented principle: further evidence from japanese and korean. *Language* 91(3). 642–675.

Kuno, Susumu. 1973. *The structure of the japanese language.* Cambridge, MA: The MIT Press.

Kuno, Susumu. 1974. The position of relative clauses and conjunctions. *Linguistic Inquiry* 5(1). 117–136.

Kuno, Susumu. 1987. *Functional syntax – anaphora, discourse and empathy.* The University of Chicago Press, Chicago & London.

Kuno, Susumu & Ken-Ichi Takami. 1993. *Grammar and discourse principles: functional syntax and GB theory.* Chicago & London: University of Chicago Press.

Kutas, Marta, Mireille Besson & Cyma Van Petten. 1988. Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Neurophysiology* 69. 218–233.

Kutas, Marta & Steven A. Hillyard. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 5947(307). 161–163.

Kynette, Donna & Susan Kemper. 1986. Aging and the loss of grammatical forms: a cross-sectional study of language performance. *Language and Communication* 6. 65–72.

Ladd, Robert D. 1996. *Intonational phonology.* Cambridge University Press.

Lakoff, George. 1986. Frame Semantic Control of the Coordinate Structure Constraint. In *Papers from the 22nd regional meeting of the chicago linguistic society.* Chicago: Chicago Linguistic Society.

Lasnik, Howard & Myung-Kwan Park. 2003. The epp and the subject condition under sluicing. *Linguistic Inquiry* 34. 649–660.

Lasnik, Howard & Mamoru Saito. 1992. *Move α: Conditions on its application and output.* MIT Press.

Lau, Ellen, Clare Stroud, Silke Plesch & Colin Phillips. 2006. The role of structural prediction in rapid syntactic analysis. *Brain and language* 1(98). 74–88.

Levin, Nancy & Ellen Prince. 1986. Gapping and clausal implicature. *Papers in Linguistics* 19. 351–364.

Levine, Robert D. 2001. The extraction riddle: just what are we missing? *Journal of Linguistics* 37. 145–174.

Levine, Robert D. N.d. *Syntactic analysis: an hpsg-based approach.* Cambridge University Press.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions.* CSLI Publications, Leland Stanford Junior University.

Levine, Robert D., Thomas E. Hukari & Michael Calcagno. 2001. Parasitic gaps in English: some overlooked cases and their theoretical implications. In Peter Culicover & Paul M. Postal (eds.), *Parasitic gaps*, 181–222. Cambridge, MA: MIT Press.

Levine, Robert D. & Ivan A. Sag. 2003. Some empirical issues in the grammar of extraction. In Stefan Müller (ed.), *Proceedings of the hpsg-2003 conference, michigan state university, east lansing*, 236–256. Stanford: CSLI Publications.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition* 3(106). 1126–1177.

Levy, Roger, Evelina Fedorenko, Mara Breen & Ted Gibson. 2012. The processing of extraposed structures in English. *Cognition* 12(1). 12–36.

Levy, Roger & Frank Keller. 2013. Expectation and locality effects in German verb-final structures. *Journal of Memory and Language* 2(68). 199–222.

Lewis, Richard L. 1993. *An architecturally-based theory of human sentence comprehension.* Carnegie Mellon University PhD Dissertation.

Mak, Willem M., Wietske Vonk & Herbert Schriefers. 2008. Discourse structure and relative clause processing. *Memory & Cognition* 36. 170–181.

May, Robert. 1985. *Logical Form: its structure and derivation.* Cambridge, MA: MIT Press.

McCawley, James D. 1981. The syntax and semantics of english relative clauses. *Lingua* 53. 99–149.

Merchant, Jason. 2001. *The syntax of silence: sluicing, islands, and the theory of ellipsis.* Oxford University Press: Oxford.

Michel, Daniel. 2014. *Individual cognitive measures and working memory accounts of syntactic island phenomena.* University of California, San Diego, CA Doctoral dissertation.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2004. Complex NPs, subjacency, and extraposition. *Snippets* 8. 10–11.

Müller, Stefan. 2007. Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. In Gisela Zifonun & Werner Kallmeyer (eds.), *Sprachkorpora – Datenmengen und Erkenntnisfortschritt* (Institut für Deutsche Sprache Jahrbuch 2006), 70–90. Berlin: Walter de Gruyter.

Munn, Alan. 1998. ATB movement without identity. In Jennifer Austin & Aaron Lawson (eds.), *Proceedings of the 14th eastern states conference on linguistics (escol-97)*, 150–160. CSC Publications.

Munn, Alan. 1999. On the identity requirement of ATB extraction. *Natural Language Semantics* 7. 421–425.

Neumann, Günter & Daniel P. Flickinger. 2003. Hpsg-dop: data-oriented parsing with HPSG. In Jong-Bok Kim & Stephen Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*. http://csli-publications.stanford.edu/HPSG/2002/, accessed 2018-2-25.

Neumann, Günter & Daniel P. Flickinger. 2015. *Learning stochastic lexicalized Tree Grammars from HPSG*. http://www.dfki.de/~neumann/publications/new-ps/sltg.ps.gz, accessed 2015-12-23.

Ni, Weijia, Stephen Crain & Donald Shankweiler. 1996. Sidestepping garden paths: assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language & Cognitive Processes* 11(3). 283–334.

Nishigauchi, Taisuke. 1999. Quantification and *wh*-constructions. In Natsuko Tsujimura (ed.), 269–296. Oxford: Blackwell.

van Noord, Gertjan & Gosse Bouma. 1996. Extraposition as a nonlocal dependency. In G.-J. Kruijff, Glynn Morrill & D. Oehrle (eds.), *Proceedings of formal grammar 96*, 1–14.

Nykiel, Joanna & Jong-Bok Kim. 2019. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 585–612. Berlin: Language Science Press. DOI:??

Oshima, David Y. 2007. On factive islands: pragmatic anomaly vs. pragmatic infelicity. In *Proceedings of the 20th annual conference on new frontiers in artificial intelligence* (JSAI'06), 147–161. Tokyo, Japan: Springer-Verlag. http://portal.acm.org/citation.cfm?id=1761528.1761544.

Partee, Barabara. 1984. Nominal and temporal anaphora. *Linguistics and Philosophy* 7(3). 243–286.

Perlmutter, Donald M. 1968. *Deep and surface structure constraints in syntax*. Doctoral dissertation, MIT dissertation.

Pesetsky, David. 1987. Wh-in-situ: movement and unselective binding. In Eric Reuland & Alice ter Meulen (eds.), *The representation of (in)definiteness*, 98–129. Cambridge: MIT Press.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge: MIT Press.

Petten, Cyma Van & Marta Kutas. 1991. Influences of semantic and syntactic context on open and closed class words. *Memory and Cognition* 19. 95–112.

Phillips, Colin. 2006. The real-time status of island phenomena. *Language* 82. 795–823.

Phillips, Colin. 2013a. On the nature of island constraints. I: language processing and reductionist accounts. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and islands effects*, 64–108. Cambridge: Cambridge University Press.

Phillips, Colin. 2013b. Some arguments and non-arguments for reductionist accounts of syntactic phenomena. *Language and Cognitive Processes* 28. 156–187.

Polinsky, Maria, Carlos Gómez Gallo, Peter Graff, Ekaterina Kravtchenko, Adam Milton Morgan & Anne Sturgeon. 2013. Subject islands are different. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and islands effects*. Cambridge: Cambridge University Press.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press & Stanford: CSLI.

Postal, Paul M. 1974. *On raising*. Cambridge, Mass.: MIT Press.

Postal, Paul M. 1998. *Three investigations of extraction*. MIT: Cambridge, Massachusetts.

Pritchett, Bradley L. 1991. Subjacency in a principle-based parser. In Robert C. Berwick (ed.), *Principle-based parsing: computation and psycholinguistics*, 301–345. Dordrecht: Kluwer.

Ritchart, Amanda, Grant Goodall & Marc Garellek. 2016. Prosody and the that-trace effect: an experimental study. In Kyeong-min Kim, Pocholo Umbal, Trevor Block, Queenie Chan, Tanie Cheng, Kelli Finney, Mara Katz, Sophie Nickel-Thompson & Lisa Shorten (eds.), *Proceedings of the 33rd west coast conference on formal linguistics*, 320–328.

Rochemont, Michael S. 1992. Bounding rightward a-bar dependencies. In H. Goodluck & M.S. Rochemont (eds.), *Island constraints: theory, acquisition and processing*. Kluwer Academic Publishers, Dordrecht.

Roland, Douglas, Gail Mauner, Carolyn O'Meara & Hongoak Yun. 2012. Discourse expectations and relative clause processing. *Journal of Memory and Language* 66. 479–508.

Ross, John R. 1967. *Constraints on variables in syntax*. MIT, Cambridge, Massachusetts. [Published in 1986 as *Infinite Syntax!* Norwood, NJ: Ablex Publishing] Ph.D. Dissertation.

Ruys, Eddy. 1993. *The scope of indefinites*. Universiteit Utrecht Doctoral Dissertation.

Saah, Kofi & Helen Goodluck. 1995. Island effects in parsing and grammar: evidence from akan. *Linguistic Review* 12. 381–409.

Sabbagh, James. 2007. Ordering and linearizing rightward movement. *Natural Language and Linguistic Theory* 25(2). 349–401.

Sag, Ivan A. 1992. Taking performance seriously. In Carlos Martin-Vide (ed.), *Vii congreso de languajes naturales y lenguajes formales. barcelona.*

Sag, Ivan A. 2000. *Another argument against wh-trace*. Jorge Hankamer Webfest. `http://ling.ucsc.edu/Jorge/sag.html`.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

Sag, Ivan A. 2012. Sign-based construction grammar: an informal synopsis. In Hans Boas & Ivan A. Sag (eds.), *Sign-based construction grammar*, 69–202. Stanford: CSLI Publications.

Sag, Ivan A. & Janet Dean Fodor. 1994. Extraction without traces. In *Proceedings of the thirtheenth west coast conference on formal linguistics*, 365–384. Stanford: CSLI.

Sag, Ivan A., Philip Hofmeister & Neal Snider. 2007. Processing complexity in subjacency violations: the complex noun phrase constraint. In *Proceedings of the 43rd annual meeting of the chicago linguistic society. chicago: cls.*

Sag, Ivan A. & Joann Nykiel. 2011. Remarks on sluicing. In Stefan Müller (ed.), *Proceedings of hpsg 2011 - 18th international conference on head-driven phrase structure grammar*. http://lingo.stanford.edu/sag/papers/sag-nykiel-hpsg11.pdf.

Sag, Ivan A. & Thomas Wasow. 2010. Performance-compatible competence grammar. In Robert Borsley & Kersti Borjars (eds.), *Non-transformational syntax*. Cambridge: Blackwells.

Sag, Ivan A. & Thomas Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguist Research* 44(1). 47–63.

Santorini, Beatrice. 2007. (Un?)expected movement. University of Pennsylvania. http://www.ling.upenn.edu/~beatrice/examples/movement.html. Accessed: Jan 2 2012.

Sauerland, Uli & Paul D. Elbourne. 2002. Total reconstruction, pf movement, and derivational order. *Linguistic Inquiry* 33(2). 283–319.

Schmerling, Susan. 1972. Apparent counterexamples to the coordinate structure constraint: a canonical constraint. *Studies in the Linguistic Sciences* 2(1). 91–104.

Sprouse, Jon. 2007. *A program for experimental syntax: finding the relationship between acceptability and grammatical knowledge*. University of Maryland Ph.D. dissertation.

Sprouse, Jon. 2009. Revisiting satiation: Evidence for an equalization response strategy. *Linguistic Inquiry* 40(1). 329–341.

Sprouse, Jon, Ivano Caponigro, Ciro Greco & Carlo Cecchetto. 2015. Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34(1). 307–344.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012a. A test of the relation between working memory capacity and syntactic island effects. *Language*.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012b. Working-memory capacity and island effects: a reminder of the issues and the facts. *Language* 88(2). 401–407.

Staub, Adrian & Charles E. Jr. Clifton. 2006. Syntactic prediction in language comprehension: evidence from *either… or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32(2). 425–436.

Staub, Adrian, Charles E. Jr. Clifton & Lyn Frazier. 2006. Heavy np shift is the parser's last resort: evidence from eye movements. *Journal of Memory and Language* 54. 389–406.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10(1). 80–126.

Stowell, Timothy Angus. 1981. *Origins of phrase structure*. MIT Cambridge, MA PhD Dissertation.

Strunk, Jan & Neal Snider. 2008. Extraposition without subjacency. http://www.linguistics.ruhr-uni-bochum.de/~strunk/DGfS2008Handout.pdf. Presented at the workshop on Rightward Movement in a Comparative Perspective at DGfS 2008 in Bamberg, March 27, 2008.

Strunk, Jan & Neal Snider. 2013. Subclausal locality constraints on relative clause extraposition. In Gert Webelhuth, Manfred Sailer & Heike Walker (eds.), *Rightward movement in a comparative perspective*. Amsterdam: John Benjamins.

Stucky, Susan U. 1987. Configurational variation in English: a study of extraposition and related matters. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Syntax and semantics 20: discontinuous constituency*, 377–404. Academic Press, Orlando, FL.

Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1. 235–284.

Tabor, Whitney & Sean Hutchins. 2004. Evidence for self-organized sentence processing: digging-in effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30(2). 431–450.

Tabor, Whitney, C Juliano & MK Tanenhaus. 1997. Parsing in a dynamical system: an attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes* 2–3(12). 211–271.

Takami, Ken-ichi. 1992. *Preposition stranding: from syntactic to functional analyses.* Berlin: Mouton de Gruyter.

Taraldsen, Knut T. 1982. Extraction from relative clauses in norwegian. readings on unbounded dependencies in scandinavian languages. In Elisabet Engdahl & Eva Ejerhed (eds.), 205–221. Stockholm: Almqvist & Wiksell International.

Traxler, Matthew J., Martin J. Pickering & Charles E. Jr. Clifton. 1998. Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of memory and language* 39. 558–592.

Truswell, Robert. 2011. *Events, phrases and questions.* Oxford.

Tsiamtsiouris, Jim & Helen Smith Cairns. 2009. Effects of syntactic complexity and sentence-structure priming on speech initiation time in adults who stutter. *Journal of Speech, Language, and Hearing Research* 52(6). 1623–1639.

Uszkoreit, Hans, Thorsten Brants, Denys Duchier, Brigitte Krenn, Lars Konieczny, Stephan Oepen & Wojciech Skut. 1998. Studien zur performanzorientierten Linguistik: Aspekte der Relativsatzextraposition im Deutschen. *Kognitionswissenschaft. Themenheft Ressourcenbeschränkungen* 7(3). 129–133.

Van Eynde, Frank. 1996. A monostratal treatment of it extraposition without lexical rules. In G. Durieux, W. Daelemans & S. Gillis (eds.), *Papers from the sixth clin meeting*, 231–248.

Van Valin, Robert D. Jr. 1986. Pragmatics, island phenomena, and linguistic competence. In *Papers from the parasession on pragmatics and grammatical theory*, vol. 22(2), 223–233. CLS.

van Schijndel, Marten, William Schuler & Peter W. Culicover. 2014. Frequency effects in the processing of unbounded dependencies. In *Proceedings of cogsci 2014.* Quebec, Quebec, Canada: Cognitive Science Society.

Vicente, Luis. 2016. Atb extraction without coordination. In Christopher Hammerly & Brandon Pickett (eds.), *Proceedings of nels 46*, 1–14.

Villata, Sandra, Luigi Rizzi & Julie Franck. 2016. Intervention effects and relativized minimality: new experimental evidence from graded judgments. *Lingua* 179. 76–96.

Wasow, Thomas. 2002. *Postverbal behavior*. CSLI Publications.

Wasow, Tom. 2019. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 775–796. Berlin: Language Science Press. DOI:??

Winter, Yoad. 2001. *Flexibility principles in boolean semantics: coordination, plurality and scope in natural language*. MIT Press, Cambridge, Massachusetts.

Witten, Edward. 1972. *Centrality*. Report No. NSF–28 to The National Science Foundation. The Computational Laboratory of Harvard University, Cambridge, Massachusetts.

Zhang, Niina Ning. 2007. The syntactic derivations of two paired dependency constructions. *Lingua* 117(12). 2134–2158.

Chapter 17

Coordination

Anne Abeillé
Laboratoire de Linguistique Formelle, University of Paris

Rui P. Chaves
Linguistics Department, University at Buffalo, The State University of New York

Coordination is a central topic for feature-based phrase structure grammars. Following GPSG which provided the first formal analysis of unlike coordination, HPSG has developed detailed analyses of different coordination constructions in a variety of unrelated languages. Central to the HPSG analyses are two main ideas : (i) coordination structures are unheaded phrases, (ii) coordinate daughters display some kind of parallelism, which is captured by feature sharing ; from these ideas, specific properties can be derived, regarding extraction and agreement for instance. Many HPSG analyses also agree that coordination is a cover term for a wide variety of different constructions, which can be viewed as different subtypes of coordinate phrases, which can be cross-classified with other subtypes of the grammar (nominal or not, with ellipsis or not etc.) We present the description of various coordination phenomena, and show that HPSG can account for their subtle properties, while integrating them in the general organization of the grammar.

1 Introduction

A great deal of research has been dedicated to the topic of coordination structures in the last 70 years, spanning a multitude of different approaches in many different theoretical frameworks. With regard to the linguistic problems, research questions abound. In the realm of syntax there is much debate concerning the role of coordination lexemes, the existence of null coordinators, the syntactic relationship between conjuncts, the peculiar extraction phenomena that certain coordination structures exhibit, the necessary properties that allow two different structures to be coordinated, the relation between coordination structures

Anne Abeillé & Rui P. Chaves. 2019. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 469–516. Berlin: Language Science Press. DOI:??

and comparative and subordination structures, peculiar ellipsis phenomena than can optionally occur, the various patterns of agreement that obtain in nominal coordination structures, the distribution and syntactic realization of the lexemes either and or, etc.. In the realm of semantics, the issues are no less complex, and the debate no less lively. There are many questions pertaining to how exactly the meaning of coordination structures is construed.

Among the first attempts to offer a precise formalization of the syntax and semantics of coordination is the seminar work of Gazdar (1980). Other seminal work soon followed, including the demonstration that phrase structure grammar offered a way to model filler-gap dependencies and certain island constraints (Gazdar 1981). In particular, Gazdar's account showed how long-distance dependencies involving multiple gaps linked to the same filler phrase could be modeled straightforwardly, something that mainstream movement-based models still struggle to this day. Finally, there was also in-depth examinations of a number of complex empirical phenomena like Gazdar et al. (1982), which proved highly influential in the genesis of Generalized Phrase-Structure Grammar, and later, of HPSG. Coordination thus has a special place in the history of HPSG, and still figures in many theoretical arguments within generative grammar given the extremely challenging phenomena it poses for linguistic theory. Nevertheless, there is no clear consensus, even within HPSG, about how to analyze coordination. For example, in some accounts the coordinator expression is a weak head, whereas in others it is a marker. Coordinate structures are binary branching in some accounts but not so in others. Finally, in some accounts non-constituent coordination involves some form of deletion, but in others no deletion operation is assumed. In this chapter we survey the empirical arguments and formal accounts of coordination, with special focus on its morphosyntax.

2 Headedness

The head of a construction is traditionally defined as the unit which determines the syntactic distribution and the meaning of the whole, and it also often the case that a dependent can be omitted, fronted, or extraposed while the head cannot (Zwicky 1985). In coordination constructions something very different occurs. First, the category and distribution of a coordinate phrase is typically collectively determined by the conjuncts, not one particular conjunct nor by the coordination particle. Thus, an S coordination yields an S, a VP coordination yields a VP and so on, for virtually all categories.[1] This is perhaps clearer in cases like (1), where expressions such as *simultaneously*, *both*, *together* can be used to show that the entire bracketed string is interpreted as a complex unit denoting a plurality.

(1) a. [[Tom sang]$_S$ and [Mia danced]$_S$]$_S$ simultaneously.
b. Often [[Kim goes to the beach]$_{VP}$ and [Sue goes to the city]$_{VP}$]$_{VP}$.
c. Sue [[read the instructions]$_{VP}$ and [dried her hair]$_{VP}$]$_{VP}$, in twenty seconds.
d. You can't simultaneously [[drive a car]$_{VP}$ and [talk on the phone]$_{VP}$]$_{VP}$.
e. Simultaneously [[shocked]$_A$ and [saddened]$_A$]$_A$, Robin decided to go home.
f. Robin is both [[tall]$_A$ and [thin]$_A$]$_A$.
g. [[Tom]$_{NP}$ and [Mia]$_{NP}$]$_{NP}$ agreed to jump into the water together.

Generally, a coordinate structure has the same grammatical function and category as the conjuncts: given a number of conjuncts of category X, the distribution of the coordinate constituent that is obtained is again the same as of an X constituent, what Pullum & Zwicky (1986) refer as 'Wasow's Generalization'. In particular, this is what allows coordination to apply recursively:

(2) a. [[Tom and Mary]$_{NP}$ or [Mia and Sue]$_{NP}$]$_{NP}$ got married.
b. I can either [[sing and dance]$_{VP}$ or [sing and play the guitar]$_{VP}$]$_{VP}$.
c. Either [[John went to Paris and Kim went to Brussels]$_S$ or [none of them ever left home]$_S$]$_S$.

Another piece of evidence in favor of a non-headed analysis comes from the fact that there is no typological correlation between the position of the coordinator and the head directionality Zwart (2005). For example, in Zwart's survey of 136 languages where half are verb-final and half are verb-initial languages, verb-final languages overwhelmingly employ initial conjunction strategies. In particular, 119 of these languages have exclusively initial conjunctions, 12 languages exhibit both initial and final conjunctions, and only 4 have exclusively final conjunctions.

[1] The exceptions include expletive pronouns (e.g. **It and there is no use*), which cannot be coordinated, and in fact coordinator expressions themselves, e.g. **You ordered a coffee and or or a tea?* The oddness of the former is presumably due to the fact that expletives are devoid of any meaning, and the oddness of the latter may be due to the conjuncts being of the wrong semantic type. See §5 for more on lexical coordination.

Finally, Coordination is also special in that the relationship between conjuncts is unlike adjunction (Levine 2001). Whereas adjuncts can in principle be displaced, conjuncts do not have any mobility, as (3) illustrates.

(3) a. Because/Since Jane likes music, Tom learned to play the piano.
 b. * And Jane likes music, Tom learned to play the piano.

Thus, no conjunct can usually be said to be a dependent, or secondary. For example, reversing the order of the conjuncts in (4) causes no major change in meaning. Neither daughter can be said to be the head because no subordination dependency is established between conjuncts.

(4) a. Sam ordered a burger and Robin ordered a pizza.
 b. Robin ordered a pizza and Sam ordered a burger.

To be sure, there are certain coordination structures which do not have such symmetric interpretations, as noted by Ross (1967); see also Goldsmith (1985), Lakoff (1986), and Levin & Prince (1986). Regardless, such constructions retain many of the properties that characterize coordinate structures, and therefore are likely coordinate just the same (Kehler 2002: Ch.5).

(5) a. Robin jumped on a horse and rode into the sunset.
 b. Robin rode into the sunset and jumped on a horse.

For these reasons, HPSG adopts a rather traditional non-headed analysis of coordination, an approach going back to Bloomfield (1933: 195) and Ross (1967), and later adopted in many other frameworks such as Pesetsky (1982), Gazdar (1980), Huddleston et al. (2002a: 1275), among many others. See Borsley (1994), Borsley & Jones (2005) and Chaves (2007: Ch.2) for more discussion about previous claims in the literature that coordination structures are headed. Finally, we note that the HPSG account is in agreement with Chomsky (1965: 196), who argued against postulating complex syntactic representations without direct empirical evidence:[2]

> It has sometimes been claimed that the traditional coordinate structures are necessarily right-recursive (Yngve, 1960) or left-recursive (Harman, 1963, p. 613, rule 3i). These conclusions seem to me equally unacceptable. Thus to

[2] In more recent times Chomskyan theorizing has assumed that all structures should be binary branching purely on conceptual ‘economy’ grounds; see Johnson & Lappin (1999) for criticism.

> assume (with Harman) that the phrase "a tall, young, handsome, intelligent man" has the structure [[[[tall young] handsome] intelligent] man] seems to me no more justifiable than to assume that it has the structure [tall [young [handsome [intelligent man]]]]. In fact, there is no grammatical motivation for any internal structure, […] The burden of proof rests on one who claims additional structure beyond this (Chomsky 1965: 196,197)

As we shall see, the empirical evidence suggests that the simplest and most parsimonious structure for coordination is neither left- nor right-recursive. In this chapter we discuss how coordination is analyzed in HPSG, as well as issues that it faces.

3 On the Syntax of Coordinate Structures

In this paper we refer to expressions like *and*, *either*, *or*, *but*, *let alone*, etc. as **coordinators** and the phrases that a coordinator can combine with as **coordinands**. Thus, in 'A or B' both A and B are coordinands and 'or' is the coordinator.

There are a wide range of coordination strategies in the languages of the world Haspelmath (2007). In some languages no coordinand is accompanied by any coordinator (syndenton coordination; as in *We came, we saw, we conquered*), or one of the conjuncts is accompanied by a coordinator (monosyndenton coordination; as in *We came, we saw, and we conquered*). Other strategies involve marking multiple coordinands with a coordinator (polysindenton coordination; *We came, and we saw, and we conquered*) or all coordinands (omnysyndenton coordination; *Either you come or you go*). All of these are schematically depicted in (6); Drellishak & Bender (2005a) for more discussion about how to accommodate such typological patterns in a computational HPSG platform.

(6)
a. A, B, C (asyndenton)
b. A, B, *coord* C (monosyndenton)
c. A *coord* B *coord* C (polysyndenton)
d. *coord* A *coord* B *coord* C (omnisyndenton)

Finally, a single coordination strategy often serves to coordinate all types of constituent phrases, but in many languages different coordination strategies only cover a subset of the types of phrases in the language. For example, in Japanese the suffix *to* is used for nominal coordination and *te* is used for other coordinations.

In what follows we start by focusing on monosyndenton coordination. There are three possible structures one can assign to such coordinations, as Figure 1 illustrates. The binary branching approach goes back to Yngve (1960), and is used in HPSG work such as Pollard & Sag (1994), Yatabe (2002a), Crysmann (2003), Beavers & Sag (2004), Drellishak & Bender (2005a), Abeillé (2005), and Borsley & Jones (2005), Chaves (2007), Chaves (2012b), among others. The flat branching approach has also been assumed in HPSG, albeit less frequently. See for example Sag et al. (2003) and Sag (2003).

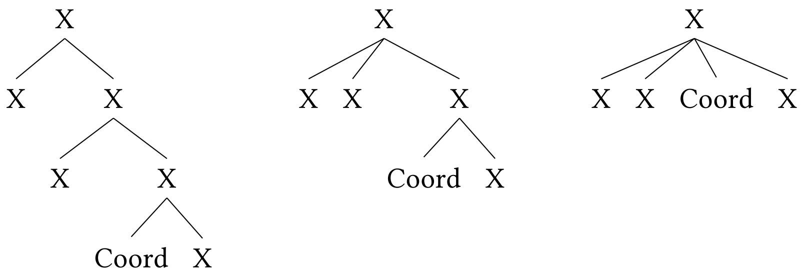


Figure 1: Three possible headless analyses of coordination

The binary branching analysis requires two different rules, informally depicted in (7), and a special feature to prevent the coordinator to recursively apply to the last coordinand, e.g. ***Robin and and and Kim*. Otherwise, the two rules are unremarkable and are handled by the grammar like any other immediate dominance schema. See for example Beavers & Sag (2004) for a formalization.

(7) a. $X_{crd+} \rightarrow$ *Coord* X_{crd-}
  b. $X \rightarrow X_{crd-}\ \ X_{crd+}$

Notice that the hierarchical structure is not the same as the one advocated by Kayne (1994), Johannessen (1998): they both argue that coordination follows X-bar theory and that the conjunction is the head of the construction (see HPSG and Minimalism chapter for discussion). In HPSG, even though one of the conjuncts (or more) may combine with a conjunction, this subconstituent is not the head of the construction, which is considered as unheaded. The two analyses are contrasted in Figure 2.

Similarly, the flat analysis where the coordinator and the coordinand attach to each other requires two rules well (where $n \geq 1$):

(8) a. $X_{crd+} \rightarrow$ *Coord* X_{crd-}

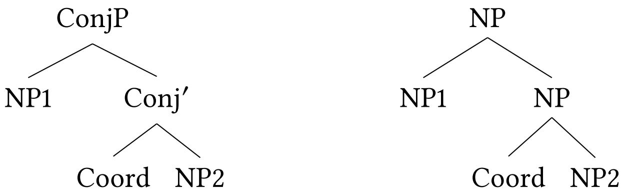


Figure 2: Binary-branching analyses of coordination, headed and non-headed

b. $\mathrm{X} \rightarrow \mathrm{X}^1_{crd-} \ \ldots \ \mathrm{X}^n_{crd-} \ \mathrm{X}_{crd+}$

However, the flat analysis requires only one rule, and no special features at all, as (3) illustrates.

(9) $\mathrm{X} \rightarrow \mathrm{X}^1 \ldots \mathrm{X}^n$ *Coord* X_{n+1}

That said, there are some reasons for assuming that the coordinator does in fact combine with the coordinand, as in (3). First, in some languages of the world the coordinator is a bound morpheme instead of a free morpheme. For example, in Abelam (Papua New Guinea), verbs can be coordinated by adding one of a set of suffixes to one of the coordinands, usually the prior one. And similarly, in Kanuri (Nilo-Saharan) verb phrases are coordinated by marking the first verb with the a conjunctive form affix. In languages like Telugu (Dravidian) the coordination of proper names is marked by the lengthening of their final vowels (Drellishak & Bender 2005b):

(10) kamalaa wimalaa poDugu
Kamala Vimala tall
'Kamala and Vimala are tall'

Second, as Ross (1967) originally noted, the natural intonation break occurs before the coordination lexeme rather than between the coordinator and the coordinand, so that a prosodic constituent is formed. Although prosodic phrasing is not generally believed to always align with syntactic phrasing, the fact that the coordinator prosodifies with the coordinand suggests that it forms a unit with it.

The analysis in Figure 3 can be formalized in HPSG as shown in (11), using parametric lists (Pollard & Sag 1994) to enforce that all conjuncts structure-share the morphosyntactic information. The type *n(on)e(mpty)-list* corresponds to a list that has at least one member, and when used parametrically as in (11) it in addition requires that every member of the list must bear the features [SYNSEM [CAT $\boxed{1}$]].

(11) COORDINATION CONSTRUCTION (preliminary)

$$\textit{coord-phr} \rightarrow \begin{bmatrix} \text{SYNSEM} \begin{bmatrix}\text{CAT } \boxed{1}\end{bmatrix} \\ \text{DTRS} \left\langle \begin{bmatrix}\text{SYNSEM} \begin{bmatrix}\text{CAT } \boxed{1}\end{bmatrix}\end{bmatrix} \right\rangle \oplus \textit{ne-list}\left(\begin{bmatrix}\text{SYNSEM} \begin{bmatrix}\text{CAT } \boxed{1}\end{bmatrix}\end{bmatrix} \right) \end{bmatrix}$$

The constraint forcing all daughters to be of the same category is excessive, as we shall see below, and this will have to undergo a revision. For now, we are focusing on standard coordinations.

In order to account for the fact that different kind of coordination strategies are possible, Mouret (2005) defines three subtypes of *coord-phr*, assuming a lexical feature COORD to distinguish between coordination types:[3]

(12)
$$\textit{simple-coord-phr} \rightarrow \begin{bmatrix}\text{DTRS} \left\langle \begin{bmatrix}\text{COORD} \;\; \textit{nil}\end{bmatrix} \right\rangle \oplus \textit{ne-list}\left(\begin{bmatrix}\text{COORD } \boxed{1}\textit{crd}\end{bmatrix} \right)\end{bmatrix}$$
$$\textit{omnisyndetic-coord} \rightarrow \begin{bmatrix}\text{DTRS } \textit{ne-list}\left(\begin{bmatrix}\text{COORD } \boxed{1}\textit{crd}\end{bmatrix} \right)\end{bmatrix}$$
$$\textit{asyndetic-coord} \rightarrow \begin{bmatrix}\text{DTRS } \textit{ne-list}\left(\begin{bmatrix}\text{COORD} \;\; \textit{none}\end{bmatrix} \right)\end{bmatrix}$$

Here, we assume that the value of COORD must be typed as *coord*, and that the latter has various sub-types as shown in Figure 3.

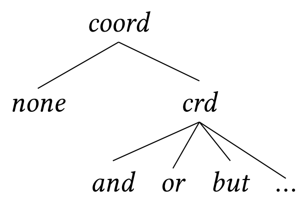


Figure 3: Coordinator sub-types

Thus, simple (monosyndenton and polysyndenton) coordinations are those where all but the first coordinand are allowed combine with a coordinator, omnisyndenton coordinations are those where all coordinands have combined with a coordinator, and likewise, asyndenton coordinations are those where none of the coordinands have combined with a coordinator.

[3] Mouret's formulation is slightly different in that the relevant feature is instead called CONJ, and a slightly different type hierarchy is assumed, with negative constraints like CONJ $\neq$ *nill* are employed instead of COORD *crd*. The current formulation is more in line with standard practice of avoiding negative constraints. Similar liberty is taken in subsequent constraints, for exposition purposes.

We turn to the analysis of coordinators. In other words, what exactly are words like *and*, *or*, and others, and how do they combine with coordinands?

3.1 The status of coordinator expressions

In HPSG, coordinators are sometimes analyzed as markers (Beavers & Sag 2004; Drellishak & Bender 2005a). In such a view, the coordinator's lexical entry does not select any arguments, since it has no arguments. In (3.1) we show the lexical entry for the conjunction, using current HPSG feature geometry. Note that the MRKG value of of the coordinator is the same as the coordinand's. Thus, if *and* coordinates S nodes that are MRKG *that* (i.e. CPs) then the result will be an S that is also MRKG *that*, and so on, for any given value of MRKG.[4]

(13)
$$\begin{bmatrix} \textit{coord-lexeme} & \\ \text{PHON} & \langle \textit{and} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{coord} & \\ \text{SEL} & \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{COORD} & \textit{none} \\ \text{MRKG} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{COORD} & \textit{and} \\ \text{MRKG} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

This sign imposes constraints on the head sign it combines with via the feature SEL, the same feature that allows other markers and adjuncts in general to combine with their hosts. The syntactic construction that allows such elements with their selected heads is the Head-Functor Construction in (3.1). Since the second daughter is the head, the value of the mother's HEAD feature will have to be the same as the head daughter's, as per the Head Feature Principle.

[4]The semantics and pragmatics of conjunction or of coordination more broadly is a particularly complex topic to which we cannot to justice here, specially when it comes to interactions with other phenomena such as scope and collective, distributive, and reciprocal readings. See the Koenig & Richter (2019), Chapter C of this volume for more discussion and in particular Copestake et al. (2005), Fast (2005), Chaves (2007: ch.4–6), Chaves (2012b), Chaves (2012a), Chaves (2009), and Park (2019) for HPSG work that specifically focuses on the semantics of coordination.

(14)
$$head\text{-}func\text{-}cxt \rightarrow \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{VAL} & \boxed{0} \\ \text{COORD} & \boxed{1} \\ \text{MRGK} & \boxed{3} \end{bmatrix} \end{bmatrix} \\ \text{HD-DTR} & \boxed{2} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{SEL} & \boxed{2} \\ \text{COORD} & \boxed{1} \\ \text{MRKG} & \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \boxed{2}\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{VAL} & \boxed{0} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Thus, the conjunction projects an NP when combined with an NP, an AP when combined with an AP, etc., as Figure 4 illustrates.

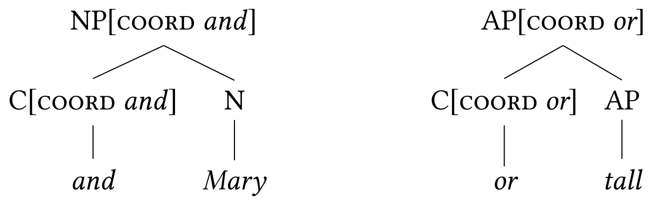


Figure 4: Coordinate marking constructions

An alternative HPSG account that yields almost the same representation through different means is adopted by Abeillé (2003), Abeillé (2005), Mouret (2007) and Bîlbîie (2017) and others, instead takes coordinators to be *weak heads*, i.e. heads which inherit most of their syntactic properties from their complement, like argument-marking prepositions, for example. Thus, the coordinator combines with coordinands via the same headed constructions that license non-coordinate structures. It preserves the Marking feature when conjuncts are themselves marked. The conjunction takes the adjacent conjunct as a complement. This captures its being first in head initial languages like English, and its final position in head final languages like Japanese.

(15) a. Kim [and Lee] (English)
 b. Lee-to Kim (Japanese)
 'Kim and Lee'

Since it is a weak head, it inherits most of its syntactic features (HEAD, MARKING, XARG) from its complement, and adds its own COORD feature. The relevant constraint over all such coordinator lexemes is shown in (16).

(16)
$$\begin{bmatrix} \textit{coord-lexeme} & \\ \text{PHON} & \langle \textit{and} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{0} \\ \text{COORD} & \textit{and} \\ \text{MRKG} & \textit{none} \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} & \boxed{1} \oplus \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \boxed{0} \\ \text{XARG} & \boxed{1} \\ \text{MRKG} & \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The ‘weak head’ analysis is illustrated in Figure 5. Here, the category of the coordinator, the conjunct and of the mother node are the same, because the coordinator’s head value is lexically required to be structure-shared with the head value of its valents.

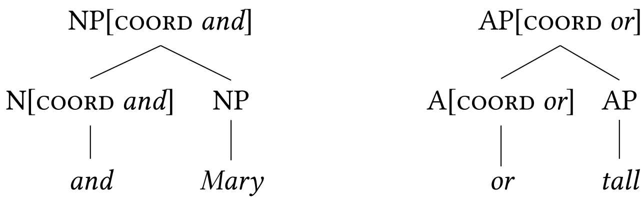


Figure 5: Coordinate weak-head constructions

Before moving on, we note that the ‘weak head’ analysis of coordinators makes certain problematic predictions that the marker analysis in (3.1) does not make. Since coordinands are selected as arguments in the former approach, additional assumptions need to be made in order to prevent Principle extraction of conjuncts as in (3.1). If coordinands are in ARG-ST then they are expected to be extractable (see Borsley & Crysmann (2019), Chapter 14 of this volume and Chaves (2019), Chapter 16 of this volume).

(17) * Which boy did you compare Robin and _ ?
(cf. with *which boy did you compare Robin with _ ?*)

For this reason, the members of ARG-ST of the coordinator are typed as *canonical* by Abeillé (2003), to prevent their extraction, analogously to how prepositions in most languages must prevent their complements from being extracted, unlike English and a few other languages. See Abeillé (2006) for a weak head analysis of certain French prepositions.

3.2 Correlative coordination

Having discussed standard coordination structures, we now move on to cases where multiple inter-dependent coordinators are present, such as *either … or …*, *neither … nor …*, and *both … and …*. Given the linearization flexibility of the first coordinator, they can be analyzed in English as adverbials rather than as true coordinators:

(18) a. Either Fred bought a cooking book or he bought a gardening magazine.
  b. Fred either bought a cooking book or he bought a gardening magazine.
  c. Fred can either buy a cooking book or he can buy a gardening magazine.

(19) a. John will read both the introduction and the conclusion.
  b. John will both read the introduction and the conclusion.

In French, as in other Romance languages, the conjunction itself can be reduplicated, and it is obligatory for some conjunctions (*soit* 'or' in French) (Mouret 2007; Bîlbîie 2017):

(20) a. Jean lira et l'introduction et la conclusion.
  Jean read.fut and the introduction and the conclusion
  b. * Jean et lira l'introduction et la conclusion.
  Jean and read.fut the introduction and the conclusion
  c. Jean lira *(soit) l'introduction soit la conclusion.
  Jean read.fut or the introduction or the conclusion

Thus, there are different structures for different types of correlative, as Figure 6 illustrates. The one on the left is for correlatives that exhibit adverbial properties and the one on the right is for correlatives that do not. See Bîlbîie (2008) for arguments that both types are attested in Romanian.

The correlative coordinate structure on the right is covered by (12), since it requires the COORD feature to be the same for all conjuncts.

3.3 Comparative correlatives

When there is no overt conjunction, it is not always clear whether a binary clause construction is coordinate or not. Comparative correlatives such as (21) have been analyzed as coordinate by Culicover & Jackendoff (1999) for English and as universally subordinate by den Dikken (2006).

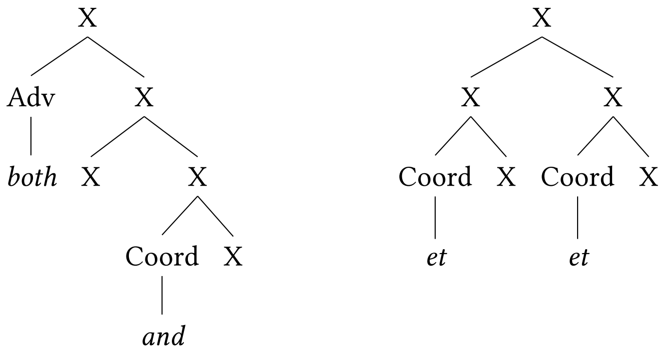


Figure 6: Two possible structures for correlative coordination

(21) The more I read, the more I understand.

On the semantic side, the interpretation is something like: 'if I read more, I understand more'. Abeillé (2006), Abeillé & Borsley (2008) propose that they are are coordinate in some languages, and subordinate in others. In English, one can add the adverb *then*, whereas in French, one can add the conjunction *et* ('and'). In English, the first clause can also be used as a standard adjunct (22).

(22) a. The more I read, then the more I understand.
  b. Plus je lis (et) plus je comprends.
  More I read (and) more I understand.
  c. I understand more, the more I read.

As shown by Culicover & Jackendoff (1999: 549-550), the second clause show matrix clause properties, not the first one:

(23) a. The more we eat, the angrier you get, don't you?
  b. * The more we eat, the angrier you get, don't we?

Syntactic parallelism seems to be stricter in French, for example clitic inversion or extraction must take place out of both clauses at the same time (Abeillé & Borsley 2008):

(24) a. Paul a peu de temps: aussi plus vite commencera-t-il, plus vite aura-t- il fini.
  Paul has little of time so more fast start.FUT-he more fast aux.FUT he finish.PPART
  'Paul has little time left: so the faster he starts, the faster he will

finish'

b. C' est un livre que plus tu lis, plus tu apprécies.
this is a book COMP more you read.2SG, more you appreciate.2SG
'This is a book that the more you read the more you like.'

In Spanish, they come in two varieties: one that can be analyzed as subordinate (a) and one that can be analyzed as coordinate (b). These are illustrated in (25a,b), respectively.

(25) a. Cuanto más leo, (tanto) más entiendo.
how-much more read.1SG, (that-much) more understand.1SG
'The more I read, the more I understand'

b. Más leo (y) más entiendo.
more read.1SG (and) more understand.1SG
'The more I read, the more I understand'
(Abeillé et al. 2006)

Be they coordinate or subordinate, they are special kinds of construction: they are binary, with a fixed order : the meaning changes if the order is reversed (a). The internal structure of each clause is also special. In English, it must start with 'the' and a comparative phrase (b), which may belong to a long distance dependency (c), and can be analyzed as extracted. Each clause must be finite and allow for copula omission (d).

(26) a. The more I understand, the more I read.
b. * I understand (the) more, I read (the) more.
c. The more I manage to read, the more I start to understand.
d. The more intelligent the students, the better the marks.

Comparative clauses are a special subtype of finite clause, starting with a comparative phrase. Abeillé et al. (2006), Borsley (2011) define a CORREL feature which is a LEFT EDGE feature (see the EDGE feature in Bonami et al. (2004) for French liaison). Assuming a degree word *the*, which can only appear as a specifier of a comparative word, Borsley (2011) defines *the*-clause as a subtype of head-filler-phrase with [CORREL *the*]; see also Sag (2010).

Comparative correlatives belong to a more general class of (binary) correlative constructions, including *as* … *so* …, and *if* … *then* … constructions in (Borsley

2004; Borsley 2011) [5] Correlative constructions can be defined as follows, where *correl-construction* is a sub-type of *declarative-clause* and the feature CORREL introduces a *correl* type hierarchy analogous to that of *coord* in Figure 3 above. The construction in (27) thus states that all correlative constructions have in common the fact that both daughters are marked by a special expression.

(27)
$$\textit{correl-construction} \rightarrow \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{finite} \\ \text{CORREL} & \textit{none} \end{bmatrix} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{array}{l} \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{CORREL} & \textit{corr-mrk} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \\ \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{CORREL} & \textit{corr-mrk} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{array} \right\rangle \end{bmatrix}$$

Naturally, *correl-construction* has various sub-types, each imposing particular patterns of correlative marking, including coordinate correlatives. More specifically, this family of constructions comes in two varieties: asymmetric (for the subordinate ones, like English comparative correlatives and Spanish type a correlatives), and symmetric for coordinate ones, like French comparative correlatives and Spanish type b correlatives). The symmetric subtype inherits from clausal-coordination-phrase, while the asymmetric one inherits from the head-adjunct-phrase as seen in Figure 7.

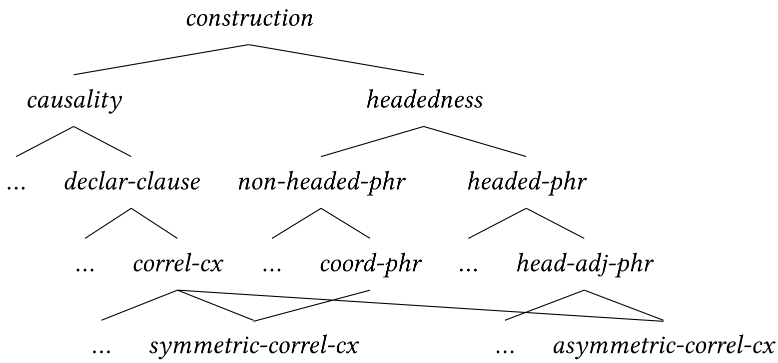


Figure 7: Type hierarchy for correlative constructions

Thus, asymmetric English comparative correlatives can be defined as in (28), where *the* is a sub-type of *corr-mrk* (i.e. is a coordinate marker).

[5] This does not handle Hindi type correlatives, which differ in that only the first clause is introduced by a correlative word, and the first clause is mobile and optional; see Pollard & Sag (1994: 228) for an analysis.

(28)
$$
\textit{asymmetric-cc-cx} \rightarrow \begin{bmatrix} \textsc{hd-dtr}\ \boxed{1} \\ \textsc{dtrs} \left\langle \begin{array}{l} \begin{bmatrix} \textsc{synsem} \begin{bmatrix} \textsc{cat} \begin{bmatrix} \textsc{correl}\ \ \textit{the} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \\ \boxed{1} \begin{bmatrix} \textsc{synsem} \begin{bmatrix} \textsc{cat} \begin{bmatrix} \textsc{correl}\ \ \textit{the} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{array} \right\rangle \end{bmatrix}
$$

Similarly, symmetric French comparative correlatives can be defined as in (29), where *et* and *compar* are subtypes of *corr-mrk* and *none* is the sub-type of *correl* indicating the absence of correlative marking.

(29)
$$
\textit{symmetric-cc-cx} \rightarrow \begin{bmatrix} \textsc{dtrs} \left\langle \begin{array}{l} \begin{bmatrix} \textsc{synsem} \begin{bmatrix} \textsc{cat} \begin{bmatrix} \textsc{correl}\ \ \textit{compar} \end{bmatrix} \end{bmatrix} \end{bmatrix}, \\ \begin{bmatrix} \textsc{synsem} \begin{bmatrix} \textsc{cat} \begin{bmatrix} \textsc{coord}\ \ \textit{none} \vee \textit{et} \\ \textsc{correl}\ \textit{compar} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{array} \right\rangle \end{bmatrix}
$$

A more complete analysis would take into account the semantics as well (Sag 2010). From a syntactic point of view, HPSG seems to be in a good position to handle both the general properties and the idiosyncrasy of the CC construction, as well as its crosslinguistic variation. See Borsley (2011) for a comparison with a tentative minimalist analysis.

4 Phrasal coordination and feature resolution

4.1 Feature sharing between coordinands

The coordination construction in (11) requires the value of CAT to be structure-shared across the coordinands and the mother node. Given the large number of features within CAT, such a constraint makes a series of predictions and mis-predictions. For example, this entails that all valence constraints are identical. Thus, in VP coordination all nodes have an empty COMPS list and share exactly the same singleton SUBJ list, as illustrated in Figure 8. Thus, nothing needs to be said from the semantic composition side: the verbs will have to share exactly the same referent for their subject. The same goes for any other combination of categories, of whatever part-of-speech.

All the unsaturated valence arguments become one and the same for all co-ordinands, and it becomes impossible to have daughters with different subcat-egorization information. For example, if one daughter requires a complement while the other does not, CAT identity is impossible. This correctly rules out a coordination of VP and V categories like the one in (30a), or S and VP as in (30b):

(30) a. * Fred [read a book]$_{\text{COMP}\langle\,\rangle}$ and [opened]$_{\text{COMP}\langle\text{NP}\rangle}$.

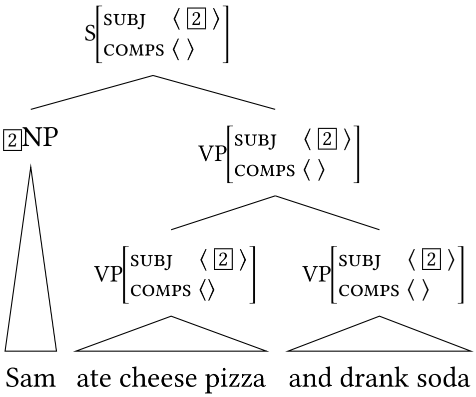


Figure 8: Valence identity in coordination

b. * Fred [she has a hat]$_{\text{SUBJ}\langle\,\rangle}$ and [smiled]$_{\text{SUBJ}\langle\text{NP}\rangle}$.

But there is other information in CAT besides valence. For example, the head feature VFORM encodes the verb form, not tense or aspectual information. The coordination of inconsistent VFORM values is ruled out as ungrammatical as seen in (31), while consistent values of VFORM are accepted as illustrated by (32).

(31) a. * Tom [whistled]$_{\text{VFORM}\ fin}$ and [walking]$_{\text{VFORM}\ prp}$.
b. * Sue [buy something]$_{\text{VFORM}\ prs}$ and [came home]$_{\text{FORM}\ fin}$.

(32) a. Tom [is married]$_{\text{VFORM}\ fin}$ and [bought a house in the suburbs]]$_{\text{VFORM}\ fin}$
b. Sue [buys groceries here]$_{\text{VFORM}\ fin}$ and [could be interested in working with us]$_{\text{VFORM}\ fin}$.
c. Dan [protested for two years]$_{\text{VFORM}\ fin}$ and [will keep on protesting]$_{\text{VFORM}\ fin}$

Yet another feature that resides in the CAT value of verbal expressions is the head feature INV, which indicates whether a given verbal expression is invertable or not. Hence, inverted structures cannot be coordinated with non-inverted ones:

(33) a. [Sue has sung in public]$_{\text{INV}-}$ and [Kim has tap-danced]$_{\text{INV}-}$?
b. * [Sue has sung in public]$_{\text{INV}-}$ and [has Kim tap-danced$_{\text{INV}+}$?

(34) a. [Elvis is alive]$_{\text{INV}-}$ and [there was a CIA conspiracy]$_{\text{INV}-}$?
b. * [Elvis is alive]$_{\text{INV}-}$ and [was there a CIA conspiracy]$_{\text{INV}+}$?

But the inverted clause precedes the non-inverted one, then such coordinations become somewhat more acceptable. In fact, Huddleston et al. (2002b: 1332-3) note attested cases like (35).

(35) Did you make your own contributions to a complying superannuation fund and your assessable income is less than $31,000?

A similar problem arises for the feature AUX, which distinguishes auxiliary verbal expressions from those that are not auxiliary:

(36) a. [I stayed home]$_{\text{AUX}-}$ but [Fred could have gone fishing]$_{\text{AUX}+}$.

b. [Tom went to NY yesterday]$_{\text{AUX}-}$ and [he will return next Tuesday]$_{\text{AUX}+}$

c. Fred [sang well]$_{\text{AUX}-}$ and [will keep on singing]$_{\text{AUX}+}$.

However, this problem vanishes in the account of the English Auxiliary System detailed in Sag et al. (2020), since in that analysis the feature AUX does not indicate whether the verb is auxiliary or not. Rather, the value of AUX for auxiliary verbs is resolved by the construction in which the verb is used. Since all the constructions in (36) are canonical VPs (e.g. non-inverted), then all the conjuncts in (36) are specified as AUX−, in the Sag et al. (2020) analysis.

Similarly, argument-marking PPs cannot be coordinated with modifying-PPs simply because the former are specified with different PFORM and SELECT values. This explains the contrast in (37). The first PP is the complement that *rely* selects but the second is a modifier. Thus, they have different CAT values and cannot be coordinated.

(37) a. Kim relied on Mia on Sunday.

b. * Kim relied on Mia and on Sunday.

Consequently, it is in general not possible to coordinate argument marking PPs headed by different prepositions, simply because they bear different PFORM values as shown in (38).

(38) a. *Kim depends [[on Sandy]$_{\text{PFORM } \textit{on}}$ or [to Fred]$_{\text{PFORM } \textit{to}}$].

b. *Kim is afraid [[of Sandy]$_{\text{PFORM } \textit{of}}$ and [to Fred]$_{\text{PFORM } \textit{to}}$].

Similarly, adjectives that are specified as PRED+ cannot be coordinated with PRED- adjectives, without stipulation:

(39) a. * I became [former]$_{\text{PRED}-}$ and [happy]$_{\text{PRED}+}$.

b. * He is [happy]$_{\text{PRED}+}$ and [Fred]$_{\text{PRED}-}$.

c. * [Mere]$_{\text{PRED}-}$ and [happy]$_{\text{PRED}+}$, Fred rode on into the sunset.

Since case information is also part of CAT, it follows that conjuncts must be consistent as in (40).[6] Many other examples of CAT mismatches exist, but the aforegoing list suffices to illustrates the breadth of predictions that follow from the feature geometry of CAT and the constraints imposed by the coordination construction.

(40) a. * I saw [her$_{acc}$ and he$_{nom}$].

b. * He likes [she$_{nom}$ and me$_{acc}$].

Of course, mispredictions also exist. We already discussed one, concerning the feature INV, but there are others. For example, requiring that the GAP value of the coordinands be the same readily predicts Coordinate Structure Constraint effects like (41) and (42), but it incorrectly rules out asymmetric coordination violation cases like (43). See Goldsmith (1985), Lakoff (1986), and Levin & Prince (1986) for more examples and discussion.

(41) a. [To him]$_{\boxed{1}NP}$ [Fred gave a football _]$_{\text{GAP}\langle\boxed{1}\rangle}$ and [Kim gave a book _]$_{\text{GAP}\langle\boxed{1}\rangle}$

b. * [To him]$_{\boxed{1}NP}$ [Fred gave a football _]$_{\text{GAP}\langle\boxed{1}\rangle}$ and [Kim gave me a book]$_{\text{GAP}\langle\,\rangle}$

c. * [To him]$_{\boxed{1}NP}$ [Fred gave a football to me]$_{\text{GAP}\langle\,\rangle}$ and [Kim gave a book _]$_{\text{GAP}\langle\boxed{1}\rangle}$

(42) a. It offers [something]$_{\boxed{1}NP}$ [that every kid wants _]$_{\text{GAP}\langle\boxed{1}\rangle}$ and [that every parent tries to help their child to achieve _]$_{\text{GAP}\langle\boxed{1}\rangle}$

b. * It offers [something]$_{\boxed{1}NP}$ [that every kid wants _]$_{\text{GAP}\langle\boxed{1}\rangle}$ and [that every parent tries to help their child to achieve it]$_{\text{GAP}\langle\,\rangle}$

c. * It offers [something]$_{\boxed{1}NP}$ [that every kid wants it]$_{\text{GAP}\langle\,\rangle}$ and [that every parent tries to help their child to achieve _]$_{\text{GAP}\langle\boxed{1}\rangle}$

(43) a. [Who]$_{\boxed{1}NP}$ did Sam [[pick up the phone]$_{\text{GAP}\langle\,\rangle}$ and [call _ $_i$]$_{\text{GAP}\langle\boxed{1}\rangle}$?

b. What was the maximum amount$_{\boxed{1}}$ that I can [contribute _ $_i$]$_{\text{GAP}\langle\boxed{1}\rangle}$ and [still get a tax deduction]$_{\text{GAP}\langle\,\rangle}$?

[6] There are nonetheless collocational cases where the distribution of pronouns defies this pattern, due to presumably prescriptive forces; see Grano (2006) and Lohmann (2014).

Chaves (2012b) argues that, since there are no independent grounds to assume that asymmetric coordination is anything other than coordination, the coordination construction must not impose GAP identity across conjuncts. Rather, the Coordinate Structure Constraint, and its asymmetric exceptions are best analyzed as pragmatic in nature, as Kehler (2002) argues. See Borsley & Crysmann (2019), Chapter 14 of this volume for more discussion. In practice this means that the coordination construction should impose identity of some of the features in CAT, though not all.

Like in the case of locally specified valents, the category of the extracted phrase is also structure-shared in coordination. Hence, case mismatches like (44) are correctly ruled out.

(44) a. * [Him]$_{acc}$, [all the critics like to praise _]$_{\text{GAP}\langle \text{NP}acc\rangle}$ but [I think _ would probably not be present at the awards]$_{\text{GAP}\langle \text{NP}nom\rangle}$

There are, however, cases where the case of the ATB-extracted phrase can be syncretic as in (4.1), due to Levine et al. (2001).

(45) a. Robin is someone who$_i$ even [good friends of _$_i$] believe _$_i$ should be closely watched.

b. We went to see [a movie]$_{\boxed{1}nom\_acc}$ [which the critics praised _]$_{\text{GAP}\langle\boxed{1}\rangle}$ but [that Fred said _ would probably be too violent for my taste]$_{\text{GAP}\langle\boxed{1}\rangle}$

The feature CASE is responsible for identifying the case of nominal expressions. Pronouns like *him* are specified as *acc*(*usative*), and pronouns like *I* are *nom*(*inative*), and expressions like *who* or *Robin* are left underspecified for case. According to Levine et al. (2001: 207), the case system of English involves the hierarchy in Figure 9.

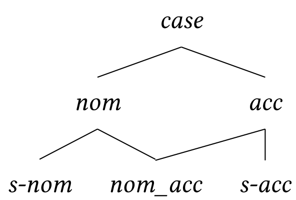


Figure 9: Type hierarchy of (structural) case assignments

Verbs subcategorize for *s-nom* NP subjects and *s-acc* NP complements. Most nouns and some pronouns like *who* and *what* are underspecified for case, and thus typed as *case*, which makes them consistent with both nominative and accusative positions. Hence, *a movie* can be simultaneously be required to be consistent with s-nom and s-acc, by resolving into syncretic type *nom_acc*, which is a subtype of both *s-nom* and *s-acc*. Pronouns like *him* and *her* are specified as *acc* and therefore are not compatible with the *nom_acc* type. The same goes for *nom* pronouns like *he* and *she*, etc. Hence, the problem of case syncretism is easily solved.

4.2 Coordination and agreement

Another thorny issue for syntactic theory and coordination structures concerns agreement. According to Pollard & Sag (1994), agreement information is introduced by the index feature in semantics, not morphosyntax. Hence, different expressions with inconsistent person, gender and number specifications are free to combine. But Wechsler & Zlatic (2003) have also argued that there should be a distinct feature called concord, which is morphosyntactic in nature (See Wechsler (2019), Chapter 6 of this volume). The motivation for this move, is that there are languages, like Serbo-Croatian, which display hybrid agreement:

(46) Ta dobra deca su doš-l-a
that.sing-fem goodsing.fem children aux3-pl comepprt.pn
'Those good children came'
(Wechsler & Zlatic 2003: 51)

The collective noun *deca* ('children') triggers feminine singular (morphosyntactic) agreement on NP-internal items, in this case the determiner ιta ('that') and the adjective *dobra* ('good'). This is a problem for a locality theory of agreement. On the other hand, there are other HPSG analyses that argue that what appears to be closest conjunct agreement is in fact agreement with the whole coordinate NP, which has additional features inherited from the first and last conjuncts. feature. (Villavicencio et al. 2005) propose two additional features: LAGR (for the left most conjunct) and RAGR (for the right most conjunct) for determiner and (attributive) adjective agreement in Romance, which involves the CONCORD feature. Semantic agreement (i.e. Concord), on the other hand, is seen in the verb *su*, which is inflected for third person plural, in agreement with the semantic properties of the subject *deca*. The two kinds of agreement are also visible in English:

(47) a. This/*These committee made a decision.

b. The committee have/has made a decision.

This predicts that inconsistent CONCORD specifications are not possible:

(48) * These committee have arrived and has made a decision.

Since (morphosyntacic) agreement is not recorded in CAT, it follows that the resolution of agreement information in coordination must be processed elsewhere in the grammar. There are usually strict and non-trivial constraints involved in determining what the agreement of the mother node is given that of the coordinands. We turn to this problem below.

4.3 Agreement

In case of conjuncts with conflicting agreement values, resolution strategies are observed crosslinguistically. For example, a coordination with a 1st person is 1st, and a coordination with 2nd person (and no 1st person) is 2nd person:

(49) a. Paul and I like ourselves.

b. Paul and you like yourselves.

In gender marking languages, coordination with conflicting gender values is often resolved to the masculine, at least for animates (Corbett 1991). This is illustrated in (50), for Portuguese.

(50) a. o homem e a mulher modernos
the.MSG man.MSG and the.FSG woman.FSG modern.MPL
'the modern man and woman'

b. morbidez e morte prematuras
mobidity.FSG and death.FSG premature.FPL premature morbidity and death
(Villavicencio et al. 2005: 433)

To keep unification as a combining operation, Sag (2003) proposes that 1st person is a supertype of 2nd person which is itself a supertype of 3rd person. Addressing gender resolution, Gabriel & Crysmann (2018) propose a list-based encoding of person and gender values, and list concatenation as a combining operation. They crucially distinguish between a feature itself and its resolution potential. Masc has a non empty list and fem has an empty list, the coordination of *masc* and *fem* yields a non empty list, hence *masc*. Person uses two lists ME and YOU, 1st person has a non empty ME list, 2d person has an empty ME list and a non empty YOU list, 3rd person has both empty lists. This enables person and gender resolution by list concatenation over conjuncts, and it has been implemented in a LKB grammar of French.

4.3.1 Closest Conjunct agreement

As observed by Corbett (1991), many languages including Romance, Celtic, Semitic and Bantu languages, also have another strategy, namely partial agreement with only one conjunct, the one closest to the target, called closest conjunct agreement (CCA). In the following example, again from Portuguese, the determiner and prenominal adjective agree with the first Noun (a) and the postnominal adjective with the last Noun (b).

(51) a. suas próprias reações ou julgamentos
his.FPL own.FPL reactions.FPL or judgements.MPL
'his own reactions or judgements'
(Villavicencio et al. 2005: 435)

b. Esta cancão anima os corações e mentes brasileiras.
This.FSG song.FSG animates the.MPL hearts.MPL and minds.FPL Brazilian.FPL
'This song animates Brazilian hearts and minds'
(Villavicencio et al. 2005: 437)

For French determiners and attributive adjectives, An & Abeillé (2017), (Abeillé et al. 2018) show on the basis of corpus data and experiments that number agreement may also obey CCA. As far as gender is concerned, prenominal adjectives always obey CCA while postnominal ones half of the time (in contemporary French). In (a), the determiner can be singular (CCA) or plural (resolution), while in (b), CCA (feminine Det) is obligatory. In (c), the postnominal adjective can be masculine (resolution) or feminine (CCA), with the same meaning.

(52) a. Votre/ Vos nom et prénom
your.SG/PL name.MSG and first name.MSG
'your name and first name'
(An & Abeillé 2017)

b. Certaines / *Certains régions et départements
certain.FPL/*MPL region.FPL and department.MPL
'certain regions and departments'
(Abeillé et al. 2018)

c. Des départements et régions importants/importantes
some department.MPL and region.FPL important.MPL/FPL
'some important departments and regions'

As proposed by Wechsler & Zlatic (2003), HPSG distinguishes two agreement features: CONCORD is used for morphosyntactic agreement and INDEX is used for semantic agreement (see Wechsler (2019), Chapter 6 of this volume). Moosally (1999) proposes an account of single conjunct predicate-argument agreement in Ndebele, which she analyses as INDEX agreement. She has a version of the following constraint that shares the Index value of the (nominal) coordination with that of the last conjunct:

(53)
$$\textit{nom-coord-phrase} \rightarrow \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{DTRS} & \left\langle [\ ], \ldots, \begin{bmatrix} \text{INDEX} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix}$$

But in other languages, such as Welsh, there is evidence that the INDEX of the coordinate structure is resolved, even though predicate-argument agreement is controlled by the closest conjunct.

(54) Dw i a Gwenllian heb gael ein talu.
be.1SG I and Gwenllian.3SG without get Cl.1PL pay
'Gwenllian and I have not been paid'
(Sadler 2003, 12)

This is why (Borsley 2009) proposes that CCA is superficial in Welsh and uses linearization domain to handle partial agreement (between the initial verb and the first conjunct, which are not sisters). On the other hand, determiner and (attributive) adjective agreement in Romance involves the CONCORD feature. (Villavicencio et al. 2005) propose two additional features: LAGR (for the left most conjunct) and RAGR (for the right most conjunct). Nouns have the same value for CONCORD, LAGR and RAGR.

(55)
$$\textit{nom-coord-ph} \rightarrow \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{LAGR} & \boxed{1} \\ \text{RAGR} & \boxed{2} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{LAGR} & \boxed{1} \end{bmatrix}, \ldots, \begin{bmatrix} \text{RAGR} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix}$$

(56) $$noun \longrightarrow \begin{bmatrix} \textsc{lagr} & \boxed{1} \\ \textsc{ragr} & \boxed{1} \\ \textsc{concord} & \boxed{1} \end{bmatrix}$$

Attributive adjectives constrain the agreement features of the noun they modify (via the MOD feature). One may distinguish two types for prenominal and postnominal adjectives, by the binary LEX ± feature (Sadler & Arnold 1994) or by the WEIGHT light/non-light feature (Abeillé & Godard 1999). In this perspective, each has its agreement pattern, which we simplify as follows, using '∨' to express a disjunction of feature values:

(57) $$\textit{prenominal-adj} \longrightarrow \begin{bmatrix} \textsc{concord} & \boxed{1} \\ \textsc{sel} & \begin{bmatrix} \textsc{lagr} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

(58) $$\textit{postnominal-adj} \longrightarrow \begin{bmatrix} \textsc{concord} & \boxed{1} \vee \boxed{2} \\ \textsc{sel} & \begin{bmatrix} \textsc{concord} & \boxed{1} \\ \textsc{ragr} & \boxed{2} \end{bmatrix} \end{bmatrix}$$

In the absence of coordination, these constraints apply vacuously, since CONCORD, LAGR and RAGR all share the same values. If CCA needs to take place to the right and to the left, it is difficult to handle in Minimalism which views agreement as a directional operation between a target (with uninterpretable features) and a c-commanded probe (with interpretable features).

5 Lexical coordination

While conjuncts have often been assumed to be phrasal (see for example Kayne (1994) and Bruening (2018) a.o.), Abeillé (2006) gives several arguments in favor of lexical coordination. In some contexts, words are allowed but not full phrases. In English, it is the case with prenominal adjectives and postverbal particles. See Abeillé (2006) for similar examples with various categories in different languages. Most English attributive adjectives are prenominal unless they have a complement. Although the prenominal position is not allowed for full AP, it is possible for coordinate adjectives.

(59)
a. A tall / proud man
b. * A [taller than you] man
c. * A [proud of his work] man
d. A [big and tall] man

As observed by Pollard & Sag (1987), a particle may project a PP after the nominal complement, but not before; but coordination is possible, at least for some speakers.

(60) a. Paul turned (*completely) off the radio.
b. Paul turned the radio (completely) off.
c. Paul was turning [on and off] the radio all the time.

While phrasal coordination can conjoin unlike categories (see below), it is not the case with lexical coordination:

(61) a. Paul is [head of the school] [and proud of it].
b. # Paul is [head and proud] of the school.

Semantically, lexical coordination is more constrained than phrasal coordination. With *and*, two lexical verbs, sharing a preverbal clitic in French, must share the same verbal root, and in Spanish, they must refer to the same event (Bosque 1986).

(62) a. Je te [dis et redis] que tu as tort.
I you tell and retell that you have wrong
'I'm telling you that you are wrong'
b. # Je te [dis et promets] que tu as tort.
I you tell and promise that you have wrong
'I'm telling and promising you that you are wrong'
c. Lo [compro y vendio] en una sola operacion.
It buy.1SG and sell.1SG in a single operation
'I buy and sell it in one single operation'
d. *Lo compro hoy y vendio mañana.
It buy.1SG today and sell.1SG tomorrow
'I buy it today and sell it tomorrow'

Some apparent cases of lexical coordination may be analyzed as Right-Node Raising (Beavers & Sag 2004). They differ semantically and prosodically, however: with typical Right-Node Raising, the two conjuncts must stand in contrast to one another, and do not have to refer to the same event. With Right-Node Raising there must be a prosodic boundary at the ellipsis site (see Chaves (2014) and Nykiel & Kim (2019), Chapter 20 of this volume). In French, the first conjunct cannot end with a clitic article or with a weak preposition as in (63b,c),

(63) a. Tout le monde dit et je te promets [que tu as tort].
'everyone says and I promise you [that you are wrong]'

b. * Paul cherche le, et Marie connaît la responsable.
'Paul looks for the.MSG and Marie knows the.FSG responsible'

c. * Paul parle de, et Marie discute avec Woody Allen.
'Paul speaks of and Marie talks with Woody Allen'

No such boundary occurs before the conjunction in lexical coordination. Thus, in French, clitic articles or weak prepositions can be conjoined, with a shared argument (Abeillé 2006):

(64) a. Paul cherche [le ou la] responsable
Paul looks for the.MSG or the.FSG responsible
'Paul is looking for the man or woman in charge'

b. Un film [de et avec] Woody Allen
a film by and with Woody Allen

c. ?? un film [de mais sans] Woody Allen
a film by but without Woody Allen

Not all conjunctions are felicitous with lexical coordination; *but*, for example is less felicitous than *and* or *or*. Analyzing the conjunction as a 'weak' head (see above), the sub-type for lexical coordination has to allow for the coordination of items waiting for complements: the conjunction (this is done by concatenation of ARG-ST lists as it is for complex predicates, see Godard & Samvelian (2019), Chapter 12 of this volume). It thus inherits all the dependents of the word it combines with.

(65)
$$\textit{lex-coord} \rightarrow \begin{bmatrix} \text{HEAD} & \boxed{0} \\ \text{WEIGHT} & \boxed{3}\,\textit{light} \\ \text{ARG-ST} & \boxed{1} \oplus \left\langle \begin{bmatrix} \text{HEAD} & \boxed{0} \\ \text{WEIGHT} & \boxed{3} \\ \text{ARG-ST} & \boxed{1} \oplus \boxed{2} \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

The construct resulting from the coordination of lexical elements has hybrid properties: as a syntactic construct, it must be a phrase, but it also behaves as a word : coordinate verbs behave as lexical heads, coordinate adjectives may occur in positions ruled out for phrases. To overcome this apparent paradox, Abeillé (2006) analyses it as an instance of "light" phrase, following the WEIGHT theory of Abeillé & Goodard (2000) and Abeillé & Goodard (2004). Light elements can be words or phrases, and can have a restricted mobility (see Müller (2019), Chapter 10 of this volume). For example prenominal modifiers can be constrained to be [WEIGHT light]. In this theory, light phrases can be coordinate phrases or head-adjunct phrases, provided all their daughters are light as Figure 10 illustrates.

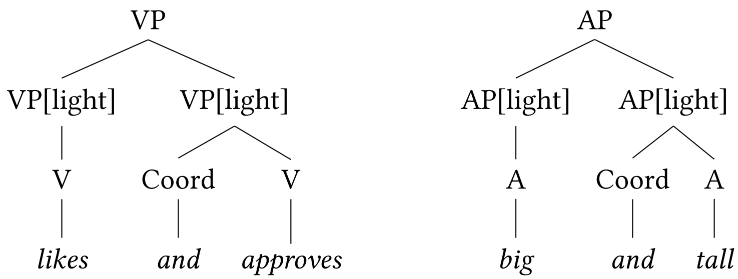


Figure 10: Examples of lexical coordination

6 Coordination of unlike categories

The coordination construction in (11) requires that the categories being coordinated are the same. However, there is some evidence that this requirement is excessive. Consider the coordinations in (66), from Gazdar et al. (1985), Bayer (1996), Huddleston et al. (2002b), among others. Such data pose a classic syntactic problem: what is the part of speech and categorial status of the bracketed constituents?

(66) a. Kim is [alone and without money].
[AP & PP]

b. Pat is [a Republican and proud of it].
[NP & AP]

c. Jack is [a good cook and always improving].
[NP & VP]

d. What I would love is [a trip to Fiji and to win $10,000].
[NP & VP]

e. That was [a rude remark and in very bad taste].
[NP & PP]

f. Chimpanzees hunt [frequently and with an unusual degree of success].
[AdvP & PP]

g. I'm planning [a four-month trip to Africa and to return to York afterwards].
[NP & VP]

As Jacobson (1987) pointed out, it is clear that the features of the mother are not simply the intersection of the features of the conjuncts. Verbs like *remain* are compatible with both AdjP and NP complements whereas *grew* is only compatible with AdjPs. This is shown in (67). Crucially, however, the information associated with the phrase *wealthy and a Republican* somehow allows *grew* to detect the presence of the nominal, as (68a) illustrates, even when the verbs are coordinated, as in (68b–d)

(67) a. Kim remained/grew wealthy.
b. Kim remained/*grew a Republican.

(68) a. Kim remained/*grew [wealthy and a Republican].
b. Kim grew and remained wealthy.
c. * Kim grew and remained a Republican.
d. * Kim grew and remained [wealthy and a Republican].

A number of influential accounts in Type-logical grammar (Morrill 1990; 1994; Bayer 1996) have used one of the rules of inference from propositional calculus in order to deal with coordination of unlikes phenomena, namely, disjunction introduction (or addition): from P one can infer $P \vee Q$. Thus, by assuming that categories like NP, PP and so on can also be disjunctive, the grammar allows an expression of type 'NP' to lead a double life as an 'NP ∨ PP' expression, or the type 'AP' to be taken as an 'AP ∨ PP ∨ NP' and so on. This kind of approach has been adopted in various forms into HPSG, see for example Daniels (2002) and Yatabe (2004). Related work aims to achieve the same result using type-underspecification, such as Sag (2002). Other, more exploratory work, views coordination of unlike categories as the result of parts-of-speech being gradient and phenomenal rather than hard-coded into the type signature (Chaves 2013). In the latter work, all coordination of unlikes boils down to coordination of like categories.

Other work, like Crysmann (2000), Yatabe (2002b), Beavers & Sag (2004), Chaves (2006) argue that coordination of unlikes can be explained by a deletion operation that omits the left periphery of non-initial conjuncts, illustrated in (69).

(69) a. Tom gave a book to Mary, and ~~gave~~ a magazine to Sue.

b. He drinks coffee with milk at breakfast and ~~drinks coffee with~~ cream in the evening.
(Hudson 1984)

c. There was one fatality yesterday, and ~~there were~~ two others on the day before.
(Chaves 2007: 339).

d. I see the music as both going backward and ~~going~~ forward.
[http://pdxjazz.com/dave-holland; 20 December 2010]

In such a view, (66) are verbal coordinations where the verb (or the verb and the subject) has been deleted (e.g. *Kim is alone and ~~is~~ without money*). The problem is that left-periphery ellipsis cannot fully explain coordination of unlikes phenomena. For example, there is no elliptical analysis of data like (70). Levine (2011) offers arguments against the coercion account of Chaves (2006), and against the existence of left-periphery ellipsis. See Yatabe (2012) for a reply.

(70) a. Simultaneously shocked and in awe, Fred couldn't believe his eyes.

b. Both tired and in a foul mood, Bob packed his gear and headed North.
(Chaves 2006)

c. Both poor and a Republican, *no one* can possibly be.

d. Dead drunk and yet in complete control of the situation, *no one* can be.
(Levine 2011)

Further problems for an ellipsis account of coordination of unlikes phenomena are posed by the position of the correlative coordinators *both*, *either* and *neither* in (71).

(71) a. Isn't this both illegal and a safety hazard?

b. It's both odd and in very poor taste to have a fake wedding.

c. Who's neither tired nor in a hurry?

d. Isn't she either drunk or on medication?

If (71a) is an elliptical coordination like *isn't this both illegal and ~~isn't this~~ a safety hazard*, then the location of *both* is unexpected. Instead of occurring before the first conjunct, it is realized inside the first conjunct. Crucially, the non-elided counterparts are not grammatical, e.g. **isn't this both illegal and isn't this a safety hazard?* The same issue is raised by (71b,c). In an elliptical account one would have to stipulate that *both* can only float in the presence of ellipsis, which is unmotivated. Finally, see Mouret (2007) for an extensive discussion in favor of a non elliptical analysis of unlike coordination, based on correlative coordination. In sum, left-periphery ellipsis does not offer a complete account of coordination of unlikes, and underspecification accounts are more promising.

7 Non-constituent Coordination

The fact that not all coordination of unlike categories can be reduced to deletion does not entail that deletion is impossible, or that no phenomena involve deletion. Consider for example the constructions in (72), all of which involve non-canonical (i.e. non-constituent) coordination, some of which were already discussed above. We refer the reader to Nykiel & Kim (2019), Chapter 20 of this volume for more discussion about other types of ellipsis.

(72) a. Tom gave a book to Mary, and a magazine to Sue.
(Argument Cluster Coordination)

b. Tom loves – and Mary absolutely hates – spinach dip.
(Right-Node Raising)

c. Tom knows how to cook pizza, and Fred – spaghetti.
(Gapping)

Some authors regard Argument Cluster Coordination as elliptical (Yatabe 2001; Crysmann 2004; Beavers & Sag 2004) others regard such phenomena as base-generated (Mouret 2006a). In the former, phonological material in the left periphery of the non-initial conjunct that is identical to phonological material in the left periphery of the initial conjunct is allowed to be not present in the mother node. This can be achieved by adding the constraints in (73) to the coordination construction, here shown in the binary-branching format, for perspicuity. Some accounts operate directly on PHON, others apply to DOM elements instead. See Müller (2019), Chapter 10 of this volume.

(73)
$$\textit{coord-phr} \Rightarrow \begin{bmatrix} \text{PHON } \boxed{1}\oplus\boxed{2}\oplus\boxed{3} \\ \text{DTRS } \left\langle \begin{bmatrix}\text{PHON } \boxed{1}\oplus\boxed{2}_{\textit{ne-list}}\end{bmatrix}, \begin{bmatrix}\text{PHON } \boxed{1}\oplus\boxed{3}_{\textit{ne-list}}\end{bmatrix} \right\rangle \end{bmatrix}$$

If $\boxed{1}$ is resolved as the empty list then no ellipsis occurs, but if $\boxed{1}$ is non-empty then ellipsis occurs, as illustrated in Figure 11. In other formulations, what is elided are morphophonological linearization domain units; see Chaves (2008).

$$\text{VP}\left[\textsc{phon}\ \boxed{1}\oplus\boxed{2}\oplus\boxed{3}\right]$$

$$\text{VP}\left[\textsc{phon}\ \boxed{1}\langle \textit{give}\rangle\oplus\boxed{2}\langle \textit{a,book,to,Mary}\rangle\right] \quad \text{VP}\left[\textsc{phon}\ \boxed{1}\langle \textit{give}\rangle\oplus\boxed{3}\langle \textit{a,magazine,to,Sue}\rangle\right]$$

Figure 11: Analysis of '*give a book to Mary and ~~give~~ a magazine to Sue*'

This approach is motivated by the existence of ambiguity in sentences like (74), from Beavers & Sag (2004) and Chaves (2006). Because (74a) involves a one-time predicate, the ellipsis must include the subject phrase, otherwise the intepretation is such that the same two trees are cut down twice. In contrast, (74b) does not involve a one-time predicate, and thus is it possible for the ellipsis to simply involve the verb.

(74) a. Two trees were cut down by Robin in July and by Alex in September.
(Two trees were cut down by Robin in July and ~~two trees were cut down~~ by Alex in September.

b. Two trees were photographed by Robin in July and by Alex in September.
(Two trees were photographed by Robin in July and ~~photographed~~ by Alex in September)

In the non-elliptical analysis of such data, the missing material is recovered from the preceding conjunct. For example, Mouret proposes a rule along the lines of (7). Here, a new head feature CLUSTER is introduced, which takes as its value the list of synsem description of the daughters.

(75)
$$\textit{ac-cx} \Rightarrow \begin{bmatrix} \textsc{head} & \begin{bmatrix}\textsc{cluster}\ \langle \boxed{1}\ldots\boxed{n}\rangle\end{bmatrix} \\ \textsc{dtrs} & \left\langle \begin{bmatrix}\textsc{synsem}\ \boxed{1}\end{bmatrix},\ldots \begin{bmatrix}\textsc{synsem}\ \boxed{n}\end{bmatrix}\right\rangle \end{bmatrix}$$

Mouret defines argument-clusters as instances of some underspecified non-headed construction *ac-cx* with one daughter or more. The construction is valence saturated. He also postulates a lexical rule that allow a ditransitive verb to take a CLUSTER as complement (this rule will also allow clusters for complements and adjuncts, assuming the latter are included in the COMPS list).

(76)
$$\begin{bmatrix}\text{COMPS}\ \langle [\text{CAT 1}]\ldots[\text{CAT N}]\rangle\end{bmatrix} \Rightarrow \begin{bmatrix}\text{COMPS}\left\langle\begin{bmatrix}\text{COORD} & + \\ \text{CLUSTER} & \langle [\text{CAT 1}]\ldots[\text{CAT N}]\rangle\end{bmatrix}\right\rangle\end{bmatrix}$$

This approach is motivated by non clausal conjunctions (*as well as, ainsi que*), which are possible in Argument Cluster Coordination, but cannot conjoin tensed VPs:

(77) a. John gave a book to Mary as well as a magazine to Sue.
b. * John gave a book to Mary as well as gave a magazine to Sue.
c. Paul offrira un disque à Marie ainsi qu'un livre à Jean.
'Paul will.offer a record to Mary as well as a book to Jean'
(Abeillé & Goodard 1996)

Another argument is the placement of correlative conjunctions: the first conjunction in (a) must be postverbal; this shows that Argument Cluster Coordination does not include the first verb. The examples below are from Mouret (2006b: 254).

(78) a. Jean a donné et un livre à Marie et un magazine à Sue.
'Jean has given both a book to Marie and a magazine to Sue'
b. Paul compte offrir et un disque à Marie et un livre à Jean.
Paul is.planning.to offer and a record to Marie and a book to Jean
c. * Paul compte et offrir un disque à Marie et un livre à Jean.
'Jean is.planning and to.offer a record to Marie and a book to Jean'

Another argument is negation placement, which is a case of constituent negation (Mouret 2006b: 253):

(79) a. Paul offrira un disque à Marie et (non) pas un livre à Jean.
'Paul will offer a record to Marie and not a book to Jean'
b. Paul gave a record to Mary and not a book to Bill.
c. * Paul gave a record to Mary and not gave a book to Bill.

A syntactic and non-elliptical account of Right-Node Raising is harder to maintain given that this phenomenon does not seem to be sensitive to syntactic structure as (80) shows. See Bresnan (1974), Wexler & Culicover (1980: 299), Grosu (1981: 45), and McCawley (1982), and Sabbagh (2007: 382,ft.30) for more data and discussion.[7]

[7] Steedman (1985; 1990; 2001) and Dowty (1988: 183) claim that Right-Node Raising is bounded, nonetheless. For example, Dowty (1988) argues that **an idea that, and a robot which* [*can solve this problem*] is evidence for islands in RNR. But as Phillips (1996: 95) points out, this oddness is explained by semantic factors: it is impossible to semantically contrast *that* (which

(80) a. I know a man who SELLS and you know a person who BUYS [pictures of Elvis Presley].

b. John wonders when Bob Dylan WROTE and Mary wants to know when he RECORDED [his great song about the death of Emmet Till].

c. Politicians WIN WHEN THEY DEFEND and LOSE WHEN THEY ATTACK [the right of a woman to an abortion].

d. Lucy CLAIMED that – but COULDN'T SAY exactly when – [the strike would take place].

e. I found a box IN which and Andrea found a blanket UNDER which [a cat could sleep peacefully for hours without being noticed].

Another source of evidence against syntactic and non-elliptical accounts of Right-Node Raising comes is that this phenomenon can involve lexical structure, as (81) illustrates

(81) a. Please list all publications of which you were the SOLE or CO-[author]. (Huddleston et al. 2002b: 1325, ft. 44).

b. It is neither UN- nor OVERLY [patriotic] to tread that path.

c. The EX- or CURRENT [smokers] had a higher blood pressure. (Chaves 2008)

d. The NEURO- and COGNITIVE [sciences] are presently in a state of rapid development (...) [opinionator.blogs.nytimes.com/2011/12/25/the-future-of-moral-machines/?hp]

e. Are you talking about A NEW or about AN EX-[boyfriend]?

Elliptical accounts of Right-Node Raising are proposed by Beavers & Sag (2004), Yatabe (2004), Chaves (2014) and others. The rule in (82) illustrates the account adopted by Chaves (2014) and Shiraïshi et al. (2019) in simplified format.[8] In a nutshell, the M(ORPHO-)P(HONOLOGY) feature introduces two list-valued features, namely PHON(OLOGY) and L(EXICAL-)ID(ENTIFIER). The former encodes phonological content, including phonological phrasing information, whereas the latter is used to individuate lexical items semantically (i.e. the value of LID is a list of semantic frames that canonically specify the meaning of a lexeme).

is semantically vacuous) with *which*. Steedman (2001: 17) argues that RNR exhibits islands effects by claiming that *I hope that I will meet the woman* WHO WROTE *and you expect to interview the consortium* THAT PUBLISHED [*that novel about the secret life of legumes*] is ungrammatical. In our experience, informants do not systematically share this judgment.

[8]See Chaves (2014) for more details about how 'cumulative' Right-Node Raising is modeled by this rule, i.e. cases like *Mia lost – and Fred spent – (a total of) $10.000.*

(82)
$$\textit{rpe-cx} \Rightarrow \begin{bmatrix} \text{PHON} & \boxed{L_1}\oplus\boxed{R_1}\oplus\boxed{R_2}\oplus\boxed{R_3} \\ \text{SYNSEM} & \boxed{0} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{MP} & \boxed{L_1}\oplus\boxed{L_2}\left\langle \begin{bmatrix}\text{PHON} & \boxed{p_1}\\ \text{LID} & \boxed{1}\end{bmatrix},\ldots,\begin{bmatrix}\text{PHON} & \boxed{p_n}\\ \text{LID} & \boxed{n}\end{bmatrix}\right\rangle\oplus \\ & \boxed{R_1}\oplus\boxed{R_2}\left\langle \begin{bmatrix}\text{PHON} & \boxed{p_1}\\ \text{LID} & \boxed{1}\end{bmatrix},\ldots,\begin{bmatrix}\text{PHON} & \boxed{p_n}\\ \text{LID} & \boxed{n}\end{bmatrix}\right\rangle\oplus\boxed{R_3} \\ \text{SYNSEM} & \boxed{0}\end{bmatrix}\right\rangle\end{bmatrix}$$

Requiring PHON identity ensures that the Right-Node Raising construction only targets strings that are phonologically independent, and have the same surface form, ruling out (83). Stressed pronouns, affixes that correspond to independent prosodic words, and compound parts can be RNRaised because they are independent prosodic units in their local domains. See Swingle (1995) for more discussion.

(83) a. He tried TO PERSUADE but he couldn't CONVINCE [THEM] / *[them].
 * I think that I'D and I know that PAT'LL [buy those portraits of Elvis].
 * They've always WANTED a – and so I've GIVEN THEM a – [coffee grinder]
 * I bought EVERY RED and Jo liked SOME BLUE [t-shirt].

Conversely, LID identity prevents homophonous strings that have fundamentally different semantics from being Right-Node Raised, as in (84). In such cases, oddness arises because in general the same phrase cannot simultaneously have two meanings, except in puns (Zaenen & Karttunen 1984: 316).

(84) a. * Randy SAW and Rene has BEEN [flying planes].
b. * Jo WILL and Sandy BUILT THE [drive].
(Milward 1994)
c. * Mary FED and Tom ENJOYED [the lamb].
(adapted from Buitelaar (1998: 64))
d. * Robin SWUNG and Leslie TAMED [an unusual bat].
(Levine & Hukari 2006: 156)
e. * We need new BLACK- and FLOOR[boards].
(Artstein 2005)
f. * We caught BUTTER- and FIRE[flies].
(Chaves 2008)
g. * There stood a ONE- and WELL-[armed man].
(Chaves 2014)

At the same time, LID identity does not go as far as requiring co-referentiality of the shared material. This is as intended given ambiguous examples like *Chris* LIKES *and Bill* LOVES [*his bike*]. The account of Right-Node Rasing is illustrated below.

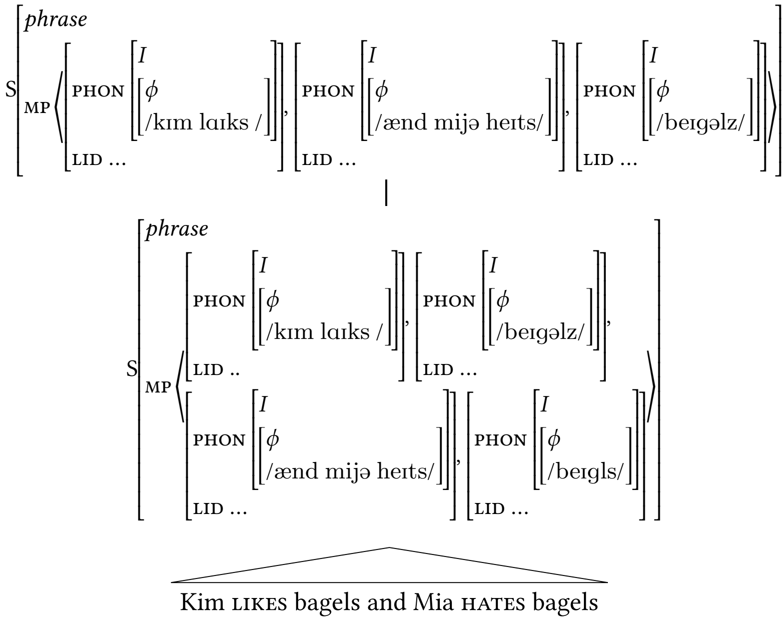


Figure 12: Analysis of *Kim likes, and Mia hates, bagels*

Note that this is a unary-branching rule, which means that it can in principle apply to any phrasal node, which allows the account to scale to non-coordinate RNR:

(85) a. It's interesting to compare the people who LIKE with the people who DISLIKE [the power of the big unions]. (Hudson 1976: 550)

b. Anyone who MEETS really comes to LIKE [our sales people]. (adapted from Williams (1990))

c. Spies who learn WHEN can be more valuable than those able to learn WHERE [major troop movements are going to occur].

d. Politicians who fought FOR may well snub those who have fought AGAINST [chimpanzee rights]. (Postal 1994)

e. Those who voted AGAINST far outnumbered those who voted FOR [my father's motion].
(Huddleston et al. 2002b: 1344)

f. If there are people who OPPOSE then maybe there are also some people who actually SUPPORT [the hiring of unqualified workers].
(Chaves 2014)

In the example in Figure 12 the sub-list $\boxed{R_3}$ is resolved as the empty list, but this need not be so. When the latter sublist is not resolved as the empty we obtain discontinuous Right-Node Raising cases like (86), where the RNRaised expression is followed by extra material.

(86) a. The blast UPENDED and <u>NEARLY SLICED</u> [an armored Chevrolet Suburban] <u>in half</u>.

b. During the War of 1982, American troops OCCUPIED and <u>BURNED</u> [the town] <u>to the ground</u>.

c. Please move from the exit rows if you are UNWILLING or <u>UNABLE</u> [to perform the necessary actions] <u>without injury</u>.

d. The troops that OCCUPIED ended up <u>BURNING</u> [the town] <u>to the ground</u>.

Finally, let us now turn our attention to Gapping, as in *Robin likes Sam and Tim – Sue*. There are elliptical accounts of Gapping (Chaves 2006) as well as direct-interpretation accounts where the missing material is recovered from the preceding linguistic context (Mouret 2006a; Mouret 2006b; Abeillé et al. 2014; Park 2019); See Nykiel & Kim (2019), Chapter 20 of this volume. The latter is illustrated in Figure 13, in simplified format. Basically, the Question Under Discussion (QUD) of the first clause is $\lambda y.\lambda x.\exists e(like(x, y))$ which is information that shared across the clausal daughters as $\boxed{1}$. This allows the second conjunct to combine the two NPs with the verbal semantics, and recover the propositional meaning.

Like RNR, Gapping is not restricted to coordinate structures as (87) illustrates, and so the rule that derives REF is not specific to coordination. Thus, the rule that allows a gapped clause to follow a non-gapped clause is not specific to coordination. See Park (2019) for more details and discussion.

(87) a. Robin speaks French better than Leslie – German.

b. My purpose here is not to resolve the crucial disagreement between two prominent theoreticians in a way that one would be declared true while the other one – false.

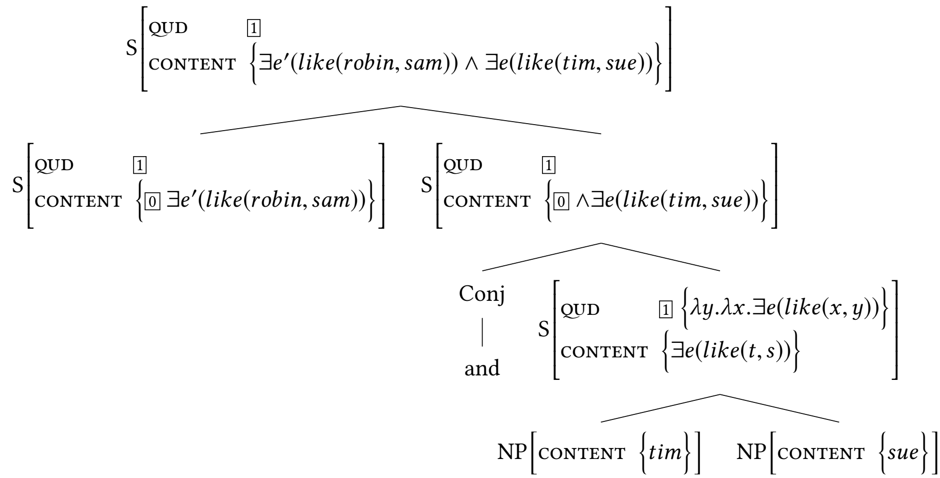


Figure 13: Analysis of *Robin likes Sam and Tim – Sue* (abbreviated)

c. The keynote of their relationship was set when Victoria, already a reigning queen, had to propose to Albert, rather than he – to her.

d. The public remembers all that and usually recognizes us before we – them.
(Park 2019: §2)

8 Conclusion

Coordination is a pervasive phenomenon in all natural languages. Despite intensive research in the last 70 years, its empirical properties continue to challenge most linguistic theories : the coordination lexemes play a crucial role but do not behave like usual syntactic heads, the conjuncts do not need to be identical but display some parallelism relations and can be in unlimited number, some non constituent sequences can be coordinated, peculiar ellipsis phenomena can optionally occur, etc. We have shown how HPSG offer precise detailed analyses of various coordinate constructions for a wide variety of languages, factoring out the common properties shared by other constructions and the properties specific to coordination.

Central to the HPSG analyses are two main ideas : (i) coordination structures are non-headed phrases and come with different subtypes ; (ii) the parallelism between coordinate daughters is captured by feature sharing ; from these ideas, specific properties can be derived, regarding extraction and agreement for instance. Nevertheless, there is no clear consensus about some remaining issues. In some accounts, the coordinator is a weak head, whereas in others it is a marker. Coordinate structures are binary branching in some accounts but not so in others. Agreement is always local (with the whole coordinate phrase) in some account, whereas locality is abandoned by others to account for Closest Conjunct agreement. Finally, in some accounts non-constituent co-ordination involves some form of deletion, but in others no deletion operation is assumed.

Acknowledgements

We are thankful to Bob Borsley, Stefan Müller, and other reviewers for comments and suggestions on earlier drafts. As usual, all errors and omissions are our own.

References

Abeillé, Anne. 2003. A lexicon- and construction-based approach to coordinations. In Stefan Müller (ed.), *Proceedings of the hpsg-2003 conference, michigan state university, east lansing*, 5–25. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/4/.

Abeillé, Anne. 2005. Les syntagmes conjoints et leurs fonctions syntaxiques. *Langages* 160. 42–66.

Abeillé, Anne. 2006. In Defense of Lexical Coordination. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical Issues in Formal Syntax and Semantics*, 7–36. CSSP.

Abeillé, Anne, Aixiu An & Aoi Shiraishi. 2018. L'accord de proximité dans le groupe nominal en français. *Discours* 22.

Abeillé, Anne, Gabriela Bîlbîie & François Mouret. 2014. A Romance perspective on gapping constructions. In Hans C. Boas & Francisco Gonzálvez-García (eds.), *Romance in Construction Grammar*, 227–267. Amsterdam: John Benjamins.

Abeillé, Anne & Robert D. Borsley. 2008. Comparative correlatives and parameters. *Lingua* 118. 1139–1157.

Abeillé, Anne, Robert D. Borsley & Maria-Teresa Espinal. 2006. The syntax of comparative correlatives in french and spanish. In Stefan Müller (ed.), *The proceedings of the 13th international conference on head-driven phrase structure grammar*, 6–26. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/7/.

Abeillé, Anne & Danièle Godard. 1999. La place de l'adjectif épithète en français : le poids des mots. *Recherches Linguistiques* 28. 9–31.

Abeillé, Anne & Danièle Goodard. 1996. La complémentation des auxiliaires en français. *Langages* 122. 32–61.

Abeillé, Anne & Danièle Goodard. 2000. French word order and lexical weight. In Robert D. Borsley (ed.), *The nature and function of syntactic categories, syntax and semantics*, 325–360. New York: Academic Press.

Abeillé, Anne & Danièle Goodard. 2004. De la légèreté en syntaxe. In *Bulletin de la société de linguistique de paris, xcix*, 69–106.

An, Aixiu & Anne Abeillé. 2017. Agreement and interpretation of binominals in french. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky,Lexington*, 26–43. Stanford, CA: CSLI Publications.

Artstein, Ron. 2005. Coordination of parts of words. *Lingua* 115(4). 359–393.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72. 579–616.

Beavers, John & Ivan A. Sag. 2004. Ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th international conference on head-driven phrase structure grammar, katholieke universiteit leuven*, 48–69. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.

Bîlbîie, Gabriela. 2008. A syntactic account of romanian correlative coordination from a romance perspective. In Stefan Müller (ed.), *Proceedings of the 15th international conference on hpsg*, 25–45. Stanford: CSLI Publications.

Bîlbîie, Gabriela. 2017. *Grammaire des constructions elliptiques : une étude comparative des phrases sans verbe en roumain et en français* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press.

Bloomfield, Leonard. 1933. *Language*. London: George Allen & Unwin.

Bonami, Olivier, Gilles Boyé & Jesse Tseng. 2004. An integrated analysis of french liaison. In *Proceedings of the 9th conference on formal grammar*. Stanford: CSLI Publications.

Borsley, Bob. 2011. Constructions, functional heads, and comparative correlatives. In O. Bonami & P. Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics*, vol. 8, 7–20. http://www.cssp.cnrs.fr/eiss8.

Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 1994. In defence of coordinate structures. *Linguistics Analysis* 24. 218–426.

Borsley, Robert D. 2004. An approach to English comparative correlatives. In Stefan Müller (ed.), *Proceedings of the HPSG-2004 conference, center for computational linguistics, katholieke universiteit leuven*, 70–92. Stanford, CA: CSLI Publications. cslipublications . stanford . edu / HPSG / 5/. cslipublications.stanford.edu/HPSG/5/.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language and Linguistic Theory* 27(2). 225–265.

Borsley, Robert & Bob Jones. 2005. *Welsh negation and grammatical theory*. University of Wales Press.

Bosque, Ignacio. 1986. Constricciones morfologicas sobre la coordinacion. In *Lingüistica española actual*, vol. 9, 83–100.

Bresnan, Joan. 1974. On the position of certain clause-particles in phrase structure. *Linguistic Inquiry* 5. 614–619.

Bruening, Benjamin. 2018. The lexicalist hypothesis: both wrong and superfluous. *Language* 49. 1–42.

Buitelaar, Paul. 1998. *Corelex: systematic polysemy and underspecification*. Brandeis University Ph.D. dissertation.

Chaves, Rui. 2019. Island phenomena and related matters. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 415–468. Berlin: Language Science Press. DOI:??

Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. In Stefan Müller (ed.), *Proceedings of the hpsg-2006 conference, bulgarian academy of sciences, varna, bulgaria*, 102–122. Stanford: Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Chaves, Rui P. 2007. *Coordinate structures – constraint-based syntax-semantics processing*. University of Lisbon Doctoral Dissertation.

Chaves, Rui P. 2008. Linearization-based Word-Part Ellipsis. *Linguistics and Philosophy* 31(3). 261–307.

Chaves, Rui P. 2009. Construction-based cumulation and adjunct extraction. In *Proceedings of the 16th international conference on head-driven phrase structure grammar, university of göttingen, germany*, 47–67. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2009/.

Chaves, Rui P. 2012a. Conjunction, cumulation and *respectively* readings. *Journal of Linguistics* 48(2). 297–344.

Chaves, Rui P. 2012b. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30(2). 465–512.

Chaves, Rui P. 2013. Grammatical alignments and the gradience of lexical categories. In Philip Hofmeister & Elisabeth Norcliffe (eds.), *The core and the periphery: data-driven perspectives on syntax inspired by ivan a. sag*, 167–220. CSLI Publications.

Chaves, Rui P. 2014. On the disunity of right node raising phenomena: extraposition, ellipsis, and deletion. 4(90). 834–886.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.

Crysmann, Berthold. 2000. Clitics and coordination in linear structure. In B. Gerlach & J. Grijzenhout (eds.), *Clitics in phonology, morphology and syntax*, vol. 36, 121–159. John Benjamins.

Crysmann, Berthold. 2003. An asymmetric theory of peripheral sharing in HPSG. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of formal grammar 2003*, 47–62.

Crysmann, Berthold. 2004. Underspecification of intersective modifier attachment: some arguments from german. In Stefan Müller (ed.), *Proceedings of the hpsg-2004 conference, center for computational linguistics, katholieke universiteit leuven*, 378–392. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/5/.

Culicover, Peter & Ray Jackendoff. 1999. The view from the periphery: the English comparative correlative. *Linguistic Inquiry* 30. 543–571.

Daniels, Michael W. 2002. On a type-based analysis of feature neutrality and the coordination of unlikes. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *8th international conference on head-driven phrase structure grammar*, 137–147, Stanford: CSLI Publications.
`http://cslipublications.stanford.edu/HPSG/2/`.

den Dikken, Marcel. 2006. Either-float and the syntax of co-or-dination. *Natural Language and Linguistic Theory* 24(3). 689–749.

Dowty, David R. 1988. Type-raising, functional composition, and nonconstituent coordination. In Emmon Bach Richard Oehrle & Dierdre Wheeler (eds.), *Categorial grammars and natural language structures*, 153–198. Dordrecht: Kluwer Academic Publishers.

Drellishak, Scott & Emily M. Bender. 2005a. A coordination module for a crosslinguistic grammar resource. In Stefan Müller (ed.), *The proceedings of the 12th* *international conference on head-driven phrase structure grammar, department of informatics, university of lisbon*, 108–128. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Drellishak, Scott & Emily M. Bender. 2005b. A coordination module for a crosslinguistic grammar resource. In Stefan Müller (ed.), *The proceedings of the 12th international conference on head-driven phrase structure grammar, department of informatics, university of lisbon*, 108–128. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Fast, Jakub. 2005. Structurally underspecified semantics for distributive plural predication - respectively constructions in lexical resource semantics. In Sylvia Blaho, Luis Vicente & Erik Schoorlemmer (eds.), *Proceedings of console xiii*.

Gabriel, Aguila-Multner & Berthold Crysmann. 2018. Feature resolution by list : the case of french coordination. In Annie Foret, Greg Kobele & Sylvain Pogodalla (eds.), *Formal grammar*, 1–15.

Gazdar, Gerald. 1980. A cross-categorial semantics for coordination. *Linguistics and Philosophy* 3(3). 407–410.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12(2). 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized phrase structure grammar*. Oxford: Blackwell, & Cambridge, Ma: Harvard University Press.

Gazdar, Gerald, Geoffrey K. Pullum, Ivan A. Sag & Thomas Wasow. 1982. Coordination and transformational grammar. *Linguistic Inquiry* 13(4). 663–677.

Godard, Danièle & Pollet Samvelian. 2019. Complex predicates. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 315–350. Berlin: Language Science Press. DOI:??

Goldsmith, John. 1985. A Principled Exception to the Coordinate Structure Constraint. In *Papers from the twenty-first annual regional meeting of the chicago linguistic society*, 133–143. Chicago: Chicago Linguistic Society.

Grano, Thomas. 2006. *"ME AND HER" MEETS "HE AND I": CASE, PERSON, AND LINEAR ORDERING IN ENGLISH COORDINATED PRONOUNS*. Honors Thesis. Honors Program Stanford University, School of Humanities and Sciences.

Grosu, Alexander. 1981. *Approaches to island phenomena*. North-Holland.

Haspelmath, Martin. 2007. Coordination. In Timothy Shopen (ed.), *Language typology and syntactic description*, 2nd, vol. 2. Cambridge University Press.

Huddleston, Rodney D., John Payne & Peter Peterson. 2002a. Coordination and supplementation. In Rodney D. Huddleston & Geoffrey K. Pullum (eds.), *The* *cambridge grammar of the English language*, chap. 15, 1273–1362. Cambridge University Press.

Huddleston, Rodney D., John Payne & Peter Peterson. 2002b. The Cambridge Grammar of the English Language. In, chap. 15: Coordination and supplementation, 1273–1362. Cambridge University Press.

Hudson, Richard. 1976. Conjunction reduction, gapping, and right-node-raising. *Language* 52. 535–562.

Hudson, Richard. 1984. *Word grammar*. Blackwell.

Jacobson, Pauline. 1987. Review of Gerald Gazdar, Ewan Klein, Geoffrey K. Pullum, and Ivan A. Sag, 1985: *Generalized Phrase Structure Grammar*. *Linguistics and Philosophy* 10. 389–426.

Johannessen, Janne Bondi. 1998. *Coordination*. New York: Oxford University Press.

Johnson, David E. & Shalom Lappin. 1999. *Local constraints vs. economy* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: The MIT Press.

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Lakoff, George. 1986. Frame Semantic Control of the Coordinate Structure Constraint. In *Papers from the 22nd regional meeting of the chicago linguistic society*. Chicago: Chicago Linguistic Society.

Levin, Nancy & Ellen Prince. 1986. Gapping and clausal implicature. *Papers in Linguistics* 19. 351–364.

Levine, Robert D. 2001. The extraction riddle: just what are we missing? *Journal of Linguistics* 37. 145–174.

Levine, Robert D. 2011. Linearization and its discontents. Presented at the 18th International Conference on HPSG, Seattle, Washington. USA.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions*. CSLI Publications, Leland Stanford Junior University.

Levine, Robert D., Thomas E. Hukari & Michael Calcagno. 2001. Parasitic gaps in English: some overlooked cases and their theoretical implications. In Peter Culicover & Paul M. Postal (eds.), *Parasitic gaps*, 181–222. Cambridge, MA: MIT Press.

Lohmann, Arne. 2014. *English coordinate constructions: A processing perspective on constituent order*. Cambridge, UK: Cambridge University Press.

McCawley, James D. 1982. Parentheticals and discontinuous constituent structure. *Linguistic Inquiry* 13. 91–106.

Milward, David. 1994. Non-constituent coordination: theory and practice. In *Proceedings of coling 94*, 935–941. Kyoto, Japan.

Moosally, Michelle J. 1999. Subject and object coordination in Ndebele: An HPSG analysis. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *Proceedings of the WCCFL 18 conference*. Somerville, MA: Cascadilla Press.

Morrill, Glynn. 1990. Grammar and logical types. In M. Stokhof & L. Torenvliet (eds.), *Proceedings of the seventh amsterdam colloquium*. Amsterdam: Institute for Logic, Language & Information.

Morrill, Glynn. 1994. *Type logical grammar*. Dordrecht: Kluwer.

Mouret, Francois. 2006a. A phrase structure approach to argument cluster coordination in French. In Kiril Simov & Petya Osenova (eds.), *Proceedings of the hpsg-2006 conference, bulgarian academy of sciences, varna, bulgaria*, 101–107. Publishing House ETO, Sofia. `http://www.bultreebank.org/HPSG06/paper16.pdf`.

Mouret, François. 2005. La syntaxe des coordinations corrélatives du français. *Langages* 160. 67–92.

Mouret, François. 2006b. A Phrase Structure Approach to Argument Cluster Coordination. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, 247–267.

Mouret, François. 2007. *Grammaire des constructions coordonnées. Coordinations simples et coordinations à redoublement en français contemporain*. University Paris Diderot dissertation.

Müller, Stefan. 2019. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

Nykiel, Joanna & Jong-Bok Kim. 2019. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 585–612. Berlin: Language Science Press. DOI:??

Park, Sang-Hee. 2019. *Gapping: a constraint-based syntax-semantics interface*. University of Buffalo dissertation.

Pesetsky, David. 1982. *Paths and categories*. Cambridge Mass. PhD. Dissertation, MIT.

Phillips, Colin. 1996. *Order and structure*. Massachusetts Institute of Technology. Distributed by MIT Working Papers in Linguistics PhD. dissertation.

Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics; volume one – fundamentals*. CSLI Lecture Notes No.13. Stanford: CSLI Publications.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press & Stanford: CSLI.

Postal, Paul M. 1994. Parasitic and pseudoparasitic gaps. *Linguistic Inquiry* 25. 63–117.

Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 72. 751–773.

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT dissertation. http://files.eric.ed.gov/fulltext/ED016965.pdf, accessed 2018-2-25. Reproduced by the Indiana University Linguistics Club and later published as Ross (1986).

Ross, John Robert. 1986. *Infinite syntax!* (Language and Being 5). Norwood, New Jersey: Ablex Publishing Corporation.

Sabbagh, James. 2007. Ordering and linearizing rightward movement. *Natural Language and Linguistic Theory* 25(2). 349–401.

Sadler, Louisa & Doug Arnold. 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30(1). 187–226.

Sag, Ivan A. 2002. Coordination and underspecification. In *Proceedings of the 9th international conference on head-driven phrase structure grammar*, Kyung Hee University, Seoul, South Korea, 267–291, Stanford, CSLI Publications. bibrangessep `http://cslipublications.stanford.edu/HPSG/3/`.

Sag, Ivan A. 2003. Coordination and Underspecification. In Jong-Bok Kim & Stephen Wechsler (eds.), *The Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 267–291. Stanford: CSLI Publications.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

Sag, Ivan A., Rui P. Chaves, Anne Abeillé, Bruno Estigarribia, Dan Flickinger, Paul Kay, Laura A. Michaelis, Stefan Müller, Geoffrey K. Pullum, Frank Van Eynde & Thomas Wasow. 2020. Lessons from the English auxiliary system. *Journal of Linguistics* 56(1). pp.69.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Shiraïshi, Aoi, Anne Abeillé, Barbara Hemforth & Philip Miller. 2019. Verbal mismatch in peripheral ellipsis. *Glossa*.

Steedman, Mark. 1985. Dependency and coordination in dutch and english. *Language* 61. 523–568.

Steedman, Mark. 1990. Gapping as constituent coordination. *Linguistics and Philosophy* 13. 207–264.

Steedman, Mark. 2001. *The syntactic process*. MIT Press.

Swingle, Kari. 1995. The role of prosody in right node raising. Ms. UCSC, Santa Cruz, CA.

Villavicencio, Aline, Louisa Sadler & Doug Arnold. 2005. An hpsg account of closest conjunct agreement in np coordination in portuguese. In Stefan Müller (ed.), *The proceedings of the 12th international conference on head-driven phrase structure grammar, department of informatics, university of lisbon*, 427–447. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/6/.

Wechsler, Stephen. 2019. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 131–154. Berlin: Language Science Press. DOI:??

Wechsler, Stephen & Larisa Zlatic. 2003. *The many faces of agreement: morphology, syntax, semantics, and discourse factors in serbo-croatian agreement*. CSLI Lecture Notes.

Wexler, Kenneth & Peter Culicover. 1980. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Williams, Edwin. 1990. The ATB-theory of parasitic gaps. *The Linguistic Review* 6. 265–279.

Yatabe, Shûichi. 2001. The syntax and semantics of left-node raising in Japanese. In Dan Flickinger & Andreas Kathol (eds.), *Proceedings of the 7th international conference on head-driven phrase structure grammar*, 325–344.

Yatabe, Shûichi. 2002a. A linearization-based theory of summative agreement in peripheral node-raising constructions. In J.-B. Kim & S. Wechsler (eds.), *Proceedings of the 9th international conference on head-driven phrase structure grammar*, 391–411. http://cslipublications.stanford.edu/HPSG/3/yatabe.pdf.

Yatabe, Shûichi. 2002b. A linearization-based theory of summative agreement in peripheral node-raising constructions. In J.-B. Kim & S. Wechsler (eds.), *Proceedings of the 9th international conference on head-driven phrase structure grammar*, 391–411.

Yatabe, Shûichi. 2004. A Comprehensive Theory of Coordination of Unlikes. In Stefan Müller (ed.), *Proceedings of the hpsg-2004 conference,* Katholieke Universiteit Leuven, 335–355.

Yatabe, Shûichi. 2012. Comparison of the ellipsis-based theory of non-constituent coordination with its alternatives. In Stefan Müller (ed.), *The proceedings of the 19th international conference on head-driven phrase structure grammar*, 454–474. Stanford: CSLI Publications. http://phiz.c.u-tokyo.ac.jp/~yatabe/hpsg2012.pdf.

Yngve, Victor. 1960. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104. 444–466.

Zaenen, Annie & Lauri Karttunen. 1984. Morphological non-distinctness and coordination. In *Escol 83*, 309–320.

Zwart, Jan-Wouter. 2005. Some notes on coordination in head-final languages. In Jenny Doetjes & Jeroen van de Weijer (eds.), *Linguistics in the netherlands 2005*, 232–241. John Benjamins, Amsterdam.

Zwicky, Arnold M. 1985. Heads. *Journal of Linguistics* 21. 1–20.

Chapter 18

Idioms

Manfred Sailer
Goethe-Unversität Frankfurt

This chapter first sketches basic empirical properties of idioms. The state of the art before the emergence of HPSG is presented, followed by a discussion of four types of HPSG-approaches to idioms. A section on future research closes the discussion.

1 Introdction

In this chapter, I will use the term *idiom* interchangeably with the broader terms such as *phraseme*, *phraseologism*, *phraseological unit*, or *multiword expression*. This means, that I will subsume under this notion expressions such as prototypical idioms (*kick the bucket* 'die'), support verb constructions (*take advantage*), formulaic expression (*Good morning!*) and many more.[1] The main focus of the discussion will, however, be on prototypical idioms, as these have been in the center of the theoretical development.

I will sketch some empirical aspects of idioms in Section 2. In Section 3, I will present the theoretical context within which idiom analyses arose in HPSG. An overview of the development within HPSG will be given in Section 4. Desiderata for future research are mentioned in Section 5, before I close with a short conclusion.

2 Empirical domain

In the context of the present handbook, the most useful characterization of idioms might be the definition of *multiword expression* from Baldwin & Kim (2010).

[1] I will provide a paraphrase for all idioms at their first mention. They are also listed in the appendix, together with their paraphrase and a remark on which aspects of the idiom are discussed in the text.

Manfred Sailer. 2019. Idioms. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 517–549. Berlin: Language Science Press. DOI:??

For them, any combination of words counts as a multiword expression if it is syntactically complex and shows some degree of *idiomaticity* (i.e., irregularity), be it lexical, syntactic, semantic, pragmatic, or statistical.[2] I speak of a "combination of words" in the sense of a *substantive* or *lexically filled idiom*, which contrasts with *formal* or *lexically open idioms* (**Fillmoreetal1988**).

Baldwin & Kim's criteria can help us structure the data presentation in this section, expanding them where it seems suitable. My expansions concern the aspect known as *fixedness* in the phraseological tradition as in Fleischer (1997).[3]

For Baldwin & Kim (2010), *lexical idiosyncracy* concerns expressions with words that only occur in an idiom, so-called *phraseologically bound words*, or *cranberry words* (Aronoff 1976). Examples include *make headway* 'make progress', *take umbrage* 'take offence', *in a trice* 'in a moment/very quickly'.[4] For such expressions, the grammar has to make sure that the bound word does not occur outside the idiom, i.e., we need to prevent combinations such as (1b).[5]

(1) a. They fixed the problem in a trice.

b. * It just took them a trice to fix the problem.

We can expand this type of idiosyncrasy to include a second important property of idioms. Most idioms have a fixed inventory of words. In their summary of this aspect of idioms, Gibbs & Colston (2007: 827–828) include the following examples: *kick the bucket* means *die*, but *kick the pail*, *punt the bucket*, or *punt the pail* do not have this meaning. However, some degree of lexical variation seems to be allowed, as the idiom *break the ice* 'relieve tension in a strained situation' can be varied into *shatter the ice*.[6] So, a challenge for idiom theories is to

[2] In the phraseological tradition the aspect of *lexicalization* is added (Fleischer 1997; Burger 1998). This means that an expression is stored in the lexicon. This criterion might have the same coverage as *conventionalization* used in Nunberg et al. (1994). These criteria are addressing the mental representation of idioms as a unit and are, thus, rather psycholinguistic in nature.

[3] Baldwin & Kim (2010) describe idioms in terms of syntactic fixedness, but they seem to consider fixedness a derived notion.

[4] See https://www.english-linguistics.de/codii/, accessed 2019-09-03, for a list of bound words in English and German (Trawiński et al. 2008).

[5] Tom Wasow (p.c.) points out that there are attested uses of many alleged bound words outside their canonical idiom, as in (i). Such uses are, however, rare and restricted.

(i) Not a trice later, the sounds of gunplay were to be heard echoing from Bad Man's Rock. (COCA)

[6] While Gibbs & Colston (2007), following Gibbs et al. (1989), present this example as a lexical variation, Glucksberg (2001: 85), from which it is taken, characterizes it as having a somewhat

guarantee that the right lexical elements are used in the right constellation.

Syntactic idiomaticity is used in Baldwin & Kim (2010) to describe expressions that are not formed according to the productive rules of English syntax, following Fillmore et al. (1988), such as *by and large* 'on the whole/everything considered', *trip the light fantastic* 'dance'.

In my expanded use of this notion, this also subsumes irregularities/restrictions in the syntactic flexibility of an idiom, i.e., the question whether an idiom can occur in the same syntactic constructions as an analogous non-idiomatic combination. In Transformational Grammar, such as Weinreich (1969) and Fraser (1970), lists of different syntactic transformations were compiled and it was observed that some idioms allow for certain transformations but not for others. This method has been pursued systematically in the framework of *Lexicon-Grammar* (Gross 1982).[7] Sag et al. (2002) distinguish three levels of fixedness: *fixed*, *semi-fixed*, and *flexible*. Completely fixed idioms include *of course*, *ad hoc* and are often called *words with spaces*. Semi-fixed idioms allow for morpho-syntactic variation, such as inflection. These include some prototypical idioms (*trip the light fantastic*, *kick the bucket*) and complex proper names. In English, semi-flexible idioms show inflection, but cannot easily be passivized nor do they allow for parts of to be topicalized, see (2).

(2) a. Alex kicked / might kick the bucket.
  b. * The bucket was kicked by Alex.
  c. * The bucket, Alex kicked.

Flexible idioms pattern with free combinations. For them, we do not only find inflection, but also passivization, topicalization or pronominalization of parts etc. Free combinations include some prototypical idioms (*spill the beans* 'reveal a secret', *pull strings* 'exert influence/use one's connections'), but also collocations (*brush one's teeth*) and light verbs (*make a mistake*).

The assumption of two flexibility classes is not uncontroversial: Horn (2003) distinguishes two types among what Sag et al. (2002) consider flexible idioms. Fraser (1970) assumes six flexibility classes, looking at a wide range of syntactic operations. Ruwet (1991) takes issue with the cross-linguistic applicability of the

different aspect of an "abrupt change in the social climate". Clear cases of synonymy under lexical substitution are found with German *wie warme Semmeln/Brötchen/Schrippen weggehen* (lit.: like warm rolls vanish) 'sell like hotcakes' in which some regional terms for rolls can be used in the idiom.

[7] See Laporte (2018) for a recent discussion of applying this method for a classification of idioms.

classification of syntactic operations. Similarly, Schenk (1995) claims that for languages such as Dutch and German, automatic/meaningless syntactic processes other than just inflection are possible for semi-fixed idioms, such as verb-second movement and some types of fronting.

The analytic challenge of syntactic idiomaticity is to capture the difference in flexibility in a non-ad hoc way. It is this aspect of idioms that has received particular attention in Mainstream Generative Grammar,[8, 9] but also in the HPSG-approaches sketched in Section 4.

Semantic idiomaticity may sound pleonastic, as, traditionally, an expression is called idiomatic if it has a conventional meaning that is different from its literal meaning. Since I use the terms idiom and idiomaticity in their broad senses of phraseological unit and irregularity, respectively, the qualification *semantic* idiom(aticity) is needed.

One challenge of the modelling of idioms is to capture the relation between the literal and the idiomatic meaning of an expression. Gibbs & Colston (2007) give an overview over psycholinguistic research on idioms. Whereas it was first assumed that speakers would compute the literal meaning of an expression and, then, derive the idiomatic meaning, evidence has been accumulated that the idiomatic meaning is accessed directly.

Wasow et al. (1984) and Nunberg et al. (1994) explore various semantic relations for idioms, in particular *decomposability* and *transparency*. An idiom is *decomposable* if its idiomatic meaning can be distributed over its component parts in such a way that we would arrive at the idiomatic meaning of the overall expression if we interpreted the syntactic sturcture on the basis of such a meaning assignment. The idiomatic meaning of the expression *pull strings* can be decomposed by interpreting *pull* as *exploit/use* and *strings* as *connections*. The expressions *kick the bucket* and *saw logs* 'snore' are not decomposable.

An idiom is *transparent* if there is a synchronically accessible relation between the literal and the idiomatic meaning of an idiom. For some speakers, *saw logs* is transparent in this sense, as the noise produced by this activity is similar to a snoring noise. For *pull strings*, there is an analogy to a puppeteer controlling the puppets' behavior by pulling strings. A non-transparent idiom is called *opaque*.

Some idioms do not show semantic idiomaticity at all, such as collocations (*brush one's teeth*) or support verb constructions (*take a shower*). Many body-part expressions such as *shake hands* 'greet' or *shake one's head* 'decline/negate'

[8] I follow Culicover & Jackendoff (2005: 3) in using the term *Mainstream Generative Grammar* (MGG) to refer to work in Minimalism and the earlier Government & Binding framework.

[9] See the references in Corver et al. (2019) for a brief up-to-date overview of MGG work.

constitute a more complex case: They describe a conventionalized activity and denote the social meaning of this activity.[10]

In addition, we might need to assume a *figurative* interpretation. For some expressions, in particular proverbs or cases like *take the bull by the horns*) we might get a figurative reading rather than an idiomatic reading. Glucksberg (2001) explicitly distinguishes between idiomatic and figurative interpretations. In his view, the above-mentioned case of *shatter the ice* would be a figurative use of the idiom *break the ice*. While there has been a considerable amount of work on figurativity in psycholinguistics, the integration of its results into formal linguistics is still a desideratum.

Pragmatic idiomaticity covers expressions that have a *pragmatic point* in the terminology of Fillmore et al. (1988). These include complex formulaic expressions (*Good morning!*). There has been little work on this aspect of idiomaticity in formal phraseology.

The final type of idiomaticity is *statistical idiomaticity*. Contrary to the other idiomaticity criteria, this is a usage-based aspect. If we find a high degree of co-occurrence of a particular combination of words that is idiosyncratic for this combination, we can speak of a statistical idiomaticity. This category includes *collocations*. Baldwin & Kim (2010) mention *immaculate performance* as an example. Collocations are important in computational linguistics and in foreign-language learning, but their status for theoretical linguistics and for a competence-oriented framework such as HPSG is unclear.

This discussion of the various types of idiomaticity shows that idioms do not form a homogeneous empirical domain but are rather defined negatively. This leads to the basic analytical challenge of idioms: while the empirical domain is defined by absence of regularity in at least one aspect, idioms largely obey the principles of grammar. In other words, there is a lot of regularity in the domain of idioms, while any approach still needs to be able to model the irregular properties.

3 Predecessors to HPSG analyses of idioms

In this section, I will sketch the theoretical environment within which HPSG and HPSG analyses of idioms have emerged.

The general assumption about idioms in Mainstream Generative Grammar is that they must be represented as a complex phrasal form-meaning unit. Such

[10] The basic reference for the phraseological properties of body-part expressions is Burger (1976).

units are inserted *en bloc* into the structure rather than built by syntactic operations. This view goes back to Chomsky (1965: 190). With this unquestioned assumption, arguments for or against particular analyses can be constructed. To give just one classical example, Chomsky (1981) uses the passivizabilty of some idioms as an argument for the existence of Deep Structure, i.e. a structure on which the idiom is inserted holistically. Ruwet (1991) and Nunberg et al. (1994) go through a number of such lines of argumentation showing their basic problems.

The holistic view on idioms is most plausible for idioms that show many types of idiomaticity at the same time, though it becomes more and more problematic if only one or only a few types of idiomaticity are attested. HPSG is less driven by analytical pre-decisions than other frameworks, see Borsley & Müller (2019), Chapter L of this volume. Nonetheless, idioms have been used to motivate assumptions about the architecture of linguistic signs in HPSG as well.

Wasow et al. (1984) and Nunberg et al. (1994) are probably the two most influential papers in formal phraseology in the last decades. While there are many aspects of Nunberg et al. (1994) that have not been integrated into the formal modelling of idioms, there are at least two insights that have been widely adapted in HPSG. First, not all idioms should be represented holistically. Second, the syntactic flexibility of an idiom is related to its semantic decomposability. In fact, Nunberg et al. (1994) state this last insight even more generally:[11]

> We predict that the syntactic flexibility of a particular idiom will ultimately be explained in terms of the compatibility of its semantics with the semantics and pragmatics of various constructions. (Nunberg et al. 1994: 531)

Wasow et al. (1984) and Nunberg et al. (1994) propose a simplified first approach to a theory that would be in line with this quote. They argue that, for English, there is a correlation between syntactic flexibility and semantic decomposability in that non-decomposable idioms are only semi-flexible, whereas decomposable idioms are flexible, to use our terminology from Section 2. This idea has been directly encoded formally in the idiom theory of Gazdar et al. (1985: Chapter 7), who define the framework of *Generalized Phrase Structure Grammar* (GPSG).

Gazdar et al. (1985) assume that non-decomposable idioms are inserted into sentences *en bloc*, i.e. as fully specified syntactic trees which are assigned the idiomatic meaning holistically. This means that the otherwise strictly context-free grammar of GPSG needs to be expanded by adding a (small) set of larger

[11] Aspects of this approach are already present in Higgins (1974) and Newmeyer (1974).

trees. Since non-decomposable idioms are inserted as units, their parts cannot be accessed for syntactic operations, such as passivization or movement. Consequently, the generalization about semantic non-decomposability and syntactic fixedness of English idioms from Wasow et al. (1984) is implemented directly.

Decomposable idioms are analyzed as free combinations in syntax. The idiomaticity of such expressions is achieved by two assumptions: First, there is lexical ambiguity, i.e. for an idiom like *pull strings*, the verb *pull* has both a literal meaning and an idiomatic meaning. Similarly for *strings*. Second, Gazdar et al. (1985) assume that lexical items are not necessarily translated into total functions but can be partial functions. Whereas the literal meaning of *pull* might be a total function, the idiomatic meaning of the word would be a partial function that is only defined on elements that are in the denotation of the idiomatic meaning of *strings*. This analysis predicts syntactic flexibility for decomposable idioms, just as proposed in Wasow et al. (1984).

Nunberg et al. (1994: 511–514) show that the connection between semantic decomposability and syntactic flexibility is not as straightforward as suggested. They say that, in German and Dutch, "noncompositional idioms are syntactically versatile" (Nunberg et al. 1994: 514). Similar observations have been brought forward for French in Ruwet (1991). Bargmann & Sailer (2018), and Fellbaum (2019) argue that even for English, passive examples are attested for non-decomposable idioms, such as (3).

(3) Live life to the fullest, you never know when the bucket will be kicked. (Fellbaum 2019: 756)

The current state of our knowledge of the relation between syntactic and semantic idiosyncrasy is that the semantic idiomaticity of an idiom does have an effect on its syntactic flexibility, though the relation is less direct than assumed in the literature based on Wasow et al. (1984) and Nunberg et al. (1994).

4 HPSG-analyses of idioms

HPSG does not make a core-periphery distinction, see Müller (2014). Consequently, idioms belong to the empirical domain to be covered by an HPSG grammar. Nonetheless, idioms are not discussed in Pollard & Sag (1994) and their architecture of grammar does not have a direct place for an analysis of idioms.[12]

[12] This section follows the basic structure and argument of Sailer (2012) and Richter & Sailer (2014).

They situate all idiosyncrasy in the lexicon, which consists of lexical entries for basic words. Every word has to satisfy a lexical entry and all principles of grammar, see Davis & Koenig (2019), Chapter 4 of this volume.[13] All properties of a phrase can be inferred from the properties of the lexical items occurring in the phrase and the constraints of grammar.

In their grammar, Pollard & Sag (1994) adhere to the *Strong Locality Hypothesis* (SLH), i.e., all lexical entries describe leaf nodes in a syntactic structure, all phrases are constrained by principles that only refer to local (i.e., *synsem*) properties of the phrase and to local properties of its immediate daughters. This hypothesis is summarized in (4).

(4) Strong Locality Hyphothesis (SLH)
The rules and principles of grammar are statements on a single node of a linguistic structure or on nodes that are immediately dominated by that node.

This precludes any purely phrasal approaches to idioms. Following the heritage of GPSG, we would assume that all regular aspects of linguistic expressions can be handled by mechanisms that follow the SLH, whereas idiomaticity would be a range of phenomena that may violate it. It is, therefore, remarkable that a grammar framework that denies a core-periphery distinction would start with a strong assumption of locality, and, consequently, of regularity.

This is in sharp contrast to the basic motivation of Construction Grammar, which assumes that constructions can be of arbitrary depth and of an arbitrary degree of idiosyncrasy. Fillmore et al. (1988) use idiom data and the various types of idiosyncrasy discussed in Section 2 as an important motivation for this assumption. To contrast this position clearly with the one taken in Pollard & Sag (1994), I will state the *Strong Non-locality Hypothesis* (SNH) in (5).

(5) Strong Non-locality Hypothesis (SNH)
The internal structure of a construction can be arbitrarily deep and show an arbitrary degree of irregularity at any substructure.

The actual formalism used in Pollard & Sag (1994), King (1989) – see Richter

[13] I refer to the lexicon in the technical sense as the collection of lexical entries, i.e. as *descriptions*, rather than as a collection of lexical items, i.e. linguistic signs. Since Pollard & Sag (1994) do not discuss morphological processes, their lexical entries describe full forms. If there is a finite number of such lexical entries, the lexicon can be expressed by a *Word Principle*, a constraint on words that contains a disjunction of all such lexical entries. Once we include morphology, lexical rules, and phrasal lexical entries in the picture, we need to refine this simplified view.

(2019), Chapter 3 of this volume – does not require the strong versions of the locality and the non-locality hypotheses, but is compatible with weaker versions. I will call these the *Weak Locality Hypothesis* (WLH), and the *Weak Non-locality Hypothesis* (WNH), see (6) and (7) respectively.

(6) Weak Locality Hypothesis (WLH)
At most the highest node in a structure is licensed by a rule of grammar or a lexical entry.

According to the WLH, just as in the SLH, each sign needs to be licensed by the lexicon and/or the grammar. This precludes any *en bloc*-insertion analyses, which would be compatible with the SNH. According to the WNH, in line with the SLH, a sign can, however, impose further constraints on its component parts, that may go beyond local (i.e., *synsem*) properties of its immediate daughters.

(7) Weak Non-locality Hypothesis (WNH)
The rules and principles of grammar can constrain – though not license – the internal structure of a linguistic sign at arbitrary depth.

This means that all substructures of a syntactic node need to be licensed by the grammar, but the node may impose idiosyncratic constraints on which particular well-formed substructures it may contain.

In this section, I will review four types of analyses developed within HPSG, in a mildly synchronic order: First, I will discuss a conservative extension of Pollard & Sag (1994) for idioms (Krenn & Erbach 1994) that sticks to the SLH. Then, I will look at attempts to incorporate constructional ideas more directly, i.e., ways to include a version of the SNH. The third type of approach will exploit the WLH. Finally, I will summarize recent approaches, which are, again, emphasizing the locality of idioms.

4.1 Early lexical approaches

Krenn & Erbach (1994), based on Erbach (1992), present the first comprehensive HPSG account of idioms. They look at a wide variety of different types of German idioms, including support verb constructions. They only modify the architecture of Pollard & Sag (1994) marginally and stick to the Strong Locality Hypothesis. They base their analysis on the apparent correlation between syntactic flexibility and semantic decomposability from Wasow et al. (1984) and Nunberg et al. (1994). Their analysis is a representational variant of the analysis in Gazdar et al. (1985).

To maintain the SLH, Krenn & Erbach (1994) assume that the information available in syntactic selection is slightly richer than what has been assumed in Pollard & Sag (1994): First, they use a lexeme-identification feature, LEXEME, which is located inside the INDEX value and whose value is the semantic constant associated with a lexeme. Second, they include a feature THETA-ROLE, whose value indicates which thematic role a sign is assigned in a structure. In addition to standard thematic roles, they include a dummy value *nil*. Third, as the paper was written in the transition phase between Pollard & Sag (1987) and Pollard & Sag (1994), they assume that the selectional attributes contain complete *sign* objects rather than just *synsem* objects. Consequently, they can assume selection for phonological properties and internal constituent structure, which we could consider a violation of the SLH.

The effect of these changes in the analysis of idioms can be seen in (8) and (9). In (8), I sketch the analysis of the syntactically flexible, decomposable idiom, *spill the beans*. There are individual lexical items for the idiomatic words.[14]

(8) a.
$$\begin{bmatrix} \textsc{phon} & \langle \textit{spill} \rangle \\ \textsc{synsem} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{subcat} & \left\langle \text{NP, NP}\begin{bmatrix} \textsc{lexeme} & \textit{beans\_i} \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{rel} & \textit{spill\_i} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

b.
$$\begin{bmatrix} \textsc{phon} & \langle \textit{beans} \rangle \\ \textsc{synsem} & \begin{bmatrix} \textsc{content} & \begin{bmatrix} \textsc{index} & \begin{bmatrix} \textsc{lexeme} & \textit{beans\_i} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The LEXEME values of these words can be used to distinguish them from their ordinary, non-idiomatic, homonyms. Each idiomatic word comes with its idiomatic meaning, which models the decomposability of the expression. The lexical items satisfying the entries in (8) can undergo lexical rules such as passive.

The idiomatic verb *spill* selects an NP complement with the LEXEME value *beans_i*. The lexicon is built in such a way that no other word selects for this LEXEME value. This models the lexical fixedness of the idiom.

The choice of putting the lexical identifier into the INDEX guarantees that it is shared between a lexical head and its phrase, which allows for syntactic flexibility inside the NP. Similarly, the information shared between a trace and its

[14] We do not need to specify the REL value the noun *beans*, as the LISTEME and the REL value are usually identical.

antecedent contains the INDEX value. Consequently, participation in unbounded dependency constructions is equally accounted for. Finally, since a pronoun has the same INDEX value as its antecedent, pronominalization is also possible.

I sketch the analysis of a non-decomposable, fixed idiom, *kick the bucket*, in (9). In this case, there is only a lexical entry of the syntactic head of the idiom, the verb *kick*. It selects the full phonology of its complement. This blocks any syntactic processes inside this NP. It also follows that the complement cannot be realized as a trace, which blocks extraction.[15] The special THETA-ROLE value *nil* will be used to restrict the lexical rules that can be applied. The passive lexical rule, for example, would be specified in such a way that it cannot apply if the NP complement in its input has this theta-role.

(9)
$$\left[\begin{array}{ll} \textsc{phon} & \langle \textit{kick} \rangle \\ \textsc{synsem} & \left[\begin{array}{ll} \textsc{cat} & \left[\begin{array}{ll} \textsc{subcat} & \left\langle \text{NP, NP}\left[\begin{array}{ll} \textsc{phon} & \langle \textit{the, bucket} \rangle \\ \textsc{theta-role} & \textit{nil} \end{array}\right] \right\rangle \end{array}\right] \\ \textsc{cont} & \left[\begin{array}{ll} \textsc{rel} & \textit{die} \end{array}\right] \end{array}\right] \end{array}\right]$$

With this analysis, Krenn & Erbach (1994) capture both the idiosyncratic aspects and the regularity of idioms. They show how it generalizes to a wide range of idiom types. I will briefly mention some problems of the approach, though.

There are two problems for the analysis of non-decomposable idioms. First, the approach is too restrictive with respect to the syntactic flexibility of *kick the bucket*, as it excludes cases such as *kick the social/figurative bucket*, which are discussed in Ernst (1981). Second, it is built on equating the class of non-decomposable idioms with that of semi-fixed idioms. As shown in my discussion around example (3), this cannot be maintained.

There are also some undesired properties of the LEXEME-value selection. The index-identity between a pronoun and its antecedent would require that the subject of the relative clause in (10) has the same INDEX value as the head noun *strings*. However, the account of the lexical fixedness of idioms is built on the assumption that no verb except for the idiomatic *pull* selects for an argument with LEXEME value *strings_i*.[16]

[15] See Borsley & Crysmann (2019), Chapter 14 of this volume for details on the treatment of extraction in HPSG.

[16] Pulman (1993) discusses the analogous problem for the denotational theory of Gazdar et al. (1985).

(10) Parky pulled the strings that got me the job. (McCawley 1981: 137)

Notwithstanding these problems, the analytic ingredients of Krenn & Erbach (1994) constitute the basis of later HPSG analyses. In particular, a mechanism for lexeme-specific selection has been widely assumed in most approaches. The attribute THETA-ROLE can be seen as a simple form of an *inside-out* mechanism, i.e., as a mechanism of encoding information about the larger structure within which a sign appears.

4.2 Phrasal approach

With the advent of constructional analyses within HPSG, starting with Sag (1997), it is natural to expect phrasal accounts of idioms to emerge as well, as idiomaticity is a central empirical domain for Construction Grammar, see Müller (2019), Chapter P of this volume. In this version of HPSG, there is an elaborate type hierarchy below *phrase*. Sag (1997) also introduces *defaults* into HPSG, which play an important role in the treatment of idioms in constructional HPSG. The clearest phrasal approach to idioms can be found in Riehemann (2001), which incorporates insights from earlier publications such as Riehemann (1997) and Riehemann & Bender (1999). The overall framework of Riehemann (2001) is constructional HPSG with *Minimal Recursion Semantics* (Copestake et al. 1995; 2005), see also Koenig & Richter (2019), Chapter C of this volume.

For Riehemann, idioms are phrasal units. Consequently, she assumes a subtype of *phrase* for each idiom, such as *spill-beans-idiomatic-phrase* or *kick-bucket-idiomatic-phrase*. The proposal in Riehemann (2001) is at the same time phrasal and obeys the SLH. To achieve this, Riehemann (2001) assumes an attribute WORDS, whose value contains all words dominated by a phrase. This makes it possible to say that a phrase of type *spill-beans-idiomatic-phrase* dominates the words *spill* and *beans*. This is shown in the relevant type constraint for the idiom *spill the beans* in (11).[17]

[17] The percolation mechanism for the feature WORDS is rather complex. In fact, in Riehemann (2001: Section 5.2.1) the idiom-specific words appear within a C-WORDS value, the other words dominated by the idiomatic phrase in the value of an attribute OTHER-WORDS, which together form the value of WORDS. While all the values of these features are subject to local percolation principles, the fact that entire words are percolated undermines the locality intuition behind the SLH.

(11) Constraint on the type *spill-beans-idiomatic-phrase* from Riehemann (2001: 185):

$$\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{spill-beans-ip}} \\ \textsc{words} & \left\{\begin{array}{l} \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{i\_spill}} \\ \ldots\ \textsc{liszt} & \left\langle\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{i\_spill\_rel}} \\ \textsc{undergoer} & \boxed{1} \end{array}\right]\right\rangle \end{array}\right] \overset{<}{\sqcap} \left[\ldots\ \textsc{liszt}\ \left\langle \textit{\_spill\_rel} \right\rangle\right], \\ \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{i\_beans}} \\ \ldots\ \textsc{liszt} & \left\langle\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{i\_beans\_rel}} \\ \textsc{inst} & \boxed{1} \end{array}\right]\right\rangle \end{array}\right] \overset{<}{\sqcap} \left[\ldots\ \textsc{liszt}\ \left\langle \textit{\_beans\_rel} \right\rangle\right], \ldots \end{array}\right\} \end{array}\right]$$

The WORDS value of the idiomatic phrase contains at least two elements, the idiomatic words of type *i_spill* and *i_beans*. The special symbol “ $\overset{<}{\sqcap}$ ” used in this constraint expresses a default. It says that the idiomatic version of the word *spill* is just like its non-idiomatic homonym, except for the parts specified in the left-hand side of the default. In this case, the type of the words and the type of the semantic predicate contributed by the words are changed. Riehemann (2001) only has to introduce the types for the idiomatic words in the type hierarchy but need not specify type constraints on the individual idiomatic words, as these are constrained by the default statement within the constraints on the idioms containing them.

As in the account of Krenn & Erbach (1994), the syntactic flexibilty of the idiom follows from its free syntactic combination and the fact that all parts of the idiom are assigned an independent semantic contribution. The lexical fixedness is a consequence of the requirement that particular words are dominated by the phrase, namely the idiomatic versions of *spill* and *beans*.

The appeal of the account is particularly clear in its application to non-decomposable, semi-flexible idioms such as *kick the bucket* (Riehemann 2001: 212). For such expressions, the constituting idiomatic words are assumed to have an empty semantics and the semantics of the idiom is contributed as a constructional semantic contribution only by the idiomatic phrase. Since the WORDS list contains entire words, it is also possible to require that the idiomatic word *kick* be in active voice and/or that it takes a complement compatible with the description of the idiomatic word *bucket*. This analysis captures the syntactically regular internal structure of this type of idioms, and is compatible with the occurrence of modifiers such as *proverbial*. At the same time, it prevents passivization and excludes extraction of the complement.

Riehemann's approach clearly captures the intuition of idioms as phrasal units much better than any other approach in HPSG. However, it faces a number of problems. Frist, the integration of the approach with constructional HPSG is done in such a way that the phrasal types for idioms are cross-classified in complex type hierarchies with the various syntactic constructions in which the idiom can appear. This allows Riehemann to account for idiosyncratic differences in the syntactic flexibility of idioms, but the question is whether such an explicit encoding misses generalizations that should follow from independent properties of the components of an idiom and/or of the syntactic construction – in line with the quote from Nunberg et al. (1994) on page 522.

Second, the mechanism of percolating dominated words to each phrase is not compatible with the intuitions of most HPSG researchers. Since no empirical motivation for such a mechanism aside from idioms is provided in Riehemann (2001), this idea has not been pursued in other papers.

Third, the question of how to block the free occurrence of idiomatic words, i.e. the occurrence of an idiomatic word without the rest of the idiom, is not solved in Riehemann (2001). While the idiom requires the presence of particular idiomatic words, the occurrence of these words is not restricted.[18] Note that idiomatic words may sometimes be found without the other elements of the idiom – evidenced by expressions such as in *bucket list* 'list of things to do before one dies'. Such data may be considered as support of Riehemann's approach, however, the extent to which we find such free occurrences of idiomatic words is extremely small.[19]

Before closing this subsection, I would like to point out that Riehemann (2001) and Riehemann & Bender (1999) are the only HPSG papers on idioms that address the question of statistical idiomaticity, based on the variationist study in Bender (2000). In particular, Riehemann (2001: 297–301) proposes phrasal constructions for collocations even if these do not show any lexical, syntactic, semantic, or prag-

[18]Since the problem of free occurrences of idiomatic words is not an issue for parsing, versions of Riehemann's approach have been integrated in practical parsing systems (Villavicencio & Copestake 2002), see Bender & Emerson (2019), Chapter H of this volume. Similarly, the approach to idioms sketched in Flickinger (2015) is part of a system for parsing and machine translation. Idioms in the source language are identified by bits of semantic representation – analogous to the elements in the WORDS set. This approach, however, does not constitute a theoretical modelling of idioms, it does not exclude ill-formed uses of idioms but identifies potential occurrences of an idiom in the output of a parser.

[19]See the discussion around (1) for a parallel situation with bound words.

matic idiosyncrasy but just a statistical co-occurrence preference. She extends this into a larger plea for an *experience-based HPSG*. Bender (2000) discusses the same idea under the notions of *minimal* versus *maximal* grammars, i.e., grammars that are as free of redundancy as possible to capture the grammatical sentences of a language with their correct meaning versus grammars that might be open to a connection with usage-based approaches to language modelling. Bender (2000: 292) sketches a version of HPSG with frequencies/probabilities attached to lexical and phrasal types.[20]

4.3 Mixed lexical and phrasal approaches

While Riehemann (2001) proposes a parallel treatment of decomposable and non-decomposable idioms – and of flexible and semi-flexible idioms – the division between fixed and non-fixed expressions is at the core of another approach, the *two-dimensional theory of idioms*. This approach was first outlined in Sailer (2000) and referred to under this label in Richter & Sailer (2009; 2014). It is intended to combine constructional and collocational approaches to grammar.

The basic intuition behind this approach is that signs have internal and external properties. All properties that are part of the feature structure of a sign are called *internal*. Properties that relate to larger feature structures containing this sign are called its *external* properties. The approach assumes that there is a notion of *regularity* and that anything diverging from it is *idiosyncratic* – or idiomatic, in the terminology of this chapter.

This approach is another attempt to reify the GPSG analysis within HPSG. Sailer (2000) follows the distinction of Nunberg et al. (1994) into non-decomposable and non-flexible idioms on the one hand and decomposable and flexible idioms on the other. The first group is considered internally irregular and receives a constructional analysis in terms of a *phrasal lexical entry*. The second group is considered to consists of independent, smaller lexical units that show an external irregularity in being constrained to co-occur within a larger structure. Idioms of the second group receive a collocational analysis. The two types of irregularity are connected by the *Predictability Hypothesis*, given in (12).

(12) Predictability Hypothesis (Sailer 2000: 366):

[20] A so-far unexplored solution to the problem of free occurrence of idiomatic words within an experience-based version of HPSG could be to assign the type *idiomatic-word* an extremely low probability of occurring. This might have the effect that such a word can only be used if it is explicitly required in a construction. However, note that neither defaults nor probabilities are well-definied part of the formal foundations of theoretical work on HPSG, see Richter (2019), Chapter 3 of this volume.

For every sign whose internal properties are fully predictable, the distributional behavior of this sign is fully predictable as well.

In the most recent version of this approach, Richter & Sailer (2009; 2014), there is a feature COLL defined on all signs. The value of this feature specifies the type of internal irregularity. The authors assume a cross-classification of regularity and irregularity with respect to syntax, semantics, and phonology – ignoring pragmatic and statistical (ir)regularity in their paper. Every basic lexical entry is defined as completely irregular as its properties are not predictable. Fully regular phrases such as *read a book* have a trivial value of COLL. A syntactically internally regular but fixed idiom such as *kick the bucket* is classified as having only semantic irregularity, whereas a syntactically irregular expression such as *trip the light fantastic* is of an irregularity type that is a subsort of syntactic and semantic irregularity, but not of phonological irregularity. Following the terminology of Fillmore et al. (1988), this type is called *extra-grammatical-idiom*. The phrasal lexical entry for *trip the light fantastic* is sketched in (13), adjusted to the feature geometry of Sag (1997).

(13) Phrasal lexical entry for the idioms *trip the light fantastic*:

$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{headed-phrase}} \\
\text{PHON} & \boxed{1} \oplus \langle \textit{the, light, fantastic} \rangle \\
\text{SYNS} & \left[\begin{array}{ll}
\text{HEAD} & \boxed{4} \\
\text{LISTEME} & \textit{trip-the-light-fantastic} \\
\text{VAL} & \boxed{5} \left[\text{SUBJ} \left\langle \boxed{2} \left[\text{...INDEX} \ \boxed{3} \right] \right\rangle \right] \\
\text{CONT} & \left[\begin{array}{l} \textit{trip-light-fant} \\ \text{DANCER} \ \boxed{3} \end{array}\right]
\end{array}\right] \\
\text{HEAD-DTR} & \left[\begin{array}{ll}
\text{PHON} & \boxed{1} \\
\text{SYNS} & \left[\begin{array}{ll}
\text{HEAD} & \boxed{4} \ \textit{verb} \\
\text{LISTEME} & \textit{trip} \\
\text{VAL} & \boxed{5} \left[\begin{array}{ll} \text{SUBJ} & \langle \boxed{2} \rangle \\ \text{COMPS} & \langle \rangle \end{array}\right]
\end{array}\right]
\end{array}\right] \\
\text{COLL} & \textit{extra-grammatical-idiom}
\end{array}\right]$$

In (13), the constituent structure of the phrase is not specified, but the phonology is fixed, with the exception of the head daughter's phonological contribution. This accounts for the syntactic irregularity of the idiom. The semantics of the idiom is not related to the semantic contributions of its components, which accounts for the semantic idiomaticity.

Soehn (2006) applies this theory to German. He solves the problem of the relatively large degree of flexibility of non-decomposable idioms in German by using underspecified descriptions of the constituent structure dominated by the idiomatic phrase.

For decomposable idioms, the two-dimensional theory assumes a collocational component. This component is integrated into the value of an attribute REQ, which is only defined on *coll* objects of one of the irregularity types. This encodes the Predictability Hypothesis. The most comprehensive version of this collocational theory is given in Soehn (2009), summarizing and extending ideas from Soehn (2006) and Richter & Soehn (2006). Soehn assumes that collocational requirements can be of various types: a lexical item can be constrained to co-occur with particular *licensers* (or collocates). These can be other lexemes, semantic operators, or phonological units. In addition, the domain within which this licensing has to be satisfied is specified in terms of syntactic barriers, i.e., syntactic nodes dominating the externally irregular item.

To give an example, the idiom *spill the beans* would be analyzed as consisting of two idiomatic words *spill* and *beans* with special LISTEME values *spill-i* and *beans-i*. The idiomatic verb *spill* imposes a lexeme-selection on its complement. The idiomatic noun *beans* has a non-empty REQ value, which specifies that it must be selected by a word with LISTEME value *spill-i* within the smallest complete clause dominating it.

The two-dimensional approach suffers from a number of weaknesses. First, it presupposes a notion of regularity. This assumption is not shared by all linguists. Second, the criteria whether an expression should be treated constructionally or collocationally are not always clear. Idioms with irregular syntactic structure need to be analyzed constructionally, but this is less clear for non-decomposable idioms with regular syntactic structure such as *kick the bucket*.

4.4 Recent lexical approaches

Kay et al. (2015) marks an important re-orientation in the analysis of idioms: the lexical analysis is extended to all syntactically regular idioms, i.e., to both decomposable (*spill the beans*) and non-decomposable idioms (*kick the bucket*).[21]

[21]This idea has been previously expressed within a Minimalist perspective in Everaert (2010).

Kay et al. (2015) achieve a lexical analysis of non-decomposable idioms by two means: (i) an extension of the HPSG selection mechanism, (ii) the assumption of semantically empty idiomatic words.

As in previous accounts, the relation among idiom parts is established through lexeme-specific selection, using a feature LID (for: *lexical identifier*). The authors assume that there is a difference between idiomatic and non-idiomatic LID values. Only heads that are part of idioms themselves can select for idiomatic words.

For the idiom *kick the bucket*, Kay et al. (2015) assume that all meaning is carried by the lexical head, an idiomatic version of *kick*, whereas the other two words, *the* and *bucket* are meaningless. This meaninglessness allows Kay et al. to block the idiom from occurring in constructions which require meaningful constituents, such as questions, *it*-clefts, middle and others. To exclude passivization, the authors assume that English passive cannot apply to verbs selecting a semantically empty direct object.

The approach in Kay et al. (2015) is a recent attempt to maintain the SLH as much as possible. Since the SLH has been a major conceptual motivation for Sign-based Construction Grammar, Kay et al.'s paper is an important contribution to show the empirical robustness of this assumption.

Bargmann & Sailer (2018) propose a similar lexical approach to non-decomposable idioms. They take as their starting point the syntactic flexibility of semantically non-decomposable idioms in English and, in particular, in German. There are two main differences between Kay et al.'s paper and Bargmann & Sailer's: (i), Bargmann & Sailer assume a collocational rather than a purely selectional mechanism to capture lexeme restrictions of idioms, and (ii), they propose a redundant semantics rather than an empty semantics for idiom parts in non-decomposable idioms. In other words, Bargmann & Sailer (2018) propose that both *kick* and *bucket* contribute the semantics of the idiom *kick the bucket*. Bargmann & Sailer argue that the semantic contiributions of parts of non-decomposable, syntactically regular idioms are the same across languages, whereas the differences in syntactic flexibility are related to the different syntactic, semantic, and pragmatic constraints imposed on various constructions. To give just one example, whereas there are barely any restrictions on passive subjects in German, there are strong discourse-structural constraints on passive subjects in English.

Both Kay et al. (2015) and Bargmann & Sailer (2018) attempt to derive the (partial) syntactic inflexibility of non-decomposable idioms from independent properties of the relevant constructions. As such, they subscribe to the programmatic statement of Nunberg et al. (1994) quoted on page 522. In this respect, the extension of the lexical approach from decomposable idioms to all syntactically regular expressions has been a clear step forward.

Findlay (2017) provides a recent discussion and criticism of lexical approaches to idioms in general, which applies in particular to non-decomposable expressions. His reservations comprise the following points. First, there is a massive proliferation of lexical entries for otherwise homophonous words. Second, the lexical analysis does not represent idioms as units, which might make it difficult to connect their theoretical treatment with processing evidence. Findlay refers to psycholinguistic studies, such as Swinney & Cutler (1979), that point to a faster processing of idioms than of free combinations. While the relevance of processing arguments for an HPSG analysis are not clear, I share the basic intuition that idioms, decomposable or not, are a unit and that this should be part of their linguistic representation.

5 Where to go from here?

The final section of this article contains short overviews over research that has been done in areas of phraseology that are outside the main thread of this chapter. I will also identify desiderata.

5.1 Neglected phenomena

Not all types of idioms and not all types of idiomaticity mentioned in Section 2 have received an adequate treatment in the (HPSG) literature. I will briefly look at three empirical areas that deserve more attention: neglected types of idiom variation, phraseological patterns, and the literal and non-literal meaning components of idioms.

Most studies on idiom variation have looked at verb- and sentence-related syntactic constructions, such as passive and topicalization. However, not much attention has been paid to lexical variation in idioms. This variation is illustrated by the following examples from Richards (2001: 184, 191).

(14) a. The Count gives everyone the creeps.
  b. You get the creeps (just looking at him).
  c. I have the creeps.

In (14), the alternation of the verb seems to be very systematic – and has been used by Richards (2001) to motivate a lexical decomposition of these verbs. A similar argument has been made in Mateu & Espinal (2007) for similar idioms in Catalan. We are lacking systematic, larger empirical studies of this type of substitution, and it would be important to see how it can be modeled in HPSG. One option would be to capture the *give*–*get*–*have* alternation(s) with lexical rules. Such lexical rules would be different from the standard cases, however, as they would change the lexeme itself rather than just alternating its morphosyntactic properties or its semantic contribution.

In the case mentioned in footnote 6, the alternation consists in substituting a word with a (near) synonym and keeping the meaning of the idiom intact. Again, HPSG seems to have all the required tools to model this phemonenon – for example, by means of hierarchies of lexical-id values. However, the extent of this phenomenon across the set of idioms is not known empirically.

In the domain of syntactic variation, the nominal domain has not received the attention it might deserve yet. There is a well-known variation with respect to the marking of possession within idioms. This has been documented for English in Ho (2015), for Modern Hebrew in Almog (2012), for Modern Greek and German in Markantonatou & Sailer (2016). In German, we find a relatively free alternation between a plain definite and a possessive, see (15a). This is, however, not possible with all idioms, (15b).

(15) a. Alex hat den / seinen Verstand verloren.
Alex has the his mind lost
'Alex lost his mind.'

b. Alex hat *den / ihren Frieden mit der Situation gemacht.
Alex has the her peace with the situation made
'Alex made her peace with the situation.'

We can also find a free dative in some cases, expressing the possessor. In (16a), a dative possessor may co-occur with a plain definite or a coreferential possessive determiner, in (16b) only the definite article but not the possessive determiner is possible.

(16) a. Alex hat mir das / mein Herz gebrochen.
Alex has me.DAT the my heart broken
'Alex broke my heart.'

b. Alex sollte mir lieber aus den / *meinen Augen gehen.
Alex should me.DAT rather out of the my eyes go
'Alex should rather disappear from my sight.'

While they do not offer a formal encoding, Markantonatou & Sailer (2016) observe that a particular encoding of possession in idioms is only possible if it would also be possible in a free combination. However, an idiom may be idiosyncratically restricted to a subset of the realizations that would be possible in a corresponding free combination. A formalization in HPSG might consist of a treatment of possessively used definite determiners, combined with an analysis of free datives as an extension of a verb's argument structure.

Related to the question of lexical variation are *phraseological patterns*, i.e., very schematic idioms in which the lexical material is largely free. Some examples of phraseological patterns are the *Incredulity Response Construction* as in *What, me worry?* (Akmajian 1984; Lambrecht 1990), or the *What's X doing Y?*-construction (Kay & Fillmore 1999). Such patterns are of theoretical importance as they typically involve a non-canonical syntactic pattern. The different locality and non-locality hypotheses introduced above make different predictions. Fillmore et al. (1988) have presented such constructions as a motivation for the non-locality of constructions, i.e., as support of a SNH. However, Kay & Fillmore (1999) show that a lexical analysis might be possible for some cases at least. Kay & Fillmore provide a detailed lexical analysis of the *What's X doing Y?*-construction.

Borsley (2004) looks at another phraseological pattern, the *the X-er the Y-er*-construction, or *comparative correlative construction*. Borsley analyzes this construction by means of two special (local) phrase structure types: one for the comparative *the*-clauses, and one for the overall construction. He shows that (i), the idiosyncrasy of the construction concerns two levels of embedding and is, therefore, non-local, however, (ii), a local analysis is still possible. This approach raises the question as to whether the WNH is empirically vacuous since we can always encode a non-local construction in terms of a series of idiosyncratic local construction. Clearly, work on more phraseological patterns is needed to assess the various analytical options and their consequences for the architecture of grammar.

A major charge for the conceptual and semantic analysis of idioms is the interaction between the literal and the idiomatic meaning. I presented the basic empirical facts in Section 2. All HPSG approaches to idioms so far basically ignore the literal meaning. This position might be justified, as an HPSG grammar should just model the structure and meaning of an utterance and need not worry about the meta-linguistic relations among different lexical items or among different readings of the same (or a homophonous) expression. Nonetheless, this issue addresses an important conceptual point. Addressing it might immediately provide possibilities to connect HPSG research to other disciplines and/or frameworks, such as cognitive linguistics, such as Dobrovol'skij & Piirainen (2005), and psycholinguistics.

5.2 Challenges from other languages

The majority of work on idioms in HPSG has been done on English and German. As discussed in Section 4.4, the recent trend in HPSG idiom research necessitates a detailed study of individual syntactic structures. Consequently, the restriction on two closely related languages limits the possible phenomena that can be studied on idioms. It would be essential to expand the empirical coverage of idiom analyses in HPSG to as many different languages as possible. The larger degree of syntactic flexibility of French, German, and Dutch idioms (Ruwet 1991; Nunberg et al. 1994; Schenk 1995) has led to important refinements of the analysis in Nunberg et al. (1994) and, ultimately, to the lexical analyses of all syntactically regular idioms.

Similarly, the above-mentioned data on possessive alternations only become prominent when languages beyond English are taken into account. Modern Greek, German, and many others show the type of external possessor classified as a European areal phenomenon in Haspelmath (1999). It would be important to look at idioms in languages with other types of external possessors.

In a recent paper, Sheinfux et al. (2019) provide data from Modern Hebrew that show that opacity and figurativity of an idiom are decisive for its syntactic flexibility rather than decomposability. This result stresses the importance of the literal reading for an adequate account of the syntactic behavior of idioms. It shows that the inclusion of other languages can cause a shift of focus to other types of idioms or other types of idiomaticity.

To add just one more example, HPSG(-related) work on Persian such as Müller (2010) and Samvelian & Faghiri (2016) establishes a clear connection between complex predicates and idioms. Their insights might also lead to a reconsideration of the similarities between light verbs and idioms, as already set out in Krenn & Erbach (1994).

As far as I can see, the following empirical phenomena have not been addressed in HPSG-approaches to idioms as they do not occur in the main object languages for which we have idiom analyses, i.e. English and German. They are, however, common in other languages: the occurrence of clitics in idioms (found in Romance and Greek); aspectual alternations in verbs (Slavic and Greek); argument alternations other than passive and dative alternation (such as anti-passive, causative, inchoative etc (in part found in Hebrew and addressed in Sheinfux et al. (2019)); displacement of idiom parts into special syntactic positions (focus position in Hungarian).

Finally, so far, idioms have usually been considered as either offering irregular structures or as being more restricted in their structures than free combinations.

In some languages, however, we find archaic syntactic structures and function words in idioms that do not easily fit these two analytic options. To name just a few, Lødrup (2009) argues that Norwegian used to have an external possessor construction similar to that of other European languages, which is only conserved in some idioms. Similarly, Dutch has a number of archaic case inflections in multiword expressions (Kuiper 2018: 129), and there are archaic forms in Modern Greek multiword expressions. It is far from clear what the best way would be to integrate such cases into an HPSG grammar.

6 Conclusion

Idioms are among the topics in linguistics for which HPSG-related publications have had a clear impact on the field and have been widely quoted across frameworks. This handbook article aimed at providing an overview over the development of idiom analyses in HPSG. There seems to be a clear development towards ever more lexical analyses, starting from the holistic approach for all idioms in Chomsky's work, to a lexical account for all syntactically regular expressions. Notwithstanding the advantages of the lexical analyses, I consider it a basic problem of such approaches that the unit-status of idioms is lost. Consequently, I think that the right balance between phrasal and lexical aspects in the analysis of idioms has not been fully achieved yet. .

The sign-based character of HPSG seems to be particularly suited for a theory of idioms as it allows to take into consideration syntactic, semantic, and pragmatic aspects and to use them to constrain the occurrence of idioms appropriately.

Appendix: List of used idioms

Some idioms do not show semantic idiomaticity at all, such as collocations (*brush one's teeth*) or support verb constructions (*take a shower*). Many body-part expressions such as *shake hands* 'greet' or *shake one's head* 'decline/negate' constitute a more complex case. They describe a conventionalized activity and denote the social meaning of this activity (Burger 1976).

English

idiom	paraphrase	comment
break the ice	relieve tension in a strained situation	non-decomposable
brush one's teeth	clean one's teeth with a tooth brush	collocation, no idiomaticity
give so the creeps	make so feel uncomfortable	systematic lexical variation
Good morning!	(morning greeting)	formulaic expression
immaculate performance	perfect performance	statistical idiomaticity
in a trice	in a moment	bound word: *trice*
kick the bucket	die	non-decomposable
make headway	make progress	bound word: *headway*
pull strings	exert influence/use one's connections	flexible
saw logs	snore	transparent, non-decomposable, semi-flexible
shake hands	greet	body-part expression
shake one's head	decline/negate	body-part expression, possessive idiom
take a shower	clean oneself using a shower	collocation, light verb construction
take umbrage	take offence	bound word: *umbrage*
shake hands	greet	body-part expression
shake one's head	decline/negate	body-part expression
shit hit the fan	there is trouble	subject as idiom component, transparent/figurative, non-decomposable
shoot the breeze	chat	non-decomposable
spill the beans	reveal a secret	flexible
take a shower	to shower	collocation
take the bull by the horns	approach a problem directly	figurative expression
trip the light fantactic	dance	syntactically irregular

German

idiom	gloss	translation	comment
den/seinen Verstand verlieren	the/one’s mind lose	lose one’s mind	alternation of possessor marking
jm das Herz brechen	so the heart break	break s.o.’s heart	dative possessor and possessor alternation
jm aus den Augen gehen	so out of the eyes go	disappear from s.o.’s sight	dative possessor, restricted possessor alternation
seinen Frieden machen mit	one’s peace make with	make one’s peace with	no possessor alternation possible
wie warme Semmeln/ Brötchen/Schrippen weggehen	like warm rolls vanish	sell like hotcakes	parts can be exchanged by synonyms

Abbreviations

GPSG	Generalized Phrase Structure Grammar (Gazdar et al. 1985)
MWE	multiword expression
SLH	Strong Locality Hypothesis, see page 524
SNH	Strong Non-locality Hypothesis, see page 524
WLH	Weak Locality Hypothesis, see page 525
WNH	Weak Non-locality Hypothesis, see page 525

Acknowledgements

I have perceived Ivan A. Sag and his work with various colleagues as a major inspiration for a lot of my own work on idioms and multiword expressions. This is clearly reflected in the structure of this paper, too. I apologize for this bias, but I think it is legitimate within an HPSG handbook. I am grateful to Stefan Müller and Tom Wasow for comments on the outline and the first version of this chapter. I would not have been able to format this chapter without the support of Sebastian Nordhoff.

References

Akmajian, Adrian. 1984. Sentence types and the form-function fit. *NLLT* 2(1). 1–23.

Almog, Lior. 2012. *The formation of idioms: Evidence from possessive datives in Hebrew*. http://humanities1.tau.ac.il/linguistics_eng/images/stories/Lior_Ordentlich_MA_2012.pdf, accessed 2018-10-13. MA thesis, Tel Aviv University.

Aronoff, Mark. 1976. *Word formation in Generative Grammar*. Cambridge, MA: MIT Press.

Baldwin, Timothy & Su Nam Kim. 2010. Multiword expressions. In Nitin Indurkhya & Fred J. Damerau (eds.), *Handbook of natural language processing*, 2nd edn., 267–292. Boca Raton: CRC Press.

Bargmann, Sascha & Manfred Sailer. 2018. The syntactic flexibility of semantically non-decomposable idioms. In Manfred Sailer & Stella Markantonatou (eds.), *Multiword expressions: Insights from a multi-lingual perspective* (Phraseology and Multiword Expressions 1), 1–29. Berlin: Language Science Press. DOI:10.5281/zenodo.1182587

Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford University dissertation. http://faculty.washington.edu/ebender/dissertation/, accessed 2018-2-25.

Bender, Emily M. & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 797–840. Berlin: Language Science Press. DOI:??

Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 2004. An approach to English comparative correlatives. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 70–92. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

Borsley, Robert D. & Stefan Müller. 2019. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 937–1012. Berlin: Language Science Press. DOI:??

Burger, Harald. 1976. *Die Achseln zucken*: Zur sprachlichen Kodierung nichtsprachlicher Kommunikation. *Wirkendes Wort* 26. 311–339.

Burger, Harald. 1998. *Phraseologie: Eine Einführung am Beispiel des Deutschen*. Berlin: Erich Schmidt Verlag.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Copestake, Ann, Daniel P. Flickinger, Robert Malouf, Susanne Riehemann & Ivan A.Sag. 1995. Translation using Minimal Recursion Semantics. In *Proceedings of the sixth International Conference on Theoretical and Methodological Issues in Machine Translation (TMI95), Leuven, Belgium.*

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Corver, Norbert, Jeroen van Craenenbroeck, William Harwood, Marko Hladnik, Sterre Leufkens & Tanja Temmermann. 2019. Introduction: The compositionality and syntactic flexibilty of verbal idioms. *Journal of Linguistics* 57(4). 725–733. DOI:https://doi.org/10.1515/ling-2019-0021

Culicover, Peter W. & Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Davis, Anthony & Jean-Pierre Koenig. 2019. The nature and role of the lexicon in HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 47–92. Berlin: Language Science Press. DOI:??

Dobrovol'skij, Dmitrij & Elisabeth Piirainen. 2005. *Figurative language: Cross-cultural and cross-linguistic perspectives* (Current Research in the Semantics/Pragmatics Interface 13). Amsterdam: Elsevier.

Erbach, Gregor. 1992. Head-driven lexical representation of idioms in HPSG. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Proceedings of the International Conference on Idioms*. Tilburg, The Netherlands: ITK.

Ernst, Thomas. 1981. Grist for the linguistic mill: Idioms and 'extra' adjectives. *Journal of Linguistic Research* 1(3). 51–68.

Everaert, Martin. 2010. The lexical encoding of idioms. In Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds.), *Lexical semantics, syntax, and event structure* (Oxford Studies in Theoretical Linguistics 27), 76–97. Oxford: Oxford University Press.

Fellbaum, Christiane. 2019. How flexible are idioms? A corpus-based study. 57(4). 735–767. DOI:https://doi.org/10.1515/ling-2019-0015

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538.

Findlay, Jamie Y. 2017. Multiword expressions and lexicalism. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'17 Conference*, 209–229. Stanford, CA: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2017/lfg2017-findlay.pdf, accessed 2018-6-22.

Fleischer, Wolfgang. 1997. *Phraseologie der deutschen Gegenwartssprache*. 2nd edn. Tübingen: Max Niemeyer Verlag.

Flickinger, Dan. 2015. *Multi-word expressions in HPSG*. https://ufal.mff.cuni.cz/events/parseme-1st-training-school, accessed 2018-8-14. Slides to a course at the 1st Parseme Training School, Prague, January 2015.

Fraser, Bruce. 1970. Idioms within a transformational grammar. *Foundations of Language* 6(1). 22–42.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Gibbs, Raymond W. Jr. & Herbert L. Colston. 2007. Psycholinguistic aspects of phraseology: The American tradition. In Harald Burger, Dmitrij Dobrovol'skij, Peter Kühn & Neal R. Norrick (eds.), *Phraseologie/phraseology*, vol. 2 (Ein internationales Handbuch der zeitgenössischen Forschung/An International Handbook of Contemporary Research), 819–836. Berlin: De Gruyter.

Gibbs, Raymond W., Nandini P. Nayak, John L. Bolton & Melissa E. Keppel. 1989. Speakers' assumptions about the lexical flexibility of idioms. *Memory and Cognition* 17(1). 58–68.

Glucksberg, Sam. 2001. *Understanding figurative language: From metaphors to idioms* (Oxford Psychology Series 36). Oxford: Oxford University Press.

Gross, Maurice. 1982. Une classification des phrases “figées” du français. *Revue Québecoise de Linguistique* 11(2). 151–185. DOI:10.7202/602492ar

Haspelmath, Martin. 1999. External possession in a European areal perspective. In Doris L. Payne & Immanuel Barshi (eds.), *External possession*, 109–135. Amsterdam: Benjamins.

Higgins, Francis Roger. 1974. On the use of idioms as evidence for movement. A cautionary note. Amherst, MA. Unpublished manuscript of a talk given at LSA 1974, New York.

Ho, Jia Qian. 2015. *Losing one's mind over meaning: analysing the behaviour of English possessive idioms*. Tech. rep. Linguistics & Multilingual Studies, Nanyang Technological University, Singapore. http://compling.hss.ntu.edu.sg/pdf/2015-fyp-ho-jia-qian.pdf, accessed 2019-9-3.

Horn, George M. 2003. Idioms, metaphors and syntactic mobility. *Journal of Linguistics* 39(2). 245–273.

Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The What's X Doing Y? Construction. *Language* 75(1). 1–33.

Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. *A lexical theory of phrasal idioms*. Ms. CSLI Stanford.

King, Paul. 1989. *A logical formalism for Head-Driven Phrase Structure Grammar*. University of Manchester dissertation.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Krenn, Brigitte & Gregor Erbach. 1994. Idioms and support verb constructions. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 365–396. Stanford, CA: CSLI Publications.

Kuiper, Koenraad. 2018. Multiword expressions and the Law of Exception. In Manfred Sailer & Stella Markantonatou (eds.), *Multiword expressions: Insights from a multi-lingual perspective* (Phraseology and Multiword Expressions 1), 121–141. Berlin: Language Science Press. DOI:10.5281/zenodo.1182595

Lambrecht, Knud. 1990. ‘‘What, me worry?– ‘Mad Magazine sentences’ revisited. In *Proceedings of the 16th annual meeting of the Berkeley Linguistics Society*, 215–228. Berkeley, CA: Berkley Linguistics Society.

Laporte, Éric. 2018. Choosing features for classifying multiword expressions. In Manfred Sailer & Stella Markantonatou (eds.), *Multiword expressions: Insights from a multi-lingual perspective* (Phraseology and Multiword Expressions 1), 143–186. Berlin: Language Science Press. DOI:DOI:10.5281/zenodo.1182597

Lødrup, Helge. 2009. Looking possessor raising in the mouth: Norwegian possessor raising with unergatives. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG09 conference*, 420–440. Stanford, CA: CSLI Publications. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/14/papers/lfg09lodrup.pdf, accessed 2018-10-15.

Markantonatou, Stella & Manfred Sailer. 2016. *Affectees in MWEs: German and Modern Greek*. https://typo.uni-konstanz.de/parseme/images/Meeting/2016-04-07-Struga-meeting/WG1-MARKANTONATOU-SAILER-poster-1.pdf, accessed 2018-10-14. Poster presented at PARSEME's 6th general meeting in Struga.

Mateu, Jaume & Teresa Espinal. 2007. Argument structure and compositionality in idiomatic constructions. *The Linguistic Review* 24(1). 33–59. DOI:https://doi-org.proxy.ub.uni-frankfurt.de/10.1515/TLR.2007.002

McCawley, James D. 1981. The syntax and semantics of English relative clauses. *Lingua* 53. 99–149.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655.

Müller, Stefan. 2014. Kernigkeit: Anmerkungen zur Kern-Peripherie-Unterscheidung. In Antonio Machicao y Priemer, Andreas Nolda & Athina Sioupi (eds.), *Zwischen Kern und Peripherie* (studia grammatica 76), 25–39. Berlin: de Gruyter.

Müller, Stefan. 2019. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

Newmeyer, Frederick J. 1974. The regularity of idiom behavior. *Lingua* 34(4). 327–342.

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Pulman, Stephen G. 1993. The recognition and interpretation of idioms. In Cristina Cacciari & Patrizia Tabossi (eds.), *Idioms: Processing, structure, and interpretation*, 249–270. Hillsdale, NJ: Lawrence Erlbaum Associates.

Richards, Norvin. 2001. An idiomatic argument for lexical decomposition. *Linguistic Inquiry* 32(1). 183–192.

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 45–46. Berlin: Language Science Press. DOI:??

Richter, Frank & Manfred Sailer. 2009. Phraseological clauses in constructional HPSG. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, Göttingen 2009*, 297–317. Stanford, CA: CSLI Publications. cslipublications.stanford.edu/HPSG/2009/richter-sailer.pdf.

Richter, Frank & Manfred Sailer. 2014. Idiome mit phraseologisierten Teilsätzen: Eine Fallstudie zur Formalisierung von Konstruktionen im Rahmen der HPSG. In Alexander Lasch & Alexander Ziem (eds.), *Grammatik als Netzwerk von Konstruktionen*, 291–312. Berlin: De Gruyter.

Richter, Frank & Jan-Philipp Soehn. 2006. *Braucht niemanden zu scheren*: a survey of NPI licensing in German. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna*, 421–440. http://cslipublications.stanford.edu/HPSG/2006/richter-soehn.pdf.

Riehemann, Susanne. 1997. Idiomatic constructions in HPSG. http://doors.stanford.edu/~sr/idioms.ps, accessed 2018-6-28. Paper presented at the 1997 HPSG conference, Ithaca.

Riehemann, Susanne Z. 2001. *A Constructional approach to idioms and word formation*. Stanford University Dissertation. http://mally.stanford.edu/~sr/sr-diss.pdf, accessed 2018-3-20.

Riehemann, Susanne Z. & Emily Bender. 1999. Absolute constructions: On the distribution of predicative idioms. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *WCCFL 18 Proceedings*, 476–489. Somerville, MA: Cascadilla Press. http://doors.stanford.edu/~sr/wccfl.ps, accessed 2018-2-25.

Ruwet, Nicolas. 1991. On the use and abuse of idioms in syntactic argumentation. In *Syntax and human experience*, 171–251. Chicago: University of Chicago Press. Edited and translated by John Goldsmith.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake & Dan Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. In *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLING 2002)*, 1–15. Mexico City, Mexico. http://lingo.stanford.edu/pubs/WP-2001-03.pdf, accessed 2018-3-20.

Sailer, Manfred. 2000. *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar*. Eberhard-Karls-Universität Tübingen Dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46191, accessed 2018-2-25.

Sailer, Manfred. 2012. Phraseologie und Syntax (Head-Driven Phrase Structure Grammar). In Michael Prinz & Ulrike Richter-Vapaatalo (eds.), *Idiome, Konstruktionen, „verblümte rede". Beiträge zur Geschichte der germanistischen Phraseologieforschung*, vol. 3 (Beiträge zur Geschichte der Germanistik), 241–262. Stuttgart: S. Hirzel Verlag.

Samvelian, Pollet & Pegah Faghiri. 2016. Re-thinking compositionality in Persian complex predicates. In *Proceedings of Berkeley Linguistic Society 39th annual meeting, February 16-17 2013*, 212–227. Berkeley: Berkeley Linguistics Society.

Schenk, André. 1995. The syntactic behavior of idioms. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: structural and psychological perspectives*, 253–271. Hillsdale NJ: Erlbaum Assoc.

Sheinfux, Livnat Herzig, Tali Arad Greshler, Nurit Melnik & Shuly Wintner. 2019. Verbal MWEs: Idiomaticity and flexibility. In Yannick Parmentier & Jakub Waszczuk (eds.), *Representation and parsing of multiword expressions: Current trends* (Phraseology and Multiword Expressions 2), 35–68. Berlin: Language Science Press.

Soehn, Jan-Philipp. 2006. *Über Bärendienste und erstaunte Bauklötze – Idiome ohne freie Lesart in der HPSG* (Deutsche Sprache und Literatur 1930). Frankfurt/Main: Peter Lang.

Soehn, Jan-Philipp. 2009. *Lexical licensing in formal grammar*. Universität Tübingen. http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-42035, accessed 2018-8-15. Online publication. Universität Tübingen.

Swinney, David A. & Anne Cutler. 1979. The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior* 18(3). 523–534.

Trawiński, Beata, Jan-Philipp Soehn, Manfred Sailer, Frank Richter & Lothar Lemnitzer. 2008. Cranberry expressions in English and in German. In Nicole Grégoire, Stefan Evert & Brigitte Krenn (eds.), *Proceedings of the LREC workshop towards a shared task for multiword expressions (MWE 2008)*, 35–38. Marrakech, Morokko.

Villavicencio, Aline & Ann Copestake. 2002. *On the nature of idioms*. Tech. rep. 2002-04. Stanford, CA. http://lingo.stanford.edu/working-papers.html.

Wasow, Thomas, Geoffrey Nunberg & Ivan A. Sag. 1984. Idioms: An interim report. In Hattori Shiro & Inoue Kazuko (eds.), *Proceedings of the XIIIth International Congress of Linguistics*. The Hague: CIPL.

Weinreich, Uriel. 1969. *Problems in the analysis of idioms*. Reprinted in Weinreich (1980), 208–264.

Weinreich, Uriel. 1980. *On semantics*. William Labov & Beatrice S. Weinreich (eds.). Philadelphia: University of Pennsylvania Press.

Chapter 19

Negation

Jong-Bok Kim

Kyung Hee University, Seoul

Each language has a way to express (sentential) negation that reverses the truth value of a certain sentence, but employs language particular expressions as well as grammatical strategies. There are four main types of negative in expressing sentential negation: adverbial negative, morphological negative, auxiliary negative verb, and preverbal negative. This chapter discusses HPSG analyses for these four strategies in marking sentential negation.

1 Modes of expressing negation

There are four different types of negative markers in expressing negation in languages: morphological negative, auxiliary negative verb, adverbial negative, and clitic-like preverbal negative (Dahl 1979; Payne 1985; Dryer 2005: see). Each of these types is illustrated in the following:[1]

(1) a. Ali elmalar-i ser-me-di-∅. (Turkish)
Ali apples-ACC like-NEG-PST-3SG
'Ali didn't like apples.'

b. sensayng-nim-i o-ci anh-usi-ess-ta. (Korean)
teacher-HON-NOM come-CONN NEG-HON-PST-DECL
'The teacher didn't come.'

[1]The abbreviations used here include ACC (accusative), CAUS (causative), COND (conditional), CONN (connective), COP (copula), DECL (declarative), DEL (delimiter), FUT (future), GEN (genitive), HON (honorific), INF (infinitival), NEG (negative), NOM (nominative), NPST (non-past), PASS (passive), PERF (perfective), PRES (present), PST (past), PROG (progressive), SG (singular), and so forth.

Jong-Bok Kim. 2019. Negation. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 551–583. Berlin: Language Science Press. DOI:??

c. Dominique (n')écrivait pas de lettre. (French)
Dominique wrote NEG of letter
'Dominiquedid not write a letter.'

d. Gianni non legge articoli di sintassi. (Italian)
Gianni NEG reads articles of syntax
'Gianni doesn't read syntax articles.'

As shown in (1a), languages like Turkish have typical examples of morphological negatives where negation is expressed by an inflectional category realized on the verb by affixation. Meantime, languages like Korean employ a negative auxiliary verb as in (1b).[2] The negative auxiliary verb here is marked with the basic verbal categories such as agreement, tense, aspect, and mood, while the lexical, main verb remains in an invariant, participle form. The third major way of expressing negation is to use an adverbial negative. This type of negation, forming an independent word, is prevalent in English and French, as given in (1c). In these languages, negatives behave like adverbs in their ordering with respect to the verb. The fourth type is to introduce a preverbal negative in expressing sentential negation. The negative marker in Italian in (1d), preceding a finite verb like other types of clitics in the language, belongs to this type.

In analyzing these four main types of sentential negation, there have been two main strands: derivational vs. nonderivational views. The derivational view has claimed that the positioning of all of the four types of negatives is basically determined by the interaction of movement operations, a rather large set of functional projections including NegP, and their hierarchically fixed organization.[3] In particular, to account for the fact that unlike English, only French allows main or lexical verb inversion as in (1c), Pollock (1989; 1994) and a number of subsequent researchers have interpreted these contrasts as providing critical motivation for the process of head movement and the existence of functional categories such as MoodP, TP, AgrP, and NegP (Belletti 1990; Zanuttini 1991; 1997; 2001; Chomsky 1991; 1993; 1995; Lasnik 1995; Haegeman 1995; 1997; Vikner 1997; Zeijlstra 2015: see). Within the derivational view, it has thus been widely accepted that the variation between French and English can be explained only in terms of the respective properties of verb movement and its interaction with a view of clause structure organized around functional projections.

[2] Korean is peculiar in that it has two ways to express sentential negation: a negative auxiliary (a long form negation) and a morphological negative (a short form negation) for sentential negation. See Kim (2000; 2016) and references therein for detail.

[3] The term 'negator' or 'negative marker' is a cover term for any linguistic expression functioning as sentential negation.

Departing from the derivational view, the non-derivational, lexicalist view requires no uniform syntactic category (e.g., Neg or NegP) for the different types of negatives. The view allows negation to be realized in different grammatical categories, e.g., a morphological suffix, an auxiliary verb, or an adverbial expression. For instance, the negative *not* in English is taken to be an adverb as other negative expressions in English (e.g., *never, barely, hardly*). This view has been suggested by Jackendoff (1972), Baker (1991), Ernst (1992), Abeillé & Godard (1997), Kim (2000), and Warner (2000). In particular, Kim & Sag (1996), Abeillé & Godard (1997), Kim (2000), and Kim & Sag (2002) develop analysis of sentential negation in English, French, Korean, and Italian within the framework of HPSG, showing that the postulation of Neg and its projection NegP creates more empirical and theoretical problems than it solves (see Newmyer (2006) for this point). In addition, there have been substantial work for negation in other languages within the framework HPSG, which does not resort to the postulation of functional projections as well as movement operations to account for the various distributional possibilities of negation (Przepiórkowski & Kupść 1999; Borsley & Jones 2000; Przepiórkowski 2000; Kupść & Przepiórkowski 2002; De Swart & Sag 2002; Borsley & Jones 2005; Crowgey 2010; Bender & Lascarides 2013a: see).

In what follows, we review the HPSG analyses of these four main types of negation, focusing on the distributional possibilities of these four types of negatives in relation to other main constituents of the sentence. When necessary, the chapter also discusses implications for the theory of grammar.[4] We start the HPSG analyses of adverbial negatives in English and French, which have been most extensively studied in transformational grammars. The chapter then moves to the discussion of morphological negatives, negative auxiliary verbs, and preverbal negatives. The chapter also reviews the HPSG analyses of phenomena like genitive negation and negative concord which are sensitive to the presence of negative expressions. The final section concludes this chapter.

2 Adverbial Negative

2.1 Two Key Factors

The most extensively studied type of negation is adverbial negative, which we find in English and French. There are two main factors that determine the position of an adverbial negative: finiteness of the verb and its intrinsic properties,

[4]This chapter grows out of **18**; Kim (2000).

namely, whether it is an auxiliary or lexical verb (Kim 2000; Kim & Sag 2002: see).[5]

The first crucial factor that affects the position of adverbial negatives in English and French concerns the finiteness of the lexical verb. English shows us how the finiteness of a verb influences the surface position of the adverbial negaitve *not*:

(2) a. Kim does not like Lee.
b. * Kim not likes Lee.
c. * Kim likes not Lee.

(3) a. Kim is believed [not [to like Mary]].
b. * Kim is believed to [like not Mary].

As seen from the data above, the negation *not* precedes an infinitive verb, but cannot follow a finite lexical verb (Baker 1989; 1991; Ernst 1992: see). French is not exceptional in this respect. The finiteness also affects the distributional possibilities of the French negative *pas* (Abeillé & Godard 1997; Kim & Sag 2002; Zeijlstra 2007: see):

(4) a. Robin n'aime pas Stacy.
Robin (n')likes NEG Stacy
'Robin does not like Stacy.'
b. * Robin ne pas aime Stacy.

(5) a. Ne pas parler Français est un grand désavantage en ce cas.
ne NEG to.speak French is a great disadvantage in this case
'Not to speak French is a great disadvantage in this case.'
b. * Ne parler pas Français est un grand désavantage en ce cas.

The negator *pas* cannot precede the finite verb but must follow it. But its placement with respect to the nonfinite verb is the reverse image. The negator *pas* should precede the infinitive verb.

The second important factor that determines the position of adverbial negatives concerns the presence of an auxiliary or lexical verb. Modern English displays a clear example where this intrinsic property of the verb influences the position of the English negator *not*: the negator cannot follow a finite lexical verb but when the finite verb is an auxiliary verb, this ordering is possible.

[5] German also employs an adverbial negative *nicht*, which behaves quite differently from the negative in English and French. See Müller (2016) for a detailed review of the previous, theoretical analyses of German negation.

(6) a. * Kim left not the town.
 b. Kim has not left the town.
 c. Kim is not leaving the town.

The placement of *pas* in French infinitival clauses also illustrates that the intrinsic property of the verb affects the position of the adverbial negative *pas*:

(7) a. Ne pas avoir de voiture dans cette ville rend la vie difficile.
 'Not to have a car in this city makes life difficult.'
 b. N'avoir pas de voiture dans cette ville rend la vie difficile.

(8) a. Ne pas être triste est une condition pour chanter des chansons.
 'Not to be sad is a prerequisite condition for sining songs.'
 b. N'être pas triste est une condition pour chanter des chansons.

The negator *pas* can either follow or precede the infinitive auxiliary verb in French, though the acceptability of the ordering in (7b) and (8b) is restricted to certain conservative varieties.

In capturing the distributional behavior of such adverbial negatives in English and French, as noted earlier, the derivational view (exemplified by Pollock (1989) and Chomsky (1991)) has relied on the notion of verb movement and functional projections. The most appealing aspect of this view (initially at least) is that it can provide an analysis of the systematic variation between English and French. By simply assuming that the two languages have different scopes of verb movement – in English only auxiliary verbs move to a higher functional projection whereas all French verbs undergo the same process, the derivational view could explain why the French negator *pas* follows a finite verb, unlike the English negator. In order for this system to succeed, nontrivial complications are required in the basic components of the grammar, e.g rather questionable subtheories (see Kim (2000) and Kim & Sag (2002) for detailed discussion).

Meantime, the nonderivational, lexicalist analyses of HPSG license all surface structures by the system of phrase types and constraints. That is, the position of adverbial negatives is taken to be determined not by the respective properties of verb movement, but by their lexical properties, the morphosyntactic (finiteness) features of the verbal head, and independently motivated Linear Precedence (LP) constraints, as we will see in the following discussion.

2.2 Constituent Negation

When English *not* negates an embedded constituent, it behaves much like the negative adverb *never*. The similarity between *not* and *never* is particularly clear in nonfinite verbal constructions (participle, infinitival and bare verb phrases), as illustrated in (30) and (31) (Klima 1964; Baker 1989; 1991).

(9)
   a. Kim regrets [never [having read the book]].
   b. We asked him [never [to try to read the book]].
   c. Duty made them [never [miss the weekly meeting]].

(10)
   a. Kim regrets [not [having read the book]].
   b. We asked him [not [to try to read the book]].
   c. Duty made them [not [miss the weekly meeting]].

French *ne-pas* is no different in this regard. *Ne-pas* and certain other adverbs precede an infinitival VP:

(11)
   a. [Ne pas [repeindre sa maison]] est une négligence.
      ne not paint one's house is a negligence
      'Not to paint one's house is negligent.'
   b. [Régulièrement [repeindre sa maison]] est une nécessité.
      regularly to.paint one's house is a necessity

To account for these properties, Kim (2000) and Kim & Sag (2002) regard *not* and *ne-pas* not as heads of their own functional projection, but rather as adverbs that modify nonfinite VPs. The lexical entries for *ne-pas* and *not* include the information shown in (12).[6]

(12)
$$\begin{bmatrix} \text{FORM} & \langle \textit{not/ne-pas} \rangle \\ \text{SYN}|\text{HEAD} & \begin{bmatrix} \textit{adv} & \\ \text{MOD} & \langle \text{VP}[\textit{nonfin}]: \boxed{2} \rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{RESTR} & \left\langle \begin{bmatrix} \text{PRED} & \textit{neg-rel} \\ \text{ARG1} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The lexical information in (12) specifies that *not* and *ne-pas* modifies a nonfinite VP and that this modified VP serves as the semantic argument of the negation. This simple lexical specification correctly describes the distributional similarities between English *not* and French *ne-pas*, licensing the structure in Figure 1.

[6] Here we assume that both languages distinguish between *fin(ite)* and *nonfin(ite)* verb forms, but that certain differences exist regarding lower levels of organization. For example, *prp* (*present participle*) is a subtype of *fin* in French, whereas it is a subtype of *nonfin* in English.

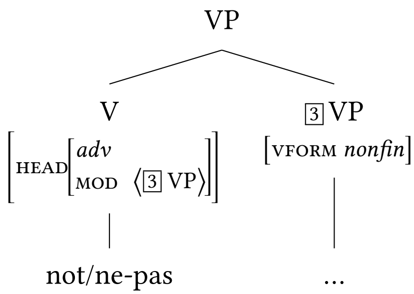


Figure 1: Structure of constituent negation

This structure implies that with the LP (linear precedence) rule specificying a modifier precedes the head modifies, both *ne-pas* and *not*, like other adverbs of this type, precede the VPs that they modify while not separating an infinitival verb from its complements, as observed in the following data:

(13) a. [Not [speaking English]] is a disadvantage.

b. * [Speaking not English] is a disadvantage.

(14) a. [Ne pas $_{\text{VP}[inf]}$[parler français]] est un grand désavantage en ce cas.
ne not $_{\text{VP}[inf]}$[to.speak French is a great disadvantage in this case

b. * [Ne parler pas français] est un grand désavantage en ce cas.

Interacting with the LP constraint, the lexical specification in (12) ensures that the constituent negation precede the VP it modifies. This predicts the grammaticality of (13a) and (35), where *ne-pas* and *not* are used as VP[*nonfin*] modifiers. (13b) and (14b) are ungrammatical, since the modifier fails to appear in the required position—i.e. before all elements of the nonfinite VP.

The HPSG analyses sketched here have thus recognized the fact that finiteness plays a crucial role in determining the distributional possibilities of negative adverbs. Its main explanatory resource has basically come from the proper lexical specification of these negative adverbs. The lexical specification that *pas* and *not* both modify nonfinite VPs has sufficed to predict their occurrences in nonfinite clauses.

2.3 Sentential Negation

With respect to negation in finite clauses, there are important difference between English and French. As we have noted earlier, it is a general fact of French that *pas* must follow a finite verb, in which case the verb optionally bears negative morphology (*ne*-marking):

(15) a. Dominique (n')aime pas Alex.
  b. * Dominique pas aime Alex.

In English, *not* must follow a finite auxiliary verb, not a lexial (or main) verb:

(16) a. Dominique does not like Alex.
  b. * Dominique not does like Alex.
  c. * Dominique likes not Alex.

In contrast to the distribution of *not* in nonfinite clauses as constituent negation, its distribution in finite clauses concerns sentential negation. The need to distinguish the two types of negation can be found from scope possibilities in an example like (17) (Klima 1964; Baker 1989; Warner 2000: see).[7]

(17) The president could not approve the bill.

Negation here could have the two different scope readings paraphrased in (43).

(18) a. It would be possible for the president not to approve the bill.
  b. It would not be possible for the president to approve the bill.

The first interpretation is constituent negation; the second is sentential negation. As noted, sentential *not* may not modify a finite VP, different from the adverb *never*:

(19) a. Lee never/*not left. (cf. Lee did not leave.)
  b. Lee will never/not leave.

The contrast in these two sentences shows one clear difference between *never* and *not*. The negator *not* cannot precede a finite VP though it can freely occur as a nonfinite VP modifier.

Another distributional difference between *never* and *not* is found in the VP ellipsis construction. Observe the following contrast (Kim & Sag 2002):

[7] Warner (2000) and Bender & Lascarides (2013b) discuss scopal interactions of negation with auxiliaries (modals) and quantifiers within the system of Minimal Recursion Semantics (MRS).

(20) a. Mary sang a song, but Lee never could _.

b. * Mary sang a song, but Lee could never _.

c. Mary sang a song, but Lee could not _.

The data here indicate that *not* behaves differently from adverbs like *never* in finite contexts, even though the two behave alike in nonfinite contexts. The adverb *never* is a true diagnostic of a VP-modifier, and we use contrasts between *never* and *not* to reason about what the properties of the negator *not* must be.

We saw the lexical representation for constituent negation *not* in (12) above. Unlike the constituent negator, the sentential negator *not* typically follows a finite auxiliary verb. In this respect, *too, so* and *indeed* also behave alike:

(21) a. Kim will not read it.

b. Kim will too/so/indeed read it.

These expressions are used to reaffirm the truth of the sentence in question and follow a finite auxiliary verb. This implies that the sentential *not* in English form a group of adverbs that combine with a preceding auxiliary verb (Kim 2000: see).

With the observation of such properties, the HPSG analyses of Abeillé & Godard (1997), Kim (2000), and Warner (2000) have taken this group of adverbs ($\mathrm{Adv}_{\mathrm{I}}$) including the sentential negation *not* to function as the complement of a finite auxiliary verb via the following lexical rule:

(22) NEGATIVE ADVERB-COMPLEMENT LEXICAL RULE:

$$\begin{bmatrix} \textit{fin-aux} \\ \textsc{syn} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{aux} & + \\ \textsc{vform} & \textit{fin} \end{bmatrix} \\ \textsc{val}|\textsc{comps} & L \end{bmatrix} \end{bmatrix} \mapsto \begin{bmatrix} \textit{neg-fin-aux} \\ \textsc{syn} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{aux} & + \\ \textsc{vform} & \textit{fin} \\ \textsc{neg} & + \end{bmatrix} \\ \textsc{val}|\textsc{comps} & \left\langle \textsc{Adv}_{\mathrm{I}} \right\rangle \oplus L \end{bmatrix} \end{bmatrix}$$

This lexical rule specifies that when the input is a finite auxiliary verb, the output is a neg-finite auxiliary (*fin-aux* ⟶ *neg-fin-aux*) that selects $\mathrm{Adv}_{\mathrm{I}}$ as an additional complement. This then would license a structure like in Figure 2.

As shown in Figure 2, the negative finite auxiliary verb *could* selects two complements, the negator *not* and the VP *approve the bill.* The finite auxiliary then combines with these two complements, forming a well-formed Head-Complement construct. By treating *not* as both a modifier (constituent negation) and a lexical complement (sentential negation), we thus can account for the scope differences in (17) with the following two possible structures:

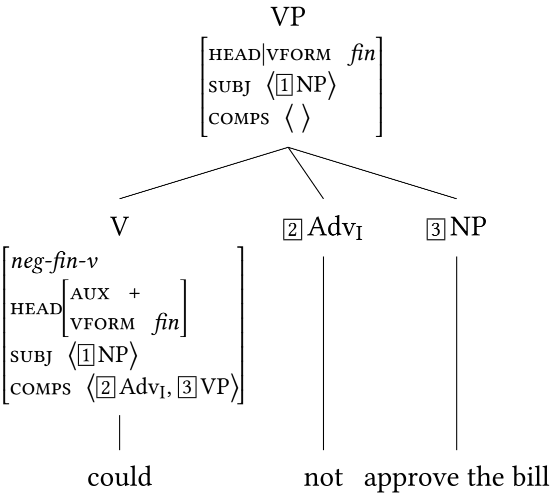


Figure 2: Structure of sentential negation

(23) a. The president could [not [approve the bill]].

b. The president [could] [not] [approve the bill].

In (23a), *not* functions as a modifier to the base VP, while in (23b) whose partial structure is given in Figure 2, it is a sentential negation that functions as the complement of the finite auxiliary with a wide scope reading.

The present analysis can also offer us a simple way to account for various other phenomena including VP ellipsis we discussed in (20). The point was that unlike *never*, the sentential negation can host a VP ellipsis. The VP ellipsis after *not* is possible, given that any VP complement of an auxiliary verb can be unexpressed, as specified by the following lexical rule (**kim:00**):

(24) PREDICATE ELLIPSIS LEXICAL RULE:

$$\begin{bmatrix} \textit{aux-v-lxm} \\ \text{ARG-ST} \ \langle \boxed{1}\text{XP}, \boxed{2}\text{YP} \rangle \end{bmatrix} \mapsto \begin{bmatrix} \textit{aux-ellipsis-wd} \\ \text{COMPS} \ \langle \ \rangle \\ \text{ARG-ST} \ \langle \boxed{1}\text{XP}, \boxed{2}\text{YP}[\textit{pro}] \rangle \end{bmatrix}$$

What the rule in (24) tells us is that an auxiliary verb selecting two arguments can be projected into an elided auxiliary verb whose second argument is realized as a small *pro*. Note that this argument is not mapped onto the syntactic grammatical function COMPS. This then would license the structure in Figure 3:

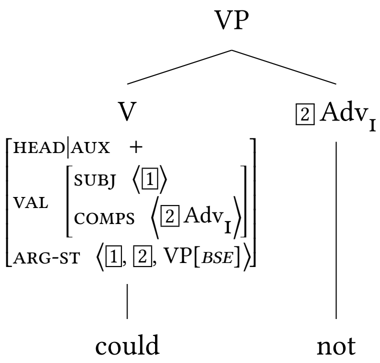


Figure 3: A licensed VPE structure

As represented in Figure 3, the auxiliary verb *could* forms a well-formed head-complement construct with *not* while its VP[*bse*] is unrealized (see Kim (2000); Kim & Sells (2008) for detail). The sentential negator *not* can 'survive' VPE because it can be licensed in the syntax as the complement of an auxiliary, independent of the following VP. However, an adverb like *never* is only licensed as a modifier of VP. Thus if the VP were elided, we would have the hypothetical structure like the one in Figure 4: In Figure 4, the adverb *never* modifies a VP

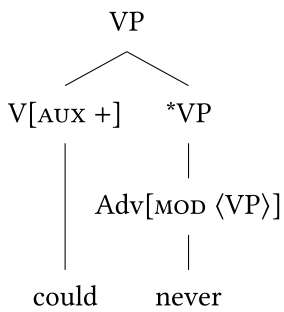


Figure 4: Ill-formed structure of the Head-Mod Construction

through the feature MOD, which guarantees that the adverb requires the head VP that it modifies. In an ellipsis structure, the absence of such a VP means that there is no VP for the adverb to modify. In other words, there is no rule licensing such a combination – predicting the ungrammaticality of **has never*, as opposed to *has not*.

The HPSG analysis just sketched here can be easily extended to French negation, whose data we repeat here.

(25) a. * Robin ne pas aime Stacy.
Robin ne NEG likes Stacy
'Robin does not like Stacy.'

b. Robin (n')aime pas Stacy.
Robin likes NEG Stacy
'Robin does not like Stacy.'

Unlike the English negator *not*, *pas* must follow the finite verb. Such a distributional contrast has motivated verb movement analyses (Pollock 1989; Zanuttini 2001: see). By contrast, the present HPSG analysis is cast in terms of a lexical rule that maps a finite verb into a verb with a certain adverb like *pas* as an additional complement. The idea of converting modifiers in French into complements has been independently proposed by Miller (1992) and Abeillé & Godard (1994) for French adverbs including *pas*. Building upon this previous work, Abeillé & Godard (1997) and Kim (2000) allow the adverb *pas* to function as a syntactic complement of a finite verb in French.[8] This output verb *neg-fin-v* then allows the negator *pas* to function as the complement of the verb *n'aime*, as represented in Figure 5.

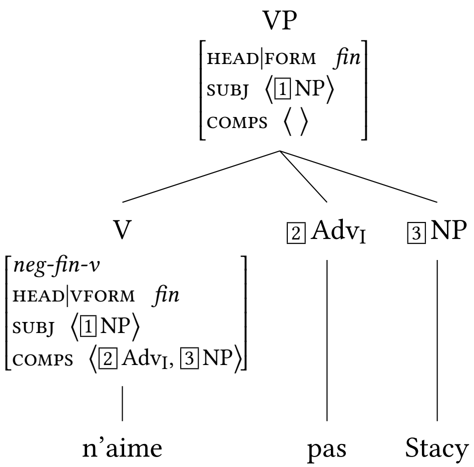


Figure 5: Partial structure of (25b)

[8]Following Abeillé & Godard (1994), we could assume *ne* to be an inflectional affix which can be optionally realized in the output of the lexical rule in Modern French.

The analysis also explains the position of *pas* in finite clauses. The placement of *pas* before a finite verb in (25a) is unacceptable since *pas* here is used not as a nonfinite VP modifier, but as a finite VP modifier. But due to the present analysis which allows *pas*-type negative adverbs to function as the complement of a finite verb, *pas* in (25b) can function as the sister of the finite verb *aime*.

Given that the conditional, imperative, and subjunctive, and even present participle verb forms in French are finite, the construction analysis also predicts that *pas* cannot precede any of these verb forms:

(26) a. Si j'avais de l'argent, je ne achèterais pas.
'If I had money, I would not buy a car.'
b. * Si j'avais de l'argent, je ne pas achèterais.

(27) a. Ne mange pas ta soupe.
'Don't eat your soup!'
b. * Ne pas mange ta soupe.

(28) a. Il est important que vous ne répondiez pas.
'It is important that you not answer.'
b. * Il est important que vous ne pas répondiez.

(29) a. Ne parlant pas Français, Stacy avait des difficultés.
'Not speaking French, Stacy had difficulties.'
b. * Ne pas parlant Français, Stacy avait des difficultés.

Note that this analysis allows us to reduce the differences between French and English negation to be a matter of lexical properties. The negators *not* and *pas* are identical in that they both are VP[*nonfin*] modifying adverbs. But they are different with respect to which verbs can select them as complements: *not* can be the complement of a finite auxiliary verb, whereas *pas* can be the complement of any finite verb. So the only difference is the morphosyntactic value [AUX +] and this induces the difference in positioning the negators in English and French.

3 Morphological Negative

As noted earlier, languages like Turkish and Japanese employ morphological negation in which the negative marker behaves like a suffix. Consider Turkish and Japanese examples in the following:

(30) a. Git-me-yeceğ-∅-im
go-NEG-FUT-COP-1SG
'(I) will not come.'

b. kare-wa kinoo kuruma-de ko-na-katta.
he-TOP yesterday car-INST come-NEG-PST
'He did not come by car yesterday.'

As the examples illustrate, the sentential negation of Turkish and Japanese employ morphological suffixes *-me* and *-na*, respectively. It is possible to state the ordering of these morphological negative markers in configurational terms with assigning an independent syntactic status to them. But it is too strong a claim to take the negative suffix *-me* or *-na* to be an independent syntactic element, and to attribute its positional possibilities to syntactic constraints such as verb movement and other configurational notions (see Kelepir (2001) for Turkish and Kato (1997; 2000) for Japanese). In these languages, the negative affix acts just like other verbal inflections in numerous respects. The morphological status of these negative markers comes from their morphophonemic alternation. For example, the vowel of the Turkish negative suffix *-me* shifts from open to closed when followed by the future suffix, as in *gel-mi-yecke* 'come-NEG-FUT'. Their strictly fixed position also indicates their morphological constituenthood. Though these languages allow rather a free permutation of syntactic elements (scrambling), there exist strict ordering restrictions among verbal suffixes including the negative suffix, as can be seen from the following examples:

(31) a. tabe-sase-na-i/*tabe-na-sase-i
eat-CAUS-NEG-NPST

b. tabe-rare-na-katta/*tabe-na-rare-katta
eat-PASS-NEG-PST

c. tabe-sase-rare-na-katta/*tabe-sase-na-rare-katta
eat-CAUS-PASS-NEG-PST

The ordering of the negative affix is a matter of morphology. If it were a syntactic concern, then the question would arise as to why there is an obvious contrast in the ordering principles of morphological and syntactic constituents, i.e., why the ordering rules of morphology are distinct from the ordering rules of syntax. The simplest explanation for this contrast is to accept the view that morphological constituents including the negative marker are formed in the lexical component and hence have no syntactic status (see Kim (2000) for detailed discussion).

This being noted, it is more reasonable to assume that the placement of a negative affix is regulated by morphological principles, i.e. by the properties of the morphological negative affix itself. Przepiórkowski & Kupść (1999) and Przepiórkowski (2000; 2001) discuss aspects of Polish negation which is realized as the prefix *nie* to a verbal expression. Przepiórkowski (2000) focuses on non-local genitive of negation in Polish where the object argument is not accusative but genitive marked with the presence of negative marker as in (32a). The assignment of genitive case to the object is also effective in the unbounded relation as shown in (32b):

(32) a. Nie lubię Marii/*Marię.
not like-1st.SG Mary-GEN/Mary-ACC
'I don't like Mary.'

b. Mogę nie chcieć tego napisać.
may-1st.SG not want-INF this-GEN write-INF
'I may not want to write this.'

To account for this kind of phenomena, Przepiórkowski (2000) develops an HPSG-based analysis with the assumption that the combination of the negative morpheme with the verb stem introduces the feature NEG. This feature tightly interacts with the mechanism of argument composition and construction-based case assignment (or satisfaction).

The process of adding a negative morpheme to a lexeme can be modelled straightforwardly by a lexical rule given in the following (Kim 2000; Crowgey 2012: see):

(33) NEGATIVE WORD FORMATION LEXICAL RULE:

$$\begin{bmatrix} \textit{v-lxm} & \\ \text{FORM} & \langle \boxed{1} \rangle \\ \text{SEM} & \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \textit{neg-v-lxm} & \\ \text{FORM} & \langle \text{F}_{neg}(\boxed{1}) \rangle \\ \text{SYN|HEAD|POL} & \textit{neg} \\ \text{SEM} & \begin{bmatrix} \text{RESTR} & \left\langle \begin{bmatrix} \text{PRED} & \textit{neg-rel} \\ \text{ARG1} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

As shown here, any verb lexeme can be turned into a verb with the negative morpheme attached. The language particular rule will ensure an appropriate negative morpheme to be attached to the lexeme. For instance, the suffix *-ma* for Turkish and *-na* for Japanese will be attached to the verb lexeme, generating the verb forms in (30a).

This morphological analysis can be extended to the negation of languages like Libyan Arabic, as discussed in Borsley & Krer (2012). The language has a bipartite realization of negation, the proclitic *ma-* and the enclitic -š:

(34) la-wlaad ma-mšuu-š li-l-madrsa.
the-boys NEG-go.PST.3.PL-NEG to-the-school
'The boys didn't go to the school.'

As did Borsley & Krer (2012), we could take these clitics as affixes and generate a negative word as following:[9]

(35) $$\begin{bmatrix} \text{FORM } \langle \text{mšuu} \rangle \\ \text{SEM}|\ \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \text{FORM } \langle \text{ma-mšuu-š} \rangle \\ \text{SYN}|\text{HEAD}|\text{POL } \textit{neg} \\ \text{SEM} \begin{bmatrix} \text{RESTR} \left\langle \begin{bmatrix} \text{PRED} & \textit{neg-rel} \\ \text{ARG1} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The only thing we need to define here is the function $\mathbf{F}_{neg}$ in the language that allows the attachment of the prefix *ma-* and the suffix *š* to the verb stem *mšuu*.[10]

The lexicalist HPSG analyses sketched here have been built upon the thesis that autonomous (i.e. non-syntactic) principles govern the distribution of morphological elements (Bresnan & Mchombo 1991). The position of the morphological negation is simply defined in relation to the verb stem it attaches to. There are no syntactic operations such as head-movement or multiple functional projections in forming a verb with the negative marker.[11]

4 Negative Auxiliary Verb

Another way of expressing sentential negation, as noted earlier, is to employ a negative auxiliary verb. Some head-final languages like Korean and Hindi employ negative auxiliary verbs. Consider a Korean example:

(36) John-un ku chayk-ul ilk-ci anh-ass-ta.
John-TOP that book-ACC read-CONN NEG-PST-DECL
'John did not read the book.'

[9] The formulation given in Borsley & Krer (2012) is slightly different from the one given here, but both have the same effects.

[10] Borsley & Krer (2012) note that the suffix *š* is not realized when a negative clause includes an n-word or an NPI (negative polarity item). See Borsley & Krer (2012) for further details.

[11] The lexical-rule based approached here can be extended to a construction-based HPSG approach or a constructionist approach. See Sag (2012) and Hilpert (2016) for a construction-based approach to morphological processes.

The negative auxiliary in head-final languages like Korean typically appears clause-finally, following the invariant form of the lexical verb. In head-initial SVO languages, however, the negative auxiliary almost invariably occurs immediately before the lexical verb (Payne 1985). Finnish also exhibits this property (Mitchell 1991):

(37) Minä e-n puhu-isi.
I.NOM NEG-1SG speak-COND
'I would not speak.'

These negative auxiliaries have syntactic status: they can be inflected, above all. Like other verbs, they can be marked with verbal inflections such as agreement, tense, and mood.

In dealing with auxiliary negative constructions, most of the derivational approaches have followed Pollock's and Chomsky's analyses in factoring out grammatical information (such as tense, agreement, and mood) carried by lexical items into various different phrase-structure nodes (see, among others, Hagstrom (2002), and Han et al. (2007) for Korean and Vasishth (2000) for Hindi). This derivational view has been appealing, in that the configurational structure for English-type languages could be applied even for languages with different types of negation. However, issues arise how to address the grammatical properties of auxiliary negative which are quite different from the other negative forms.

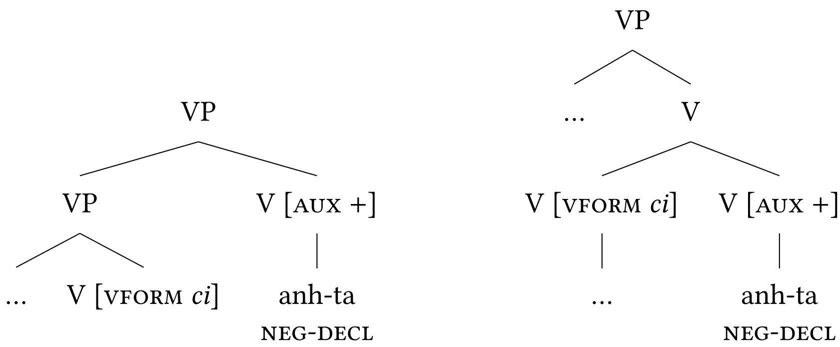
The Korean negative auxiliary displays all the key properties of auxiliary verbs in the language. For instance, the typical auxiliary verbs as well as the negative auxiliary all require the preceding lexical verb to be marked with a specific verb form (VFORM), as illustrated in the following:

(38) a. ilk-ko/*ci siph-ta.
read-CONN/CONN would.like-DECL
'(I) would like to read.'

b. ilk-ci anh-ass-ta.
read-CONN not-PST-DECL
'(I) did not read.'

The auxiliary verb *siph-* in (38a) requires the *-ko* marked lexical verb while the negative auxiliary verb *anh-* in (38b) asks for the *-ci* marked lexical verb.

In terms of syntactic structure, there are at least possible analyses. One is to assume that the negative auxiliary takes a VP complement and the other is to



(a) VP-structure

(b) Verb-complex structure

Figure 6: Two possible structures with the negative auxiliary

claim that it forms a verb complex with the preceding lexical verb, as represented in Figures 6a and 6b, respectively (Chung 1998a; Kim 2016).

The distributional properties of the negative auxiliary in the language, however, support a complex predicate structure (cf. Figure 6b) in which the negative auxiliary verb forms a syntactic/semantic unit with the preceding lexical verb. For instance, no adverbial expression, including a parenthetical adverb, can intervene between the main and auxiliary verb, as illustrated by the following Korean example:

(39) Mimi-nun (yehathun) tosi-lul (yehathun) ttena-ci (*yehathun)
Mimi-TOP anyway city-ACC anyway leave-CONN anyway
anh-ass-ta.
NEG-PST-DECL
'Anyway, Mimi didn't leave the city.'

Further, in an elliptical construction, the elements of a verb complex always occur together. Neither the lexical verb nor the auxiliary verb alone can serve as the fragment answer to the polar question. The two verbs must occur together.

(40) a. Kim-i hakkyo-eyse pelsse tolawa-ss-ni?
Kim-NOM school-SRC already return-PST-QUE
'Did Kim return from school already?'

b. ka-ci-to anh-ass-e.
go-CONN-DEL NOTP-PST-DECL
'(He) didn't even go.'

c. * ka-ci-to. go-CONN-DEL

d. * anh-ass-e NEG-PST-DECL

As shown in (40c) and (40d) here, neither the lexical verb nor the auxiliary verb alone can serve as an indepedent fragment answer. The two verbs must appear together as given in (40b). These constituent tests indicate that the negative auxiliary forms a syntactic unit with a preceding lexical verb in Korean. Following the HPSg analyses of Bratt (1996), Chung (1998b), and Kim (2016), we then could assume that an auxiliary verb forms a complex predicate, licensed by the following construction:

(41) HEAD-LIGHT CONSTRUCTION:

$$\begin{bmatrix} \textit{hd-light-cxt} \\ \text{COMPS } L \\ \text{LIGHT } + \end{bmatrix} \rightarrow \boxed{1}\begin{bmatrix} \text{LIGHT } + \\ \text{COMPS } L \end{bmatrix}, \text{H}\begin{bmatrix} \text{LIGHT } + \\ \text{COMPS } \langle \boxed{1} \rangle \oplus L \end{bmatrix}$$

This construction rule means that a LIGHT head expression combines with a LIGHT complement, yielding a light, quasi-lexical constituent (Bonami & Webelhuth 2013). When this combination happens, there is a kind of argument composition: the COMPS value (L) of this lexical complement is passed up to the resulting mother. The constructional constraint thus induces the effect of argument composition in syntax, as illustrated by the following example:

The auxiliary verb *anh-ass-ta* 'NEG-PST-DECL' combines with the matrix verb *ilk-ci* 'read-CONN', forming a well-formed Head-Light construct. Note that the resulting construction metaphorically inherits the COMPS value from that of the lexical complement *ilk-ci* 'read-CONN' in accordance with the structure-sharing emposed on by the HEAD-LIGHT CONSTRUCTION in (41). That is, the HEAD-LIGHT CONSTRUCTION licenses the combination of an auxiliary verb with its lexical verb, while inheriting the lexical verb's complement value as argument composition. The present system thus allows the argument composition at the syntax level, rather than in the lexicon.

The HPSG analyses have taken the negative auxiliary in Korean to select a lexical verb, whose combination forms a verb complex structure. The present analysis implies that there is no upper limit for the number of auxiliary verbs to occur in sequence, as long as each combination observes the morphosyntactic constraint on the preceding auxiliary expression. Consider the following:

(42) a. sakwa-lul [mek-ci anh-ta].
apple-ACC eat-CONN NEG-DECL
'do not eat the apple'

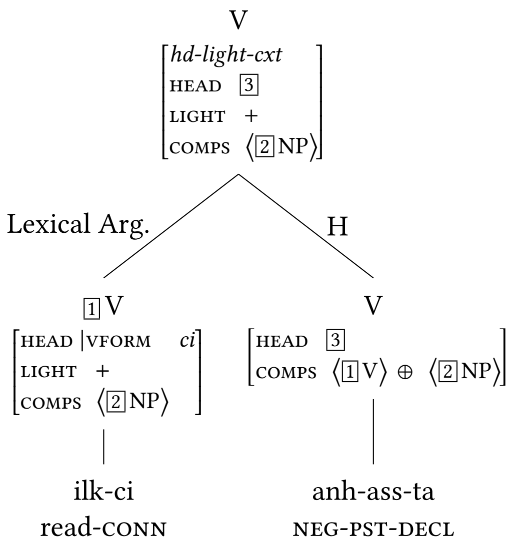


Figure 7: An instance (construct) of the HEAD-LIGHT CONSTRUCTION

b. sakwa-lul [[mek-ko siph-ci] anh-ta].
   apple-ACC eat-CONN wish-CONN NEG-DECL
   'would not like to eat the apple'

c. sakwa-lul [[[mek-ko siph-e] ha-ci] anh-ta].
   apple-ACC eat-CONN wish-CONN do-CONN NEG-DECL
   'do not like to eat the apple'

d. sakwa-lul [[[[mek-ko siph-e] ha-key] toy-ci] anh-ta].
   apple-ACC eat-CONN wish-CONN do-CONN become-CONN NEG-DECL
   'do not become to like to eat the apple.'

As seen from the bracketed structures, we can add one more auxiliary verb to an existing Head-Light construct with the final auxiliary bearing an appropriate connective marker. there is no upper limit to the possible number of auxiliary verbs we can add (see Kim (2016) for detailed discussion).

The present analysis in which the auxiliary negative forms a complex predicate structure with a lexical verb can be applied for languages like Basque, as suggested by Crowgey & Bender (2011). They explore the interplay of sentential negation and word order in Basque. Consider a Basque example from Crowgey & Bender (2011):

(43) ez-ditu irakurri liburuak.
NEG-3PLO.PRES.3SGS read.PERF book.ABS.PL
'has not read books'

Unlike Korean, the negative auxiliary *ez-ditu* precedes the main verb. Other than this ordering difference, just like Korean, the two form a verb complex structure, as represented in Figure 8 (adopted from Crowgey & Bender (2011)):

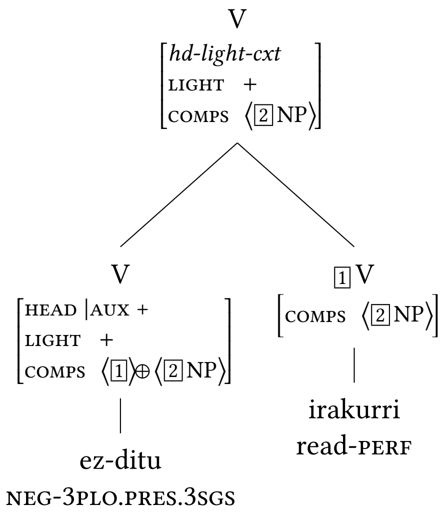


Figure 8: Partial structure of (43)

In the treatment of negative auxiliary verbs, the HPSG analyses have taken the negative auxiliary to be an independent lexical verb whose grammatical (syntactic) information is not distributed over different phrase structure nodes, but incorporated into its precise lexical specifications. In particular, the negative auxiliary forms a verb complex structure in many languages, whose constituenthood is motivated by other independent phenomena.

5 Preverbal Negative

The final type of sentence negation is preverbal negatives, which we can observe in languages like Italian and Welsh:

(44) a. Gianni non telefona a nessuno. (Italian)
Gianni NEG telephones to noone
'Gianni does not call anyone.'

b. Dw i ddim wedi gweld neb. (Welsh)
am I NEG PERF see nobody
'I haven't seen anybody.'

As seen here, in Italian, the preverbal negative *non*, also called negative particle or clitic, always precedes a lexical verb, whether finite or non-finite, as further attested by the following examples:

(45) a. Gianni vuole che io non legga articoli di sintassi.
Gianni wants that I NEG read articles of syntax.
'Gianni hopes that I do not read syntax articles.'

b. Non leggere articoli di sintassi è un vero peccato.
NEG to.read articles of syntax is a real shame
'Not to read syntax articles is a real shame.'

c. Non leggendo articoli di sintassi, Gianni trova la linguistica noiosa.
NEG reading articles of syntax Gianni finds linguistics boring
'Not reading syntax articles, Gianni finds linguistics boring.'

The derivational view again attributes the distribution of such a preverbal negative to the reflex of verb movement and functional projections (Belletti 1990; Zanuttini 1991: see). This line of analysis also appears to be persuasive in that the different scope of verb movement application could explain the observed variations among typologically and genetically related languages. Such an analysis, however, fails to capture unique properties of preverbal negative from morphological negative, auxiliary negative, and adverb negative.

Kim (2000) offers an HPSG analysis of Italian and Spanish negation. His analysis takes *non* to be an independent lexical head element though it is a clitic. This claim follows the analyses sketched by Monachesi (1993) and Monachesi (1998) assuming that there are two types of clitics, affix-like clitics and word-like clitics: pronominal clitics belong to the former, whereas the bisyllabic clitic *loro* 'to-them' to the latter. Kim's analysis suggests that *non* also belongs to the latter group.[12] One key difference from pronominal clitics is thus that it functions

[12]But one main difference between *non* and *loro* is that *non* is a head element, whereas *loro* is a complement XP. See Monachesi (1998) for further discussion of the behavior of *loro* and its treatment.

as an independent word. Treating *non* as a word-like element will allow us to capture its word-like properties such as the possibility of stress on the negator and its separation from the first verbal element. But it is not a phrasal modifier, but an independent particle (or clitic), which combines with the following lexical verb.[13]

(46) Lexical specifications for *non* in Italian:

$$\begin{bmatrix} \textsc{form} \ \langle \text{non} \rangle \\ \textsc{syn} \begin{bmatrix} \textsc{head} \ \boxed{1} \\ \textsc{comps} \left\langle \text{V}\begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{comps} & L \end{bmatrix} \right\rangle \oplus L \end{bmatrix} \\ \textsc{sem} \begin{bmatrix} \textsc{restr} \left\langle \begin{bmatrix} \textsc{pred} & \textit{neg-rel} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

This lexical entry roughly corresponds to the entry for Italian auxiliary verbs (and restructuring verbs with clitic climbing), in that the negator *non* selects a verbal complement and further for the complement list (L). One key property of *non* is its HEAD value: its HEAD value is in a sense undetermined, but structure-shared with the HEAD value of its verbal complement. The value is thus determined by what it combines with. When *non* combines with a finite verb, it will be a finite verb. When it combines with an infinitive verb, it will be an infinitive verb.

In order to see how this system works, let us consider an Italian example where the negator combines with a transitive verb as in the following:

(47) Gianni non legge articoli di sintassi.
Gianni NEG reads articles of syntax
'Gianni doesn't read syntax articles.'

When the negator *non* combines with the finite verb *legge* that selects an NP object, the resulting combination will form a verb complex structure given in Figure 9.

Borsley (2006), adopting Kathol (2000) topological approach, provides a linear-based HPSG approach to capture the distributional possibilities of negation in Italian and Welsh, which we have seen in (44a) and In Italian, negation is realized in a preverbal position as shown in (44b), respectively. Borsley (2006) assumes that a negative clause in these languages has a negative element (bearing the feature NEG) in the specified field. Different from **BJ:05** selectional approach where a negative expression selects its own complement, this linear-based approach

[13] See Kim (2000) for detailed discussion.

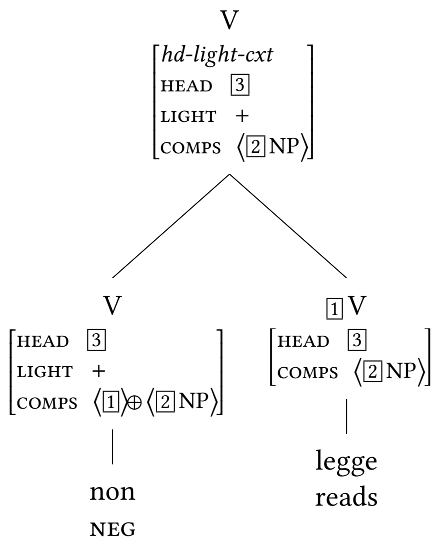


Figure 9: Verb complex structure of (47)

tries to offer a unified approach to the distributional possibilities in three different languages. by assigning the negative expression in each language to a specified topological field. The linear-based approach assumes that constituents have an order domain to which ordering constituents apply. Borsley (2006), accepting the analysis of Kim (2000) where *non* is taken to be a type of clitic-auxiliary, posits the following order domain:

(48) $$\begin{bmatrix}\textsc{dom}\left\langle\begin{bmatrix}\textit{first}\\ \langle\text{Gianni}\rangle\end{bmatrix}, \begin{bmatrix}\textit{second}\\ \textsc{neg}\ +\\ \langle\text{non}\rangle\end{bmatrix}, \begin{bmatrix}\textit{third}\\ \langle\text{telephona}\rangle\end{bmatrix}, \begin{bmatrix}\textit{third}\\ \textsc{neg}\ +\\ \langle\text{a nessuno}\rangle\end{bmatrix}\right\rangle\end{bmatrix}$$

The only required constraint that Borsley (2006) needs to postulate is that a negative element bearing the positive NEG feature is either in the second or the third field. One advantage of this direction is to allow Borsley (2006) to attribute the properties of Welsh negation to the difference in the domain value of the same NEG element. Unlike Italian, the NEG bearing negative is in the second or the third field, as illustrated in the following domain for the sentence (44b) (from (49) Borsley (2006)):[14]

[14]Different from Borsley (2006), Borsley & Jones (2000) offer a selectional analysis of Welsh negation. That is, the finite negative verb selects two complements (e.g., subject and object) while the nonfinite negative verb selects a VP. See Borsley & Jones (2000) for details.

(49) $\begin{bmatrix}\textsc{dom}\left\langle \begin{bmatrix}\textit{second}\\ \langle \text{dw}\rangle\end{bmatrix}, \begin{bmatrix}\textit{third}\\ \langle \text{i}\rangle\end{bmatrix}, \begin{bmatrix}\textit{third}\\ \textsc{neg}\ +\\ \langle \text{ddim}\rangle\end{bmatrix}, \begin{bmatrix}\textit{third}\\ \langle \text{wedi gweld neb}\rangle\end{bmatrix}\right\rangle\end{bmatrix}$

6 Other related phenomena

In addition to these work focusing on the distributional possibilities of negation, there have also been work on genitive negation and negative concord.

Przepiórkowski (2000) focuses on non-local genitive of negation in Polish where the object argument is not accusative but genitive marked with the presence of negative marker as in (50ab). The assignment of genitive case to the object is also effective in the unbounded relation as shown in (51a) (data from Przepiórkowski (2000)):

(50) a. Lubię Marię
like-1st.SG Mary-ACC
'I like Mary.'

b. Nie lubię Marii/*Marię
not like-1st.SG Mary-GEN/Mary-ACC
'I don't like Mary.'

(51) a. Janek wydawał się lubić Marię.
John seemed RM like-INF Mary-ACC
'John seemed to like Mary.'
Janek nie wydawał się lubić Marii/Marię.
John not seemed RM like-INF Mary-GEN/Mary-ACC
'John did not seem to like Mary.'

To account for this kind of phenomena, Przepiórkowski (2000) develops an HPSG-based analysis with the assumption that the combination of the negative morpheme *nie* with the verb stem introduces the feature NEG.[15] The case assignment constraint such that a NEG verbal expression assigns the GEN to its non-initial argument ensures the object NP in (50ab) to be GEN-marked (adopted from Przepiórkowski (2000)):

(52) POLISH CASE ASSIGNMENT RULE:

[15] In Polish, negation is realized as the prefix *nie* to a verbal expression. Przepiórkowski & Kupść (1999); Przepiórkowski (2000; 2001: see).

$$\begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{neg} & + \end{bmatrix} \\ \textsc{arg-st} & \langle \text{XP, YP[}\textsc{case}\ \textit{str}] \rangle \end{bmatrix} \mapsto \begin{bmatrix} \textsc{arg-st} & \langle \text{XP} \rangle \oplus \langle \text{NP[}\textsc{case}\ \textit{gen}] \rangle \oplus \textit{L} \end{bmatrix}$$

This type constraint will ensure that the object complement of (50a) to be GEN-marked due to the negative word *nie lubię*. As for the long distance GEN in (51a), Przepiórkowski (2000) allows the VP complement of raising verbs like *seem* to optionally undergo the lexical argument composition, yielding the following for the matrix verb in (51a):

(53)
$$\begin{bmatrix} \textsc{form} & \langle \text{nie wydawał się} \rangle \\ \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{neg} & + \end{bmatrix} \\ \textsc{arg-st} & \left\langle \text{NP, VP}\begin{bmatrix} \textsc{comps}\ \langle \boxed{1}\text{NP} \rangle \end{bmatrix} \right\rangle \oplus \langle \boxed{1}\text{NP[}\textit{str}] \rangle \end{bmatrix}$$

This lexical specification allows the object NP of the verb to get GEN-marked in accordance with the constraint in (52).[16] In his analysis, the feature NEG thus tightly interacts with the mechanism of argument composition and construction-based case assignment (or satisfaction).

Negative concord also also concerns negation that we often find in languages like French, Italian, Polish, and so forth. De Swart & Sag (2002) investigates negative concord in French where multiple occurrences of phonologically negative constituents express either double negation or single negation:

(54) Personne n'aime personne.
None likes no.one
'No one is such that they love no one.' (DN)
'No one likes anyone.' (NC)

The double negation reading in (54) has two quantifiers while the single negation reading is an instance of negation concord where the two quantifiers merge into one. De Swart & Sag (2002), assuming that the information of each quantifier is stored in QSTORE and retrieved at the lexical level in accordance with constraints on the verb's arguments and semantic content. For instance, the verb *n'aime* in (54) will have two different ways of retrieving the QSTORE value as in the following:[17]

[16] When there is no argument composition, the positive verb *lubić* assigns ACC to the object NP.

[17] The QSTORE value contains information roughly equivalent to first order logic expressions like *NO*x[Person(x)]. See De Swart & Sag (2002).

(55) a.
$$\begin{bmatrix} \textsc{form} & \langle \textit{n'aime} \rangle \\ \textsc{arg-st} & \left\langle \text{NP}[\textsc{store}\ \{\boxed{1}\}], \text{NP}[\textsc{store}\ \{\boxed{2}\}] \right\rangle \\ \textsc{quants} & \langle \boxed{1}, \boxed{2} \rangle \end{bmatrix}$$

b.
$$\begin{bmatrix} \textsc{form} & \langle \textit{n'aime} \rangle \\ \textsc{arg-st} & \left\langle \text{NP}[\textsc{store}\ \{\boxed{1}\}], \text{NP}[\textsc{store}\ \{\boxed{2}\}] \right\rangle \\ \textsc{quants} & \langle \boxed{1} \rangle \end{bmatrix}$$

In (55ba), the two quantifiers are retrieved, inducing double negation (¬∃x¬∃y[Love(x,y)]) while in (55bb), the two have a resumptive interpretation in which the two are merged into one (¬∃x∃y[Love(x,y)]).[18] This analysis, coupled with the complement treatment of *pas* as a lexically stored quantifier, can account for why *pas* does not induce a resumptive interpretation with a quantifier (from De Swart & Sag (2002)):

(56) Il ne va pas nulle part, il va á son travail.
'He does not go nowhere, he goes to work.'

In this standard French example, De Swart & Sag (2002), accepting the analysis of **kim:00** as *pas* as a complement, specifies the adverbial complement *pas* to be included the negative quantifier in the QUANTS value. This means there would be no resumptive reading for standard French, inducing double negation as in (57):[19]

(57)
$$\begin{bmatrix} \textsc{form} & \langle \textit{ne va} \rangle \\ \textsc{arg-st} & \left\langle \textit{Adv}_{\text{I}}[\textsc{store}\ \{\boxed{1}\}], \text{NP}[\textsc{store}\ \{\boxed{2}\}] \right\rangle \\ \textsc{quants} & \langle \boxed{1}, \boxed{2} \rangle \end{bmatrix}$$

Przepiórkowski & Kupść (1999) and Borsley & Jones (2000) also investigate negative concord in Polish and Welsh and offer HPSG analyses. Consider a Welsh example from Borsley & Jones (2000):

(58) Nid oes neb yn yr ystafell
NOT is no.one in the room
'There is no one in the room.'

Borsley & Jones (2000), identifying n-words with the feature NC (negative concord), takes the verb *nid oes* to bear the positive NEG value, and specifies the subject *neb yn* to carry the positive NC (negative concord) feature. This selectional

[18] See De Swart & Sag (2002) for detailed formulation of the retrieval condition of stored value.
[19] See De Swart & Sag (2002) for cases where *pas* induces negative concord.

approach, interacting with well-defined features, tries to capture how more than one negative elements are corresponding to a single semantic negation.[20]

7 Conclusion

One of the most attractive consequences of the derivational perspective has been that one uniform category, given other syntactic operations and constraints, explains the derivational properties of all types of negation in natural languages, and further can provide a surprisingly close and parallel structure among languages, whether typologically related or not. However, this line of thinking, first of all, runs the risk of missing the peculiar properties of each type of negation. Each individual language has its own way of expressing negation, and further has its own restrictions in the surface realizations of negation which can hardly be reduced to one uniform category.

In the nonderivational HPSG analyses for the four main types of sentential negation that we have reviewed in this chapter, there is no uniform syntactic element, though a certain universal aspect of negation does exist, viz. its semantic contribution. Languages appear to employ various possible ways of negating a clause or sentence. Negation can be realized as different morphological and syntactic categories. By admitting morphological and syntactic categories, we have been able to capture their idiosyncratic properties in a simple and natural manner. Further this theory has been built upon the lexical integrity principle, the thesis that the principles that govern the composition of morphological constituents are fundamentally different from the principles that govern sentence structures. The obvious advantage of this perspective is that it can capture the distinct properties of morphological and syntactic negation, and also of their distribution, in a much more complete and satisfactory way.

Acknowledgements

I thank reviewers of this chapter for comments and suggestions, which helped improve the quality of this chapter. I also thank Bob Borsley and Stefan Müller for detailed comments on the earlier version. My thanks also go to Rok Sim and Jungsoo Kim for helpful feedback.

[20]See Borsley & Jones (2000) for detailed discussion.

References

Abeillé, Anne & Danièle Godard. 1994. The syntax of French auxiliaries. Ms. UFRL and CNRS, Université Paris 7.

Abeillé, Anne & Danièle Godard. 1997. The syntax of French negative adverbs. In M.-L. Rivero (ed.), *Negation and polarity: Syntax and semantics*, 1–27. Amsterdam: John Benjamins.

Baker, Curtis L. 1989. *English syntax*. Cambridge, MA: The MIT Press.

Baker, Curtis L. 1991. The syntax of English *not*: The limits of core grammar. *Linguistic Inquiry* 22. 387–429.

Belletti, Adriana. 1990. *Generalized verb movement*. Torino: Rosenberg & Sellier.

Bender, Emily & Alex Lascarides. 2013a. On modeling scope of inflectional negation. In Elisabeth Norcliffe & Philip Hofmeister (eds.), *The core and the periphery: data-driven perspectives on syntax inspired by ivan a. sag*, 101–124. Stanford: CSLI Publications.

Bender, Emily & Alex Lascarides. 2013b. On modelling scope of inflectional negation. In Philp Hofmeister & Elisabeth Norcliffe (eds.), chap. On modelling scope of inflectional negation, 101–124. Stanford, CA: CSLI Publication.

Bonami, Olivier & Gert Webelhuth. 2013. The phrase-structural diversity of periphrasis: a lexicalist account. In Marina Chumakina & Greville G. Corbett (eds.), *Periphrasis : the role of syntax and morphology in paradigms*, 141–167. Oxford University Press.

Borsley, Robert. 2006. A linear approach to negative prominence. In Stefan Müller (ed.), *Proceedings of the hpsg06 conference*, 60–80. Stanford, CA: CSLI Publication.

Borsley, Robert & Bob Jones. 2000. The syntax of welsh negation. *Transactions of the Philological Society* 98(1). 15–47.

Borsley, Robert & Bob Jones. 2005. *Welsh negation and grammatical theory*. University of Wales Press.

Borsley, Robert & Mohamed Krer. 2012. An hpsg approach to negation in libyan arabic. *Essex Research Reports in Linguistics* 61(2). 1–24.

Bratt, Elizabeth Owen. 1996. *Argument composition and the lexicon: lexical and periphrastic causatives in korean*. Stanford University, MA PhD Dissertation.

Bresnan, Joan & Sam A. Mchombo. 1991. The lexical integrity principle: evidence from bantu. *Natural Language and Linguistic Theory* 13. 181–254.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In R. Friedin (ed.), *Principles and parameters in comparative grammar*, 417–454. Cambridge, MA: The MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Keyser (eds.), *The view from building*, 1–52. Cambridge, MA: MIT Press. Reprinted in Chomsky 1995.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: The MIT Press.

Chung, Chan. 1998a. Argument composition and long-distance scrambling in Korean. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 159–220. San Diego: Academic Press.

Chung, Daeho. 1998b. Wh-scope identification via quasi-binding. *Japanese/Korean Linguistics* 7. 293–309.

Crowgey, Berthold. 2010. Discontinuous negation in hausa. In Stefan Müller (ed.), *Proceedings of the hpsg10 conference*, 269–287. Stanford, CA: CSLI Publication.

Crowgey, Joshua. 2012. An a priori typology of sentential negation from an hpsg perspective. In Stefan Müller (ed.), *The proceedings of the 19th international conference on head-driven phrase structure grammar*.

Crowgey, Joshua & Emily Bender. 2011. Analyzing interacting phenomena: word order and negation in basque. In Stefan Müller (ed.), *Proceedings of the hpsg 2011 conference*, 46–59. Stanford, CA: CSLI Publication.

Dahl, Östen. 1979. Typology of sentence negation. *Linguistics* 17. 79–106.

De Swart, Henriette & Ivan Sag. 2002. Negation and negative concord in romance. *Linguistics and Philosophy* 25. 373–417.

Dryer, Matthew S. 2005. Negative morphemes. *The World Atlas of Language Structures*. 454–457.

Ernst, Thomas. 1992. The phrase structure of english negation. *Linguistic Review* 9. 109–144.

Haegeman, Liliane. 1995. *The syntax of negation*. Cambridge University Press, Cambridge.

Haegeman, Liliane. 1997. The syntax of n-words and the neg criterion. In F. Martineau D. Forget P. Hirschbuhler & M.-L.Rivero (eds.), *Negation and polarity: syntax and semantics*, 115–137. John Benjamins, Amsterdam.

Hagstrom, Paul. 2002. Implications of child errors for the syntax of negation in korean. *Journal of East Asian Linguistics* 11(3). 211–242.

Han, Chung-hye, Lidz Jeffrey & Julien Musolino. 2007. V-raising and grammar competition in korean: evidence from negation and quantifier scope. *Linguistic Inquiry* 38(1). 1–47.

Hilpert, Martin. 2016. Change in modal meanings. *Constructions and Frames* 8(1). 66–85.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar.* The MIT Press.

Kathol, Andreas. 2000. *Linear syntax.* New York, Oxford: Oxford University Press.

Kato, Yasuhiko. 1997. Review of the syntax of negation by l.haegeman. *Language* 73. 391–394.

Kato, Yasuhiko. 2000. Interpretive asymmetries of negation. In L. R. Horn & Y. Kato (eds.), *Negation and polarity: syntactic and semantic perspective*, 62–87. Oxford: Oxford University Press.

Kelepir, Meltem. 2001. *Topics in turkish syntax: clausal structure and scope.* MIT dissertation.

Kim, Jong-Bok. 2000. *The grammar of negation: a constraint-based grammar.* Stanford: CSLI Publications.

Kim, Jong-Bok. 2016. *The syntactic structures of korean: a construction based perspective.* Cambridge: Cambridge University Press.

Kim, Jong-Bok & Ivan Sag. 1996. French and english negation: a lexicalist alternative to head movement.

Kim, Jong-Bok & Ivan A. Sag. 2002. Negation without head movement. *Natural Language and Linguistic Theory* 20(2). 339–412.

Kim, Jong-Bok & Peter Sells. 2008. *English syntax: anintroduction.* Stanford: CSLI Publications.

Klima, Edward S. 1964. Negation in english. In J. A. Fodor & J. J. Katz (eds.), *The structure of language: readings in the philosophy of language*, 246–323. Englewood Cliffs, NJ: Prentice-Hall.

Kupść, Anna & Adam Przepiórkowski. 2002. Morphological aspects of verbal negation in polish. In Peter Kosta & Jens Frasek (eds.), *Current approaches to formal slavic linguistics*, 337–346. Frankfurt: Peter Lang.

Lasnik, Howard. 1995. Verbal morphology: syntactic structures meets the minimalist program. In H. Campos & P. Kempchinsky (eds.), *Evolution and revolution in linguistic theory: essays in honor of charlos otero*, 251–275.

Miller, Philip H. 1992. *Clitics and Constituents in Phrase Structure Grammar.* New York. Published version of 1991 Ph. D. thesis, University of Utrecht, Utrecht.

Mitchell, Erika. 1991. Evidence from finnish for pollock's theory of ip. *Linguistic Inquiry* 22. 373–379.

Monachesi, Paola. 1993. Object clitics and clitic climbing in italian hpsg grammar. In *Proceedings of the sixth conference on european chapter of the association for computational linguistics*, 437–442. Association for Computational Linguistics.

Monachesi, Paola. 1998. Italian restructuring verbs: a lexical analysis. complex predicates in nonderivational syntax. *Syntax and Semantics* 30. 313–368.

Müller, Stefan. 2016. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Newmyer, Fredrick. 2006. Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn. In Betty Birner & Gregory Ward (eds.), chap. Negation and Modularity, 247–268. John Benjamins.

Payne, John R. 1985. Negation. In Timothy Shopen (ed.), *Language typology and syntactic description 1: clause structure*, 197–242. Cambridge: Cambridge University Press.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of ip. *Linguistic Inquiry* 20. 365–424.

Pollock, Jean-Yves. 1994. Checking theory and bare verbs studies in honor of richard s. kayne. In Gulielmo et al Cinque (ed.), *Paths towards universal grammar*, 293–310. Georgetown University Press.

Przepiórkowski, Adam. 2000. Long distance genitive of negation in polish. *Journal of Slavic Linguistics* 8. 151–189.

Przepiórkowski, Adam. 2001. Optional and multiple long distance genitive of negation in polish. In Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn & Luka Szucsich (eds.), *Current issues in formal slavic linguistics*, 337–346. Frankfurt: Peter Lang.

Przepiórkowski, Adam & Anna Kupść. 1999. Eventuality negation and negative concord in polish and italian. In Robert Borsley & Adam Przepiórkowski (eds.), *Slavic in head-driven phrase structure grammar*, 211–246. Stanford, CA: CSLI Publication.

Sag, Ivan A. 2012. Sign-based construction grammar: an informal synopsis. *Sign-based construction grammar* 193. 69–202.

Vasishth, Shravan. 2000. Word order, negation and negative polarity in hindi. *OSU Working Papers in Linguistics* 53. 108–131.

Vikner, Sten. 1997. V^0 to I^0 movement and inflection for person in all tenses. In Liliane Haegeman (ed.), *The new comparative syntax*, 189–213. London: Longman.

Warner, Anthony. 2000. English auxiliaries without lexical rules. In Robert D. Borsley (ed.), *Syntax and semantics*, vol. 32, 167–220. San Diego: Academic Press. DOI:10.1163/9781849500098

Zanuttini, Raffaella. 1991. *Syntactic properties of sentential negation: a comparative studies of romance languages*. University of Pennsylvania PhD Dissertation.

Zanuttini, Raffaella. 1997. Negation and verb movement. In L. Hageman (ed.), *The new comparative syntax*, 214–245. London & New York: Longman.

Zanuttini, Raffaella. 2001. *Negation and clausal structure*. Oxford: Oxford University Press.

Zeijlstra, Hedde. 2007. On french negation. *Berkeley Linguistic Society* 35. 447–458.

Zeijlstra, Hedde. 2015. The morpho-syntactic realization of sentential negation. In A. Alexiadou & T. Kiss (eds.), *The international syntax handbook*, 275–310. Berlin: Mouton de Gruyter.

Chapter 20

Ellipsis

Joanna Nykiel

Kyung Hee University, Seoul

Jong-Bok Kim

Kyung Hee University, Seoul

This chapter provides an overview of HPSG analyses of ellipsis. The structure of the chapter follows three types of ellipsis, nonsentential utterances, predicate ellipsis, and nonconstituent coordination, with three types of analyses applied to them. The analyses characteristically don't admit silent syntactic material for any ellipsis phenomena with the exception of certain types of nonconstituent coordination.

1 Introduction

Ellipsis is a phenomenon that involves a noncanonical mapping between syntax and semantics. What appears to be a syntactically incomplete utterance still receives a semantically complete representation, based on the features of the surrounding context, be the context linguistic or nonlinguistic. The goal of syntactic theory is thus to account for how the complete semantics can be reconciled with the apparently incomplete syntax. One of the key questions here relates to the structure of the ellipsis site, that is, whether or not we should assume the presence of invisible syntactic material. This chapter begins by introducing three types of ellipsis (nonsentential utterances, predicate ellipsis, and nonconstituent coordination) that have attracted considerable attention and received treatment within HPSG. We next overview existing evidence for and against the so-called WYSIWYG ('What You See Is What You Get') approach to ellipsis, where no invisible material is posited at the ellipsis site. Finally, we walk the reader through three types of HPSG analyses applied to the three types of ellipsis presented in section 2.

Joanna Nykiel & Jong-Bok Kim. 2019. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 585–612. Berlin: Language Science Press. DOI:??

2 Three types of ellipsis

Depending on the type of analysis by means of which HPSG handles them, elliptical phenomena can be broadly divided into three types: nonsentential utterances, predicate ellipsis, and nonconstituent coordination. We overview the key features of these types here before discussing in greater detail how they have been brought to bear on the question of whether there is invisible syntactic structure at the ellipsis site or not. We begin with stranded XPs, which HPSG treats as nonsentential utterances, and then move on to predicate and argument ellipsis, followed by phenomena known as nonconstituent coordination.

2.1 Nonsentential utterances

This section introduces utterances smaller than a sentence, which we refer to as nonsentential utterances (NUs). These range from Bare Argument Ellipsis (BAE) (1), including fragment answers (2), to direct or embedded sluicing (3)-(4). Sluicing hosts stranded wh-phrases and has the function of an interrogative clause, while BAE hosts XPs representing various syntactic categories and typically has the function of a declarative clause.[1]

(1) A: You were angry with them.
B: Yeah, angry with them and angry with the situation.

(2) A: Where are we?
B: In Central Park.

(3) A: So what did you think about that?
B: About what?

(4) A: There's someone at the door.
B: Who?/I wonder who.

The theoretical question NUs raise is whether they are parts of larger sentential structures or rather nonsentential structures whose semantic and morphosyntactic features are licensed by the surrounding context. To adjudicate between these views researchers have looked for evidence that NUs in fact behave as if they were fragments of sentences. As we will see in section 3, there is evidence to support both of these views. However, HPSG doesn't assume that NUs are underlyingly sentential structures.

[1]Several subtypes of nonsentential utterances can be distinguished, based on their contextual functions, which we don't discuss here (for a recent taxonomy, see Ginzburg 2011).

2.2 Predicate ellipsis and argument ellipsis

The section looks at three constructions whose syntax includes null, hence non-canonical, elements. They are Verb Phrase Ellipsis (VPE), Null Complement Anaphora (NCA), and argument drop (or pro drop). VPE features stranded auxiliary verbs (5) and NCA is characterized by omission of complements to some lexical verbs (6). Argument drop refers to omission of a pronominal subject or an object argument, as illustrated in (7) for Polish.

(5) A: I didn't ask George to invite you.
B: Then who did?

(6) Some mornings you can't get hot water in the shower, but nobody complains.

(7)

Pia	późno	wróciła	do domu.	Od razu	poszła	spać.
PIA	LATE	GOT	TO HOME	RIGHT AWAY	WENT	SLEEP

'Anna got home late. She went straight to bed.'

Here the question is whether these null elements should be assumed to be underlyingly present in the syntax of theses constructions, and the answer is no.

2.3 Nonconstituent coordination

We focus on two instances of nonconstituent coordination, Right Node Raising (RNR) and gapping (Ross 1967), illustrated in (8) and (9) respectively. In RNR, a single constituent located in the right-peripheral position is associated with both conjuncts. In gapping, a finite verb is associated with both (or more) conjuncts but only present in the leftmost one. This results in what appears to be coordination of standard constituents and elements not normally defined as constituents (a stranded transitive verb in (8) and a cluster of subject NP and object NP in (9)).

(8) Ethan sold and Rasmus gave away all his CDs.

(9) Ethan gave away his CDs and Rasmus his old guitar.

To handle such constructions the grammar must either be permitted to coordinate noncanonical constituents, to generate coordinated constituents parts of which can undergo deletion, or to coordinate nonsentential utterances. As we will see, HPSG makes use of the latter two options.

3 Evidence for and against invisible material at the ellipsis site

This section is concerned with NUs and VPE since this is where the contentious issues arise of where ellipsis is licensed (sections 3.3. and 3.4) and whether there is invisible syntactic material in an ellipsis site (sections 3.1 and 3.2). Below we consider evidence for and against invisible structure found in the ellipsis literature; the evidence is based not only on intuitive judgments, but also experimental and corpus data.

3.1 Connectivity effects

Connectivity effects refer to parallels between NUs and their counterparts in sentential structures, thus speaking in favor of the existence of silent sentential structure. We focus on two kinds here: case-matching effects and preposition-stranding effects (for other examples of connectivity effects, see Ginzburg & Miller 2018). It's been known since Ross (1969) that NUs exhibit case-matching effects, that is, they are typically marked for the same case that is marked on their counterparts in sentential structures. (10) illustrates this for German, where case matching is seen between a wh-phrase functioning as an NU and its counterpart in the antecedent.

(10) Er will jemandem schmeicheln, aber sie wissen nicht
he will someone.DAT flatter, but they know not
wem/*wen.
who.DAT/*who.ACC
'He wants to flatter someone, but they don't know whom.'

Case-matching effects are crosslinguistically robust in that they are found in the great majority of languages with overt case marking systems, and therefore, they have been taken as strong evidence for the reality of silent structure. The argument is that the pattern of case matching follows straightforwardly from the silent structure that embeds an NU and matches the structure of the antecedent clause. However, a language like Hungarian poses a problem for this reasoning (see Jacobson 2016). Hungarian has verbs that assign one of two cases to their object NPs with no meaning difference, but case matching is still required between an NU and its counterpart, whichever case is marked on the counterpart. To see this, consider (11) from Jacobson (2016: 356). The verb *hasonlit* assigns either sublative (SUBL) or allative (ALL) case to its object, but if SUBL is selected for an NU's counterpart, the NU must match this case.

(11) A: Ki-re hasonlit Péter?
WHO.SUBL RESEMBLES PETER
A: 'Who does Peter resemble?'
B: János-ra/*János-hoz.
JÁNOS.SUBL/*JÁNOS.ALL
B: 'János.'

Jacobson (2016) notes that there is some speaker variation regarding the (un)acceptability of case mismatch here at the same time that all speakers agree that either case is fine in a corresponding nonelliptical response to (11A). This last point is important, because it shows that the requirement of—or at least a preference for—matching case features applies to NUs to a greater extent than it does to their nonelliptical equivalents, challenging connectivity effects.

Similarly problematic for case-based parallels between NUs and their sentential counterparts are some Korean data. Korean NUs can drop case markers more freely than their counterparts in nonelliptical clauses can, a point made in Morgan (1989). Observe the following exchanges:

(12) A: Nwukwu-ka ku chaek-ul sa-ass-ni?
WHO.NOM THE BOOK.ACC BUY.PST.QUE
A: 'Who bought the book?'
B: Yongsu-ka/Yongsu/*Yongsu-lul.
YONGSU.NOM/YONGSU/*YONGSU.ACC
B: 'Yongsu.'
B′: Yongsu-ka/*Yongsu ku chaek-ul sa-ass-e
YONGSU.NOM/*YONGSU THE BOOK.ACC BUY.PST.DECL
B′: 'Yongsu bought the book'

When an NU corresponds to a nominative subject in the antecedent (as in 12B), it can be either marked for nominative or caseless. However, replacing the same NU with a full sentential answer, as in (12B′), rules out case drop from the subject. This strongly suggests that the case-marked and caseless NUs couldn't have identical source sentences if they were to derive via PF-deletion.[2] Data like these

[2]Nominative differs in this respect from three other structural cases, dative, accusative and genitive, in that the latter may also be dropped from nonelliptical clauses (see Morgan 1989, Lee 2016, Kim 2016).

led Morgan (1989) to propose that not all NUs have a sentential derivation, an idea later picked up in Barton (1998).

The same pattern is associated with semantic case. That is, in (13), an NU can optionally be marked for comitative like its counterpart in the A-sentence or be caseless. But being caseless is not an option for the NU's counterpart.

(13) A: Nwukwu-wa/*nwukwu hapsek-ul ha-yess-e?
who.COM/*who sitting.together.ACC do.PST.QUE

A: 'With whom did you sit together?'

B: Mimi-wa/*Mimi.
Mimi.SRC/Mimi

B: 'With Mimi/*Mimi.'

The generalization for Korean is then that NUs may be optionally realized as caseless but may never be marked for a different case than is marked on their counterparts.

Overall, case-marking facts show that there is some morphosyntactic identity between NUs and their antecedents, though not to the extent that NUs have exactly the features that they would have if they were constituents embedded in sentential structures. The Hungarian facts also suggest that aspects of the argument structure of the antecedent relating to case licensing are relevant for an analysis of NUs.[3]

The second kind of connectivity effects goes back to Merchant (2001, 2004) and highlights apparent links between wh- and focus movement and the features of NUs. The idea is that prepositions behave the same under wh- and focus movement as they do under clausal ellipsis, that is, they pied-pipe or strand in the same environments. If a language (e.g., English) permits preposition stranding under wh- and focus movement (*With what did Harvey paint the wall?* vs *What did Harvey paint the wall with?*), then NUs may surface with or without prepositions, as illustrated in (14) for sluicing and BAE.

(14) A: I know what Harvey painted the wall with.
B: (With) what?/(With) primer.

If there indeed was a link between between preposition stranding and NUs, then we would not expect prepositionless NUs in languages without preposition stranding. This expectation is disconfirmed by an ever-growing list of nonpreposition-stranding languages that do feature prepositionless NUs: Brazilian Portuguese

[3]Hungarian and Korean are in fact not the only problematic languages; for a list, see Vicente (2015).

(Almeida & Yoshida 2007), Spanish and French (Rodrigues et al. 2009), Greek (Molimpakis 2018), Bahasa Indonesia (Fortin 2007), Emirati Arabic (Leung 2014), Russian (Philippova 2014), Polish (Szczegielniak 2008, Nykiel 2013, Sag & Nykiel 2011), Czech (Caha 2011), Bulgarian (Abels 2017), and Serbo-Croatian (Stjepanović 2008, 2012). A few of these studies have presented experimental evidence that prepositionless NUs are acceptable, though — for reasons still poorly understood — they typically don't reach the same level of acceptability as their variants with prepositions do (see Nykiel 2013 for Polish and Molimpakis 2018 for Greek).

It is evident from this research that there is no grammatical constraint on NUs that keeps track of what preposition-stranding possibilities exist in any given language. On the other hand, it doesn't seem sufficient to assume that NUs can freely drop prepositions, given examples of sprouting like (15), in which prepositions are not omissible (see Chung 2006, 2013 on the inomissibility of prepositions under sprouting). The difference between (14) and (15) is that there is an explicit phrase the NU corresponds to (in the HPSG literature this phrase is termed a Salient Utterance (Ginzburg & Sag 2000) or a Focus-Establishing Constituent (Ginzburg 2012)) in the former but not in the latter.

(15) A: I know Harvey painted the wall.
B: *(With) what?/Yeah, *(with) primer.

This issue has not received much attention in the HPSG literature, though see Kim (2015).

3.2 Island effects

One of the predictions of the view that NUs are underlyingly sentential is that they should respect island constraints on long-distance movement. But as illustrated below, NUs (both sluicing and BAE) exhibit island-violating behavior.[4]

(16) A: Harriet drinks scotch that comes from a very special part of Scotland.
B: Where? (Culicover & Jackendoff 2005: 245)

(17) A: The administration has issued a statement that it is willing to meet with one of the student groups.
B: Yeah, right—the Gay Rifle Club. (Culicover & Jackendoff 2005: 245)

[4]Merchant (2004) argued that BAE, unlike sluicing, does respect island constraints, an argument that was later challenged (see e.g, Culicover & Jackendoff 2005, Griffiths & Lipták 2014). However, Merchant (2004) focused specifically on pairs of wh-questions and answers like (2) and ran into the difficulty of testing for island-violating behavior, since a well-formed antecedent couldn't be constructed.

Among Culicover & Jackendoff's (2005: 245) examples of well-formed island-violating NUs are also sprouted NUs (those that correspond to implicit Salient Utterances) like (18)-(19).

(18) A: John met a woman who speaks French.
B: With an English accent?

(19) A: For John to flirt with at the party would be scandalous.
B: Even with his wife?

Other research assumes that sprouted NUs are one of the two kinds of NUs that respect island constraints, the other kind being contrastive NUs, illustrated in (20) (Chung, Ladusaw & McCloskey 1995, Merchant 2001, Griffiths & Lipták 2014).

(20) A: Does Abby speak the same Balkan language that Ben speaks?
B: *No, Charlie. (Merchant 2001)

Schmeh et al. (2015) explore the acceptability of NUs like (20) compared to NUs introduced by the particle *Yes* depicted in (21). (20) differs from (21) in terms of discourse function in that it corrects rather than supplement the antecedent, which is signaled by a different response particle.

(21) A: John met a guy who speaks a very unusual language.
B: Yes, Albanian. (Culicover & Jackendoff 2005: 245)

Schmeh et al. (2015) find that corrections lower acceptability ratings compared to supplementations and propose that this follows from the fact that corrections induce greater processing difficulty than supplementations do, and hence the acceptability difference between (20) and (21). This finding makes it plausible that the perceived degradation of island-violating NUs could ultimately be attributed to nonsyntactic factors, e.g., the difficulty of successfully computing a meaning for them.

In contrast to NUs, many instances of VPE appear to obey island constraints, as would be expected if there was unpronounced structure from which material was extracted. An example is depicted in (22) (note that the corresponding sluicing NU is fine).

(22) *They want to hire someone who speaks a Balkan language, but I don't remember which they do [~~want to hire someone who speaks t~~].
(Merchant 2001: 6)

(22) contrasts with well-formed examples like (23) and (24), from Ginzburg & Miller (2018).

(23) He managed to find someone who speaks a Romance language, but a Germanic language, he didn't [~~manage to find someone who speaks t~~].

(24) He was able to find a bakery where they make good baguette, but croissants, he couldn't [~~find a bakery where they make good t~~].

As Ginzburg & Miller (2018) rightly point out, we don't yet have a complete understanding of when or why island effects show up in VPE. Its behavior is at best inconsistent, failing to provide convincing evidence for silent structure.

3.3 Structural mismatches

Because structural mismatches are not permitted by NUs (see Merchant 2005, 2013),[5] this section focuses on VPE and developments surrounding the question of which contexts license it. In a seminal study of anaphora, Hankamer and Sag (1976) classified VPE as a surface anaphor with syntactic features closely matching those of an antecedent present in the linguistic context. They argued in particular that VPE is not licensed if it mismatches its antecedent in voice. Compare (25) and (26) from Sag & Hankamer (1984: 327).

(25) * The children asked to be squirted with the hose, so we did.

(26) The children asked to be squirted with the hose, so they were.

[5] Ginzburg & Miller (2018) cite examples—originally from Beecher (2008)—of sprouting NUs with nominal, hence mismatched, antecedents, e.g., (i).

(i) We're on to the semi-finals, though I don't know who against.

Somewhat similar examples, where NUs appear to take APs as antecedents, appear in COCA:

(ii) A: Well, it's a defense mechanism. B: Defense against what?

(iii) Our Book of Mormon talks about the day of the Lamanite, when the church would make a special effort to build and reclaim a fallen people. And some people will say, Well, fallen from what?

The NUs in (ii)–(iii) repeat the lexical heads whose complements are being sprouted (*defense* and *fallen*), that is, they contain more material than is usual for NUs (cf. (i)). It seems that without this additional material it would be difficult to integrate the NUs into the propositions provided by the antecedents and hence to arrive at the intended interpretations.

This proposal places tighter structural constraints on VPE than on other verbal anaphors (e.g., *do it/that*) in terms of identity between an ellipsis site and its antecedent and has prompted extensive evaluation in a number of corpus and experimental studies in the decades following Hankamer & Sag (1976). Below are examples of acceptable structural mismatches reported in the literature, ranging from voice mismatch (27) to nominal antecedents (28) to split antecedents (29).

(27) This information could have been released by Gorbachev, but he chose not to [release it]. (Hardt 1993: 37)

(28) Mubarak's survival is impossible to predict and, even if he does [survive], his plan to make his son his heir apparent is now in serious jeopardy. (Miller & Hemforth 2014:)

(29) Wendy is eager to sail around the world and Bruce is eager to climb Mt. Kilimanjaro, but neither of them can [do the things they want], because money is too tight. (Webber 1978:)

There are two opposing views that have emerged from the empirical work. The first view takes mismatches to be grammatical and connects degradation in acceptability to violation of certain independent discourse (Kehler 2002, Miller 2011, Kertz 2013, Miller & Hemforth 2014, Miller & Pullum 2014) or processing constraints (Kim et al. 2011). Two types of VPE have been identified on this view through corpus work—auxiliary choice VPE and subject choice VPE—each with different discourse requirements with respect to the antecedent (Miller 2011, Miller & Hemforth 2014, Miller & Pullum 2014). The second view assumes that there is a grammatical ban on structural mismatch but violations thereof may be repaired under certain conditions; repairs are associated with differential processing costs compared to matching ellipses and antecedents (Arregui et al. 2006, Grant et al. 2012). If we follow the first view, it is perhaps unexpected that voice mismatch should consistently incur a greater penalty under VPE than when no ellipsis is involved, as recently reported in Kim & Runner (2017). Kim & Runner (2017) stop short of drawing firm conclusions regarding the grammaticality of structural mismatches, but one possibility is that the observed mismatch effects reflect a construction-specific constraint on VPE. HPSG analyses take structurally mismatched instances of VPE to be unproblematic and fully grammatical, while also recognizing construction-specific constraints: discourse or processing constraints formulated for VPE may or may not extend to other elliptical constructions, such as NUs (see Ginzburg & Miller 2018 for this point).

3.4 Nonlinguistic antecedents

Like structural mismatches, the availability of nonlinguistic antecedents for an ellipsis points to the fact that it needn't be interpreted by reference to and licensed by a structurally identical antecedent. Although this option is somewhat limited, VPE does tolerate nonlinguistic antecedents, as shown in (30)–(31) (see also Hankamer & Sag 1976, Schachter 1977).

(30) Mabel shoved a plate into Tate's hands before heading for the sisters' favorite table in the shop. "You shouldn't have." She meant it. The sisters had to pool their limited resources just to get by. (Pullum & Miller 2014, ex. 23)

(31) Once in my room, I took the pills out. "Should I?" I asked myself. (Pullum & Miller 2014, ex. 22a)

Pullum & Miller (2014) provide an extensive critique of the earlier work on the ability of VPE to take nonlinguistic antecedents, arguing for a streamlined discourse-based explanation that neatly captures the attested examples as well as examples of structural mismatch like those discussed in section 3.3. The important point here is again that VPE is subject to construction-specific constraints which limit its use with nonlinguistic antecedents.

NUs appear in various nonlinguistic contexts as well. Ginzburg & Miller (2018) distinguish three classes of such NUs: sluices (32), exclamative sluices (33), and declarative fragments (34).

(32) (In an elevator) What floor? (Ginzburg & Sag 2000: 298)

(33) It makes people "easy to control and easy to handle," he said, "but, God forbid, at what a cost!" (Ginzburg & Miller To appear, ex. 34a)

(34) BOBADILLA turns, gestures to one of the other men, who comes forward and gives him a roll of parchment, bearing the royal seal. "My letters of appointment." (COCA)

In addition to being problematic from the licensing point of view, NUs like these have been put forward as evidence against the idea that they are underlyingly sentential, because it is unclear what the structure that underlies them would be (see Ginzburg & Sag 2000, Culicover & Jackendoff 2005, Stainton 2006).[6]

[6] This is not to say that a sentential analysis of fragments without linguistic antecedents hasn't been attempted. For details of a proposal involving a 'limited ellipsis' strategy, see Merchant (2004, 2010).

4 Analyses of NUs

It is worth noting at the outset that the analyses of NUs within the framework of HPSG are based on an elaborate theory of dialog (Ginzburg 1994, Ginzburg & Cooper 2004, 2014, Larsson 2002, Purver 2006, Fernández 2006, Fernández & Ginzburg 2002, Fernández et al. 2007, Ginzburg & Fernández 2010, Ginzburg et al. 2014, Ginzburg 2012, 2013) and on a wider range of data than is common practice in the ellipsis literature. Existing analyses of NUs go back to Ginzburg & Sag (2000), who recognize declarative fragments (35) and two kinds of sluicing NUs, direct sluices (36) and reprise sluices (37) (the relevant fragments are bolded). The difference between direct and reprise sluices lies in the fact that the latter are requests for clarification of any part of the antecedent. For instance, in (37) the referent of *that* is unclear to the interlocutor.

(35) "I was wrong." Her brown eyes twinkled. "Wrong about what?" "**That night**." (COCA)

(36) "You're waiting," she said softly. "**For what?**" (COCA)

(37) "Can we please not say a lot about that?" "**About what?**" (COCA)

Ginzburg and Sag (2000: 304) make use of the constraint shown in (38) (we have added information about the MAX-QUD) to generate NUs.

(38)

Head-Fragment Construction

$$\begin{bmatrix} \text{CAT} & S \\ \text{CTXT} & \begin{bmatrix} \text{MAX-QUD} & \lambda\{\pi^{i}\} \\ \text{SAL-UTT} & \left\{\begin{bmatrix}\text{CAT} & \boxed{2} \\ \text{CONT} & [\text{IND} \;\; i]\end{bmatrix}\right\}\end{bmatrix}\end{bmatrix} \Rightarrow \begin{bmatrix}\text{CAT} & \boxed{2} \\ \text{CONT} & [\text{IND} \;\; i]\end{bmatrix}$$

Let us see how this constructional constraint allows us to license NUs and capture their properties, including the connectivity effects we discussed in section 3.1. Note first that any phrasal category can function as an NU, that is, can be mapped onto a sentential utterance as long as it corresponds to a Salient Utterance (SAL-UTT). This means that the head daughter's syntactic category must match that of a SAL-UTT, which is an attribute supplied by the surrounding context as a (sub)utterance of another contextual attribute—the Maximal Question under Discussion (MAX-QUD). The two contextual attributes SAL-UTT and MAX-QUD are introduced specifically for the purpose of analyzing NUs. The context gets updated with every new question-under-discussion, and MAX-QUD represents the most recent question-under-discussion, while SAL-UTT is the (sub)utterance with the widest scope within MAX-QUD. To put it informally, SAL-UTT represents a (sub)utterance of a MAX-QUD that has not been resolved yet. Its feature CAT supplies information relevant for establishing morphosyntactic identity with an NU, that is, syntactic category and case information, and (38) requires that an NU match this information. Meanwhile, MAX-QUD provides the propositional semantics for an NU and is, typically, a unary question. The content of MAX-QUD can be supplied by linguistic or nonlinguistic context. In the prototypical case, MAX-QUD arises from the most recent wh-question uttered in a given context (39), but can also arise (via accommodation) from other forms found in the context, such as constituents bearing focal accent (40) and constituents in need of clarification (41), or from a nonlinguistic context (42).[7]

(39) A: What did Barry break?
B: The mike.

(40) A: Barry broke the MIKE.
B: Yes, the only one we had.

(41) A: Barry broke the mike.
B: Who?

(42) (Cab driver to passenger on the way to airport) A: Which airline?

The existing analyses of NUs (Ginzburg 2012, Sag & Nykiel 2011, Kim 2015, Abeillé et al. 2014, Abeillé & Hassamal 2017) are based on Ginzburg and Sag's (2000) constraint. Below we illustrate how it is applied to the declarative fragment in (39).

[7] Ginzburg (2012) uses the notion of the Dialog Game Board (DGB) to keep track of all information relating to the common ground between interlocutors. The DGB is also the locus of contextual updates arising from each new question-under-discussion that is introduced.

(43)

$$\begin{array}{c}
\text{S} \\
\begin{bmatrix}
\text{CAT} & \begin{bmatrix}\text{HEAD v}\end{bmatrix} \\
\text{CTXT} & \begin{bmatrix}\text{MAX-QUD} & \lambda\{\pi^{i}\}[\textit{BREAK}(B,I)] \\ \text{SAL-UTT} & \left\{\begin{bmatrix}\text{CAT} & \boxed{2} \\ \text{CONT} & [\text{IND } i]\end{bmatrix}\right\}\end{bmatrix}
\end{bmatrix} \\
| \\
\text{NP} \\
\begin{bmatrix}\text{CAT} & \boxed{2} \\ \text{CONT} & [\text{IND } i]\end{bmatrix} \\
| \\
\text{The mike}
\end{array}$$

This construction-based analysis, in which dialogue updating plays a key role in the licensing of NUs, can also offer a simple account of sprouting examples like (36).[8] As discussed in Kim (2015), we could take an unrealized oblique argument of the verb *wait* as an instance of indefinite null instantiation (INI) (see Ruppenhofer and Michaelis 2014):

(44) Lexical entry for *wait*:

$$\begin{bmatrix}
\text{FORM} & \langle \textit{WAIT} \rangle \\
\text{ARG-ST} & \left\langle \text{NP}_{I}, \text{PP}_{X} \right\rangle \\
\text{SYN} & \begin{bmatrix}\text{SUBJ} & \langle \text{NP}[\textit{OVERT}] \rangle \\ \text{COMPS} & \langle \text{PP}[\textit{INI}] \rangle\end{bmatrix} \\
\text{SEM} & \textit{WAIT}(I, X)
\end{bmatrix}$$

The lexical information specifies that the second argument of *wait* can be an unrealized indefinite NP while the first argument needs to be an overt one. Now consider the dialogue in (36). Uttering the sentence *You're waiting?* would then update the DGB with a SAL-UTT represented by the unrealized NP:

[8] Ginzburg and Sag (2000: 330) suggest a way of analyzing sprouted NUs, such as (i). The implied direct object of *eat* functioning as SAL-UTT here would appear as a noncanonical synsem on the verb's ARG-ST list, but not on the COMPS list, and thereby be able to provide appropriate morphosyntactic identity information.

(i) A: And what did you do then?
B: I ate.
A: What?

(45)
$$\left[\text{DGB}\left[\text{SAT-UTT}\left[\begin{array}{l}\text{SYN PP}\left[\begin{array}{l}\textit{INI}\\ \text{PFORM } \textit{FOR}\\ \text{IND } x\end{array}\right]\\ \text{SEM } \textit{WAIT.FOR}(i,x)\end{array}\right]\right]\right]$$

The fragment *for what?*, matching this SAL-UTT, projects a well-formed NU in accordance with the Head-Fragment Construction.[9]

The advantages of the nonsentential analyses sketched here follow from their ability to capture limited morphosyntactic parallelism between NUs and SAL-UTT without having to account for why NUs don't behave like constituents of sentential structures. The island-violating behavior of NUs is unsurprising on this analysis, as are attested cases of structural mismatch and situationally controlled NUs.[10] However, some loose ends still remain. (38) currently has no means of capturing certain connectivity effects: it can't rule preposition drop out under sprouting and it incorrectly rules out case mismatch in languages like Hungarian for speakers that do accept it (see discussion around example (11)).[11]

5 Analyses of predicate/argument ellipsis

The first issue in the analysis of VPE is the status of an elided VP. It is assumed to be a *pro* element due to its pronominal properties (see Lobeck 1995, Lopez 2000, Kim 2006, Aelbrecht and Harwood 2015, Ginzburg and Miller 2018). For instance, VPE applies only to phrasal categories (46–47), can cross utterance boundaries (48), can override island constraints (49–50), is subject to the Backwards Anaphora Constraint (51–52).

(46) *Mary will meet Bill at Stanford because she didn't _ John.

(47) Mary will meet Bill at Stanford because she didn't _ at Harvard.

(48) A: Tom won't leave Seoul soon.
B: I don't think Mary will _ either.

(49) John didn't hit a home run, but I know a woman who did. (CNPC)

[9] See the detailed analysis of such sprouting examples in Kim (2015).

[10] The rarity of NUs with nonlinguistic antecedents can be understood as a function of how easily a situational context can give rise to a MAX-QUD and thus license ellipsis (see Miller & Pullum 2014 for this point with regard to VPE).

[11] See, however, Kim (2015) for proposals envoking a case hierarchy specific to Korean to explain case mismatch and introducing an additional constraint to block preposition drop under sprouting.

(50) That Betsy won the batting crown is not surprising, but that Peter didn't know she did __ is indeed surprising. (SSC)

(51) *Sue didn't [e] but John ate meat.

(52) Because Sue didn't [e], John ate meat.

Argument ellipsis we find in languages like Polish and Korean can also be taken to be ellipsis of a pronominal expression, as in (53).

(53) Mimi-ka *pro* po-ass-ta.
Mimi.NOM pro see.PST.DECL
'Mimi saw (him)'

In accounting for *pro*-drop phenomena, we do not need to posit a phonologically empty pronoun if the level of argument structure is available (see Bresnan 1982). We simply encode the required pronominal properties in the argument structure. In the framework of HPSG, we represent this as the following Argument Realization Constraint allowing mismatch between argument-structure and syntactic-valence features:[12]

(54) Argument Realization Constraint (ARC):

$$\textit{v-word} \Rightarrow \begin{bmatrix} \text{SYN|VAL} \begin{bmatrix} \text{SUBJ} & \boxed{\text{A}} \\ \text{COMPS} & \boxed{\text{B}} \ominus \textit{list(pro)} \end{bmatrix} \\ \text{ARG-ST} \ \boxed{\text{A}} \oplus \boxed{\text{B}} \end{bmatrix}$$

The argument realization here tells us that a *pro* element in the argument structure need not be realized in the syntax. For example, as represented in (55), the transitive verb *po-ass-ta* 'see.PST.DECL' takes a *pro* object NP as its argument and thus the *pro* NP is not instantiated as the syntactic complement of the verb.

(55) Lexical entry for *po-ass-ta*

$$\begin{bmatrix} \textit{v-word} \\ \text{FORM} \ \langle \text{PO-ASS-TA} \rangle \\ \text{SYN} \begin{bmatrix} \text{HEAD|VFORM} & \textit{fin} \\ \text{VAL} \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \ \rangle \end{bmatrix} \end{bmatrix} \\ \text{ARG-ST} \ \langle \boxed{1}\text{NP}, \text{NP}[\textit{pro}] \rangle \end{bmatrix}$$

[12] Expressions have two subtypes: overt and covert ones, the latter of which has two subtypes, *pro* and *gap*. See Sag (2012) for details.

Similarly, English VPE is analyzed as a language-particular VP *pro* drop phenomenon normally licensed by auxiliary verbs. This idea can be formalized as in (56):

(56) Aux-Ellipsis Construction:

$$\begin{bmatrix} \textit{aux-ellipsis-wd} \\ \text{ARG-ST } \langle \boxed{1}\text{XP}, \boxed{2}\text{YP}[\textit{pro}] \rangle \end{bmatrix}$$

$$|$$

$$\begin{bmatrix} \textit{aux-v-lxm} \\ \text{ARG-ST } \langle \boxed{1}\text{XP}, \boxed{2}\text{YP} \rangle \end{bmatrix}$$

What this tells us is that an auxiliary verb selecting two arguments can be projected into an elided auxiliary verb whose second argument is realized as a small *pro*. Note that this argument is not mapped onto the syntactic grammatical function COMPS. (56) will also project a structure like the one in Figure 1.

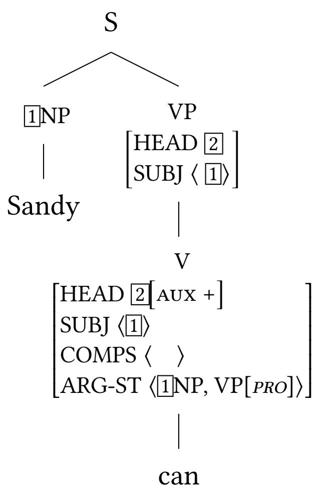


Figure 1: add caption

The head daughter's COMPS list (VP[bse]) is empty because the second element in the ARG-ST is a *pro*.

6 Analyses of nonconstituent coordination and gapping

In this section, we focus on RNR and gapping, whose analyses we address in separate subsections below.

6.1 Right Node Raising

A characteristic property of RNR is that it's the only phenomenon where seemingly incomplete structure has consistently attracted HPSG analyses involving deletion of silent material. All existing analyses of RNR (Abeillé et al. 2016, Beavers & Sag 2004, Chaves 2008, 2014, Crysmann 2003, Yatabe 2001, 2012) agree on this point, although some of this work proposes more than one mechanism for accounting for coordination of nonconstituents (Chaves 2014, Yatabe 2001, 2012, Yatabe & Tam 2018).

The RNR literature engages with the question of what kind of deletion it is that targets shared material, based on the kind of material that may be RNRaised and the range of mismatches permitted between the left and right conjuncts.[13] For instance, Chaves (2014: 839–840) concludes that RNR can't be syntactic deletion because it exhibits various argument-structure mismatches (57–58) and can target material below the word level (59–60).

(57) Sue gave me—but I don't think I will ever read—[a book about relativity].

(58) Never let me—or insist that I—[pick the seats].

(59) We ordered the hard- but they got us the soft-[cover edition].

(60) Your theory under- and my theory over[generates].

Furthermore, RNR can target strings that are not subject to any known syntactic operations, such as rightward movement (Chaves 2014: 865).

(61) I thought it was going to be a good but it ended up being a very bad [reception].

(62) Tonight a group of men, tomorrow night he himself, [would go out there somewhere and wait].

[13] Although we refer to the material on the left and right as conjuncts, it is been known since Hudson (1976, 1984) that RNR extends to other syntactic environments than coordination (see Chaves (2014) for stressing this point).

(63) They were also as liberal or more liberal [than any other age group in the 1986 through 1989 surveys].

RNRaised material can also be discontinuous, as in (64–65) (Chaves 2014: 868; citing Whitman 2009: 238–240).

(64) Please move from the exit rows if you are unwilling or unable [to perform the necessary actions] without injury.

(65) The blast upended and nearly sliced [an armored Chevrolet Suburban] in half.

This evidence leads Chaves (2014) to propose that: (1) only 'true' RNR should be accounted for via the mechanism of surface-based deletion, (2) this deletion is sensitive to morph form identity, and (3) the targets of RNR are any linearized strings, whether constituents or otherwise. Chaves' (2014: 874) constraint licensing RNR is given in 54. It permits the M(orpho)P(honology) feature of the mother to contain only one instance (represented as L_3 in (66)) of the two morphophonologically identical sequences [FORM F_1],…,[FORM F_n] present in the daughters; the leftmost of these sequences undergoes deletion. The final list in the mother, L_4, may be empty or nonempty, depending on whether RNR is discontinuous.

(66) Backward periphery deletion construction

$$\begin{bmatrix} \textit{phrase} \\ \text{MP } L_1{:}\textit{ne-list} \circ L_2{:}\textit{ne-list} \circ L_3 \circ L_4 \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} \textit{phrase} \\ \text{MP } L_1 \circ \langle\, [\text{FORM } F_1],\ldots,[\text{FORM } F_n] \rangle \circ L_2 \circ L_3 : \langle [\text{FORM } F_1],\ldots,[\text{FORM } F_n] \rangle \circ L_4 \end{bmatrix}$$

Chaves' (2014) proposal reflects the idea that nonconstituent coordination is a multi-faceted phenomenon, requiring more than one kind of analysis. Indeed, the HPSG literature includes analyses based on NP-ellipsis or A(cross) T(he) B(oard) extraposition (Chaves 2014) and on phonological vs syntactic deletion (Yatabe 2001, 2012, Yatabe & Tam 2018). Abeillé et al. (2016) argue for a finer-grained analysis of French RNR based on phonological deletion. Their empirical evidence reveals a split between functional and lexical categories in French such that the former permit mismatch between the two conjuncts (where determiners or prepositions differ) under RNR, while the latter don't.

6.2 Gapping

HPSG analyses of gapping fall into two kinds: one kind draws on Beavers and Sag's (2004) deletion-based analysis of nonconstituent coordination (Chaves 2009) and the other on Ginzburg and Sag's (2000) analysis of NUs (Abeillé et al. 2014, Park et al. 2018). The latter analyses align gapping with analyses of NUs, as discussed in section 4, more than with analyses of nonconstituent coordination, and therefore gapping could be classified together with other NUs. We use the analysis due to Abeillé et al. (2014) for illustration below.

Abeillé et al. (2014) assume an identity condition on gapping requiring that gapping remnants match major constituents in the antecedent clause, which they term source clause. In other words, gapping remnants are constrained by elements of the argument structure of the verbal head present in the antecedent and absent from the rightmost conjunct, as originally proposed by Hankamer (1971). This identity condition correctly predicts that gapping remnants needn't appear in the same order as their counterparts in the antecedent (67) (see Sag et al. 1985: 156–158), nor are they required to be the same syntactic category as their counterparts (68). However, note that despite the syntactic category mismatch, the NP *an incredible bore* belongs to the subcategorization frame of the predicate *become.*

(67) A policeman walked in at 11, and at 12, a fireman.

(68) Pat has become [crazy]$_{AP}$ and Chris [an incredible bore]$_{NP}$.

Abeillé et al. (2014) offer additional evidence from Romance (e.g., case mismatch between gapping remnants and their counterparts and even more possibilities of ordering remnants than is the case in English) to strengthen their point that syntactic identity is relaxed under gapping.

The key assumption in Abeillé et al.'s (2014) analysis is that (two or more) gapping remnants form a cluster whose mother has an underspecified syntactic category, that is, is a non-headed phrase. This phrase then serves as the head daughter of a head-fragment phrase, whose syntactic category is also underspecified. This means that there is no unpronounced verbal head in the phrase to which gapping remnants belong. Furthermore, the contextual attribute SAL-UTT introduced for NUs serves to ensure syntactic identity between gapping remnants and their counterparts such that the latter are SAL-UTTs bearing the specification [Major +] as part of their Head feature and being coindexed with the remnants.

With these ingredients of the analysis in place, we reproduce the gapping construction in (69). The construction represents asymmetric coordination in the sense that the daughters include both nonelliptical verbal conjuncts and head-fragment phrases with an underspecified syntactic category. The mother only shares its syntactic category with the nonelliptical conjuncts so that its own category is specified to be verbal.

(69) Gapping construction

gapping-ph → *coord-ph* &

$$\left[\begin{array}{l}\text{HEAD}\ \boxed{\text{H}}\ \textit{verbal}\\ \text{CONTXT}\mid\text{BACKGROUND}\left\{\ldots,\ \textsc{sym-discourse-rel}(\boxed{\text{M}_1},\ldots,\boxed{\text{M}_j},\boxed{\text{M}_{j+1}},\ldots,\boxed{\text{M}_n}),\ldots\right\}\\ \text{DTRS}\ \left\langle\left[\begin{array}{l}\text{HEAD}\ \boxed{\text{H}}\left[\begin{array}{l}\textsc{verbal}\\ \text{CLUSTER}\ \textit{elist}\end{array}\right]\\ \text{CONTENT}\ \boxed{\text{M}_1}\end{array}\right],\ldots,\left[\begin{array}{l}\text{HEAD}\ \boxed{\text{H}}\left[\begin{array}{l}\textsc{verbal}\\ \text{CLUSTER}\ \textit{elist}\end{array}\right]\\ \text{CONTENT}\ \boxed{\text{M}_j}\end{array}\right]\right\rangle\end{array}\right]\oplus$$

$$\left\langle\left[\begin{array}{l}\text{HEAD}\left[\text{CLUSTER}\ \langle\boxed{1},\ldots,\boxed{\text{N}}\rangle\right]\\ \text{SOURCE}\ \boxed{\text{M}_j}\\ \text{CONTENT}\ \boxed{\text{M}_{j+1}}\end{array}\right],\ldots,\left[\begin{array}{l}\text{HEAD}\left[\text{CLUSTER}\ \langle\boxed{1'},\ldots,\boxed{\text{N}'}\rangle\right]\\ \text{SOURCE}\ \boxed{\text{M}_j}\\ \text{CONTENT}\ \boxed{\text{M}_n}\end{array}\right]\right\rangle$$

7 Summary

This chapter has reviewed three types of ellipsis, nonsentential utterances, predicate ellipsis, and nonconstituent coordination, corresponding to three kinds of analysis within HPSG. The pattern that emerges from this overview is that HPSG favors the 'what you see is what get' approach to ellipsis and limits a deletion-based approach, common in the minimalist literature on ellipsis, to a subset of nonconstituent coordination phenomena.

Abbreviations

NUs Nonsentential utterances

BAE Bare Argument Ellipsis

VPE Verb Phrase Ellipsis

NCA Null Complement Anaphora

SAL-UTT Salient Utterance

MAX-QUD Maximal-Question-under-Discussion

DGB Dialog Game Board

Acknowledgements

We thank the editors and Yusuke Kubota for helpful comments.

8 Handmade list of references: Fix!

Abeillé, Anne, Gabriela Bîlbîie, and François Mouret. 2014. A Romance perspective on gapping constructions. In Hans Boas and F. Gonzálvez Garcia, (eds.), Romance Perspectives on Construction Grammar, 227–267. Amsterdam: Benjamins.

Abeillé, Anne, Berthold Crysmann, and Aoi Shiraïshi. 2016. Syntactic mismatch in French peripheral ellipsis. In Christopher Piñn (ed.), Empirical Issues in Syntax and Semantics 11, 1–30. Paris: CSSP.

Abeillé, Anne and Shrita Hassamal. 2017. Sluicing in Mauritian: a fragment analysis. Ms.

Abels, Klaus. 2017. On the interaction of P-stranding and sluicing in Bulgarian. In Olav Mueller-Reichau and Marcel Guhl (eds.), Aspects of Slavic linguistics: Formal grammar, lexicon and communication, volume 16 of Language, Context and Cognition. Berlin: De Gruyter Mouton.

Aelbrecht, Lobke and William Harwood. 2015. To be or not to be elided: VP ellipsis revisited. Lingua 153:66–97.

Arregui, Ana, Charles Clifton, Jr., Lyn Frazier, and Keir Moulton. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. Journal of Memory and Language 55:232–246.

Barton, Ellen. 1998. The grammar of telegraphic structures: Sentential and nonsentential derivation. Journal of English Linguistics 26:37–67.

Beavers, John and Ivan A. Sag. 2004. Coordinate ellipsis and apparent nonconstituent coordination. In Stefan Müller, (ed.), Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, HPSG Online Proceedings, 48–69. Stanford: CSLI Publications.

Beecher, Henry. 2008. Pragmatic inference in the interpretation of sluiced prepositional phrases. San Diego Linguistics Papers 3. http://escholarship.org/uc/item/2

Bresnan, Joan (ed.). 1982. The Mental Representation of Grammatical Relations. Cambridge, Massachusetts: MIT Press.

Chaves, Rui P. 2008. Linearization-based word-part ellipsis. Linguistics and Philosophy 3(31):261–307.

Chaves, Rui P. 2009. A linearization-based approach to gapping. In James Rogers, (ed.), Proceedings of FG-MoL 2005, 205–218. Stanford, CA: CSLI Publications.

Chaves, Rui P. 2014. On the disunity of right-node raising phenomena. Language 90:834–886.

Chung, Sandra, William A. Ladusaw and James McCloskey. 1995. Sluicing and logical form. Natural Language and Linguistic Theory 3:239–282.

Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerald Penn, (ed.), Proceedings of FGVienna: The 8th Conference on Formal Grammar, 45–64. Stanford, CA: CSLI Publications.

Culicover, Peter W. and Ray S. Jackendoff. 2005. Simpler Syntax. Oxford: Oxford University Press.

Fernández, Raquel. 2006. Non-Sentential Utterances in Dialogue: Classification, Resolution and Use. Ph.D. thesis, King's College, London.

Fernández, Raquel and Jonathan Ginzburg. 2002. Non-sentential utterances: A corpus study. Traitement automatique des languages Dialogue 43(2):13–42.

Fernández, Raquel, Jonathan Ginzburg, and Shalom Lappin. 2007. Classifying ellipsis in dialogue: A machine learning approach. Computational Linguistics 33(3):397–427.

Fortin, Catherine. 2007. Indonesian sluicing and verb phrase ellipsis: Description and explanation in a minimalist framework. Doctoral dissertation, University of Michigan, Ann Arbor.

Ginzburg, Jonathan. 1994. An update semantics for dialogue. In H. Bunt, (ed.), Proceedings of the 1st International Workshop on Computational Semantics. Tilburg: ITK, Tilburg University.

Ginzburg, Jonathan. 2012. The Interactive Stance: Meaning for Conversation. Oxford: Oxford University Press.

Ginzburg, Jonathan. 2013. Rethinking deep and surface: towards a comprehensive model of anaphoric processes in dialogue.

Ginzburg, Jonathan and Robin Cooper. 2004. Clarification, ellipsis, and the nature of contextual updates. Linguistics and Philosophy 27(3):297?66.

Ginzburg, Jonathan and Robin Cooper. 2014. Quotation via dialogical interaction. Journal of Logic, Language, and Information 23(3):287?11.

Ginzburg, Jonathan and Raquel Fernández. 2010. Computational models of dialogue. In Alex Clark, Chris Fox, and Shalom Lappin, (eds.), Handbook of Computational Linguistics and Natural Language. Oxford: Blackwell.

Ginzburg, Jonathan, Raquel Fernández, and David Schlangen. 2014. Disfluencies as intra-utterance dialogue moves. Semantics and Pragmatics 7(9):1?4.

Ginzburg, Jonathan and Ivan A. Sag. 2000. Interrogative Investigations: the form, meaning and use of English Interrogatives. Number 123 in CSLI Lecture Notes. Stanford: California: CSLI Publications.

Ginzburg, Jonathan and Philip Miller. 2018. Ellipsis in Head-Driven Phrase Structure Grammar. In Jeroen van Creanen and Tanja Temmerman (eds.), The Oxford handbook of ellipsis. Oxford: OUP.

Grant, Margaret, Charles Clifton, and Lyn Frazier. 2012. The role of non-actuality implicatures in processing elided constituents. Journal of Memory and Language 66:326–343.

Griffiths, James, and Anikó Lipták. 2014. Contrast and island sensitivity in clausal ellipsis. Syntax 17:189–234.

Hankamer, Jorge and Ivan A. Sag. 1976. Deep and Surface Anaphora. Linguistic Inquiry 7(3):391–428.

Hardt, Daniel. 1993. Verb Phrase Ellipsis: Form, Meaning, and Processing. Ph.D. thesis, University of Pennsylvania. Distributed as IRCS Report 93–23.

Jacobson, Pauline. 2016. The short answer: Implications for direct compositionality (and vice versa). Language 92(2):331–375.

Kehler, Andrew. 2002. Coherence, Reference, and the Theory of Grammar. Stanford: CSLI Publications.

Kertz, Laura. 2013. Verb phrase ellipsis: The view from information structure. Language 89:390–428.

Kim, Christina S., Gregory M. Kobele, Jeffrey T. Runner, and John T. Hale. 2011. The acceptability cline in VP ellipsis. Syntax 14:318–354.

Kim, Christina and Jeffrey Runner. 2017. The division of labor in explanations of verb phrase ellipsis. Linguistics and Philosophy 41:41–85.

Kim, Jong-Bok. 2006. Similarities and differences between English VP ellipsis and VP fronting: An HPSG analysis. Studies in Generative Grammar 13.3:429–459.

Kim, Jong-Bok. 2015. Syntactic and semantic identity in Korean sluicing: A direct interpretation approach. Lingua 66:260–293.

Kim, Jong-Bok. 2016. The syntactic structures of Korean: A construction-based perspective. Cambridge: Cambridge University Press.

Larsson, Steffan. 2002. Issue based Dialogue Management. Ph.D. thesis, Gothenburg University.

Lee, Hanjung. 2016. Usage probability and subject–object asymmetries in Korean case ellipsis: Experiments with subject case ellipsis. Journal of Linguistics 52:70–110.

Leung, Tommy. 2014. The Preposition Stranding Generalization and conditions on sluicing: Evidence from Emirati Arabic. Linguistic Inquiry 45:332–340.

Lobeck, Anne. 1995. Ellipsis. Functional heads, licensing, and identification. Oxford: Oxford University Press.

López, Luis. 2000. Verb Phrase Ellipsis and Discourse-Linking. Lingua 110:183–213.

Manning, Christopher and Ivan Sag. 1999. Argument structure, valence, and binding. Nordic Journal of Linguistics 21:107–144.

Melnik, Nurit. 2007. Extending partial pro-drop in modern Hebrew: A comprehensive analysis. In Stefan Müller, (ed.), Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, HPSG Online Proceedings, 173–193. Stanford: CSLI Publications. URL http://csli-publications.stanford.edu/HPSG/2007.

Merchant, Jason. 2001. The Syntax of Silence. Sluicing, Islands, and the Theory of Ellipsis. Oxford: Oxford University Press.

Merchant, Jason. 2004. Fragments and ellipsis. Linguistics and Philosophy 27:661–738.

Merchant, Jason. 2005. Revisiting syntactic identity conditions. Workshop on ellipsis, University of California, Berkeley.

Merchant, Jason. 2013. Voice and ellipsis. Linguistic Inquiry 44:77–108.

Miller, Philip. 2011. The choice between verbal anaphors in discourse. In I. Hendrickx, S. Lalitha Devi, A. Branco, and R. Mitkov, (eds.), Anaphora Processing and Applications: 8th Discourse Anaphora and Anaphor Resolution Colloquium, DAARC 2011, Volume 7099 of Lecture Notes in Artificial Intelligence, 82–95. Berlin: Springer.

Miller, Philip and Barbara Hemforth. 2014. Verb phrase ellipsis with nominal antecedents. Ms. Université Paris Diderot.

Miller, Philip and Geoffrey K. Pullum. 2014. Exophoric VP ellipsis. In Philip Hofmeister and Elisabeth Norcliffe, (eds.), The Core and the Periphery: Data-driven Perspectives on Syntax Inspired by Ivan A. Sag, 5–32. Stanford, CA: CSLI Publications.

Molimpakis, Emilia. 2018. P-stranding under sluicing: Acceptability and processing. Workshop Relating Elliptical Utterances to Information in Context. Stuttgart, Germany.

Morgan, Jerry. 1989. Sentence fragments revisited. In Bradley Music, Randolph Graczyk, and Caroline Wiltshire (eds.), CLS 25, 228–241. Chicago Linguistic Society, Chicago.

Nykiel, Joanna. 2013. Clefts and preposition omission under sluicing. Lingua 123:74–117.

Park, Sang-Hee, Koenig, Jean-Pierre, and Rui P. Chaves. 2018. Scope ambiguities and Lexical Resource Semantics: The case of gapping. Paper presented at the 25th International Conference on Head-Driven Phrase Structure Grammar.

Philippova, Tatiana. 2014. P-omission under sluicing, [P clitic] and the nature of P-stranding. In Martin Kohlberger, Kate Bellamy and Eleanor Dutton (eds.), Proceedings of ConSOLE XXII, 133–155.

Purver, Matthew. 2006. Clarie: Handling clarification requests in a dialogue system. Research on Language & Computation 4(2):259–288.

Rodrigues, Cillene, Andrew Nevins and Luis Vicente. 2009. Cleaving the interactions between sluicing and preposition stranding. In D. Torck and L. Wetzels (eds.), Romance Languages and Linguistic Theory 2006. 175–198. Amsterdam: John Benjamins Publishing Company.

Ross, John R. 1967. Constraints on variables in syntax. (Ph.D. Dissertation). MIT.

Ross, John R. 1969. Guess who. In Proceedings of the 5th Annual Meeting of the Chicago Linguistic Society, 252–286. CLS.

Ruppenhofer, Josef and Laura A. Michaelis. 2014. Frames and the interpretation of omitted arguments in English. In Katz Bourns, Stacey and Lindsy L. Myers (eds.), Perspectives on Linguistic Structure and Context: Studies in honor of Knud Lambrecht, 57–86. John Benjamins.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Boas and Sag (eds.), Sign-Based Construction Grammar, 69–202. Stanford: CSLI Publications.

Sag, Ivan A. and Joanna Nykiel. 2011. Remarks on sluicing. In Stefan Müller, (ed.), Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, HPSG Online Proceedings, 188–298. Stanford: CSLI Publications. URL http:// web.stanford.edu/group/cslipublications/cslipublications/H nykiel.pdf.

Sag, Ivan A., Gerold Gazdar, Thomas Wasow, and Steven Weisler. 1985. Coordination and how to distinguish categories. Natural Language and Linguistic Theory, 3, 117–171.

Schachter, Paul. 1977. Does she or doesn't she? Linguistic Inquiry 8:763–767.

Schmeh, Katharina, Peter W. Culicover, Jutta Hartmann and Susanne Winkler. 2015. Discourse function ambiguity of fragments: A linguistic puzzle. In S. Winkler (ed.), Ambiguity, 368–400. de Gruyter.

Stainton, Robert. 2006. Words and Thoughts. Oxford University Press.

Stjepanović, Sandra. 2008. P-stranding under sluicing in a non-P-stranding language? Linguistic Inquiry 37:179–190.

Stjepanović, Sandra. 2012. Two cases of violation repair under sluicing. In Jason Merchant and Andrew Simpson (eds.), Sluicing. Cross-linguistic perspectives. 68–82. Oxford: Oxford University Press.

Szczegielniak, Adam. 2008. Islands in sluicing in Polish. In Natasha Abner and Jason Bishop (eds.), Proceedings of the 27th West Coast Conference on Formal Linguistics. Cascadilla Proceedings Project, 404–412. Somerville, MA, USA.

Vicente, Luis. 2015. Morphological case mismatches under sluicing. Snippets 29.

Webber, Bonnie. 1979. A Formal Approach to Discourse Anaphora. Outstanding Dissertations in Linguistics. New York: Garland.

Yatabe, Shûichi. 2001. The Syntax and Semantics of Left-Node Raising in Japanese. In D. Flickinger and A. Kathol (eds.), Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar, CSLI, Stanford, pp. 325–344.

Yatabe, Shûichi. 2012. Comparison of the Ellipsis-Based Theory of Non-Constituent Coordination with Its Alternatives. In S. Müller (ed.), Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, CSLI, Stanford, 454–474.

Yatabe, Shûichi and Wai Lok Tam. 2018. In defense of an HPSG-based theory of non-constituent coordination: A reply to Kubota and Levine. Ms. University of Tokyo and Rozetta Corp.

Chapter 21

Anaphoric Binding

António Branco
University of Lisbon

The interpretation of anaphors depends on their antecedents and respective semantic value. This chapter presents the constraints on the admissible antecedents of nominal anaphors that are based on grammatical relations and structure. The integration into grammar of these anaphoric binding constraints and of the semantic representation of anaphors is also presented here.

1 Introduction

In large enough contexts, a nominal anaphoric expression may have more admissible antecedents than the antecedent that happen to be eventually selected to cospecify its interpretation. And when occurring in a given syntactic position, different anaphoric expressions may have different sets of admissible antecedents. This is illustrated in the emblematic examples below, with two anaphoric expressions from English — *herself* and *her* — occurring in the same position, each with different sets of admissible antecedents.

(1) Mary's friend knows that Paula's sister described Joan to herself / her.

For the expression *herself*, either *Joan* or *Paula's sister* is an admissible antecedent. For *her*, its set of admissible antecedents includes instead *Paula*, *Mary's friend* and *Mary*.

Further examples with two anaphoric expressions occurring in the same position and each with different sets of admissible antecedents are the following:

(2) Mary's friend knows that Paula's sister saw her / the little girl.

António Branco. 2019. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 613–665. Berlin: Language Science Press. DOI:??

For the expression *the little girl*, either *Paula* or *Mary* is an admissible antecedent. For *her*, its set of admissible antecedents includes additionally *Mary's friend*.

Such differences in terms of sets of admissible antecedents are the basis for the partition of nominal anaphoric expressions into different groups according to their anaphoric capacity. It has been an important topic of research to determine how many such groups or types of anaphoric expressions there are, what expressions belong to which type in each language, and what exactly are the sets of admissible antecedents for each type.

The results of this inquiry have been condensed in a handful of anaphoric binding constraints, or principles, which seek to capture the relative positioning of anaphors and their admissible antecedents in grammatical representations.

From an empirical perspective, these constraints stem from what appears as quite cogent generalisations and exhibit a universal character, given their cross linguistic validity.[1] From a conceptual point of view, in turn, the relations among binding constraints involve non-trivial cross symmetry, which lends them a modular nature and provides further strength to the plausibility of their universal character.[2]

Accordingly, anaphoric binding principles appears as one of the most significant subsets of grammatical knowledge, usually termed as Binding Theory.

This grammar module is presented in this chapter. In the next section, Section 2, the empirical generalisation captured in the binding constraints are introduced, together with the relevant auxiliary notions and parameterisation options.

The key ingredients for the integration of binding constraints into grammar are discussed in Section 3, and a detailed account of this integration is provided in the following Section 4 — which is further illustrated with the support of the working example in the Appendix.

Section 5 is devoted to discuss how the account of anaphoric binding presented in the previous sections ensures a neat interface of grammar with reference processing systems, and thus supports a seamlessly articulation of binding constraints with anaphora resolution.

In the penultimate section, Section 6, additional binding constraints are introduced, that hold from the perspective of the antecedents, rather from the perspective of the anaphors, together with the respective supporting empirical evidence.

The final Section 7 is devoted to underline the design features that emerge as

[1](Branco 2000c), *i.a.*

[2](Branco 2005a).

crucial for an account of the grammar of anaphoric binding constraints, and to provide an outlook to promising avenues for future research that may further enhance our understanding of anaphoric binding and the semantics of nominal anaphors.

2 Empirical generalisations

Since the so called integrative approach to anaphora resolution was set up,[3] it is common wisdom that factors determining the antecedents of anaphors divide into filters, or hard constraints, and preferences, or soft constraints. The former exclude impossible antecedents and help to circumscribe the set of admissible antecedents; the latter interact to converge on the eventual antecedent among the admissible antecedents.

So-called binding principles are a notorious subset of hard constraints on anaphora resolution: they capture generalisations concerning the constraints on the relative positioning of anaphors with respect to their admissible antecedents in the grammatical geometry of sentences.

We present below the definition of binding constraints,[4] which resorts to a few auxiliary notions — locality, o-command, o-binding —, whose definition, in turn, are presented right afterwards.

There are four such constraints on the anaphoric capacity of nominals, named Principle A, Z, B and C. They induce a partition of the set of anaphors into four classes. According to this partition, every nominal anaphor is of one of the following anaphoric types: short-distance reflexive, long-distance reflexive, pronoun, or non-pronoun.

The definition of each binding principle in (3)-(6) is paired with an illustrative example with key grammatical contrasts empirically supporting the respective generalisation. In particular, Principle A in (3) is paired with an example with the short-distance reflexive *himself*, Principle Z in (4) is paired with the Portuguese long-distance reflexive *ele próprio*, Principle B in (5) with the pronoun *him*, and Principle C in (6) with the non pronoun *the boy*. These examples will be discussed below right after the definitions of the auxiliary notions have been presented.

[3] The integrative approach to anaphora resolution was set up in (Carbonell & Brown 1988; Rich & LuperFoy 1988; Asher & Wada 1989), and its practical viability was extensively checked out in (Lappin & Leass 1994; Mitkov 1997).

[4] This is the approach to Binding Theory proposed in (Pollard & Sag 1992) and (Pollard & Sag 1994: Chap.6), and subsequent developments in (Xue et al. 1994; Branco 1996; Branco & Marrafa 1997; Manning & Sag 1999; Wechsler 1999; Koenig 1999; Branco & Marrafa 1999; Richter et al. 1999; Golde 1999; Branco 2000c; Kiss 2001; Branco 2002a,c,b) *i.a.*

(3) **Principle A**: A locally o-commanded short-distance reflexive must be locally o-bound.

...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes himself$_{*x/*i/*j/*k/l}$].

(4) **Principle Z**: An o-commanded long-distance reflexive must be o-bound.

...X_x...[O amigo do Lee$_i$]$_j$ acha [que [o irmão do Max$_k$]$_l$
the friend of.the Lee thinks that the brother of.the Max
gosta dele próprio$_{*x/*i/j/*k/l}$]. (Portuguese)
likes of.him self
'...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes him$_{*x/*i/j/*k}$ / himself$_l$].'

(5) **Principle B**: A pronoun must be locally o-free.

...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes him$_{x/i/j/k/*l}$].

(6) **Principle C**: A non-pronoun must be o-free.

...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes the boy$_{x/i/*j/k/*l}$].

2.1 Binding, coindexation, locality, command and crosslinguistic variation

The empirical generalisations presented above result from linguistic analysis supported by empirical evidence of which the respective examples above are just a few key illustrative cases. These examples will be discussed in detail in the next subsection, thus illustrating the analysis underlying the binding principles above.

The above definition of binding principles is rendered with the help of a few auxiliary notions. For many of these auxiliary notions, their final value or definition is amenable to be set according to a range of options: as briefly exemplified below, this parameterisation may be driven by the particular language at stake, by the relevant predicator selecting the anaphor, by the specific anaphoric form, etc.

These are the definitions of those auxiliary notions:

Binding *O-binding* is such that "x o-binds y iff x o-commands y and x and y are coindexed" (*o-freeness* is non o-binding).[5]

[5](Pollard & Sag 1994: 279).

Coindexation *Coindexation* is meant to represent an anaphoric link between the expressions with the same index. A starred index, in turn, indicates that the anaphoric link represented is not acceptable, as in the following examples:

(7) a. John$_i$ said that Peter$_j$ shaved himself$_{*i/j}$.
   b. John$_i$ said that Peter$_j$ shaved him$_{i/*j}$.

Turning to example (5), for instance, *him*$_k$ and *Max*$_k$ are coindexed with *k*, thus indicating their anaphoric binding and representing that *Max* is the antecedent of *him*. The starred index **l*, in turn, indicates that the coindexation between *Max's brother*$_l$ and *him*$_{*l}$ is not felicitous, and thus that *Max's brother* is not an admissible antecedent of *him* in (5).

In the examples above, '…X_x…' represents a generic, extra-sentential antecedent available from the context.

Plural anaphors with so-called split antecedents, that has concomitantly more than one antecedent, are represented with a sum of indexes as a subscript, as exemplified below by *them* being interpreted as referring to John and Mary:[6]

(8) John$_i$ told Mary$_j$ that Kim talked about them$_{i+j}$.

Locality The *local domain* of an anaphor results from the partition of sentences and associated grammatical geometry into two zones of greater or less proximity with respect to the anaphor.

Typically, the local domain coincides with the immediate selectional domain of the predicator directly selecting the anaphor. In the following example, the local domain of *him* is explicitly marked within square brackets:

(9) John knows that [Peter described him].

[6] When at least one of the antecedents in a split antecedent relation does not comply with the relevant binding principle (and there is at least one that complies with it), the acceptability of that anaphoric link degrades. Apparently, the larger the number of antecedents that violate the binding constraint the less acceptable is the anaphoric link: while both examples below are not fully acceptable, two coindexations out of three, via *j* and *k*, in violation of the Principle B render example b. less acceptable than example a., which has one coindexation only, via *k*, in violation of that binding constraint (Seeley 1993: 313):

(i) a. ? The doctor$_i$ told the patient$_j$ [that the nurse$_k$ would protect them$_{i+j+k}$ during the storm].
   b. ?? The doctor$_i$ said [that the patient$_j$ told the nurse$_k$ about them$_{i+j+k}$].

As for plural reflexives, which in turn comply with Principle A, they accept split antecedents only in exempt positions — on the notion of exemption, see Section 2.3.

In the example in (3), for instance, *Max's brother* is immediately selected by *likes*, the predicator that immediately selects *himself*, while *Lee's friend* is not. Hence, the first is in the local domain of *himself*, while the latter is not.

In some cases, there may be additional requirements that the local domain is circumscribed by the first selecting predicator that happens to be finite, bears tense or indicative features, etc.[7] One such example can be the following:[8]

(10) a. Jón$_i$ segir að [Maria$_j$ elskar sig$_{*i/j}$]. (Icelandic)
Jón says-IND that Maria loves-IND himself
'Jón$_i$ says that [Maria$_j$ loves himself$_{*i}$/herself$_j$].'

b. [Jón$_i$ segir að Maria$_j$ elski sig$_{i/j}$].
Jón says-IND that Maria loves-SUBJ himself
'[Jón$_i$ says that Maria$_j$ loves himself$_i$/herself$_j$].'

In the first sentence above, the verb in the embedded clause is Indicative and the local domain of its Direct Object is circumscribed to this clause as the reflexive cannot have the Subject of the upwards clause as its antecedent. The second sentence is identical to the first one except that the mood of the embedded verb is now Subjunctive. This leads to a change in the local domain of the reflexive: it can now have also the upwards Subject as its antecedent, thus revealing that its local domain is determined by the first selecting verb in the Indicative, which happens now to be the verb of the upwards clause.

In some other languages, there are anaphors whose local domain is the immediate selectional domain not of the directly selecting predicator but of the immediately upwards predicator, irrespective of the inflectional features of the directly or indirectly selecting predicators. This seems to be the case of the Greek *o idhios*:[9]

(11) O Yannis$_i$ ipe stin Maria [oti o Costas$_j$ pistevi [oti o Vasilis$_k$
the Yannis told the Maria that the Costas believes that the Vasilis
aghapa ton idhio$_{??i/j/*k}$]]. (Greek)
loves the same.
'Yannis$_i$ told Maria that [Costas$_j$ believes that [Vasilis$_k$ loves him$_{??i/j/*k}$]].'

Languages shows diversity concerning which of these options are materialized and which grammatical and lexical means are brought to bear.[10] Additionally, not

[7] Vd. (Manzini & Wexler 1987; Koster & Reuland 1991; Dalrymple 1993) for further details.
[8] (Manzini & Wexler 1987: 47).
[9] Alexis Dimitriadis p.c. See also (Iatridou 1986; Varlokosta & Hornstein 1993).
[10] (Dimitriadis et al. 2005).

all languages have anaphors of every one of the anaphoric types: For instance, English is not known to have long-distance reflexives.

Command *o-command* is a partial order defined on the basis of the obliqueness hierarchies of grammatical functions, possibly embedded in each other along the relation of subcategorisation:"Y o-commands Z just in case either Y is less oblique than Z; or Y o-commands some X that subcategorises for Z; or Y o-commands some X that is a projection of Z".[11]

The grammatical function Subject is less oblique than the Direct Object, the Direct Object is less oblique than the Indirect Object, etc., thus establishing a so-called obliqueness hierarchy. The obliqueness hierarchy of grammatical functions is represented in the value of ARG-ST feature in as much as the arguments are ordered from those whose grammatical function is less oblique to those whose function is more oblique. As discussed in detail in Section 4 and in connection with the working example in the Appendix, ARG-ST feature value plays a crucial role in the formalization and explicit integration of binding principles into grammar.[12]

Accordingly, the Subject o-commands the Direct Object, the Direct Object o-commands the Indirect Object, etc.; and in a multi-clausal sentence, the arguments in the upwards clauses o-command the arguments in the successively embedded clauses.

(12) [[John's friend] said that [[Peter's brother] presented [Martin's cousin] to him]].

In the example above, *John's friend* o-commands *Peter's brother*, *Peter*, *Martin's cousin*, *Martin* and *him*. *Peter's brother* locally o-commands *Martin's cousin* and *him*, and (non-locally) o-commands *Martin*. Neither *John*, *Peter*, *Martin* nor *him* is o-commanding any nominal in this example.

In the example of (4), for instance, the Portuguese long-distance reflexive *ele próprio*, which is the Object of the embedded clause, is o-commanded by *o irmão do Max* (Max's brother), the Subject of that clause, and by *o amigo do Lee* (Lee's friend), the Subject of the upwards clause. *Lee* and *Max*, in turn, do not o-command this reflexive in this example.

[11] (Pollard & Sag 1994: 279).

[12] For further discussion of the notion of obliqueness of grammatical functions as well as further references on this topic, see (Pollard & Sag 1987: Section 5.2).

2.2 Binding principles

With the definition of the auxiliary notions above in place, the definition of binding principles in (3)-(6) is now complete and it is possible to appreciate how the respective examples instantiate them.

Principle A The example in (3) shows that the anaphoric capacity of *himself* complies with the anaphoric discipline captured by Principle A: if it is locally o-commanded, it has to be locally o-bound, i.e. only locally o-commanders can be its admissible antecedents if it happens to be locally o-commanded.

Max's brother is in the local domain of *himself* because it is immediately selected by the predicator *likes* which also immediately selects *himself*. Moreover, *Max's brother*, in a Subject position, o-commands *himself*, in an Object position. *Max's brother* is thus a local o-commander of *himself*, and hence it is an admissible antecedent of *himself*.

The other nominals in this example are not local o-commanders of *himself*: both *Lee* and *Lee's friend* are selected by the main clause predicator *thinks*, not by the predicator *likes* which is immediately selecting *himself* and are thus not in its local domain; *Max* in turn, given it is embedded inside the local Subject, it is not immediately selected by the predicator *likes* that is immediately selecting *himself*. Hence, none of the nominals in the sentence other than *Max's brother* happen to be local o-commanders of *himself* and thus are not one of its admissible antecedents.

Also any other antecedent candidate eventually available in the extra-sentential context is not a local o-commander of *himself* and thus it is not one of its admissible antecedents.

Given the anaphoric capacity of *himself* complies with the anaphoric discipline captured by Principle A, it belongs to the class of short-distance reflexives.

In connection with Principle A, it is also worth signalling that it is not the case that only Subjects can be local o-commanders of short-distance reflexives, as illustrated below.

(13) a. Peter$_i$ didn't talk to John$_j$ about himself$_{i/j}$.
b. About himself$_{i/j}$, Peter$_i$ didn't talk to John$_j$.

In the examples in (13), *John* and *himself* are in the same local domain. Moreover, *John*, in the Object position, is less oblique than *himself*, in the Indirect Object position. Hence, *John* is a local o-commander of the reflexive and qualifies as one of its admissible antecedents, together with *Peter*, in the Subject position.

The absence of contrast between (13a) and (13b) is a central piece of evidence that the command relation for anaphoric binding is based on the obliqueness hierarchy of grammatical functions (o-command) rather than on a configurational hierarchy based on surface syntactic structure (c-command).[13]

Principle Z The example in (4) shows that the anaphoric capacity of the Portuguese nominal *ele próprio* complies with the anaphoric discipline captured by Principle Z: if it is o-commanded, it has to be o-bound, i.e. only o-commanders can be its admissible antecedents if it happens to be o-commanded.

ele próprio is (locally) o-commanded by *o irmão do Max* because both are selected by the predicator *gosta* and *o irmão do Max* is less oblique than *ele próprio*. *ele próprio* is also o-commanded by *o amigo do Lee* because *o amigo do Lee* is selected by the predicator of the upwards clause *acha*, which selects the embedded clause whose predicator selects *ele próprio*, and *o amigo do Lee* is thus less oblique than *ele próprio* in the composite obliqueness hierarchy.

The other nominals in this example are not o-commanders of *ele próprio*: both *Lee* and *Max* are embedded inside arguments of the relevant predicators *acha* and *gosta* but are not arguments of them. Hence, none of the nominals in the sentence other than *o amigo do Lee* and *o irmão do Max* happen to be o-commanders of *ele próprio* and thus are not one of its admissible antecedents.

Also any other antecedent candidate eventually available in the extra-sentential context is not an o-commander of *ele próprio* and thus it is not one of its admissible antecedents.

Given the anaphoric capacity of *ele próprio* complies with the anaphoric discipline captured by Principle Z, it belongs to the class of long-distance reflexives.

Principle B The example in (5) shows that the anaphoric capacity of *him* complies with the anaphoric discipline captured by Principle B: it has to be locally o-free, i.e. its local o-commanders cannot be its admissible antecedents.

In that example, *Max's brother* is the only local o-commander of *him* because *Max's brother* is the only argument of the predicator *likes* other than *him*, and is less oblique than *him*.

The other nominals in this example are not local o-commanders of *him*: neither

[13]Binding principles based on o-command, rather than on c-command as proposed in (Chomsky 1980) and (Chomsky 1986), is a hallmark of the analysis in (Pollard & Sag 1992).

The analysis based on c-command incorrectly predicts that anaphoric links like those in (13b) would not be acceptable, because the admissible antecedents of *himself* do not c-command it (it is *himself* that c-commands them instead).

The analysis based on c-command also incorrectly predicts that the anaphoric link between *himself* and *John* in (13a) would not be acceptable, because *John* does not c-command *himself*.

For a detailed discussion, see (Pollard & Sag 1994: Chap.6).

Lee's friend, *Lee* or *Max* are immediately selected by *likes*.

Also any other antecedent candidate eventually available in the extra-sentential context is not a local o-commander of *him*.

Hence, in this example all antecedent candidates, sentential and non sentential, are admissible antecedents of *him* except *Max's brother*.

Given the anaphoric capacity of *him* complies with the anaphoric discipline captured by Principle B, it belongs to the class of pronouns.

Principle C The example in (6) shows that the anaphoric capacity of *the boy* complies with the anaphoric discipline captured by Principle C: it has to be o-free, i.e. its o-commanders are not admissible antecedents.

In that example, *Lee's friend* and *Max's brother* are the only o-commanders of *the boy*: *Lee's friend* is selected by the predicator of the upwards clause *likes*, which selects the embedded clause whose predicator selects *the boy*; *Max's brother*, in turn, is the only argument immediately selected by the predicator *likes* other than *the boy*; and both *Lee's friend* and *Max's brother* are less oblique than *the boy*.

The other nominals in this example are not o-commanders of *the boy*: neither *Lee* or *Max* are immediately selected by *thinks* or *likes*.

Also any other antecedent candidate eventually available in the extra-sentential context is not an o-commander of *the boy*.

Hence, in this example all antecedent candidates, sentential and non sentential, are admissible antecedents of *the boy* except *Lee's friend* and *Max's brother*.

Given the anaphoric capacity of *the boy* complies with the anaphoric discipline captured by Principle C, it belongs to the class of non pronouns.

2.3 O-bottom positions: reshuffling and exemption

For the interpretation of an anaphor to be accomplished, an antecedent has to be found for it. Such an antecedent is to be picked from the set of its o-commanders, if the anaphor is a long-distance reflexive, or from the set of its local o-commanders, if it is a short-distance reflexive.

This requirement may not be satisfied in some specific cases, namely when the reflexive occurs in a syntactic position such that it is the least element of its o-command order, in an o-bottom position for short. In such circumstances, it has no o-commander (other than itself, if the o-command relation is formally defined as a reflexive relation) to qualify as its antecedent.

That is the motivation for the conditional formulation of Principles A and Z, in (3) and (4) respectively: a (short/) long-distance reflexive has to be (locally/) o-bound if it is (locally/) o-commanded. In case it is not (locally/) o-commanded, there is no imposition concerning their admissible antecedents following from Principles A and Z.

Reshuffling As a consequence, in some cases, the binding domain for the reflexive which happens to be the least element of its local obliqueness order may be reshuffled, being reset as containing the o-commanders of the reflexive in the domain circumscribed by the immediately upwards predicator.[14] One such case for a nominal domain can be found in the following example:[15]

(14) a. Gernot$_i$ dachte, dass Hans$_j$ dem Ulrich$_k$ [Marias$_l$ Bild von
Gernot thought that Hans the Ulrich Maria's picture of
sich$_{*i/*j/*k/l}$] überreichte. (German)
self gave
'Gernot$_i$ thought that Hans$_j$ gave Ulrich$_k$ [Maria$_l$'s picture of himself$_{*i/*j/*k}$/herself$_l$].'

b. Gernot$_i$ dachte, dass [Hans$_j$ dem Ulrich$_k$ ein Bild von sich$_{*i/j/k}$
Gernot thought that Hans the Ulrich a picture of self
überreichte].
gave
'Gernot$_i$ thought that [Hans$_j$ gave Ulrich$_k$ [a picture of himself$_{*i/j/k}$]].'

In the first sentence above, the short-distance reflexive is locally o-commanded by *Maria* and only this nominal can be its antecedent. In the second sentence, the reflexive is the first element in its local obliqueness hierarchy and its admissible antecedents, which form now its local domain, are the nominals in the obliqueness hierarchy of the immediately upwards predicator.

The null subject in languages like Portuguese is another example of a short-distance reflexive that is in an o-bottom position and whose local domain is reshuffled:[16]

(15) O médico$_i$ disse-me [que [o director do Pedro$_j$]$_k$ ainda não
the doctor told-me that the director of.the Pedro yet not
reparou [que $\emptyset_{*i/*j/k}$ cometeu um erro]]. (Portuguese)
noticed that made a mistake.
'The doctor$_i$ told me [that [Pedro$_j$'s director]$_k$ didn't notice yet [that he$_{*i/*j/k}$ made a mistake]].'

[14](Branco 2005b).

[15]Tibor Kiss p.c., which is a development with regards to his data in (Kiss 2001).

[16](Branco 2007).

In the example above, as the null reflexive is in an o-bottom position, its local domain gets reshuffled to include the immediately upwards o-commander *Pedro's director*. Once it is thus o-commanded, in accordance do Principle A, the null reflexive cannot take other nominal in the sentence, viz. *the doctor* or *Pedro*, as its admissible antecedent given none of these o-commands it.

Exemption In some other cases, this resetting of the binding domain is not available. In such cases, the reflexive is in the bottom of its local obliqueness order and is observed to be exempt of its typical binding regime: the reflexive may take antecedents that are not its o-commanders or that are outside of its local or immediately upward domains,[17] as illustrated in the following example:[18]

(16) Mary$_i$ thought the artist had done a bad job, and was sorry that her parents came all the way to Columbus just to see the portrait of herself$_i$.

In an exempt position, a reflexive can even have so-called split antecedents, as illustrated in the following example with a short-distance reflexive:[19]

(17) Mary$_i$ eventually convinced her sister Susan$_j$ that John had better pay visits to everybody except themselves$_{i+j}$.

That is an option not available for reflexives in non exempt positions:

(18) Mary$_i$ described$_j$ John to themselves$_{*(i+j)}$.

Some long-distance reflexives may also be exempt from their binding constraint if they occur in the bottom of their o-command relation. In such cases, they can have an antecedent in the previous discourse sentences or in the context, or a deictic use, as illustrated in the following example:

(19) [O Pedro e o Nuno]$_i$ também conheceram ontem a Ana. Eles
the Pedro and the Nuno also met yesterday the Ana. They
próprios$_i$ ficaram logo a gostar muito dela. (Portuguese)
próprios stayed immediately to liking much of.her
'[Pedro and Nuno]$_i$ also met Ana yesterday. They$_i$ liked her very much right away.'

Such options are not available in non exempt positions:[20]

[17](Pollard & Sag 1994: 263).
[18](Golde 1999: 73).
[19](Zribi-Hertz 1989: 42).

[20]For further details, vd. (Branco & Marrafa 1999).

(20) A Ana também conheceu ontem [o Pedro e o Nuno]$_i$. Ela
The Ana also met yesterday the Pedro and the Nuno. She
ficou logo a gostar muito deles próprios$_{*i}$. (Portuguese)
stayed immediately to liking much of.them *próprios*
'Ana also met [Pedro and Nuno]$_i$ yesterday. She liked them$_{*i}$ very much right away.'

Admittedly, an overarching interpretability condition is in force in natural languages requiring the "meaningful" anchoring of anaphors to antecedents. Besides this general requirement, anaphors are concomitantly ruled by specific constraints concerning their particular anaphoric capacity, including the sentence-level constraints in (3)-(6), i.e. the binding principles.

When reflexives are in o-bottom positions, an o-commander (other than the reflexive itself) may not be available to function as antecedent and anchor their interpretation. Hence, such specific binding constraints, viz. Principle A and Z, cannot be satisfied in a "meaningful" way and the general interpretability requirement may supervene them. As a consequence, in cases displaying so-called exemption from binding constraints, o-bottom reflexives appear to escape their specific binding regime to comply with such general requirement and its interpretability be rescued.

The anaphoric links of exempt reflexives have been observed to be governed by a range of non sentential factors (from discourse, dialogue, non linguistic context, etc.), not being determined by the sentence-level binding principles in (3)-(6).[21]

2.4 o-command: alternations and subject-orientedness

Alternations In languages like English, the o-command order can be established over the obliqueness hierarchies of active and passive sentences alike:[22]

(21) a. John$_i$ shaved himself$_i$.
b. John$_i$ was shaved by himself$_i$.

The obliqueness hierarchy of grammatical functions is represented in ARG-ST and in both ARG-ST values of (21a) and of (21b), *John* appears as the Subject and qualifies as a local o-commander of *himself*, and thus as an admissible antecedent of this reflexive.

[21] For further details, vd. (Kuno 1987; Zribi-Hertz 1989; Golde 1999) among others.
[22] (Jackendoff 1972; Pollard & Sag 1994).

In some other languages, only the obliqueness hierarchy of a given syntactic alternation is available to support the o-command order relevant for binding constraints in both alternations.

This is the case, for example, of the alternation active/objective voice in Toba Batak. In this language, a reflexive in Object position of an active voice sentence can have the Subject as its antecedent, but not vice-versa:[23]

(22) a. mang-ida diri-na$_i$ si John$_i$. (Toba Batak)
[ACTIVE-saw himself$_{\text{OBJECT}}$]VP PM John$_{\text{SUBJECT}}$
'John$_i$ saw himself$_i$.'

b. mang-ida si John$_i$ diri-na$_{*i}$.
[ACTIVE-saw PM John$_{\text{OBJECT}}$]VP himself$_{\text{SUBJECT}}$

Taking the objective voice paraphrase corresponding to the active sentence above, the binding pattern is inverted: a reflexive in Subject position can have the Object as its antecedent, but not vice-versa, thus revealing that the obliqueness hierarchy relevant for the verification of its binding constraint remains the hierarchy of the corresponding active voice sentence above:

(23) a. di-ida diri-na$_{*i}$ si John$_i$.
[OBJECTIVE-saw himself$_{\text{OBJECT}}$]VP PM John$_{\text{SUBJECT}}$

b. di-ida si John$_i$ diri-na$_i$.
[OBJECTIVE-saw PM John$_{\text{OBJECT}}$]VP himself$_{\text{SUBJECT}}$
'John$_i$ saw himself$_i$.'

Subject-orientedness o-command may take the shape of a linear or non linear order depending on the specific obliqueness hierarchy upon which it is realised.

In a language like English, the arguments in the subcategorisation frame of a predicator are typically arranged in a linear obliqueness hierarchy.

In some other languages, the obliqueness hierarchy upon which the o-command order is based may happen to be non linear: in the subcategorisation frame of a predicator, the Subject is less oblique than any other argument while the remaining arguments are not comparable to each other under the obliqueness relation. As a consequence, in a clause, a short-distance reflexive with an Indirect Object grammatical function, for instance, may only have the Subject as its antecedent, its only local o-commander.[24]

[23](Manning & Sag 1999: 72).

[24]For a thorough argument and further evidence motivated independently of binding facts see (Branco 1996; Branco & Marrafa 1997; Branco 2000c). In some languages, there can be an addi-

This Subject-orientedness effect induced on the anaphoric capacity of reflexives by the non linearity of the o-command relation can be observed in contrasts like the following:[25]

(24) a. Lars$_i$ fortalte Jon$_j$ om seg selv$_{i/*j}$. (Norwegian)
Lars told Jon about self *selv*
'Lars$_i$ told Jon$_j$ about himself$_{i/*j}$.'

b. Lars$_i$ fortalte Jon$_j$ om ham selv$_{*i/j}$.
Lars told Jon about him *selv*
'Lars$_i$ told Jon$_j$ about him$_{*i/j}$.'

In the first sentence above, the reflexive cannot have the Direct Object as its antecedent given that the Subject is its only local o-commander in the non linear obliqueness hierarchy. In the second sentence, under the same circumstances, a pronoun presents the symmetric pattern: it can have any co-argument as its antecedent except the Subject, its sole local o-commander.[26]

3 Binding Constraints at the Syntax-Semantics Interface

Like other sorts of constraints on semantic composition, binding constraints impose grammatical conditions on the interpretation of certain expressions — anaphors, in the present case — based on syntactic geometry.[27] This should not be seen, however, as implying that they express grammaticality requirements. By replacing, for instance, a pronoun by a reflexive in a sentence, we are not turning a grammatical construction into an ungrammatical one, even if we assign to the reflexive the antecedent adequately selected for the pronoun. In that case, we are just asking the hearer to try to assign to that sentence a meaning that it cannot express, in the same way as what would happen if we asked someone whether he could interpret *The red book is on the white table* as describing a situation where a white book is on a red table.

tional requirement that the Subject be animate to qualify as a commander to certain anaphors. On this, see (Huang & Tang 1991; Xue et al. 1994) about Chinese *ziji*, among others.

[25] Lars Hellan p.c. See also (Hellan 1988: 67).

[26] For an analysis of the Subject-orientedness of French *se* resorting to a notion of s-command, see (Abeillé, Godard, Miller, et al. 1998; Abeillé, Godard & Sag 1998).

[27] For a discussion of proposals in the literature that have tried to root binding principles on non-grammatical, cognitive search optimisation mechanisms, and their pitfalls, see (Branco 2000a; 2003; 2004).

In this example, given how they happen to be syntactically related, the semantic values of *red* and *table* cannot be composed in a way that this sentence could be used to describe a situation concerning a red table, rather than a white table.

Likewise, if we take the sentence *John thinks Peter shaved him*, given how they happen to be syntactically related, the semantic values of *Peter* and *him* cannot be composed in a way that this sentence could be used to describe a situation where John thinks that Peter shaved himself, i.e. Peter, rather than a situation where John thinks that Peter shaved other people, e.g. Paul, Bill, etc., or even John himself.

The basic difference between these two cases is that, while in the first the composition of the semantic contributions of *white* and *table* (for the interpretation of their NP *white table*) is constrained by local syntactic geometry, in the latter the composition of the semantic contributions of *John* and *him* (for the interpretation of the NP *him*) is constrained by non-local syntactic geometry.

These grammatical constraints on anaphoric binding should thus be taken as conditions on semantic interpretation given that they delimit (non-local) aspects of meaning composition, rather than aspects of syntactic wellformedness.[28]

These considerations leads one to acknowledge that, semantically, an anaphor should be specified in the lexicon as a function whose argument is a suitable representation of the context – providing a semantic representation of the NPs available in the discourse vicinity –, and delivers both an update of its anaphoric potential – which is instantiated as the set of its grammatically admissible antecedents – and an update of the context, against which other NPs are interpreted.[29] Naturally, all in all, there will be four of such functions available to be lexically associated to anaphors, each corresponding to one of the different four classes of anaphors, in accordance with the four binding constraints A, Z, B or C.[30]

[28] This approach is in line with (Gawron & Peters 1990), and departs from other approaches where binding constraints have been viewed as wellformedness conditions, thus belonging to the realm of Syntax: "[they] capture the distribution of pronouns and reflexives" (Reinhart & Reuland 1993: 657).

[29] (Branco 1998b; 2000b; 2002a).

[30] This is in line with (Johnson & Klein 1990) concerning the processing of the semantics of nominals, and also the spirit (but by no means the letter) of the dynamic semantics framework – vd. (Chierchia 1995) and (Stalnaker 1998) *i.a.*

3.1 Semantic patterns

For an anaphoric nominal *w*, the relevant input context may be represented in the form of a set of three lists of reference markers,[31] **A**, **Z** and **U**. List **A** contains the reference markers of the local o-command order where *w* is included, ordered according to their relative grammatical obliqueness; **Z** contains the markers of the (local and non local) o-command order where *w* is included, i.e. reference markers organised in a possibly multi-clausal o-command relation, based upon successively embedded clausal obliqueness hierarchies; and **U** is the list of all reference markers in the discourse context, possibly including those not linguistically introduced.

The updating of the context by an anaphoric nominal *w* may be seen as consisting simply in the incrementing of the representation of the context, with a copy of the reference marker of *w* being added to the three lists above.

The updating of the anaphoric potential of *w*, in turn, delivers a representation of the contextualised anaphoric potential of *w* in the form of the list of reference markers of its admissible antecedents. This list results from the binding constraint associated to *w* being applied to the relevant representation of the context of *w*.

Given this setup, the algorithmic verification of binding constraints consists of a few simple operations, and their grammatical specification will consist thus in stating each such sequence of operations in terms of the grammar description formalism.

If the nominal *w* is a short-distance reflexive, its semantic representation is updated with **A'**, where **A'** contains the reference markers of the o-commanders of *w* in **A**.

If *w* is a long-distance reflexive, its semantic representation includes **Z'**, such that **Z'** contains the o-commanders of *w* in **Z**.

If *w* is a pronoun, its semantics should include the list of its non-local o-commanders, that is the list $\mathbf{B}=\mathbf{U}\backslash(\mathbf{A'}\cup[\text{r-mark}_w])$ is encoded into its semantic representation, where r-mark_w is the reference marker of *w*.

Finally if *w* is a non-pronoun, its updated semantics keeps a copy of list $\mathbf{C}=\mathbf{U}\backslash(\mathbf{Z'}\cup[\text{r-mark}_w])$, which contains the non-o-commanders of *w*.

[31] See (Karttunen 1976; Kamp 1981; Heim 1982; Seuren 1985; Kamp & Reyle 1993) for the notion of reference marker.

3.2 Binding principles and other constraints for anaphora resolution

These lists **A'**, **Z'**, **B** and **C** collect the reference markers that are antecedent candidates at the light only of the relevant binding constraints, which are relative positioning filters in the process of anaphora resolution.[32] The elements in these list have to be submitted to the other constraints and preferences of this process so that one of them ends up being chosen as the antecedent.

In particular, some of these markers may eventually turn up not being admissible antecedent candidates due to the violation of some other constraints — e.g. those requiring similarity of morphological features or of semantic type — that on a par with binding constraints have to be complied with. For example, in this example *John described Mary to himself*, by the sole constraining effect of Principle A, [r-mark$_{John}$, r-mark$_{Mary}$] is the list of antecedent candidates of *himself*, which will be narrowed down to [r-mark$_{John}$] when all the other filters for anaphora resolution have been taken into account, including the one concerning similarity of morphological features, as *Mary* and *him* do not have the same gender feature value.

In this particular case, separating these two type of filters — similarity of morphological features and binding constraints — seems to be the correct option, required by plural anaphors with so called split antecedents. In an example of this type, such as *John$_i$ told Mary$_j$ they$_{i+j}$ would eventually get married*, where *they* is resolved against *John* and *Mary*, the morphological features of the anaphor are not identical to the morphological features of each of its antecedents, though the relevant binding constraint applies to each of them.[33]

When a plural anaphor takes more than one antecedent, as in the example above, its (plural) reference marker will end up being semantically related with a plural reference marker resulting from some semantic combination of the markers of its antecedents. Separating binding constraints from other constraints on the relation between anaphors and their antecedents are thus compatible with and justified by proposals for plural anaphora resolution that take into account

[32] See Branco (1999: Chap.2) for an overview of filters and preferences for anaphora resolution proposed in the literature.

[33] This was noted by (Higginbotham 1983). In this respect, this approach improves on the proposal in (Pollard & Sag 1994), where the token-identity of indices — internally structured in terms of Person, Number and Gender features — is meant to be forced upon the anaphor and its antecedent in tandem with the relevant binding constraint.

For further reasons why token-identity between the reference markers of the anaphor and the corresponding antecedent is not a suitable option for every anaphoric dependency, see the discussion below in Section 5 on the semantic representation of different modes of anaphora.

split anaphora.[34]

3.3 Computational tractability

It is also worth noting that the computational tractability of the grammatical compliance with binding principles is ensured given the polynomial complexity of the underlying operations described above.

Let n be the number of words in an input sentence to be parsed, which for the sake of the simplicity of the argument, and of the worst case scenario, it is assumed to be made only of nominal anaphors, that is every word in that sentence is a nominal anaphor. Assume also that the sets **A**, **Z** and **U**, thus of length n at worst, are available at each node of the parsed tree via copying or via list appending (more details about these two operations in the next sections), which is a process of constant time complexity.

At worst, the operations involved at each one of the n leaf nodes of the tree to obtain one of the sets **A'**, **Z'**, **B** or **C** are: list copying and list appending operations, performed in constant time; extraction of the predecessors of an element in a list, which is of linear complexity; or at most one list complementation, which can be done in time proportional to $nlog(n)$. The procedure of verifying binding constraints in a sentence of length n is thus of tractable complexity, namely $O(n^2 \log n)$ in the worst case.[35]

[34] That is the case e.g. of (Eschenbach et al. 1989). According to this approach, the set of antecedent candidates of a plural anaphor which result from the verification of binding constraints has to receive some expansion before subsequent filters and preferences apply in the anaphora resolution process. The reference markers in that set, either singular or plural, will be previously combined into other plural reference markers: it is thus from this set, closed under the semantic operation of pluralisation (e.g. i-sum a la (Link 1983)), that the final antecedent will be chosen by the anaphor resolver.

[35] For a thorough discussion of alternative procedures for the compliance with binding principles and their drawbacks, see (Branco 2000d), very briefly summarised here:

The verification of binding constraints proposed in (Chomsky 1980; 1981) requires extra-grammatical processing steps of non tractable computational complexity (Correa 1988; Fong 1990), which, moreover, are meant to deliver a forest of indexed trees to anaphor resolvers.

In Lexical Functional Grammar, the account of binding constraints requires special purpose extensions of the description formalism (Dalrymple 1993), which ensures only a partial handling of these constraints.

For accounts of binding principles in the family of Categorial Grammar frameworks, see (Szabolcsi 1989; Hepple 1990; Morril 2000), and for a critical overview, see (Jäger 2001).

4 Binding Constraints in the Grammar

In this section, the binding constraints receive a principled integration into formal grammar. For the sake of brevity, we focus on the English language. Given the discussion in the previous sections, the parameterisation for other languages will follow from this example by means of seamless adaptation.

We show how the module of Binding Theory is specified with the description language of HPSG, as an extension of the grammar fragment in the Annex of the foundational HPSG book,[36] following the feature geometry in Ivan Sag's proposed extension of this fragment to relative clauses,[37] and adopting a semantic component for HPSG based on Underspecified Discourse Representation Theory (UDRT).[38]

As exemplified in (25), this semantic component is encoded as the value of the feature CONT(ENT). This value, of sort *udrs*, has a structure permitting that the mapping into underspecified discourse representations be straightforward.[39]

The value of subfeature CONDS is a set of labeled semantic conditions. The hierarchical structure of these conditions is expressed by means of a subordination relation of the labels identifying each condition, a relation that is encoded as the value of SUBORD. The attribute LS defines the distinguished labels, which indicate the upper (L-MAX) and lower (L-MIN) bounds for a semantic condition within the overall semantic representation to be constructed.

ANAPH(ORA) subfeature of CONT(ENT) The integration of Binding Theory into formal grammar consists of a simple extension of this semantic component for the *udrs* of nominals, enhancing it with the subfeature ANAPH(ORA). This new feature keeps information about the anaphoric potential of the corresponding anaphor *w*.

Its subfeature ANTEC(EDENTS) keeps record of how this potential is realised when the anaphor enters a grammatical construction: its value is the list with the antecedent candidates of *w* which comply with the relevant binding constraint for *w*.

And its subfeature R(EFERENCE)-MARK(ER) indicates the reference marker of *w*, which is contributed by its referential force to the updating of the context.

BIND(ING) subfeature of LOC(AL) On a par with this extension of the LOC value, also the NONLOC value is extended with a new feature, BIND(ING), with subfeatures LIST-A, LIST-Z, and LIST-U. These lists provide a specification of the relevant

[36](Pollard & Sag 1994: Annex).
[37](Sag 1997).
[38](Frank & Reyle 1995).
[39](Reyle 1993).

context and correspond to the lists **A**, **Z** and **U** in the sections above. Subfeature LIST-LU is a fourth, auxiliary list encoding the contribution of the local context to the global, non local context, as explained in the next sections.[40]

4.1 Handling the anaphoric potential

Pronouns: lexical entry Given this adjustment to the grammatical geometry, the lexical definition of a pronoun, for instance, will include the following SYNSEM value:

[40] For the sake of readability, the working example in (25) displays only the more relevant features for the point at stake. The NONLOC value has this detailed definition in (Pollard & Sag 1994):

$$ {}_{\textit{NONLOC}}\left[\begin{array}{ll}\textsc{to-bind} & \textit{NONLOC1}\\ \textsc{inherited} & \textit{NONLOC1}\end{array}\right] $$

And these are the details of the extension we are using, where the information above is coded now as a *udc* object, which keeps record of the relevant non local information for accounting to *u(nbounded) d(ependency) c(onstructions)*:

$$ {}_{\textit{NONLOC}}\left[\begin{array}{ll}\textsc{udc} & {}_{\textit{udc}}\left[\begin{array}{ll}\textsc{to-bind} & \textit{nonloc1}\\ \textsc{inherited} & \textit{nonloc1}\end{array}\right]\\ \textsc{bind} & {}_{\textit{bind}}\left[\begin{array}{ll}\textsc{list-a} & \textit{list(refm)}\\ \textsc{list-z} & \textit{list(refm)}\\ \textsc{list-u} & \textit{list(refm)}\\ \textsc{list-lu} & \textit{list(refm)}\end{array}\right]\end{array}\right] $$

Given this extension, HPSG principles constraining NONLOC feature structure, or part of it, should be fine-tuned with adjusted feature paths in order to correctly target the intended (sub)feature structures.

(25)
$$\begin{bmatrix} \textsc{loc}|\textsc{cont} & \begin{bmatrix} \textsc{ls} & \begin{bmatrix} \textit{L-MAX} & \boxed{1} \\ \textit{L-MIN} & \boxed{1} \end{bmatrix} \\ \textsc{subord} & \{\} \\ \textsc{conds} & \left\{ \begin{bmatrix} \textsc{label} & \boxed{1} \\ \textsc{dref} & \boxed{2} \end{bmatrix} \right\} \\ \textsc{anaph} & \begin{bmatrix} \textit{R-MARK} & \boxed{2} \\ \textit{ANTEC} & \boxed{5}\ \textit{principleB}\left(\boxed{4},\boxed{3},\boxed{2}\right) \end{bmatrix} \end{bmatrix} \\ \textsc{nonloc}|\textsc{bind} & \begin{bmatrix} \textsc{list-a} & \boxed{3} \\ \textsc{list-z} & \textit{list(refm)} \\ \textsc{list-u} & \boxed{4} \\ \textsc{list-lu} & \left\langle \boxed{2} \right\rangle \end{bmatrix} \end{bmatrix}$$

In this feature structure, the semantic condition in CONDS associated to the pronoun corresponds simply to the introduction of the discourse referent $_2$ as the value of DREF.

This semantic representation is expected to be further specified as the lexical entry of the pronoun gets into the larger representation of the relevant utterance. In particular, the CONDS value of the sentence will be enhanced with a condition specifying the relevant semantic relation between this reference marker $_2$ and one of the reference markers in the value $_5$ of ANTEC. The latter will be the antecedent against which the pronoun will happen to be resolved, and the condition where the two markers will be related represents the relevant type of anaphora assigned to the anaphoric relation between the anaphor and its antecedent.[41]

The anaphoric binding constraint associated to pronouns, in turn, is specified as the relational constraint *principleB*/3 in the value of ANTEC. This is responsible for the realisation of the anaphoric potential of the pronoun as it enters a grammatical construction. When the arguments of this relational constraint are instantiated, it returns list **B** as the value of ANTEC.

As discussed in Section 3.1, this relational constraint *principleB*/3 is defined to take all markers in the discourse context (in the first argument and given by the LIST-U value), and remove from them both the local o-commanders of the pronoun (included in the second argument and made available by the LIST-A value) and the marker corresponding to the pronoun (in the third argument and given by the DREF value).

[41]More details on the interface with anaphora resolvers and on the semantic types of anaphora in Section 5.

Finally, the contribution of the reference marker of the pronoun to the context is ensured via token-identity between R-MARK and a LIST-LU value.

The piling up of this reference marker in the global LIST-U value is determined by a new HPSG principle specific to Binding Theory, to be detailed in the next Section 4.2.

Non pronouns and reflexives: lexical entries The SYNSEM of other anaphors — ruled by principles A, C or Z — are similar to the SYNSEM of pronouns above. The basic difference lies in the relational constraints to be stated in the ANTEC value.

Such constraints — *principleA*/2, *principleC*/3 and *principleZ*/2 — encode the corresponding binding principles and return the realised anaphoric potential of anaphors according to the surrounding context, coded in their semantic representation under the form of a list in the ANTEC value. Such lists — **A'**, **C** or **Z'**, respectively — are obtained by these relational constraints along the lines discussed in Section 3.1.

Non lexical anaphoric expressions Note that, for non-lexical anaphoric nominals in English, namely those ruled by Principle C, the binding constraint is stated in the lexical representation of the determiners contributing to the anaphoric capacity of such NPs. Also the reference marker corresponding to an NP of this kind is brought into its semantic representation from the R-MARK value specified in the lexical entry of its determiner.

Accordingly, for the values of ANAPH to be visible in the signs of non lexical anaphors, Clause I of the Semantics Principle in UDRT[42] is extended with the requirement that the ANAPH value is token-identical, respectively, with the ANAPH value of the specifier daughter, in an NP, and with the ANAPH value of the nominal complement daughter, in a subcategorised PP.

Exemption Note also that for short-distance reflexives, exemption from the constraining effect of the corresponding Principle A occurs when *PRINCIPLEA*(3,2) returns the empty list as the value of feature ANTEC:[43]

[42](Frank & Reyle 1995: 12).

[43]This account applies also to exempt occurrences of long-distance reflexives.

(26)
$$\left[\begin{array}{ll}
\text{LOC}|\text{CONT} & \left[\begin{array}{ll}
\text{LS} & \left[\begin{array}{ll} \textit{L-MAX} & \boxed{1} \\ \textit{L-MIN} & \boxed{1} \end{array}\right] \\
\text{SUBORD} & \{\} \\
\text{CONDS} & \left\{\left[\begin{array}{ll} \text{LABEL} & \boxed{1} \\ \text{DREF} & \boxed{2} \end{array}\right]\right\} \\
\text{ANAPH} & \left[\begin{array}{ll} \textit{R-MARK} & \boxed{2} \\ \textit{ANTEC} & \boxed{4}\ \textit{principleA}\left(\boxed{3},\boxed{2}\right) \end{array}\right]
\end{array}\right] \\
\text{NONLOC}|\text{BIND} & \left[\begin{array}{ll}
\text{LIST-A} & \boxed{3} \\
\text{LIST-Z} & \textit{list(refm)} \\
\text{LIST-U} & \textit{list(refm)} \\
\text{LIST-LU} & \left\langle\boxed{2}\right\rangle
\end{array}\right]
\end{array}\right]$$

This happens if the reference marker of the reflexive $\boxed{2}$ is the first element in the relevant obliqueness hierarchy, i.e. it is the first element in the LIST-A value in $\boxed{3}$, thus o-commanding the other possible elements of this list and not being o-commanded by any of them.

As discussed in Section 2.3, given its essential anaphoricity, a reflexive has nevertheless to be interpreted against some antecedent. As in the exempt occurrences no antecedent candidate is identified by virtue of Principle A activation, the anaphora resolver – which will operate then on the empty ANTEC list[44] – has thus to resort to antecedent candidates outside the local domain of the reflexive: this implies that it has to find antecedent candidates for the reflexive which actually escape the constraining effect of Principle A. The anaphora resolver will then be responsible for modelling the behaviour of reflexives in such exempt occurrences, in which case the anaphoric capacity of these anaphors appears as being exceptionally ruled by discourse-based factors.

4.2 Handling the context representation

Turning now to the representation of the context, this consists in the specification of the constraints on the values of the attributes LIST-A, LIST-Z, LIST-U and LIST-LU. This is handled by adding an HPSG principle to the grammar, termed the Binding Domains Principle (BDP). This principle has three clauses constraining signs with respect to these four lists of reference markers. A full understanding of their details, presented below, will be facilitated with the working example discussed in detail in the Appendix.

[44] More details of the interface between grammar and reference processing systems in Section 5.

Binding Domains Principle, Clause I Clause I of BDP is responsible for ensuring that the values of LIST-U and LIST-LU are appropriately setup at the different places in a grammatical representation:

(27) **Binding Domains Principle**, Clause I

i. The LIST-LU value is identical to the concatenation of the LIST-LU values of its daughters in every sign;

ii. the LIST-LU and LIST-U values are token-identical in a sign of sort *discourse*;

iii. i. the LIST-U value is token-identical to each LIST-U value of its daughters in a non-NP sign;

ii. in an NP sign *k*:

- in Spec-daughter, the LIST-U value is the result of removing the elements of the LIST-A value of Head-daughter from the LIST-U value of *k*;
- in Head-daughter, the LIST-U value is the result of removing the value of R-MARK of Spec-daughter from the LIST-U value of *k*.

By virtue of (i.), LIST-LU collects up to the outmost sign in a grammatical representation — which is of sort *discourse* — the markers contributed to the context by each NP. Given (ii.), this list with all the markers is passed to the LIST-U value at this outmost sign. And (iii.) ensures that this list with the reference markers in the context is propagated to every NP.

Subclause (iii.ii) prevents self-reference loops due to anaphoric interpretation, avoiding what is known in the literature as the i-within-i effect — recall that the R-MARK value of non lexical NPs is contributed by the lexical representation of their determiners, in Spec-daughter position, as noted above.

The HPSG top ontology is thus extended with the new subsort *discourse* for signs: $sign \equiv word \vee phrase \vee discourse$. This new type of linguistic object corresponds to sequences of sentential signs. A new Schema 0 is also added to the Immediate Dominance Principle, where the Head daughter is a phonologically null object of sort *context(ctx)*, and the Text daughter is a list of phrases.

As the issue of discourse structure is out of the scope of this chapter, we adopted a very simple approach to the structure of discourses which suffices for the present account of Binding Theory. As discussed in the next Section 7.3, this object of sort *ctx* helps representing the contribution of the non linguistic context to the interpretation of anaphors.

Binding Domains Principle, Clause II As to the other two Clauses of the Binding Domains Principle, Clause II and Clause III, they constrain the lists LIST-A and LIST-Z, respectively, whose values keep a record of o-command relations.

BDP-Clause II is responsible for constraining LIST-A:

(28) **Binding Domains Principle**, Clause II

i. Head/Arguments: in a phrase, the LIST-A value of its head, and of its nominal (or nominal preceded by preposition) or trace Subject or Complement daughters are token-identical;

ii. Head/Phrase:

  i. in a non-nominal and non-prepositional sign, the LIST-A values of a sign and its head are token-identical;

  ii. in a prepositional phrase,

    - if it is a complement daughter, the LIST-A values of the phrase and of its nominal complement daughter are token-identical;
    - otherwise, the LIST-A values of the phrase and its head are token-identical;

  iii. in a nominal phrase,

    - in a maximal projection, the LIST-A value of the phrase and its Specifier daughter are token-identical;
    - in other projections, the LIST-A values of the phrase and its head are token-identical.

This clause ensures that the LIST-A value is shared between a head-daughter and its arguments, given (i.), and also between the lexical heads and their successive projections, by virtue of (ii.).

o-command On a par with this Clause II, it is important to make sure that at the lexical entry of any predicator *p*, LIST-A includes the R-MARK values of the subcategorised arguments of *p* specified in its ARG-ST value. Moreover, the reference markers appear in the LIST-A value under the same partial order as the order of the corresponding *synsem* in ARG-ST. This is ensured by the following constraints on the lexical entries of predicators:

(29)

$$\left[_{synsem}\ \text{LOC}|\text{CONT}|\text{ARG-ST}\left\langle \cdots, \left[\text{LOC}|\text{CONT}|\text{ANAPH}|\text{R-MARK}\ \boxed{i}\right], \cdots \right\rangle\right]$$
$$\longrightarrow \left[_{synsem}\ \text{NONLOC}|\text{BIND}|\text{LIST-A}\left\langle \cdots, \boxed{i}, \cdots \right\rangle\right]$$

$$\left[_{synsem}\ \text{LOC}|\text{CONT}|\text{ARG-ST}\left\langle \begin{array}{l} \cdots, \left[\text{LOC}|\text{CONT}|\text{ANAPH}|\text{R-MARK}\ \boxed{k}\right], \\ \cdots, \left[\text{LOC}|\text{CONT}|\text{ANAPH}|\text{R-MARK}\ \boxed{l}\right], \cdots \end{array} \right\rangle\right]$$
$$\longrightarrow \left[_{synsem}\ \text{NONLOC}|\text{BIND}|\text{LIST-A}\left\langle \cdots, \boxed{k}, \cdots, \boxed{l}, \cdots \right\rangle\right]$$

In case a subcategorised argument is quantificational, it contributes also with its VAR value to the make up of LIST-A:[45]

(30)

$$\left[_{synsem}\ \text{LOC}|\text{CONT}|\text{ARG-ST}\left\langle \cdots, \left[\text{LOC}|\text{CONT}|\text{ANAPH}\begin{bmatrix}\text{R-MARK} & \boxed{r} \\ \text{VAR} & \boxed{v}\end{bmatrix}\right], \cdots \right\rangle\right]$$
$$\longrightarrow \left[_{synsem}\ \text{NONLOC}|\text{BIND}|\text{LIST-A}\left\langle \cdots, \boxed{v}, \boxed{r}, \cdots \right\rangle\right]$$

[45] More details on this and on the e-type anaphora vs. bound-variable anaphora distinction are discussed in the next sections.

Binding Domains Principle, Clause III Finally, BDP-Clause III ensures that LIST-Z is properly constrained:

(31) **Binding Domains Principle**, Clause III
For a sign F:

i. in a Text daughter, the LIST-Z and LIST-A values are token-identical;
ii. in a non-Text daughter,
    i. in a sentential daughter, the LIST-Z value is the concatenation of the LIST-Z value of F with the LIST-A value;
    ii. in a Head daughter of a non-lexical nominal, the LIST-Z value is the concatenation of L with the LIST-A value, where L is the list which results from taking the list of o-commanders of the R-MARK value, or instead of VAR value when this exists, of its Specifier sister from the LIST-Z value of F;
    iii. in other, non-filler, daughters of F, the LIST-Z value is token-identical to the LIST-Z value of F.

By means of (i.), this Clause III ensures that, at the top node of a grammatical representation, LIST-Z is set up as the LIST-A value of that sign.

Moreover, given (ii.), it is ensured that LIST-Z is successively incremented at suitable downstairs nodes — those defining successive locality domains for binding, as stated in (ii.i) and (ii.ii) — by appending, in each of these nodes, the LIST-A value to the LIST-Z value of the upstairs node.

Locality From this description of the Binding Domains Principle, it follows that the locus in grammar for the parameterisation of what counts as a local domain for a particular language is the specification of BDP–Clauses II and III for that language.

5 Interface with Reference Processing Systems

The appropriateness of the grammatical constraints on anaphoric binding presented above extends to its suitable accounting of the division of labor between grammars and reference processing systems, and of the suitable interfacing between them.

5.1 Anaphora Resolution

While the grammatical anaphoric binding constraints are specified and verified as part of the global set of grammatical constraints, they provide also for a suitable hooking up of the grammar with modules for anaphora resolution.

Feature ANTEC is the neat interface point between them: its value with a list of antecedent candidates that comply with Binding Theory requirements is easily made accessible to anaphor resolvers. This list will be then handled by a resolver where further non grammatical soft and hard constraints on anaphora resolution will apply and will filter down that list until the most likely candidate will be determined as the antecedent.

5.2 Reference Processing

The anaphoric binding constraints also provide a convenient interface for anaphoric links of different semantic types — exemplified below — to be handled and specified by reference processing systems:

(32)
a. John$_i$ said that he$_i$ would leave soon. (coreference)
b. Kim$_i$ was introduced to Lee$_j$ and a few minutes later they$_{i+j}$ went off for dinner. (split anaphora)
c. Mary could not take [her car]$_i$ because [the tyre]$_i$ was flat. (bridging anaphora)
d. [Fewer than twenty Parliament Members]$_i$ voted against the proposal because they$_i$ were afraid of riots in the streets. (e-type anaphora)
e. [Every sailor in the Bounty]$_i$ had a tattoo with [his mother's]$_i$ name on the left shoulder. (bound anaphora)

Example (32a) displays a coreference relation, where *he* has the same semantic value as its antecedent *John*.

A case of split antecedent can be found in (32b) as *they* has two syntactic antecedents and it refers to an entity comprising the two referents of the antecedents.

The referent of *the tyre* is part of the referent of its antecedent *his car* in (32c), thus illustrating a case of so called bridging anaphora (also know as indirect or associative anaphora), where an anaphor may refer to an entity that is e.g. an element or part of the denotation of the antecedent, or an entity that includes the denotation of the antecedent, etc.[46]

In (32d) *they* has a so called non-referential antecedent, *fewer than twenty Parliament Members*, from which a reference marker is inferred to serve as the semantic value of the plural pronoun: *they* refer to those Parliament Members, who

[46]See (Poesio & Vieira 1998) for an overview.

are fewer than twenty in number, and who voted against the proposal. Example (32d) illustrates a case of e-type anaphora,[47] and this inference mechanism to obtain an antecedent marker from a non referring nominal is described in Section 7.3.

Finally in (32e), though one also finds a quantificational antecedent for the anaphoric expression, the relation of semantic dependency differs to the one in the previous example. The anaphoric expression *his mother* does nor refer to the mother of the sailors of the Bounty. It acts rather in the way of a bound variable of logical languages — for each sailor *s*, *his mother* refers to the mother of *s* — thus exemplifying a case of so called bound anaphora.[48]

Given that the semantic relation between antecedent marker and anaphor marker can be specified simply as another semantic condition added to the CONDS value, a DRT/HPSG representation for the resolved anaphoric link under the relevant semantic type of anaphora is straightforward and the integration of the reference processing outcome into grammatical representation is seamlessly ensured.

For the sake of the illustration of this point, assume that a given reference marker **x** turns out to be identified as the antecedent for the anaphoric nominal Y, out of the set of antecedent candidates for Y in its ANTEC value. This antecedent **x** can be related to the reference marker **y** of anaphor Y by means of an appropriate semantic condition in its CONDS value. Such a condition will be responsible for modelling the specific mode of anaphora at stake.

For instance, coreference will require the expected condition $\mathbf{y}{=}_{coref}\mathbf{x}$, as exemplified below with the CONT value of the pronoun in (25) extended with a solution contributed by an anaphor resolver, where $\boxed{7}$ would be the marker picked up as the plausible antecedent.

[47] (Evans 1980).
[48] (Reinhart 1983).

(33)
$$\begin{bmatrix} \text{LS} & \begin{bmatrix}\text{L-MAX} & \boxed{1}\\ \text{L-MIN} & \boxed{1}\end{bmatrix}\\ \text{SUBORD} & \left\{\boxed{1}=\boxed{6}\right\}\\ \text{CONDS} & \left\{\left(\begin{bmatrix}\text{LABEL} & \boxed{1}\\ \text{DREF} & \boxed{2}\end{bmatrix}, \begin{bmatrix}\text{LABEL} & \boxed{6}\\ \text{REL} & =_{coref}\\ \text{ARG1} & \boxed{2}\\ \text{ARG2} & \boxed{7}\end{bmatrix}\right)\right\}\\ \text{ANAPH} & \begin{bmatrix}\text{R-MARK} & \boxed{2}\\ \text{ANTEC} & \boxed{5}\left\langle \dots, \boxed{7}, \dots\right\rangle\end{bmatrix}\end{bmatrix}$$

An instance of bridging anaphora, in turn, may be modelled by *bridg*(**x**, **y**), where *bridg* stands for the relevant bridging function between **y** and **x**, and similarly for the other semantic anaphora types.

5.3 Coreference Transitivity

It is also noteworthy that the interfacing of grammar with reference processing systems ensured by anaphoric binding constraints provides a neat accommodation of coreference transitivity.

If as a result of the process of anaphora resolution, a given anaphor N and another anaphor B end up being both coreferent with a given antecedent A, then they end up being coreferent with each other. That is, in addition to having marker r_a as an admissible antecedent in its set of candidate antecedents, that anaphor N has also to eventually have marker r_b included in that set.

This is ensured by including, in the CONDS value in (25), semantic conditions that follow as logical consequences from this overall coreference transitivity requirement that is operative at the level of the reference processing system with which grammar is interfaced: $\forall r_a, r_b((\boxed{2} =_{coref} r_b \wedge r_a =_{coref} r_b) \Rightarrow (\langle r_a \rangle \cup \boxed{5} = \boxed{5}))$.

An important side effect of this overall constraint is that “accidental” violations of Principle B are prevented, as illustrated with the help of the following example.

(34) * The captain$_{i/j}$ thinks he$_i$ loves him$_j$.

Given that the Subject of the main clause, *the captain*, does not locally o-command any one of them, either the pronoun *he* or the pronoun *him* can have the nominal phrase *the captain* as antecedent, in compliance with Principle B. By transitivity of anaphoric coreference though, the reference marker of *he* is made to belong to the admissible set of antecedents of *him*, which violates Principle

B. Hence, by the conjoined effect of coreference transitivity and of Principle B, that "accidental" violation of Principle B that would make *he* an (o-commanding) antecedent of *him* in this example is (correctly) blocked.

By the same token, "accidental" violations of Principle C with an analogous pattern as above, but for non pronouns, are prevented:

(35) * When John$_{i/j}$ will conclude his therapy, [the boy$_i$ will stop believing [that the patient$_j$ is a Martian]].

Separately, *the boy* and the *the patient* can have *John* as antecedent, in accordance to Principle C. But *the patient* — because is o-commanded by *the boy* — cannot have *the boy* as antecedent, which, also here, is (correctly) ensured by a conjoined effect of the coreference transitivity requirement and the relevant Principle C.

Accordingly, when the semantic type of anaphora is not one of coreference, no coreference transitivity holds, and there happens no "accidental" violation of Principle C. This is illustrated in the following example with bridging anaphora instead, where two non pronouns, though occurring in the same clause, like in (35) , can be (correctly) resolved against the same antecedent — in contrast with that example (35) above, where such possibility is blocked.

(36) Quando [o robot]$_i$ concluiu a tarefa, o operador viu que [a
when the robot concluded the task, the operator saw that the
roda]$_i$ estava a esmagar [o cabo de alimentação]$_i$. (Portuguese)
wheel was to smash the cord of power
'When [the robot]$_i$ concluded the task, the operator saw that [his$_i$ wheel] was smashing [his$_i$ power cord].'

Another range of examples where the semantic type of anaphora is not one of coreference — also with no coreference transitivity holding — and thus also where (correctly) there happens no "accidental" violation of the respective binding principle can be found for reflexives, as illustrated in the following example.

(37) The captain$_i$ thinks he$_{i/j}$ loves himself$_{*i/j}$.

The reflexive *himself* can have *he* as antecedent, because the later locally o-commands it, but cannot have *the captain* as antecedent because the later does not locally o-command it. But while the semantic anaphoric relation between *the captain* and *he* is one of coreference, the semantic anaphoric relation between *he* and *himself* is not, being rather one of bound anaphora.[49] Hence, the coreference transitivity requirement does not apply and the referent of *the captain* does not land into the set of possible antecedents of the reflexive, thus not inducing an "accidental" violation of Principle A. Example (37) can thus felicitously be interpreted as the captain thinking that the agent of loving him is himself, resulting from *himself* having *him* as antecedent and *him* having *the captain* as antecedent.

6 Binding Constraints for Antecedents

The Binding Theory presented in this chapter is also serendipitous in terms of improving the accuracy of empirical predictions offered by a formal grammar with respect to anaphoric binding restrictions that are outside the realm of the binding principles in (3)-(6).

Note first that a reference marker introduced by a non quantificational NP can be the antecedent either of an anaphor that it o-commands, as in (38a), or of an anaphor that it does not o-command, as in (38b):

(38) a. [The captain who knows this sailor]$_i$ thinks Mary loves him$_i$.

b. [The captain who knows [this sailor]$_i$] thinks Mary loves him$_i$.

Differently from a non quantificational NP, which contributes one reference marker to the representation of the context, a quantificational NP contributes two markers that exhibit symmetric features with respect to each other in several respects. The fact that one of them can serve as an antecedent in e-type anaphora, while the other can serve as an antecedent in bound-variable anaphora is certainly one of such symmetries.[50] But there are more.

Let us take a quantificational NP, introduced for instance by the quantifier *every*, acting as an antecedent. This imposes different Number requirements on its anaphors depending on the type of anaphora relation at stake — e-type or bound-variable anaphora — so that the underlying occurrence of each one of the corresponding two markers can be tracked down.

For ease of reference, let us term the marker ensuring e-type anaphora as the e-marker, and the marker ensuring bound anaphora as the v-marker.[51]

[49] Confluent evidence that reflexives entertain a bound anaphora relation with their antecedents was also observed when their inability to enter split anaphora relations in non exempt positions was noted in Section 2.3.

[50] Extensive discussion of this difference is presented in the Appendix.

[51] In the formalisation presented in the Appendix, an e-marker is the marker in the R-MARK value, introduced by Σ-abstraction, and a v-marker is the marker in the the VAR value, introduced by

The contrast below illustrates that, in an e-type anaphoric link, the e-marker stands for a plurality:

(39) Every sailor$_i$ has many girlfriends. They$_i$/He$_{*i}$ travel(s) a lot.

And the next contrast illustrates that, in a bound-variable anaphoric link, the v-marker is singular:

(40) Every sailor$_i$ shaves themselves$_{*i}$/himself$_i$.

The following contrasts can now be considered. An e-marker can be the antecedent of anaphors that it does not o-command, in (41b), but cannot be the antecedent of anaphors that it o-commands, in (41a):

(41) a. * [Every captain who knows this sailor]$_i$ thinks Mary loves them$_i$.
b. [The captain who knows [every sailor]$_i$] thinks Mary loves them$_i$.

This contrast is symmetric to the contrast for the other reference marker: a v-marker can be the antecedent of anaphors that it o-commands, in (42a), but cannot be the antecedent of anaphors that it does not o-command, in (42b):

(42) a. [Every captain who knows this sailor]$_i$ thinks Mary loves him$_i$.
b. * [The captain who knows [every sailor]$_i$] thinks Mary loves him$_i$.

As these contrasts are empirically observed as patterns holding for quantificational NPs in general (not only for those introduced by *every*), constraints emerge on which anaphors different markers can be the antecedents of, in case such markers are contributed by quantificational NPs.

E-markers and v-markers of a given quantificational NP induce a partition of the space of their possible anaphors when that NP is acting as an antecedent: a v-marker is an antecedent for anaphors in the set of its o-commanded anaphors, while an e-marker is an antecedent for anaphors in the complement of such set, i.e. in the set of its non o-commanded anaphors.

This implies that on a par with the grammatical constraints on *the relative positioning of antecedents with respect to anaphors* in (3)-(6), there are also grammatical constraints on *the relative positioning of anaphors with respect to their antecedents* when the corresponding markers are introduced by quantificational NPs. Building on the same auxiliary notions, these "reverse" binding constraints receive the following definition as R-Principles E and V:

the restrictor argument of the determiner.

(43) **R-Principle E**: An antecedent cannot o-bind its anaphor (in e-type anaphora).

[Every captain who knows [every sailor]$_i$]$_j$ thinks Mary loves them$_{i/*j}$.

R-Principle V: An antecedent must o-bind its anaphor (in bound-anaphora)

[Every captain who knows [every sailor]$_i$]$_j$ thinks Mary loves him$_{*i/j}$.

It is worth noting that these principles account also for what has been observed in the literature as the weak crossover effect.[52] In the example below, displaying a case of weak crossover, the anaphoric link is ruled out by R-Principle V since the quantificational NP *every sailor* does not o-command the pronoun *him*, which is singular and could thus enter only into a bound-anaphora relation.

(44) * [The captain who knows him$_i$] thinks Mary loves every sailor$_i$.

Weak crossover constructions appear thus as a sub-case of the class of constructions ruled out by the binding constraints for antecedents.[53]

7 Consolidation and Outlook

With the material presented in the sections above, it emerges that the grammar of anaphoric binding constraints builds on the following key ingredients:

- Interpretation: binding constraints are grammatical constraints on interpretation contributing to the contextually determined semantic value of anaphors — rather than syntactic wellformedness constraints.

[52] See (Jacobson 2000: Section 2.1) for an extensive overview of accounts of weak crossover. For an account of strong crossover in HPSG see (Pollard & Sag 1994: 279).

[53] To the best of our knowledge, the integration of the reverse anaphoric constraints E and V in (43) into HPSG — like what is obtained in Section 4 above for Principles A-Z in (3)-(6) — was not worked out yet in the literature.

Besides an explicit formal specification of (43) in terms of HPSG, there are also empirical aspects that ask to be worked out in future work. For weak crossover, for instance, it is interesting to note Jacobson's remarks: "... it is well known that weak crossover (WCO) is indeed weak, and that the effect can be ameliorated in a variety of configurations. To list a few relevant observations: WCO violations are much milder if the offending pronoun is within a sentence rather than in an NP; the more deeply one embeds the offending pronoun the milder the WCO effect; WCO effects are ameliorated or even absent in generic sentences; they are milder in relative clauses than in questions [...] For example, the possibility of binding in *Every man's$_i$ mother loves him$_i$* remains to be accounted for." (Jacobson 2000: 120).

- Lexicalisation: binding constraints are properties of anaphors determining how their semantic value can be composed or co-specified, under a non-local syntactic geometry, with the semantic value of other expressions — rather than properties of grammatical representations of sentences as such: accordingly, the proper place of these constraints in grammar is at the lexical description of the relevant anaphoric units (e.g. the English pronoun *him*, or the Portuguese multiword long distance reflexive *ele próprio*) or the anaphora inducing items (e.g. the English definite article *the* that introduces non-pronouns).

- Underspecification: binding constraints delimit how the anaphoric potential of anaphors can be realised when they enter a grammatical construction — rather than determining the eventual antecedent: on the one hand, this realisation of anaphoric potential is not a final solution in terms of circumscribing the elected antecedent, but a space of grammatically admissible solutions; on the other hand, this realisation of anaphoric potential has to be decided, locally, in terms of non-local information: accordingly, an underspecification-based strategy is required to pack ambiguity and non-locality.

- Articulation: binding constraints are grammatical constraints — rather than anaphora resolvers: accordingly, grammars, where grammatical anaphoric constraints reside, and reference processing systems, where further constraints on the resolution of anaphora reside, are autonomous with respect to each other, and their specific contribution gains from them being interfaced, rather than being mixed up.

Binding principles capture the relative positioning of anaphors and their admissible antecedents in grammatical representations. As noted at the introduction of the present chapter, together with their auxiliary notions, they have been considered one of the most outstanding modules of grammatical knowledge.

From an empirical perspective, these constraints stem from quite cogent generalisations and exhibit a universal character, given the hypothesis of their parameterised validity across anaphoric expressions and natural languages.

From a conceptual point of view, in turn, the relations among binding constraints involve non-trivial cross symmetry that lends them a modular nature and provides further strength to the plausibility of their universal character.

7.1 Symmetries

The recurrent complementary distribution of the admissible antecedents of a pronoun and of a short-distance reflexive in the same, non exempt syntactic position, in different languages from different language families, has perhaps been the most emblematic symmetry.

For the sake of convenience, the examples in (3)-(6) are copied to (45)-(48) below. The pair (45) vs. (47), with the anaphoric expressions in the same syntactic position of the same syntactic construction, illustrates the symmetry just mentioned, between reflexives and pronouns, suggestively grasped by comparing the starred and non starred indexes.

(45) ...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes himself$_{*x/*i/*j/*k/l}$].

(46) ...X_x...[O amigo do Lee$_i$]$_j$ acha [que [o irmão do Max$_k$]$_l$
the friend of.the Lee thinks that the brother of.the Max
gosta dele próprio$_{*x/*i/j/*k/l}$]. (Portuguese)
likes of.him self
'...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes him$_{*x/*i/j/*k}$ / himself$_l$].'

(47) ...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes him$_{x/i/j/k/*l}$].

(48) ...X_x...[Lee$_i$'s friend]$_j$ thinks [[Max$_k$'s brother]$_l$ likes the boy$_{x/i/*j/k/*l}$].

But given also the complementary distribution of the admissible antecedents of a long-distance reflexive and of a non pronoun in the same, non exempt syntactic position, a similar symmetry is also found between these two other types of anaphors. This is illustrated by the complementarity of the indexes in (46) vs. (48).

Another double "symmetry" worth noting is the one between short- and long-distance reflexives, on the one hand, and non pronouns and pronouns on the other hand.

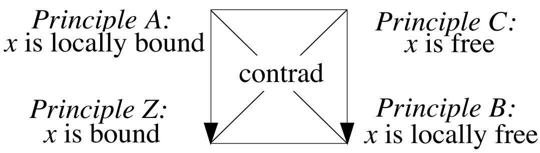
Both sorts of reflexives present the same binding regime but over o-command orders whose length is possibly different: the set of admissible antecedents of a short-distance reflexive is a subset of the set of admissible antecedents of a long-distance reflexive in the same, non exempt syntactic position. For a given non-exempt position, the admissible antecedents of a short-distance reflexive are the antecedents that are in the set of admissible antecedents of a long-distance reflexive in that same position and that are local, i.e. are in the local domain. The felicitous (non starred) indexes in (45) are a subset of the felicitous indexes in (46), which illustrates this symmetry.

A “symmetry” similar to this one is displayed by non pronouns and pronouns with respect to a given syntactic position: the set of admissible antecedents of a non pronoun is a subset of the set of admissible antecedents of a pronoun. For a given position, the admissible antecedents of a non pronoun are the antecedents that are in the set of admissible antecedents of a pronoun in that same position and that are not o-commanding the pronoun (or non pronoun). The felicitous (non starred) indexes in (48) are a subset of the felicitous indexes in (47).

7.2 Quantificational Strength

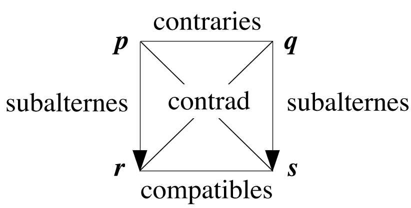
When these symmetries are further explored, the intriguing observation that emerges with respect to the empirical generalisations in (3)-(6) is that when stripped away from their procedural phrasing and non-exemption safeguards, they instantiate a square of logical oppositions:

(49)



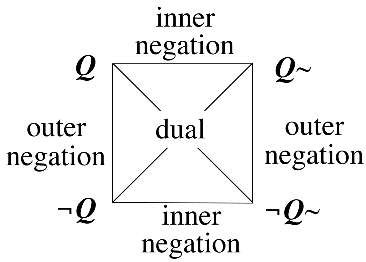
Like in the Aristotelian square of opposition, depicted in (50), there are two pairs of *contradictory* constraints, which are formed by the two diagonals, (Principles A, B) and (C, Z). One pair of *contrary* constraints (they can be both false but cannot be both true) is given by the upper horizontal edge (A, C). One pair of *compatible* constraints (they can be both true but cannot be both false) is given by the lower horizontal edge (Z, B). Finally two pairs of *subcontrary* constraints (the first coordinate implies the second, but not vice-versa) are obtained by the vertical edges, (A, Z) and (C, B).

(50)



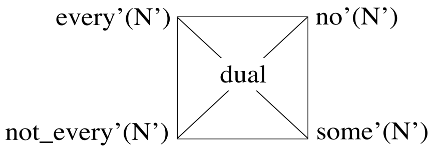
The empirical emergence of a square of oppositions for the semantic values of natural language expressions naturally raises the question about the possible existence of an associated square of duality — and importantly, about the quantificational nature of these expressions.

(51)



It is of note that the classical square of oppositions in (50) is different and logically independent from the square of duality in (51) — with the semantic values of the English expressions *every N*, *no N*, *some N* and *not every N*, or their translational equivalents in other natural languages, providing the classical example of an instantiation of the latter:

(52)



The difference lies in the fact that inner negation, outer negation and duality (concomitant inner and outer negation) are third order concepts, while compatibility, contrariness and implication are second order concepts. As a consequence, it is possible to find instantiations of the square of oppositions without a corresponding square of duality, and vice-versa.[54]

Logical duality has been a key issue in the study of natural language and, in particular, in the study of quantification as this happens to be expressed in natural language. It is a pattern noticed in the semantics of many linguistic expressions and phenomena, ranging from the realm of determiners to the realm of temporality and modality, including topics such as the semantics of the adverbials *still*/*already* or of the conjunctions *because*/*although*, etc.[55]

Under this pattern, one recurrently finds groups of syntactically related expressions whose formal semantics can be rendered as one of the operators arranged in a square of duality. Such a square is made of operators that are interdefinable by means of the relations of outer negation, inner negation, or duality. Accordingly, the emergence of a notoriously non trivial square of logical duality between the semantic values of natural language expressions has been taken as a major empirical touchstone to ascertain their quantificational nature.[56]

[54] Vd. (Löbner 1987) for examples and discussion.
[55] (Löbner 1987; 1989; 1999; ter Meulen 1988; Koning 1991; Smessaert 1997).
[56] Vd. (Löbner 1987; van Benthem 1991). While noting that the ubiquity of the square of duality

By exploring these hints, and motivated by the intriguing square of opposition in (49), the empirical generalisations captured in the binding principles were shown to be the effect of four quantifiers that instantiate a square of duality like (51).[57]

For instance, Principle A is shown to capture the constraining effects of the existential quantifier that is part of the semantic value of short-distance reflexives. Like the existential quantifier expressed by other expressions, such as the adverbial *already*,[58] this a phase quantifier. What is specific here is that the quantification is over a partial order of reference markers, the two relevant semi-phases over this order include the local o-commanders and the other reference markers that are not local o-commanders, respectively for the positive and the negative semi-phases, and the so-called parameter point in phase quantification is the reference marker of the eventual antecedent for the anaphoric nominal at stake.

Accordingly, the other three quantifiers — corresponding to the other three binding Principles B, C and Z — are defined by means of this existential one being under external negation (quantifier expressed by pronouns), internal negation (by non pronouns) or both external and internal negation (by long-distance reflexives).

7.3 Doubly Dual Nominals

While these findings deepen the rooting of binding constraints into the semantics of anaphoric nominals,[59] more importantly, they also point towards promising research directions with the potential to advance our understanding of the grammar of anaphoric binding, in particular, and more widely, to further our insights into the semantics of nominals, in general.

A shared wisdom is that nominals convey either quantificational or referential force.

The findings introduced above imply that nominals with "primary" referential force (e.g. *John*, *the book*, *he*,…) have also a certain "secondary" quantificational force: they express quantificational requirements — over reference markers, i.e. entities that live in linguistic representations —, but do not directly quantify over

may be the sign of a semantic invariant possibly rooted in some cognitive universal, (van Benthem 1991: 23) underlined its heuristic value for research on quantification inasmuch as "it suggests a systematic point of view from which to search for comparative facts".

[57] (Branco 1998a; 2001; 2005a; 2006).

[58] (Löbner 1987).

[59] Their fully-fledged discussion and justification are outside the scope of the present chapter. A thorough presentation can be found in (Branco 2005a).

extra-linguistic entities, like the other "primarily" quantificational nominals (e.g. *every man*, *most students*,…) do.

This duality of semantic behaviour, however, turns out not to be that much surprising if one takes into account a symmetric duality with regards to "primarily" quantificational nominals, which is apparent when they are able to act as antecedents in e-type anaphora. Nominals with "primary" quantificational force have also a certain "secondary" referential force: they have enough referential strength to evoke and introduce reference markers in the linguistic representation that can be picked as antecedents by anaphors — and thus support the referential force of the latter —, but they cannot be used to directly refer to extra-linguistic entities, like the other "primarily" referential terms do.

As a result, the duality quantificational vs. referential nominals appears thus as less strict and more articulated than it has been assumed. Possibly taking indefinite descriptions aside, every nominal makes a contribution in both semantic dimensions of quantification and reference but with respect to different universes. Primarily referential nominals have a dual semantic nature — they are primarily referential (to extra-linguistic entities) and secondarily quantificational (over linguistic entities) —, which is symmetric of the dual semantic nature of primarily quantificational ones — these are primarily quantificational (over extra-linguistic entities) and secondarily referential (to linguistic entities).

Appendix

In order to illustrate the combined effect of the binding constraints specified in Section 4, as well as the outcome obtained from a grammar that integrates Binding Theory, we work through the example below and the corresponding grammatical representation in Figure 1.

(53) Every student said [he likes himself].

This is a multi-clausal sentence with two anaphoric nominals in the embedded clause, a pronoun (*he*) and a short-distance reflexive (*himself*), and with a quantificational NP (*every student*) in the upper clause. In this sentence, the reflexive has the pronoun as the only admissible antecedent, and the pronoun, in turn, can either have the quantificational NP as antecedent or be resolved against an antecedent not introduced in the sentence.

Figure 1 presents an abridged version of the grammatical representation produced by the grammar for a discourse that contains only this sentence. The feature structures below the constituency tree correspond to partial grammatical representations of the leave constituents, while the ones above the tree correspond to partial representations of some of its non terminal nodes.

Circumscribing the Anaphoric Context

Let us start by considering the representation of the context.

Taking the representation of obliqueness hierarchies first, one can check that in the upper nodes of the matrix clause, due to the effect of BDP–Clause III, the LIST-Z value is obtained from the value of LIST-A, with which it is token-identical, thus comprising the list $\langle \boxed{54}, \boxed{247} \rangle$. In the nodes of the embedded clause, in turn, the LIST-Z value is the concatenation of that upper LIST-Z value and the LIST-A value in the embedded clause $\langle \boxed{24}, \boxed{392} \rangle$, from which the list $\langle \boxed{54}, \boxed{247}, \boxed{24}, \boxed{392} \rangle$ is obtained.

In any point of the grammatical representation, the LIST-A values are obtained from the subcategorisation frames of the local verbal predicators, as constrained by BDP–Clause II and the lexical constraints in (29) and (30). Therefore, $\langle \boxed{24}, \boxed{392} \rangle$ is the LIST-A value of *likes*, and $\langle \boxed{54}, \boxed{247} \rangle$ is the LIST-A value of *said*.

Taking into account LIST-LU, as one ascends in the representation of the syntactic constituency, the list gets longer since, by the effect of BDP–Clause I, the LIST-LU value at a given node gathers the reference markers of the nodes dominated by it. Consequently, at the discourse top node, LIST-LU ends up as a list including all reference markers: both those introduced in the discourse by the NPs in the example sentence and $\boxed{415}$, the one available in the non linguistic context, from which the list $\langle \boxed{415}, \boxed{54}, \boxed{247}, \boxed{24}, \boxed{392} \rangle$ is the result.

Note that in cases where the discourse contains more than one sentence, BDP–ClauseI (i.) ensures that LIST-LU ends up with all reference markers from every sentence of the discourse.

BDP–Clause I also ensures that this list of all reference markers is passed to the LIST-U value of the top node, and that this LIST-U value is then percolated down to all nodes of the grammatical representation, including the nodes of anaphoric nominals.

Circumscribing the Anaphoric Potential

To consider the representation of the NPs, we should take a closer look at the leaf nodes in the constituency tree.

Contribution to the context Let us consider first how the NPs contribute to the representation of the context.

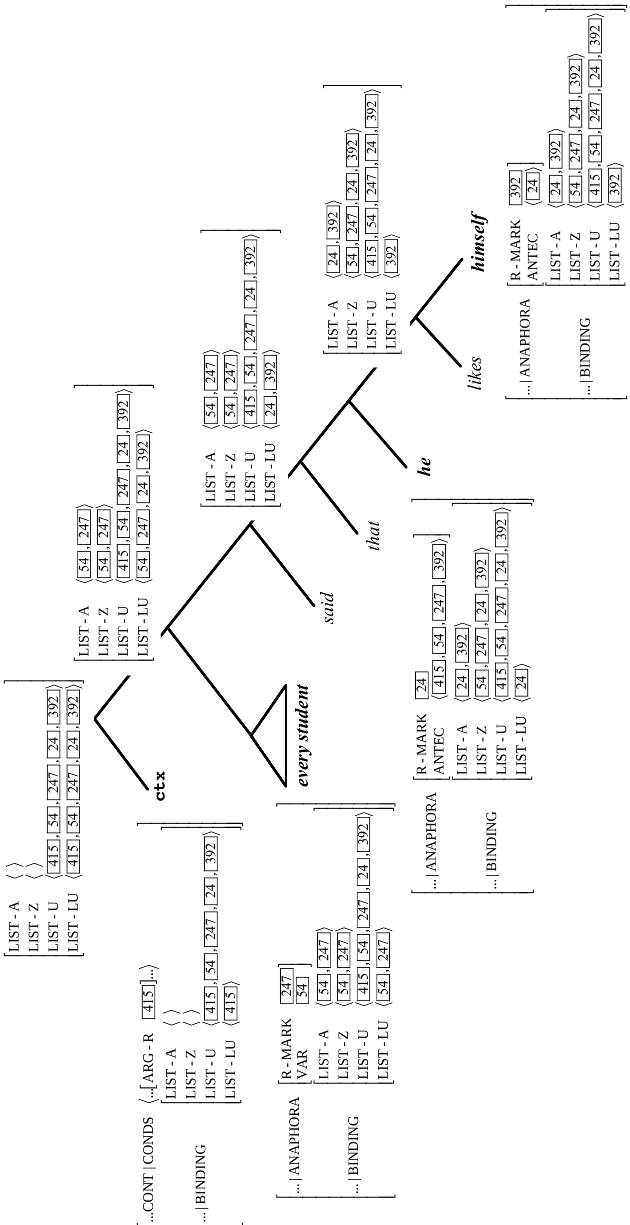


Figure 1: Partial grammatical representation of *Every student said he likes himself.*

Every phrase contributes to the global anaphoric context by passing the tag of its reference marker into its own LIST-LU.

In the case of a quantificational NP, like *every student*, two tags are passed, corresponding to the VAR value $\boxed{54}$ — token-identical with the DREF value of the restrictor and providing for bound-variable anaphora interpretations — and the R-MARK value $\boxed{247}$ — providing for e-type anaphora.

While the semantic types of anaphora — including bound-variable and e-type anaphora — are addressed in further detail in Section 5.2, it is of note at this point that a DRT account of e-type anaphora is followed here.[60] Accordingly, a quantificational NP contributes a plural reference marker to the semantic representation of the discourse that may serve as the antecedent in (e-type) anaphoric links. In a sentence like *Every bald man snores*, for instance, the quantificational NP contributes the plural reference marker which stands for the bald men that snore. Such marker is introduced in the discourse representation via the application of the DRT Abstraction operator Σ, which takes the restrictor and the nuclear scope of the determiner and introduces the plural marker that satisfies the corresponding semantic conditions.[61]

In order to incorporate such an account of e-type anaphora into Underspecified DRT,[62] the reference marker standing for the plurality satisfying the semantic condition obtained with Σ-abstraction, in the CONDS value of a determiner, is made token-identical with its R-MARK value. The *synsem* of the lexical entry for *every*, for instance, results thus as follows, where $\boxed{1}$ is the marker obtained via Σ-abstraction:

[60] (Kamp & Reyle 1993: 311ff).
[61] (Kamp & Reyle 1993: 310).
[62] (Frank & Reyle 1995).

$$
\begin{bmatrix}
\text{LOC|CONT} & \begin{bmatrix}
\text{LS} & \begin{bmatrix} \textit{L-MAX} & \boxed{4} \\ \textit{L-MIN} & \boxed{5} \end{bmatrix} \\
\text{SUBORD} & \left\{ \boxed{4} > \boxed{3}, \boxed{4} > \boxed{5}, \boxed{8} \geq \boxed{5} \right\} \\
\text{CONDS} & \left\{ \begin{matrix}
\begin{bmatrix} \text{LABEL} & \boxed{4} \\ \text{REL} & \textit{every} \\ \text{RES} & \boxed{3} \\ \text{SCOPE} & \boxed{5} \end{bmatrix}, \begin{bmatrix} \text{LABEL} & \boxed{3} \\ \text{DREF} & \boxed{2} \end{bmatrix}, \\
\begin{bmatrix} \text{LABEL} & \boxed{5} \\ \text{DREF} & \boxed{7} \end{bmatrix}, \begin{bmatrix} \text{LABEL} & \boxed{8} \\ \text{REL} & \Sigma\textit{-abstraction} \\ \text{ARG1} & \boxed{2} \\ \text{ARG2} & \boxed{7} \\ \text{DREF} & \boxed{1} \end{bmatrix}
\end{matrix} \right\} \\
\text{ANAPH} & \begin{bmatrix} \textit{R-MARK} & \boxed{1} \\ \textit{VAR} & \boxed{2} \end{bmatrix}
\end{bmatrix} \\
\text{NONLOC|BIND} & \begin{bmatrix}
\text{LIST-A} & \textit{list(refm)} \\
\text{LIST-Z} & \textit{list(refm)} \\
\text{LIST-U} & \textit{list(refm)} \\
\text{LIST-LU} & \left\langle \boxed{2}, \boxed{1} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

Contribution by the context Let us consider now at how the representation of the context is encoded in each NP.

It should be noted that the suitable values of LIST-A, LIST-Z and LIST-U at the different NP nodes are enforced by the combined effect of the three Clauses of BDP.

Due to, respectively, BDP–Clause II (iii.) and BDP–Clause I (iii.i.), LIST-Z and LIST-U values result from token-identity, respectively, with LIST-Z and with LIST-U of the immediately dominating node in the constituency tree — that is the case, for instance, with the lists $\langle \boxed{54}, \boxed{247} \rangle$ and $\langle \boxed{415}, \boxed{54}, \boxed{247}, \boxed{24}, \boxed{392} \rangle$ in the non-pronoun *every student* and in the sentential node dominating it.

Due to BDP–Clause II (i.), LIST-A value, in turn, is obtained via token-identity with LIST-A of the subcategorising predicator — that is the case, for instance, with the list $\langle \boxed{24}, \boxed{392} \rangle$ in the reflexive *himself* and in its predicator *likes*.

Realisation of anaphoric potential As to the anaphoric nominals, let us consider how their anaphoric potential is circumscribed in each specific occurrence.

The value of ANTEC is a list that records the grammatically admissible antecedents of the corresponding anaphor at the light of binding constraints.

As the result of the relational constraint *principleA*/2, the semantic representation of the reflexive *himself* includes the attribute ANTEC with the singleton list ⟨ $\boxed{24}$⟩ as value, indicating that the only antecedent candidate available in this sentence is the pronoun in the embedded clause whose reference marker is identified as $\boxed{24}$ in its own semantic representation.

The semantic representation of the pronoun *he*, in turn, includes the feature ANTEC with a value that is the list of its antecedent candidates, ⟨ $\boxed{415}$, $\boxed{247}$, $\boxed{54}$, $\boxed{392}$⟩, thus indicating that, in this sentence, the pronoun can be anaphorically linked to every nominal except itself, in line with the relational constraint *principleB*/3.

This ANTEC list includes antecedent candidates for the pronoun that will be dropped out by preferences or constraints on anaphoric links other than just the grammatical binding constraint expressed in Principle B. For instance, the plural reference marker $\boxed{247}$, which is the R-MARK value of *every student*, will eventually be excluded by the anaphora resolver given that the singular pronoun *he* cannot entertain an e-type anaphoric link with a universally quantified NP whose reference marker obtained by Σ-abstraction is a plurality.

Also the marker $\boxed{392}$ of the reflexive will be eventually discarded from this ANTEC list as a suitable antecedent by the resolver system since this would lead to an interpretive loop where the pronoun and the reflexive would be the sole antecedents of each other.

Non-linguistic context Finally, in order to illustrate how the non linguistic context may be represented in the linguistic representation of sentences, in this example, the reference marker $\boxed{415}$ was introduced in the semantic representation of the `ctx` node.

The CONDS value of this node is meant to capture the possible contribution of the non-linguistic context at stake for the interpretation of the discourse. Like in the lexical entries of nominals, in the feature representation of *ctx*, the reference marker $\boxed{415}$ is integrated in the LIST-LU value. By the effect of BDP–Clause I, this reference marker ends up added to the list of all reference markers, from both the linguistic discourse and the non-linguistic context, which is the shared value of features LIST-LU and LIST-U at the top node in Figure 1.

Abbreviations

BDP - Binding Domains Principle
DRT - Discourse Representation Theory
HPSG - Head-Driven Phrase Structure Grammar
UDRT - Underspecified Discourse Representation Theory

Acknowledgements

The research reported in this chapter was partially supported by the Research Infrastructure for the Science and Technology of Language (PORTULAN CLARIN).

References

Abeillé, Anne, Danièle Godard, Philip Miller & Ivan Sag. 1998. French bounded Dependencies. In Sergio Balari & Luca Dini (eds.), *Romance in HPSG*, 1–54. Stanford: CSLI Publications.

Abeillé, Anne, Danièle Godard & Ivan Sag. 1998. Two Kinds of Composition in French complex Predicates. In Andreas Kathol, Ehrard Hinrichs & Tsuneko Nakazawa (eds.), *Complex Predicates in non transformational Syntax*, 1–41. Cambridge MA: Academic Press.

Asher, Nicholas & Hajime Wada. 1989. A Computational Account of Syntactic, Semantic and Discourse Principles for Anaphora Resolution. *Journal of Semantics* 6. 309–344.

Branco, António. 1996. Branching Split Obliqueness at the Syntax-Semantics Interface. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING'96)*, 149–156.

Branco, António. 1998a. The Logical Structure of Binding. In *Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics (COLING-ACL'98)*, 181–187.

Branco, António. 1998b. Towards a Lean Constraint-based Implementation of Binding Theory. In Simon Botley & Tony McEnery (eds.), *New Approaches to Discourse Anaphora: Proceedings of the Second Colloquium on Discourse, Anaphora and Anaphor Resolution*, vol. 11, 54–67. Lancaster: Lancaster University.

Branco, António. 1999. *Reference Processing and its Universal Constraints*. Lisbon: University of Lisbon dissertation.

Branco, António. 2000a. Binding Classes and Search Optimisation in Anaphor Resolution. In Paul Baker, Andrew Hardie, Tony McEnery & Anna Siewierska (eds.), *Proceedings of the 3rd Discourse Anaphora and Anaphor Resolution Colloquium (DAARC'2000)*, 125–132. Lancaster: Lancaster University.

Branco, António. 2000b. Binding Constraints as Instructions of Binding Machines. In *Proceedings of the 18th International Conference on Computational Linguistics (COLING'2000)*, 104–111.

Branco, António. 2000c. *Reference Processing and its Universal Constraints*. Lisbon: Publicações Colibri.

Branco, António. 2000d. Specification and Verification of Binding Constraints: An Integrated Account. In Detmar Meurers, Shuly Wintner & Erhard Hinrichs (eds.), *Proceedings of Workshop on Linguistic Theory and Grammar Implementation*, 201–218. Birmingham: 12th European Summer School in Logic, Language & Information (ESSLLI2000).

Branco, António. 2001. Duality and Anaphora. In Robert van Roy & Martin Stokhof (eds.), *Proceedings of the 13th Amsterdam Colloquium (AC'2001)*, 49–54. Amsterdam: ILLC.

Branco, António. 2002a. Binding Machines. *Computational Linguistics* 28. 1–18.

Branco, António. 2002b. Integrating Binding Principles into Formal Grammar Processing: An HPSG Account. In Tony McEnery António Branco & Ruslan Mitkov (eds.), *Proceedings of the 4th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC'2002)*, 19–24. Lisbon: Edições Colibri.

Branco, António. 2002c. Without an Index: a Lexicalist Account of Binding Theory. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the 8th International on Head-Driven Phrase Structure Grammar Conference (HPSG'2001)*, 71–86. Stanford: CSLI Publications.

Branco, António. 2003. Anaphor Resolution: Is the Search Optimization Flawed? *Psychologica* 32. 77–89.

Branco, António. 2004. Cognitive Anaphor Resolution and the Binding Principles. In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (eds.), *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*, 629–640. Coimbra: Almedina.

Branco, António. 2005a. Anaphoric Constraints and Dualities in the Semantics of Nominals. *Journal of Language Logic and Information* 2. 149–171.

Branco, António. 2005b. Reflexives: Escaping Exemption via Domain Reshuffling. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar (HPSG2005)*, 467–481. Stanford: CSLI Publications.

Branco, António. 2006. Binding Symmetries and Nominal Dualities. In Klaus von Heusinger & Ken Turner (eds.), *Where Semantics Meets Pragmatics*, vol. 16 (Current Research in the Semantics/Pragmatics Interface Series), 225–236. Amsterdam: Elsevier.

Branco, António. 2007. Null Subjects are Reflexives, not Pronouns. *Lecture Notes in Artificial Intelligence* 4410. 59–76.

Branco, António & Palmira Marrafa. 1997. Subject-oriented and non Subject-oriented Long-distance Anaphora: an Integrated Approach. In Byung-Soo Park & Jong-Bok Kim (eds.), *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC1997)*, 21–30. Seoul: Kyung Hee University.

Branco, António & Palmira Marrafa. 1999. Long-distance Reflexives and the Binding Square of Opposition. In Gert Webelhuth, Jean-Pierre Koening & Andreas Kathol (eds.), *Lexical and Constructional Aspects of Linguistics Explanation*, chap. 11, 163–177. Stanford: CSLI Publications.

Carbonell, Jaime & Ralf Brown. 1988. Anaphora Resolution: A Multi-strategy Approach. In *Proceedings of the 12th International Conference on Computational Linguistics (COLING'88)*, 96–101.

Chierchia, Gennaro. 1995. *Dynamics of Meaning: Anaphora, Presupposition and the Theory of Grammar*. Chicago: The University of Chicago Press.

Chomsky, Noam. 1980. On Binding. *Linguistic Inquiry* 11. 1–46.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of Language*. New York: Praeger.

Correa, Nelson. 1988. A Binding Rule for Government-binding Parsing. In *Proceedings of the 12th International Conference on Computational Linguistics (COLING'88)*, 123–129.

Dalrymple, Mary. 1993. *The Syntax of Anaphoric Binding*. Stanford: CSLI Publications.

Dimitriadis, Alexis, Martin Everaert, Tanya Reinhart & Eric Reuland. 2005. *Anaphora typology database*. University of Utrectht: http://languagelink.let.uu.nl/anatyp/.

Eschenbach, Carole, Christopher Habel, Michael Herweg & Klaus Rehkaemper. 1989. Remarks on Plural Anaphora. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics (ACL'89)*, 161–167.

Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry* 2. 337–362.

Fong, Sandiway. 1990. Free Indexation: Combinatorial Analysis and a Compositional Algorithm. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics (ACL'90)*, 105–110.

Frank, Anette & Uwe Reyle. 1995. Principle Based Semantics for HPSG. In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics (EACL'95)*, 9–16.

Gawron, Jean Mark & Stanley Peters. 1990. *Anaphora and Quantification in Situation Semantics*. Stanford: CSLI Publications.

Golde, Karin. 1999. *Binding Theory and Beyond*. Columbus: Ohio State University dissertation.

Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. Amherst: University of Massachusetts dissertation.

Hellan, Lars. 1988. *Anaphora in Norwegian and the Theory of Grammar*. Dordrecht: Foris.

Hepple, Mark. 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach*. Edinburgh: University of Edinburgh dissertation.

Higginbotham, James. 1983. Logical Form, Binding and Nominals. *Linguistic Inquiry* 14. 395–420.

Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, 61–90. Stanford, CA: CSLI Publications. Republished as Höhle (2018).

Höhle, Tilman N. 2018. An architecture for phonology. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 571–607. Berlin: Language Science Press. Originally published as Höhle (1999). DOI:10.5281/zenodo.1145680

Huang, C.-T. James & C.-C. Jane Tang. 1991. The Local Nature of the Long-distance Reflexive in Chinese. In Jan Koster & Eric Reuland (eds.), *Long-Distance Anaphora*, 263–282. Cambridge: CUP.

Iatridou, Sabine. 1986. An Anaphor not Bound in Its Governing Category. *Linguistic Inquiry* 17. 766–772.

Jackendoff, Ray S. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jacobson, Pauline. 2000. Paycheck Pronouns, Bach-Peters Sentences and Variable-free Semantics. *Natural Language Semantics* 8. 77–155.

Jäger, Gerhard. 2001. *Anaphora and Ellipsis in Type Logical Grammar*. Course notes. Utrecht. Utrecht Institute of Linguistics-OTS PhD courses.

Johnson, Mark & Ewan Klein. 1990. Discourse, Anaphora and Parsing. In *Proceedings of the 28th Annual Meeting of the Association of Computational Linguistics (ACL'90)*, 197–302.

Kamp, Hans. 1981. A Theory of Truth and Discourse Representation. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal Methods in the Study of Language*, 277–322. Amsterdam: Mathematical Centre.

Kamp, Hans & Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer.

Karttunen, Lauri. 1976. Discourse Referents. In James McCawley (ed.), *Syntax and Semantics 7: Notes from the Linguistic Underground*, 363–385. New York: Academic Press.

Kiss, Tibor. 2001. Anaphora and Exemptness: A Comparative Treatment of Anaphoric Binding in German and English. In Dan Flickinger & Andreas Kathol (eds.), *Proceedings of the 7th International on Head-Driven Phrase Structure Grammar Conference (HPSG'2000)*, 182–197. Stanford: CSLI Publications.

Koenig, Jean-Pierre. 1999. Inside-Out Constraints and Description Languages for HPSG. In Gert Webelhuth, Jean-Pierre Koening & Andreas Kathol (eds.), *Lexical and Constructional Aspects of Linguistics Explanation*, chap. 17, 265–280. Stanford: CSLI Publications.

Koning, Ekkehard. 1991. Concessive Relations as the Dual of Causal Relations. In Dietmar Zäfferrer (ed.), *Semantic Universals and Universal Semantics*, 190–209. Dordrecht: Foris.

Koster, Jan & Eric Reuland (eds.). 1991. *Long-distance anaphora.* CUP: Cambridge.

Kuno, Susumu. 1987. *Functional Syntax.* Chicago: University of Chicago Press.

Lappin, Shalom & Herbert Leass. 1994. An Algorithm for Pronominal Anaphora Resolution. *Computational Linguistics* 20. 535–561.

Link, Godehard. 1983. The Logical Analysis of Plural and Mass Terms: A Lattice-theoretical Approach. In Arnim von Stechow, Christoph Schwarze & Rainer Bäuerle (eds.), *Meaning, Use and Interpretation of Language*, 302–323. Berlin: Walter de Gruyter.

Löbner, Sebastian. 1987. Quantification as a Major Module of Natural Language Semantics. In Jeroen Groenendijk, Dick de Jongh & Martin Stokhof (eds.), *Studies in DRT and the Theory of Generalized Quantifiers*, 53–85. Dordrecht: Foris.

Löbner, Sebastian. 1989. German *schon-erst-noch*: An Integrated Analysis. *Linguistics and Philosophy* 12. 167–212.

Löbner, Sebastian. 1999. Why German *schon* and *noch* are still Duals: a Reply to van der Auwera. *Linguistics and Philosophy* 22. 45–107.

Manning, Chris & Ivan Sag. 1999. Dissociations between Argument Structure and Grammatical Relations. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 63–78. Stanford: CSLI.

Manzini, Rita & Kenneth Wexler. 1987. Parameters, Binding Theory and Learnability. *Linguistic Inquiry* 18. 413–444.

Mitkov, Ruslan. 1997. Factors in Anaphora Resolution: They are not the only Things that Matter. In *Proceedings of the ACL/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution*, 36–51.

Morril, Glyn. 2000. *Type-Logical Anaphora.* Report de Recerca LSI-00-77-R. Barcelona. Departament de Llenguatges i Sistemes Informátics, Universitat Politécnica de Catalunya.

Poesio, Massimo & Renata Vieira. 1998. A Corpus-based Investigation of Definite Description Use. *Computational Linguistics* 24. 183–216.

Pollard, Carl & Ivan Sag. 1987. *Information-based Syntax and Semantics.* Stanford: CSLI Publications.

Pollard, Carl & Ivan Sag. 1992. Anaphors in English and the Scope of Binding Theory. *Linguistic Inquiry* 23. 261–303.

Pollard, Carl & Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar.* Chicago: The University of Chicago Press.

Reinhart, Tanya. 1983. Coreference and Bound Anaphora: A Restatement of the Anaphora Questions. *Linguistics and Philosophy* 6. 47–88.

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24. 657–720.

Reyle, Uwe. 1993. Dealing with Ambiguities by Underspecification: Construction, Representation and Deduction. *Journal of Semantics* 10. 123–179.

Rich, Elaine & Susann LuperFoy. 1988. An Architecture for Anaphora Resolution. In *Proceedings of the 2nd Conference on Applied Natural Language Processing (ANLP'88)*, 18–24.

Richter, Frank, Manfred Sailer & Gerald Penn. 1999. A Formal Interpretation of Relations and Quantification in HPSG. In Gert-Jan Kruijff Gosse Bouma Erhard Hinrichs & Richard Oehrle (eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, 281–298. Stanford: CSLI Publications.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Seeley, Daniel. 1993. Binding Plural Pronominals. In Katherine Beals, Gina Cooke, David Kathman, Sotaro Kita, Karl-Erik McCullough & David Testen (eds.), *CLS29: Papers from the 29th Regional Meeting of the Chicago Linguistic Society*, vol. 2, 305–317. Chicago: University of Chicago.

Seuren, Peter. 1985. *Discourse Semantics.* Oxford: Blackwell.

Smessaert, Hans. 1997. Aspectual Duality Regained. In Paul Dekker, Martin Stokhof & Yde Venema (eds.), *Proceedings of the 11th Amsterdam Colloquium*, 271–276. Amsterdam: ILLC.

Stalnaker, Robert. 1998. On the Representation of Context. *Journal of Logic, Language and Information* 7. 3–19.

Szabolcsi, Anna. 1989. Bound Variables in Syntax (are there any?) In Renate Bartsch, Johan van Benthem & Peter van Emde Boas (eds.), *Semantics and Contextual Expressions*, 295–318. Dordrecht: Foris.

ter Meulen, Alice. 1988. The Semantic Properties of English Aspectual Verbs. In *NELS*, vol. 21. Amherst: University of Massachusetts.

van Benthem, Johan. 1991. Linguistic Universals in Logical Semantics. In Dietmar Zäfferrer (ed.), *Semantic Universals and Universal Semantics*, 17–36. Dordrecht: Foris.

Varlokosta, Spyridoula & Norbert Hornstein. 1993. A Bound Pronoun in Modern Greek. *Natural Language and Linguistic Theory* 11. 175–195.

Wechsler, Stephen. 1999. HPSG, GB, and the Balinese Bind. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 179–196. Stanford: CSLI.

Xue, Ping, Carl Pollard & Ivan Sag. 1994. A New Perspective on Chinese *Ziji*. In *Proceedings of the the West Coast Conference on Formal Linguistics (WCCFL'94)*. Stanford: CSLI Publications.

Zribi-Hertz, Anne. 1989. Anaphor Binding and Narrative Point of View: English Reflexive Pronouns in Sentence and Discourse. *Language* 65. 695–727.

Part III

Other levels of description

Chapter A

Phonology

Jesse Tseng
Université Paris Diderot

Put abstract here with \abstract.

1 Introduction: PHONOLOGY in the HPSG sign

The PHONOLOGY attribute in (Pollard & Sag 1987) and (Pollard & Sag 1994):

- rudimentary PHON value
- basic Phonology Principle constrained by Linear Precedence rules: corresponds to simple terminal spell-out of the phrase structure tree
- "Phonology-Free Syntax" (Miller et al. 1997): PHON information inaccessible for selection via SYNSEM

There has been relatively little work within HPSG on phonological representation and the analysis of phonological phenomena. Most references to the PHON attribute use it simply as a lexical identifier, or they are dealing with phenomena at the phonology-syntax interface (e.g. constituent order, ellipsis). For such applications, the actual content of the PHON value is unimportant. These topics are covered in other chapters.

2 Phonological representations in HPSG

Proposals for the detailed content of PHON values:

- encoding of phonological constituents (Bird & Klein 1994; Klein 2000; Höhle 1999)

Jesse Tseng. 2019. Phonology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 669–672. Berlin: Language Science Press. DOI:??

- syllable structure Tseng (2008)
- metrical phonology (Klein 2000; Bonami & Delais-Roussarie 2006)

3 Phonological analysis in HPSG

- principles of constraint-based phonology vs derivational phonology (Bird & Klein 1994): compositionality, monotonicity
- compositional construction of prosodic structure in parallel with phrase structure (Klein 2000)

But HPSG is formally compatible with many approaches, and there is as yet no emerging consensus among practitioners.

- Finite state phonology (Bird 1992; 1995)
- need for abstract underlying forms (Skwarski 2009); phonologically empty categories
- OT in HPSG (Orgun 1996)

4 Specific phenomena and case studies

- shape conditions (Asudeh & Klein 2002)
- French (Tseng 2003; Bonami et al. 2004)
- phonological idioms [already covered in Manfred's chapter]
- …

Abbreviations

Acknowledgements

References

Asudeh, Ash & Ewan Klein. 2002. Shape conditions and phonological context. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 Conference, Norwegian University of Science and Technology, Trondheim*, 20–30. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2001/, accessed 2015-12-23.

Bird, Steven. 1992. Finite-state phonology in HPSG. In *Proceedings of coling 92*, 74–80. Nantes, France.

Bird, Steven. 1995. *Computational phonology: A constraint-based approach* (Studies in Natural Language Processing). Cambridge, UK: Cambridge University Press.

Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491.

Bonami, Olivier, Gilles Boyé & Jesse Tseng. 2004. An integrated approach to French liaison. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 29–45. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/FG/2004/, accessed 2018-2-25.

Bonami, Olivier & Elisabeth Delais-Roussarie. 2006. Metrical phonology in HPSG. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, 39–59. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2006/, accessed 2018-2-25.

Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, 61–90. Stanford, CA: CSLI Publications. Republished as Höhle (2018).

Höhle, Tilman N. 2018. An architecture for phonology. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 571–607. Berlin: Language Science Press. Originally published as Höhle (1999). DOI:10.5281/zenodo.1145680

Klein, Ewan. 2000. A constraint-based approach to English prosodic constituents. In *38th Annual Meeting of the Association for Computational Linguistics: Proceedings of the conference*, 217–224. Hong Kong: Association for Computational Linguistics.

Miller, Philip H., Geoffrey K. Pullum & Arnold M. Zwicky. 1997. The principle of Phonology-Free Syntax: Four apparent counterexamples in French. *Journal of Linguistics* 33(1). 67–90.

Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*. University of California, Berkeley dissertation.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Skwarski, Filip. 2009. Accounting for underlying forms in HPSG phonology. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 318–337. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2009/skwarski.pdf, accessed 2018-2-25.

Tseng, Jesse. 2003. Edge features and French liaison. In Jongbok Kim & Stephen Mark Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 313–333. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/3/, accessed 2018-9-25.

Tseng, Jesse. 2008. The representation of syllable structure in HPSG. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 234–252. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/, accessed 2018-2-25.

Chapter B

Morphology

Berthold Crysmann
Université Paris Diderot

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse

Berthold Crysmann. 2019. Morphology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 673–674. Berlin: Language Science Press. DOI:??

rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter C

Semantics

Jean-Pierre Koenig
University at Buffalo

Frank Richter
Goethe Universität Frankfurt

This chapter discusses the integration of theories of semantic representations. It focuses on those aspects that are specific to HPSG and, in particular, recent approaches that make use of underspecified semantic representations, as they are quite unique to HPSG.

1 Introduction

A semantic level of description is more integrated into the architecture of HPSG than in many frameworks (although, in the last couple of decades, the integration of syntax and semantics has become tighter overall; see Heim & Kratzer (1998) for minimalism, for example). Every node in a syntactic tree includes all appropriate levels of structure, phonology, syntax, semantics, and pragmatics so that *local* interaction between all these levels is in principle possible within the HPSG architecture. The architecture of HPSG thus follows the spirit of the rule-to-rule approach advocated in Bach (1976) and more specifically Klein & Sag (1985) to have every syntactic operation matched by a semantic operation (the latter, of course, follows the Categorial Grammar lead, broadly speaking). But, as we shall see, only the spirit of the rule-to-rule approach is adhered to, as there can be more than one semantic operation per class of syntactic structures, depending on the semantic properties of the expressions that are syntactically composed. The built-in interaction between syntax and semantics within HPSG

Jean-Pierre Koenig & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

is evidenced by the fact that Pollard & Sag (1987), the first book length introduction to HPSG, spends a fair amount of time on semantics and ontological issues, much more than customary in syntax-oriented books at the time.

But, despite the centrality of semantics within the HPSG architecture, not much comprehensive work on the interface between syntax and semantics was done until the late 90's, if we exclude work on the association of semantic arguments to syntactic valents in the early 90's (see the chapter on argument structure): the formal architecture was ripe for research on the interface between syntax and semantics, but few stepped in. Early work on semantics in HPSG focused on scoping issues, as HPSG surface-oriented syntax presents interesting challenges to scope. This is what Pollard & Sag (1987); Pollard & Sag (1994) focus on most. Scope of modifiers is also an area that was of importance and received attention for the same reason both in Pollard & Sag (1994) and Kasper (1997). Ginzburg & Sag (2000) is the first study not devoted to argument structure to leverage the syntactic architecture of HPSG to model the semantics of a particular area of grammar, in this case interrogatives and questions.

The real innovation HPSG brought to the interface between syntax and semantics is the use of underspecification, starting with Minimal Recursion Semantics (Copestake et al. 2001; 2005) and Lexical Resource Semantics (Richter & Sailer 2001). The critical distinction between grammars as descriptions of admissible structures and models of these descriptions makes it possible to have a new way of thinking about the meaning contributions of lexical entries and constructional entries: underspecification is the other side of descriptions.

2 A situation semantics beginning

The semantic side of HPSG was initially rooted in Situation Semantics (Pollard & Sag 1987: ch.4). The choice of Situation Semantics is probably somewhat a matter of happenstance and overall nothing too crucial depended on that choice (and other choices have been explored since, as we detail below). Our statement that the Situation Semantics underpinnings of HPSG's early approach to semantics should not be construed as implying the choice was unconsequential. There were several interesting aspects of this choice for the study of the interface between syntax and semantics that is integral to any grammatical framework. We briefly mention a few here. A first interesting aspect of this choice is that the identification of arguments was not through an ordering but via keywords standing for role names, something that made it easier to model argument structure in subsequent work (see the chapter on argument structure). A second aspect is the built-in "intentionality" of Situation Semantics. Since atomic formulas in Situation Semantics denote circumstances rather than truth values and circumstances are more finely individuated than truth-values, the need to resort to possible world semantics to properly characterize differences in the meaning of basic verbs, for example, is avoided. A third aspect of Situation Semantics that played an important role in HPSG is parameters. Parameters are akin to variables, except that there is no assumption that corresponds to the assumption that variables are bound (parameters are thus akin to discourse referents in Discourse Representation Theory, Kamp & Reyle 1993) and parameters can be of different kinds thus allowing for an easy semantic classification of types of NPs, something that HPSG's binding theory makes use of.

Parameters also play an important role in early accounts of quantification; these accounts rely on restrictions on parameters that constrain how variables are anchored, akin to predicative conditions on discourse references in DRT. Restrictions on parameters are illustrated with (1), the (non-empty) semantic content of the common noun *donkey*, where the variable $\boxed{1}$ is restricted to individuals that are donkeys, as expressed by the value of the attribute REST.

(1) $$\begin{bmatrix} \text{VAR} & \boxed{1} \\ \text{REST} & \begin{bmatrix} \text{RELN} & \textit{donkey} \\ \text{INST} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Because indices are restricted variables/parameters, the model of quantification proposed in Pollard & Sag (1987) involves restricted quantifiers. Consider the sentence *Every donkey sneezes* and its semantic representation in (2) (op.cit., p.109).

(2) $$\begin{bmatrix} \text{QUANT} & \begin{bmatrix} \text{DET} & \textit{forall} \\ \text{IND} & \begin{bmatrix} \text{VAR} & \boxed{1} \\ \text{REST} & \begin{bmatrix} \text{RELN} & \textit{donkey} \\ \text{INST} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SCOPE} & \begin{bmatrix} \text{RELN} & \textit{sneeze} \\ \text{SNEEZER} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

The subject NP contributes the value of the attribute QUANT, while the verb contributes the value of SCOPE. The quantifier includes information on the type of quantifier contributed by the determiner (a universal quantifier in this case) and the index (a parameter restricted by the common noun).

Because HPSG is a sign-based grammar, each constituent includes a PHONology and SEMantic component as well as a SYNtactic level of representation (along with other possible levels). Compositionality has thus always been directly incorporated by principles that regulate the value of the mother's SEM attribute, given the SEM values of the daughters and their mode of syntactic combination (as manifested by their syntactic properties). Different approaches to semantics within HPSG propose variants of a semantic principle that constrains this relation. The semantic principle of Pollard & Sag (1987: p.109) is stated in English in (3) (we assume for simplicity that there is a single complement daughter; Pollard and Sag define semantic composition recursively for cases of multiple complement daughters).

(3) a. If the semantic content of the head-daughter $\boxed{1}$ is of sort *circumstance* and the semantic content of the complement daughter $\boxed{2}$ is of sort *quantifier*, the semantic content of the mother is $\begin{bmatrix} \text{QUANT} & \boxed{2} \\ \text{SCOPE} & \boxed{1} \end{bmatrix}$

b. Otherwise, the semantic content of the head-daugther and the mother are identical.

The fact that the semantics principle in (3) receives a case-based definition is of note. Because HPSG is monostratal and there is only one stratum of representationn (see Ladusaw (1988) for the difference between levels and strata), the semantic contribution of complement daughters varies and the semantic principle must receive a case-based definition. In other words, syntactic combinatorics is less varied than semantic combinatorics. The standard way of avoiding violating compositionality (the fact that semantic composition is a *function*) is to have a case-based definition of the semantic effect of combining a head-daughter with its complements, a point already made in Partee (1984). As (3) shows, HPSG has followed this practice since its beginning. The reason is clear: one cannot maintain a surface-oriented approach to syntax, where syntax is "simpler" to borrow a phrase from Culicover & Jackendoff (2005), without resorting to case-based definitions of the semantic import of syntactic combinatorics.

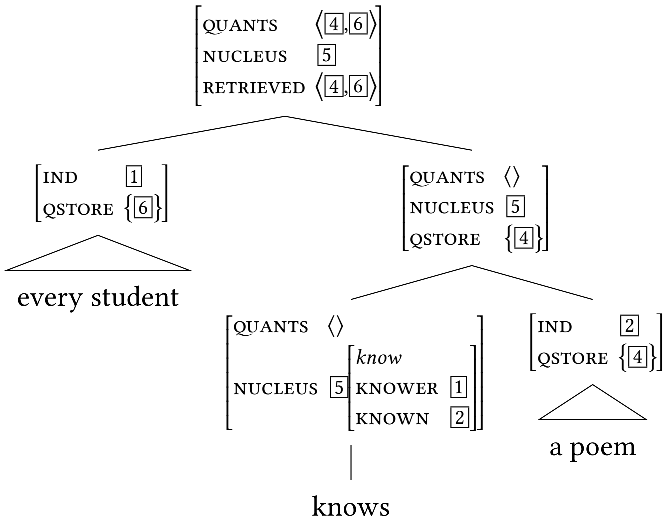
3 Scope relations in HPSG

In mainstream generative grammar, there is an assumption that syntactic constituency reflects semantic constituency at one stratum of representation. In the case of quantifier scope in works like May (1985), this means that quantified expressions are moved out of their surface position and raised to a position where they can receive their proper scope through Quantifier Raising (and/or Quantifier Lowering, see, among others Hornstein 1995). Of course, such a move requires multiple strata, as there is little evidence that quantifier scope affects surface syntactic structure. The semantic principle and the representation of quantifier meanings outlined in Pollard & Sag (1987) and briefly presented in the previous section was not flexible enough to model the relation between single syntactic structures and multiple scopal relations. As Pollard and Sag explicitly recognized, their semantic principle only models left to right scopal relations, i.e. quantifiers that are expressed by a complement that is to the left of another complement have wide scope with respect to that quantifier. So-called inverse scope, including the fact that quantifiers in object position can outscope quantifiers in subject position cannot be modeled by the kind of semantic principle they propose. Much of the discussion of semantics within HPSG in the 90's pertains to improving how scope is modeled, both the scope of quantifiers and scope of adjuncts. We discuss each in turn in this section.

3.1 Quantifier scope

HPSG's "standard" model of the interface between the syntax and semantics of phrases that contain quantifiers until the mid 2000's adapted to HPSG the approach proposed in Cooper (1975; 1983), i.e. so-called Cooper storage: When a quantified expression combines with another expression, the quantifier is put in a store and various scopal relations correspond to the various nodes at which the quantifier can be retrieved from storage. Within HPSG, quantifier storage involves a QSTORE attribute where each quantifier starts and at each node quantifiers are either retrieved (part of the RETRIEVED list) or continue to be on the mother's QSTORE. The relative scope of quantifiers itself is determined by the ordering of quantifiers on the QUANTS list. The simplified tree in (4) from Pollard & Sag (1994: p.324) illustrates the inverse scope reading of the English sentence *Every student knows a poem.*

(4)



Both subject and object quantifiers start with their quantifiers (basically, something very similar to the representation in (2)) in a QSTORE. Since the reading of interest is the one where *a poem* outscopes *every student*, the quantifier introduced by *a poem* cannot be retrieved at the VP level. This is because the value of QUANTS is the concatenation of the value of RETRIEVED with the QUANTS value of the head daughter. Were the quantifier introduced by *a poem* ($\boxed{4}$) retrieved at the VP level, the sole quantifier retrieved at the S level, the quantifier introduced by *every student*, would outscope it. So, the only way for the quantifier introduced by *a poem* to outscope the quantifier introduced by *every student* is for the former to be retrieved at the S node just like the latter. Simplifying somewhat for presentational purposes, two principles govern how quantifiers are passed on from head-daughter to mothers and how quantifier scope is assigned for retrieved quantifiers; they are stated in (5) (adapted from Pollard & Sag 1994: p.322-323).

(5) a. In a headed phrase, the RETRIEVED value is a list whose set of elements is a subset of the union of the QSTORE of the daughters; the QSTORE value is the relative complement of that set.

b. In a headed phrase (of sort *psoa* or parametrized state of affairs), the QUANTS value is the concatenation of the RETRIEVED value and the QUANTS value of the semantic head.

(5a) ensures that quantifiers in storage are passed up the tree, except for those that are retrieved; (5b) ensures that quantifiers that are retrieved outscope quantifiers that were retrieved lower in the tree. Narrow scope of quantifiers that occur in object position entails retrieval at the VP level; wide scope of quantifiers that occur in object position entails retrieval at the S level. But retrieval at the S level of quantifiers that occur in object position does not entail wide scope, as the order of two quantifiers in the same RETRIEVED list (i.e., retrieved at the same node) is unconstrained. Constraints on quantifier retrieval and scope underdetermines quantifier scope. To ensure that quantifiers are retrieved sufficiently "high" in the tree to bind bound variable uses of pronouns, e.g., *her* in (6), Pollard and Sag propose the constraint in (7).

(6) One of her$_i$ students approached [each teacher]$_i$. (Pollard & Sag 1994: ex.27a)

(7) A quantifier within a CONTENT value must bind every occurrence of that quantifier's index within that CONTENT value.

The use of Cooper storage allows for a syntactically parsimonious treatment of quantifier scope ambiguities in that no syntactic ambiguity needs be posited to account for what is a strictly semantic phenomenon. But, as Pollard and Sag note (p.328), their model of quantifier scope does not account for the possible narrow scope interpretation of the quantifier *a unicorn* in (8) (the interpretation according to which the speaker does not commit to the existence of unicorns). Raised arguments only occur once, in their surface position, and (5) ensure that quantifiers are never retrieved "lower" than their surface position.

(8) A unicorn appears to be approaching.

Pollard & Yoo (1998) is an attempt to solve that problem, as well as take into account the fact that a sentence such as (9) is ambiguous (i.e., the quantifier *five books* can have wide or narrow scope with respect to the meaning of *believe*). As Pollard and Yoo note, since quantifier storage and retrieval is a property of signs and fillers only share their LOCAL attribute values with arguments of the head (*read*) in (9), the narrow scope reading cannot be accounted for. (8) and (9), among other similar examples, illustrate some of the complexities of combining a surface-oriented approach to syntax with a descriptively adequate model of semantic composition.

(9) Five books, I believe John read. (ambiguous)

Pollard and Yoo's solution (p.419-420) amounts to making quantifier storage and retrieval a property of the LOCAL attribute and to restrict quantifier retrieval to semantically potent heads (so *to* of infinitive VPs cannot be a site for quantifier retrieval). The new feature geometry of *sign* Pollard and Yoo propose is represented in (10). The pool of quantifiers collects the quantifiers on the QSTORE of its selected arguments (members of the SUBJ, COMPS, SPR lists and value of MOD) (except for quantifying determiners and semantically vacuous heads like *to* or *be*) and the two constraints in (11) (Pollard & Yoo 1998: p.423) ensure proper percolation of quantifier store values within headed phrases as well as the semantic order of retrieved quantifiers.

(10)
$$\begin{bmatrix} \textit{sign} & \\ \textsc{phonology} & \textit{list(phon\_string)} \\ \textsc{synsem} & \begin{bmatrix} \textsc{local} & \begin{bmatrix} \textsc{category} & \textit{category} \\ \textsc{content} & \textit{content} \\ \textsc{qstore} & \textit{set(quantifier)} \\ \textsc{pool} & \textit{set(quantifier)} \end{bmatrix} \end{bmatrix} \\ \textsc{retrieved} & \textit{list(quantifier)} \end{bmatrix}$$

(11) a. The pool of the mother of a headed-phrase is identical to the quantifier store of the head daughter.

b. For a semantically nonvacuous lexical head, the QUANTS value is token-identical with the RETRIEVED value.

In a follow-up paper, Przepiórkowski (1998) proposed a strictly lexicalized retrieval mechanism which removes structural ambiguities arising from different possible retrieval sites for quantifiers along a syntactic head path, is compatible with trace-based and traceless analyses of extraction (Pollard & Yoo's analysis only covers trace-based extraction), and shifts all semantic structure under the CONTENT attribute.

3.2 Adjunct scope

HPSG phrase structure schemata are built, for a significant part, around headed structures. In the case of head-complement and head-specifier schemata, syntactic headedness and semantic headedness match. The verb is the head of VPs and clauses and the circumstance or state of affairs denoted by verbs typically takes as arguments the indices of its complements or subjects, and more generally part of the CONTENT value of the verb takes as arguments part of the CONTENT value of its dependents. But in the case of head adjunct structures, syntactic and semantic headedness do not match. The denotation of adjuncts often takes the denotation of heads as arguments. Thus, in (12), fastness is ascribed to Bob's running. Accordingly, the semantics principle distinguishes between head-adjunct structures and other structures, as shown in (13) (Pollard & Sag 1994: p.56). (The principle we cite does not consider quantifier retrieval we discussed in the previous section.)

(12) Bob runs fast.

(13) In a headed phrase, the CONTENT value is token-identical to that of the adjunct daughter if the DTRS value is of sort *head-adj-struc*, and with that of the head daughter otherwise.

Unfortunately, the hypothesis that the content of phrases "projects" from the adjunct in the case of head-adjunct structures leads to difficulties in the case of so-called recursive modification, e.g., (14), as Kasper (1997) shows. The NP in (14) denotes an existential quantifier whose restriction is a plan that is potentially controversial; intuitively speaking, what is potential is the controversiality of the plan, not it being a plan. But, the semantics principle, the syntactic selection of modified expressions by modifiers, and lexical entries for intersective and non-intersective adjectives conspire to lead to the wrong meaning for recursive modification of the kind (14) illustrates.

(14) A potentially controversial plan.

Since *controversial* selects for *plan*, combining their meaning leads to the meaning represented in (15), as *controversial* is an intersective adjective.

(15) $\begin{bmatrix} \textit{nom-obj} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \begin{bmatrix} \text{RELN} & \textit{plan} \\ \text{INST} & \boxed{1} \end{bmatrix} \& \begin{bmatrix} \text{RELN} & \textit{controversial} \\ \text{ARG} & \boxed{1} \end{bmatrix} \end{bmatrix}$

But since adjuncts are the semantic head, the meaning of *potentially controversial plan* will be projected from the meaning of *potentially*, the most deeply embedded adjunct. Now, *potentially* is a conjectural adverb, to adapt to adverbs the classification of adjectives proposed in Keenan & Faltz (1985: p.125). Within HPSG, this means that the meaning of *potentially* is a function that takes the meaning of what it modifies as argument, i.e. the meaning represented in (15). But, this leads to the meaning represented in (16), which is the wrong semantics, as a potentially controversial plan is not a potential plan as Kasper (1997: p.10-11) points out.

(16) $$\begin{bmatrix} \textit{nom-obj} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \begin{bmatrix} \text{RELN} & \textit{potential} \\ \text{ARG} & \begin{bmatrix}\text{RELN} & \textit{plan}\\ \text{INST} & \boxed{1}\end{bmatrix} \& \begin{bmatrix}\text{RELN} & \textit{controversial}\\ \text{ARG} & \boxed{1}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

The problem with Pollard and Sag's semantics principle, when it comes to recursive modification, is clear: semantic selection follows an adjunct path, so to speak, so the most deeply embedded adjunct will have widest scope.

Kasper's solution is to distinguish the inherent meaning of an expression (its internal content) from the meaning it may have in a particular construction, its combinatorial semantics (its external content). With respect to prenominal adjuncts, the internal content corresponds to the content of the adjunct's maximal projection, whereas the external content corresponds to the content of the combination of the meaning of the adjunct with what it modifies. The semantic principle is revised to reflect the distinction between internal and external contents and is provided in (17) (Kasper 1997: p.19).

(17) a. The semantic content of a head-adjunct phrase is token-identical to the MOD|ECONT value of the adjunct daughter, and the MOD|ICONT value of the adjunct daughter is token-identical with the adjunct daughter's CONT.

b. For all other headed phrases, the CONT value is token-identical with the CONT value of the head daughter.

The result of applying the revised semantic principle to *potentially controversial* is provided in (18) and the semantics of *controversial* is provided in (19).

(18) $$\begin{bmatrix} \text{HEAD} & \boxed{4}\begin{bmatrix} \text{MOD} & \begin{bmatrix} \text{ECONT} & \boxed{7}\begin{bmatrix}\text{IND} & \boxed{1}\\ \text{REST} & \boxed{2}\&\boxed{3}\end{bmatrix}\end{bmatrix} \\ \text{ARG} & \begin{bmatrix}\text{CONT} & \begin{bmatrix}\text{IND} & \boxed{1}\\ \text{RESTR} & \boxed{2}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{CONT} & \boxed{3} \\ \text{ICONT} & \boxed{3}\end{bmatrix}$$

$$\begin{bmatrix} \text{HEAD}|\text{MOD} & \begin{bmatrix}\text{ECONT} & \boxed{3}\\ \text{ARG} & \boxed{8}\\ \text{ICONT} & \boxed{3}\end{bmatrix} \\ \text{CONT} & \boxed{3}\begin{bmatrix}\text{RELN} & \textit{potential}\\ \text{ARG} & \boxed{5}\end{bmatrix}\end{bmatrix} \qquad \boxed{8}\begin{bmatrix}\text{HEAD} & \boxed{4}\\ \text{CONT} & \boxed{5}\begin{bmatrix}\text{RELN} & \textit{controversial}\\ \text{INST} & \boxed{1}\end{bmatrix}\end{bmatrix}$$

(19)
$$\begin{bmatrix} \textsc{head}|\textsc{mod} & \begin{bmatrix} \textsc{arg} & \begin{bmatrix}\textsc{cont} & \begin{bmatrix}\textsc{ind} & \boxed{1}\\ \textsc{restr} & \boxed{2}\end{bmatrix}\end{bmatrix}\\ \textsc{icont} & \boxed{3} \\ \textsc{econt} & \begin{bmatrix}\textsc{ind} & \boxed{1}\\ \textsc{restr} & \boxed{2}\&\boxed{3}\end{bmatrix}\end{bmatrix}\\ \textsc{cont} & \begin{bmatrix}\textsc{reln} & \textit{controversial}\\ \textsc{inst} & \boxed{1}\end{bmatrix}\end{bmatrix}$$

Critically, each kind of modifier specifies the ECONT value of what it modifies, i.e. specifies the combinatorial effects it has on the meaning of the modifier and modified combination. Intersective adjectives like *controversial* specify that this effect is intersective, as shown in (19); conjectural adverbs like *potentially*, on the other hand, specify that the ECONT of the adjective they modify is the result of applying their meaning to the adjective, as shown in the left daugther of (18). Now, since the MOD value of the head in a head-adjunct phrase determines the MOD value of the phrase, it means that *controversial* determines the ECONT of what it modifies and, ultimately, the CONT value of the entire phrase *potentially controversial plan*, thus ensuring that its intersectivity is preserved when it combines with a conjectural adverb. Asudeh & Crouch (2002) and Egg (2004) provide more recent solutions to the same problem through the use of a Glue semantics approach to meaning composition within HPSG and semantic underspecification, respectively.

4 Sorting semantic objects

One of the hallmarks of HPSG is that all grammatical objects are assigned a sort (see the chapter on formal foundations for details). This includes semantic objects. Sorting of semantic objects has been used profitably in models of lexical knowledge, in particular in models of argument structure phenomena. We refer the reader to the chapter on argument structure for details and only provide an illustrative example here. Consider the constraint in (20) from Koenig & Davis (2003). It says that all verbs that denote a causal change of state, i.e. are of sort *cause-rel*, link their causer argument to an NP that is the first member of the ARG-ST list. Critically, verbs like *frigthen*, *kill*, *calm* have as meanings a relation that is a subsort of *cause-rel* and are subject to this constraint. Sorting lexical semantic relations thus makes for a compact statement of linking constraints. (The chapter on argument structure provides many more instances of the usefulness of sorting semantic relations, a hallmark of HPSG semantics.)

(20) $\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} & \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \langle \text{NP}, \ldots \rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \langle \text{NP:}\boxed{1}, \ldots \rangle \end{bmatrix}$

Constructional analyses that have flourished in the late 1990's also benefited from the sorting of semantic objects. The analysis of clause types in Sag (1997) and Ginzburg & Sag (2000) makes extensive use of the sorting of semantic objects to model different kinds of clauses, as our discussion of the latter in the next section makes clear.

5 The advantages of a surface-oriented grammar

Until now we have mostly covered how semantic composition works in an approach where each node in a tree is associated with a meaning and where there is only one stratum and therefore the "location" of an expression in a syntactic tree does not necessarily correspond to where in the sentence's semantic representation its meaning is composed. Although important as a proof that semantic composition can be modeled in a surface-oriented grammar, it is fair to say that HPSG work until the late 1990's does not have too much new insight to contribute to our understanding of the interface between syntax and semantics. This is in no way a slight of that early research in the interface between syntax and semantics. Demonstrating that you can "get things right" without multiple strata is important and work on the relation between lexical meaning and argument structure (see the argument structure chapter) is also important in showing that simplicity of syntactic representation does not come at the cost of adequacy. The message was good news: you do not need to make your syntax more complex to make it possible to interface it with semantics. Of course, that was Montague's point already in the late 1960's and early 1970's (see the collected papers in Montague 1974), but that work was more a proof of concept. Carrying out what is basically the Montagovian agenda with a large scale grammar is more difficult and this is what early work in HPSG, at least retrospectively, seems to have focused on.

The development of a more constructional HPSG in the mid 90's opened up new possibilities for modeling the interface between syntax and semantics. One of them is the ability of rich phrasal constructicons to model the shared semantic combinatorics of quite distinct constructional patterns. This is for example apparent in Sag (1997) where a single modification meaning is assigned to a family of relative constructions that differ markedly syntactically. This is also what Ginzburg & Sag (2000) show with their analysis of interrogatives. But their analysis goes further in demonstrating that there may be advantages to a surface-oriented approach to syntax in that it correctly predicts an effect of the *surface* syntax onto semantics for the interpretation of interrogatives, as we now show.

The approach to interrogatives Ginzburg and Sag propose is new in that it does not rely on the traditional Hamblin semantics for questions that the meaning of questions is the set of (exhaustive) answers; see Hamblin (1973); Groenendijk & Stokhof (1997). Rather, the meaning of questions are propositional abstracts (not sets of propositions). Parameters of the kind that have been part of HPSG approaches to semantics since the beginning are used to model these propositional abstracts. Because the meaning of questions are propositional abstracts, the meaning of *wh*-phrases is not the same as that of generalized quantifiers either; rather, *wh*-phrases introduce a parameter (roughly, the equivalent of a lambda-abstracted variable). (21) and (22) provide examples of the meaning of *wh*-questions and polar questions, respectively (Ginzburg & Sag 2000: p.137), where the AVM that follows ↦ stands for the value of the CONTENT attribute of the expression that precedes ↦. Note that polar questions are modeled as zero-parameter propositional abstracts.

(21) Who left? ↦

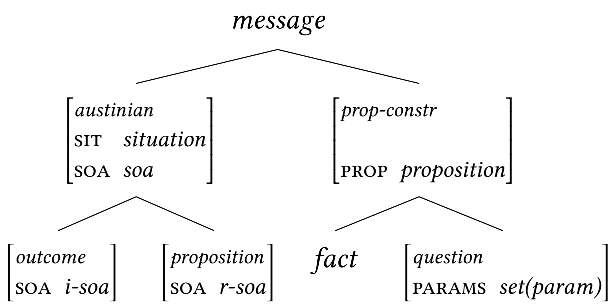
$$\begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \left\{ \begin{bmatrix} \textit{param} & \\ \text{INDEX} & \boxed{1} \\ \text{RESTR} & \{\text{PERSON}(\boxed{1})\} \end{bmatrix} \right\} \\ \text{PROP} & \begin{bmatrix} \textit{proposition} & \\ \text{SIT} & s \\ \text{SOA} & \begin{bmatrix} \text{QUANTS} & \langle\rangle \\ \text{NUCL} & \begin{bmatrix} \textit{leave-rel} & \\ \text{LEAVER} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(22) Did someone leave? ↦

$$\begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \{\} \\ \text{PROP} & \begin{bmatrix} \textit{proposition} & \\ \text{SIT} & s \\ \text{SOA} & \begin{bmatrix} \text{QUANTS} & \langle\rangle \\ \text{NUCL} & \begin{bmatrix} \textit{leave-rel} & \\ \text{LEAVER} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The meaning assigned to questions illustrated above relies on an ontology of messages (the semantic content a clause expresses) which is richer than the traditional notion of propositional content (as distinct from illocutionary force) in work such as Searle (1969). Questions in this view are not just a speech act (where the propositional content of that act remains a proposition), but rather a particular kind of propositionally constructed message, a proposition cum parameters, as shown in (23). Crucially, questions are defined as a parametrized proposition.

(23)



Of concern to us here is less the specifics of this ontology of messages (or of the introduction in the universe of discourse of place holders and other abstract objects, as is typical of situation semantics) than its role in the interface between syntax and semantics, e.g., the fact that clause types can refer to different kinds of messages. Declarative and interrogative clauses are defined as in (24), where the expression that precedes the colon indicates the sort of the phrase and what follows the colon is an informal representation of properties of the phrase's constituents (AVMs to the left of the arrow are properties of the mother node and what follows the arrow are properties of the daughters), / indicates default identity between information on the mother and daugther nodes in (24a) and …in (24b) informally indicates the absence of constraints on daughters on the maximally general *inter-cl* sort. In contrast to earlier approaches to semantics in HPSG where combining a VP with a subject amounted to nothing more than adding the relevant information in the event structure (akin to functional application), this more constructional approach associates a type-shift to the "traditional" subject-predicate construction, from a state of affair description to a proposition. In other words, the analysis of clause-types familiar from traditional grammar plays an explicit role in the grammar, as it is associated with a particular kind of semantic content. Interrogative clauses (clauses of sort *int-cl*) are partially defined by their message, i.e. as denoting questions. Different kinds of interrogatives (polar interrogatives, *wh*-interrogatives, and *in situ* interrogatives) can then be defined as subsorts of *int-cl*. Because this constructional analysis of clause types is embedded in a multiple inheritance network of constructions, an elegant model of similarities in syntax that do and do not correspond to similarities in meaning becomes possible. For example, English declaratives, like typical *wh*-interrogatives, can be inverted (and therefore some declaratives are subject-auxiliary-inversion phrases, as in *Under no circumstance will I allow Tobi to go out at night*, see Fillmore 1999) and, conversely, some interrogatives are not (*in situ* interrogatives, in particular), but some must be inverted (polar interrogatives). Embedding a constructional semantics (i.e., the association of meaning to particular kinds of clauses) in a multidimensional analysis of phrases allows a model that associates meaning to some structures, while recognizing that they need not to. It is similar to some versions of Construction Grammar, but is weaker in that it does not require phrasal constructions to be associated with an unpredictable meaning (i.e., with more than the equivalent of functional application in Categorial Grammar-like approaches).

(24) a. *decl-cl*: $\left[\textsc{cont}\ \left[\begin{array}{l}\textit{austinian}\\ \textsc{soa}\ /\boxed{1}\end{array}\right]\right] \rightarrow \text{H}\ \left[\textsc{cont}\ /\boxed{1}\right]$

b. *inter-cl*: $\left[\textsc{cont}\ \textit{question}\right] \rightarrow \ldots$

One particularly interesting aspect of the constructional semantics of Ginzburg & Sag (2000) is that it can model differences in scoping possibilities of the parameters associated with *wh*-phrases that occur as fillers of head filler structures and *wh*-phrases that occur *in situ*. Consider the sentences in (25) and the difference in interpretation they can receive. (The observation is due to Baker 1970; see Ginzburg & Sag 2000: p.242-246 for discussion.) Sentence (25a) only has interpretation (26a) and, similarly, sentence (25b) has interpretation (26b).

(25) a. Who wondered *who* saw what?

b. Who wondered *what* was seen by who?

(26) a. For which person x and thing y did x wonder who saw y

b. For which person x and person z did x wonder which z saw what

The generalization seems to be that the scope of the parameters introduced by *wh*-phrases that occur in filler position (i.e., as (part of) the filler daughter of a head-filler phrase) is constrained by its surface position, but *wh*-phrases that occur *in situ* are not so constrained. Thus, *who* in (25a) (*what* in (25b)) cannot outscope the embedded clause, but *what* (*who* in (25b)) can. The explanation for this puzzling observation runs as follows. *Wh*-interrogatives are a subsort of interrogative clauses and head-filler phrases. They are thus subject to the Filler-Inclusion Constraint in (27) that requires the WH value of the filler to be a retrieved parameter (i.e., become part of the PARAMS set). This constraint ensures that *wh*-phrases that are fillers of a head-filler phrase contribute their parameter in the clause they are fillers of. In contrast, the parameter of *wh*-phrases that remain *in situ* are not so constrained and are thus free to either be retrieved in the clause in which they occur or be retrieved in a higher clause.

(27) Filler Inclusion Constraint
wh-int-cl: $\left[\text{CONT}\ \left[\text{PARAMS}\ \{\boxed{1}\} \uplus \textit{set}\right]\right] \rightarrow \left[\text{WH}\ \{\boxed{1}\}\right]$, H

It should be noted that the combination of a constructional and surface-oriented approach to the semantics of interrogatives requires positing several unary branching constructions whose sole function is to "type-shift" the meaning of the daughter phrase to match the semantic requirements of the phrase it occurs in. Consider the discourses in (28) and (29) (from Ginzburg & Sag 2000: p.270,ex.37 and 63a), a reprise and non-reprise use, respectively, of *in situ wh*-phrases. We focus on the latter case, which involves an "ordinary" question interpretation, for simplicity. Since B's answer is syntactically a declarative subject-predicate clause, its meaning will be of sort *proposition* (as any head-subject clause that is not a *wh*-subject clause). But the meaning associated with this construction is that of a question. So, we need a unary-branching construction that maps the propositional meaning onto the question meaning, i.e. that retrieves the stored parameter contributed by the *wh*-phrase and makes the CONTENT of the head-subject phrase the value of the PROP attribute of a question. This is what is accomplished by the *is-int-cl* construction defined in (30), a construction one of whose subsorts is the construction involved in discourse (29) and defined in (31) (the Independent Clause feature value + in (30) is meant to rule out *in situ* interrogatives to be embedded interrogatives). Assigning distinct messages to different clause types while maintaining a surface oriented approach requires quite a few such unary branching constructions whose function is strictly semantic.

(28) A: Jo saw absolutely every shaman priest from East Anglia.
B: Jo saw absolutely every shaman priest from WHERE.

(29) A: Well, anyway, I'm leaving.
B: OK, so you'll be leaving WHEN exactly?

(30) *is-int-cl*:
$$\left[\text{CAT}\ \begin{bmatrix}\text{IC} & + \\ \text{SUBJ} & \langle\,\rangle \\ \text{VFORM} & \textit{fin}\end{bmatrix}\right] \rightarrow \text{H}[\ ]$$

(31) *dir-is-int-cl*:
$\left[\text{CONT|PRO}\ \boxed{1}\right] \rightarrow \text{H}\left[\text{CONT}\ \boxed{1}\right]$

6 Semantic underspecification

One of the hallmarks of constraint-based grammatical theories is the view that grammars are *descriptions* of structures and can thus be incomplete, as almost all descriptions are. This is a point that was made clear a long time ago by Martin Kay in his work on unification (see among others Kay 1979). For a long time, the distinction between (partial) descriptions (possible properties of linguistic structures, what grammars are about) and (complete) described linguistic structures was used almost exclusively in the syntactic component of grammars within HPSG. But starting in the mid 90's the importance of distinguishing between descriptions and described structures began to be appreciated in HPSG's model of semantics, as discussed for example in Copestake et al. (1995), and recent work has stressed the importance of the same distinction when modeling inflectional morphology (see Crysmann & Bonami (2016) and the chapter on morphology). Because underspecification, partiality and the like are so critical to HPSG, its inclusion in the model of the semantics of grammar has made recent work in semantics in HPSG quite distinctive from work in semantics within even conceptually related frameworks such as Lexical Functional Grammar or variants of Categorial Grammar. Two competing approaches to semantic underspecification have been developed within HPSG, Minimal Recursion Semantics (henceforth, MRS; see Copestake et al. 1995, Copestake et al. 2001, and Copestake et al. 2005 for introductions to MRS) and Lexical Resource Semantics (henceforth, LRS; see Richter & Sailer 2001; 2004 and Iordăchioaia & Richter 2015 for an introduction to LRS). MRS and LRS are not the only two "recent" approaches to assembling the meaning of phrases from lexical "meanings" (or resources). Asudeh & Crouch (2002), for example, show how to apply a glue approach to semantic interpretation to HPSG. Aside from simplification of the Semantics Principle, which , under a glue semantic approach, do esnot distinguish how to compose meaning on the basis of the semantic type of the daughters, e.g., whether one of the daughters is a quantifier, a glue approach leads to "highly efficient techniques for semantic derivation already implemented for LFG, and which target problems of ambiguity management also addressed by Minimal Recursion Semantics" (p.1). For reasons of space, we cannot detail Asudeh and Crouch's glue approach here; we concentrate on MRS and LRS as they have been the dominant approaches to semantic composition in recent years. But, the existence of yet another approach to semantic interpretation attests of the flexibility of the HPSG architecture when trying to model the interface between syntax and semantics.

6.1 Minimal Recursion Semantics

6.1.1 Why minimally recursive semantic representations

MRS developed out of computational semantic engineering considerations related to machine translation for face-to-face dialog that started in the early 90's (see Kay et al. 1992 for an overview of the the VerbMobil project). As Copestake et al. (1995) argue, syntactic differences between languages can lead to logically equivalent distinct semantic representations when using traditional "recursive" semantic representations. They point out, for example, that the English expression *fierce black cat* and Spanish *gato negro y feroz* would be given distinct semantic representations under standard assumptions, as shown in (32). These distinct semantic representations would make translating these simple nominal expressions from one language to the other difficult.

(32) a. $\lambda x(\text{fierce}(x) \wedge (\text{black}(x) \wedge \text{cat}(x)))$

b. $\lambda x(\text{cat}(x) \wedge (\text{black}(x) \wedge \text{fierce}(x)))$

Furthermore, some sentences may be similarly ambiguous in English and Spanish – for example, sentences that contain generalized quantifiers – , and requiring the semantic disambiguation of these sentences prior to translating them into sentences that contain similar ambiguities is inefficient. Semantic representations should only be as disambiguated as the source language grammar entails. For these reasons and others they detail, Copestake et al. (1995) propose to model the semantics of grammar via semantic representations that are as flat (or non-recursive) as possible. To achieve this minimal recursivity despite the fact that disambiguated scope relations among generalized quantifiers require embedding, they add additional variables or handles that serve as labels to particular relations in the flat list of relations and that can serve as "arguments" of scopal operators. (33) and its underspecified and fully disambiguated semantic representations in (34) illustrate informally and (35) more formally. Subscripts on names of relations in the informal representation stand for labels of the formulas they are part of. Thus, 1 in $every_1(x, 3, n)$ is a label for the entire formula. In the more explicit representation in (35), the label of a formula is written before it and separated from it by a colon (e.g., *h1: every(x,h3,h8)*); variables over labels are simply labels that do not correspond (yet) to labels of formulas ($h8$ and $h9$).[1]

[1] Copestake (2007) presents a Neo-Davidsonian version of MRS called R(obust)MRS where arguments of predicates (aside from their event variable) are contributed via independent elementary predications. Copestake shows that RMRS can be profitably used with shallower analyses, "including part-of-speech tagging, noun phrase chunking and stochastic parsers which operate

(33) Every dog chased some cat.

(34) a. $\text{every}_1(x, 3, n)$, $\text{dog}_3(x)$, $\text{cat}_7(y)$, $\text{some}_5(y, 7, m)$, $\text{chase}_4(e, x, y)$

b. $\text{every}_1(x, 3, 4)$, $\text{dog}_3(x)$, $\text{cat}_7(y)$, $\text{some}_5(y, 7, 1)$, $\text{chase}_4(e, x, y)$

c. $\text{every}_1(x, 3, 5)$, $\text{dog}_3(x)$, $\text{cat}_7(y)$, $\text{some}_5(y, 7, 4)$, $\text{chase}_4(e, x, y)$

(35) a. h1: every(x,h3,h8), h3:dog(x), h7:white(y), h7:cat(y), h5:some(y,h7,h9), h4:chase(x,y)

b. $h8 = h4$, $h9 = h1$

c. $h8 = h5$, $h9 = h4$

To understand the use of handles consider the expression $\textit{every}_1(x, 3, n)$. The first argument of the generalized quantifier is the handle numbered 3, which is a label for the formula $\textit{dog}_3(x)$. The formula that serves as the first argument of *every* is fixed, it is always the meaning of the nominal phrase that the determiner selects for. But to avoid embedding that relation as the restriction of the quantifier and to preserve the desired flatness of semantic representations, the second argument of *every* is not $\textit{dog}_3(x)$, but the label of that formula (indicated by the subscript 3 on the predication $\textit{dog}_3(x)$). Now, in contrast to the quantifier's restriction, which must include the content of the head noun it combines with, the nuclear scope or body of the quantifier is not as restricted. In other words, the semantic representation that is the output of the grammar of English does not fix the second argument of *every*, represented here as the variable over handles n. The same distinction applies to *some*: its first argument is fixed to the formula $\textit{cat}_7(y)$, but its second argument is left underspecified, as indicated by the variable over numbered labels, m. Resolving the scope ambiguity of the underspecified representation in (34a) amounts to deciding whether *every* takes the formula that contains *some* in its scope or the reverse; in the first case, $n = 5$, in the second, $m = 1$. Since the formula that encodes the verb meaning ($\textit{chase}_4(e, x, y)$) is outscoped by the nuclear scope or body of both generalized quantifiers, either constraint will fully determine the relative scope of all formulas in (34).

6.1.2 The nitty-gritty

We now present a brief outline of how MRS works in Typed Feature Structures. First, the content of an expression is of sort *mrs*. Structures of that sort consist in (1) a bag of relations or elementary predications (EPs) (the value of RELS), (2)

without detailed lexicon" (p.73).

a HOOK, which groups together the labels or handles that correspond to elementary predications that have widest local and global scope and the expression's index (these three semantic objects are what is visible to semantic functors), and (3) a set of constraints on handles that restrict or determine the scope of scope-relevant elementray predications (the value of HCONS). Each constraint in the value of HCONS consists of a greater or equal relation between handles. A representation of the structure of an object of sort *mrs* is provided in (36).

(36)
$$\begin{bmatrix} \textit{mrs} & \\ \textsc{hook} & \begin{bmatrix} \textsc{index} & \textit{index} \\ \textsc{ltop} & \textit{handle} \\ \textsc{gtop} & \textit{handle} \end{bmatrix} \\ \textsc{rels} & \textit{list(relations)} \\ \textsc{hcons} & \textit{list}\left(\begin{bmatrix} \textit{qeq} & \\ \textsc{harg} & \textit{handle} \\ \textsc{larg} & \textit{handle} \end{bmatrix}\right) \end{bmatrix}$$

Sentence (37) and its (underspecified) *mrs* representation in (38) illustrate how *mrs* structures can be used to capture scope underspecification (see Copestake et al. 2005: p.306).

(37) Every dog probably sleeps.

(38)
$$\begin{bmatrix} \textit{mrs} & \\ \textsc{hook} & \begin{bmatrix} \textit{hook} & \\ \textsc{gtop} & \boxed{1} \\ \textsc{ltop} & \boxed{7} \end{bmatrix} \\ \textsc{rels} & \left\langle \begin{bmatrix} \textit{every\_rel} & \\ \textsc{lbl} & \boxed{2} \\ \textsc{arg0} & \boxed{3} \\ \textsc{rstr} & \boxed{4} \\ \textsc{body} & \textit{handle} \end{bmatrix}, \begin{bmatrix} \textit{dog\_rel} & \\ \textsc{lbl} & \boxed{6} \\ \textsc{arg0} & \boxed{3} \end{bmatrix}, \begin{bmatrix} \textit{prbly\_rel} & \\ \textsc{lbl} & \boxed{7} \\ \textsc{arg1} & \boxed{8} \end{bmatrix}, \begin{bmatrix} \textit{sleep\_rel} & \\ \textsc{lbl} & \boxed{9} \\ \textsc{arg1} & \boxed{3} \end{bmatrix} \right\rangle \\ \textsc{hcons} & \left\langle \begin{bmatrix} \textit{qeq} & \\ \textsc{harg} & \boxed{1} \\ \textsc{larg} & \boxed{7} \end{bmatrix}, \begin{bmatrix} \textit{qeq} & \\ \textsc{harg} & \boxed{4} \\ \textsc{larg} & \boxed{6} \end{bmatrix}, \begin{bmatrix} \textit{qeq} & \\ \textsc{harg} & \boxed{8} \\ \textsc{larg} & \boxed{9} \end{bmatrix} \right\rangle \end{bmatrix}$$

Members of RELS correspond to the content of lexical entries while members of HCONS constrain the relative scope of semantic arguments of members of RELS. Now, although the grammar of English leaves the meaning of (37) underspecified, it *does* constrain some scope relations and the *mrs* in (38) therefore constrains how some elementary predications relate to each other. First, the identity between the value of ARG0 for both the *every_rel* and *dog_rel* elementary predications indicates that *every* in (37) quantifies over dogs; $\boxed{3}$ is the variable bound by the quantifier. And similarly, the value of ARG1 of *sleep_rel* is lexically constrained to corresponds to the index of the subject, itself constrained to be identical to the value of ARG0 for the *dog_rel* predication (i.e, $\boxed{3}$). Second, *prbly_rel* is required to outscope *sleep_rel* (a *qeq* constraint either identifies its HARG or LARG or it constrains its HARG to outscope its ARG). Similarly, the restriction of *every_rel* is constrained to outscope *dog_rel* as $\boxed{4} =_{qeq} \boxed{6}$. Finally, the global top (the value of GTOP) is constrained to outscope the local top (the value of LTOP). (To simplify, the local top is the handle of the elementary predication that is not a quantifier with the widest scope.) The semantic representation that the grammar of English motivates remains underspecified, as it does not specify what the value of the BODY of *every_rel* is, in particular whether it is the handle of the *prbly_rel* or *sleep_rel* elementary predications. Resolving this scope ambiguity amounts to adding an HCONS that identifies the value of BODY with either handle, i.e., $\boxed{7}$ or $\boxed{9}$.

Examples that include multiple quantifiers work in a similar way. Take the sentence in (39) and the elementary predications for *every*, *chases*, and *some* (we only include relevant elementary predications and attributes for simplicity). We know that the body of *every_rel* and *some_rel* each outscope *chase_rel* (so $\boxed{1} =_{qeq} \boxed{3}$ and $\boxed{1} =_{qeq} \boxed{3}$, where the left hand side of the equality corresponds to the HARG and the right hand side to the LARG). But, we do not know if *every_rel* outscopes *some_rel* or the reverse; adding either HCONS $\boxed{1} =_{qeq} \boxed{2}$ or $\boxed{2} =_{qeq} \boxed{1}$ specifies which is the case. (This example illustrates that $=_{qeq}$ is not commutative, as it is meant to encode greater or equal scope.)

(39) Every dog chases some cat.

(40) $\begin{bmatrix} \textit{every\_rel} & \\ \text{RESTR} & \textit{handle} \\ \text{BODY} & \boxed{1} \end{bmatrix} \quad \begin{bmatrix} \textit{some\_rel} & \\ \text{RESTR} & \textit{handle} \\ \text{BODY} & \boxed{2} \end{bmatrix} \quad \begin{bmatrix} \textit{chase\_rel} & \\ \text{LBL} & \boxed{3} \end{bmatrix}$

Semantic composition within MRS is relatively simple and is stated in (41) (Copestake et al. 2005: p. 313–314); the second clause of this semantic composition rule amounts to a case-based definition, as is true of all Semantics Principles since Pollard & Sag (1987), as different constructions determine differently the HOOK of the head-daughter (Copestake et al. 2005 only discuss intersective and scopal constructions in their paper). (A slot in (41) is defined as "a semantic argument position in a word or phrase A that is associated with syntactic constraints on the word or phrase B whose semantics will supply that argument when the relevant grammar rule combines A and B" (op.cit., p.313).)

(41)
1. The RELS value of a phrase is the concatenation (append) of the RELS values of the daughters.
2. The HCONS of a phrase is the concatenation (append) of the HCONS values of the daughters.
3. The HOOK of a phrase is the HOOK of the semantic head daughter, which is determined uniquely for each construction type.
4. One slot of the semantic head daughter of a phrase is identified with the HOOK in the other daughter.

This quite brief description of MRS illustrates what is attractive about it from an engineering point of view. Semantic composition is particularly simple, concatenation of lists (lists of elementary predications and constraints), percolation of the HOOK from the semantic head, and some general constraint on connectedness between the head daughter and the non-head daughter. Furthermore, resolving scope means adding $=_{qeq}$ constraints to a list of $=_{qeq}$, thus avoiding traversing the semantic tree to check on scope relations. Furthermore, a flat representation makes translation easier, as argued in Copestake et al. (1995), and has several other advantages from an engineering perspective as detailed in Copestake (2009). The ease flat representations provide comes at a cost, though, namely semantic representations are cluttered with uninterpretable symbols (handles) and, more generally, do not correspond to sub-pieces of a well-formed formula. For example, we would expect the value of a quantifier restriction and nuclear scope to be, say, formulas denoting sets (as per Barwise & Cooper 1981), not pointers to or labels of predications. This is not to say that a compositional, "standard" interpretation of MRS structures is not possible (see, for example, Copestake et al. 2001); it is rather that the model-theoretic interpretation of MRS requires adding to the model hooks and holes, abstract objects of dubious semantic import. While it is true, as Copestake et al. point out, that abstract objects have been included in the models of other semantic approaches, DRT in particular (Zeevat 1989), abstract objects in compositional specifications of DRT and other such dynamic semantic approaches are composed of semantically interpretable objects. In the case of DRT, the set of variables (discourse referents) that form the other component of semantic representations (aside from predicative conditions) are anchored to individuals in the "traditional" model theoretic sense. Holes and hooks, on the other hand, are not necessarily so anchored, as labels (handles) do not have any interpretation in the universe of discourse.

(42) $\begin{bmatrix} \textit{every\_rel} & \\ \textsc{restr} & \textit{handle} \\ \textsc{body} & \textit{handle} \end{bmatrix}$

An example of the model-theoretic opacity of handles is provided by the compositional semantics of intersective attributive adjectives. The RELS value of *white horse*, for example, is as shown in (43) (after identification of handles due to the meaning composition performed by the *intersective_phrase* rule that (intersective) adjectival modification is a subsort of).

(43) $\left[\text{RELS}\ \left\langle \begin{bmatrix} \textit{white\_rel} \\ \text{LBL} & \boxed{1} \\ \text{ARG0} & \boxed{2} \end{bmatrix}, \begin{bmatrix} \textit{horse\_rel} \\ \text{LBL} & \boxed{1} \\ \text{ARG0} & \boxed{2} \end{bmatrix} \right\rangle \right]$

The fact that the value of ARG0 is the same for both elementary predications ($\boxed{2}$) is model-theoretically motivated: both properties are predicated of the same individual. The fact that the value of LBL is identical ($\boxed{1}$) is also motivated if labels are used to help determine the scope of quantifiers; in a quantifier like *every white horse*, the content of *white* and *horse* have the same scope, they conjunctively serve as the restriction of *every_rel*. But, the identity of the two elementary predications' labels is not *directly* model-theoretically motivated. It is a consequence of the semantic representation language that is used to model the meaning of sentences, not a consequence of the sentences truth-conditions.

6.2 Lexical Resource Semantics

Whereas MRS emphasizes underspecification in semantic representations and expresses the syntax of underspecified representations in HPSG as typed feature structures, LRS focuses primarily on fine-grained linguistic analyses with explicit higher-order logics for meaning representation and utilizes underspecification prominently in the architecture of the syntax-semantics interface. Instead of encoding underspecified representations as denotations of grammar principles, it uses the feature logic itself as a tool for underspecifying fully specific logical representations in the symbolic languages of the literature on formal semantics. This means that a grammar with LRS semantics denotes sets of syntactic structures that comprise fully explicit meaning representations in a standard logical language, but it does so with means of underspecification in the grammar principles, enlisting the techniques which HPSG developed for writing general descriptions in grammar principles in the definition of the relationship between syntactic structure and semantic representation. Grammar principles may admit a large number of structures, which in this case can be multiple semantic representations compatible with one and the same syntactic structure. An LRS analysis may then represent the readings of (33) with two generalized quantifiers as value of a semantic feature as shown in (44).

(44) a. $\forall\left(\lambda x.dog_w(x), \lambda x.\exists\left(\lambda y.cat_w(y), \lambda y.chase_w(x,y)\right)\right)$
b. $\exists\left(\lambda y.cat_w(y), \lambda y.\forall\left(\lambda x.dog_w(x), \lambda x.chase_w(x,y)\right)\right)$

The syntactic format of semantic representations is flexible and can be adapted to the purposes of the linguistic analysis at hand. While (44) chooses predicates with an argument for possible worlds, lambda abstraction over the unary predicates which translate the nominal arguments, and categorematic quantifiers of type $\langle\langle et\rangle\langle\langle et\rangle t\rangle\rangle$, in many contexts less elaborate representations will suffice, and the two readings would be rendered in a notational variant of first order languages. Other phenomena might necessitate more semantic structure. The LRS framework makes a selection of choices available to linguists to decide what is most adequate to spell out a semantic analysis.

6.2.1 Basic architecture

Lexical items contribute semantic resources to utterances; every semantic representation of an utterance must use up all and only the semantic resources provided by the lexical items in the utterance in all their legitimate combinations.[2] What is legitimate is determined by semantic principles which restrict at each phrase how the semantic resources of its daughters may be combined. What these restrictions do not rule out is permitted. Scope ambiguities between co-arguments of a verb can be seen as arising from the lack of a principled restriction to the effect that one outscope the other. In the absence of restrictions, LRS expects ambiguity. As a special property setting LRS apart from other semantic underspecification frameworks, LRS semantics exploits HPSG's notion of structure sharing in its semantic representations by permitting that semantic contributions of different lexemes may in fact be identical. For example, if two words in a clause contribute negation in their meaning, the two negations may in fact turn out to be the same negation, in which case we observe a negative concord reading. The implementation of this idea is based on the fundamental structure-sharing mechanism of HPSG, which is available throughout all levels of grammatical description.

The combinatorial semantics of phrases is encoded with structures of sort *lrs*:

(45) $\left[\text{SEM} \begin{bmatrix} \textit{lrs} & \\ \text{EXCONT} & \textit{me} \\ \text{INCONT} & \textit{me} \\ \text{PARTS} & \textit{list(me)} \end{bmatrix}\right]$

[2] Lexical items may be phrasal.

Signs have an attribute SEMantics with value *lrs*. External content (EXCONT) and internal content (INCONT) designate two prominent aspects of the semantics of signs. Both of these attributes have values of sort *meaningful_expression*, short *me*. The attribute EXCONT contains a term that represents the meaning of the maximal syntactic projection of the sign and is built from semantic material contributed within the projection. The INCONT is that part of a lexical sign's representation which is outscoped by any scope-taking operator that it combines with within its syntactic projection. The PARTS list records all semantic resources contributed by a given sign. The LRS Projection Principle governs the percolation of these attribute values along the syntactic head path of phrases, whereas the EXCONT and INCONT principles determine the relationship of the respective attribute values to other semantic attribute values within local syntactic trees. The most important relationships are those of term identity and of subtermhood of one term relative to another or to some designated part of another term. Subterm restrictions are in essence similar to the *qeq* constraints of MRS.

(46) a. LRS Projection Principle
In each phrase,
1. the EXCONT values of syntactic head and mother are identical,
2. the INCONT values of syntactic head and mother are identical,
3. the list in the PARTS value contains all and only the elements of the PARTS values of the daughters.

b. INCONT Principle
In each *lrs*, the INCONT value is an element of the PARTS list and a component of the EXCONT value.

c. EXCONT Principle
First, in every phrase, the EXCONT value of the non-head daughter is an element of the non-head daughter's PARTS list. Second, in every utterance, every subexpression of the EXCONT value of the utterance is an element of its PARTS list, and every element of the utterance's PARTS list is a subexpression of the EXCONT value.

The Projection Principle guarantees the percolation of EXCONT and INCONT values along the head path of syntactic phrases, and it records the semantic resources available at each phrase based on the semantic contributions of their daughters (46a). The INCONT Principle and the EXCONT Principle manage the properties of the respective attribute values. The term with minimal scope of each lexeme must be contributed by the lexeme itself and must be semantically realized within the representation of the maximal syntactic head projection (46b).

The maximal semantic meaning contribution of a maximal syntactic projection must originate from within that maximal projection, and an utterance (as a distinguished maximal projection) consists of all and only those pieces of semantic representation which are contributed by some lexeme in the utterance (46c). The meaning of an utterance is given by the semantic representation which is its EXCONT value. An ambiguous utterance receives structural analyses that are potentially only distinguished by different EXCONT values of their root node.

The constraints in (46) take care of the integrity of the semantic combinatorics. The task of the clauses of the Semantics Principle is to regulate the semantic restrictions on specific syntactic constructions (as in all previously discussed versions of semantics in HPSG). A quantificational determiner, represented as a generalized quantifier, which syntactically combines as non-head daughter with a nominal projection, integrates the INCONT of the nominal projection as a subterm into its restrictor and requires that its own INCONT (containing the quantificational expression) be identical with the EXCONT of the nominal projection. This clause makes the quantifier take wide scope in the noun phrase and forces the semantics of the nominal head into the restrictor. In (44) we observe the effect of this clause by the placement of the predicate dog in the restrictor of the universal and the predicate cat in the restrictor of the existential quantifier.

Another clause of the Semantics Principle governs the combination of quantificational NP arguments with verbal projections. If the non-head of a verbal projection is a quantificational NP, the INCONT of the verbal head must be a subexpression of the scope of the quantifier. Since this clause does not require immediate scope, other quantificational NPs which combine in the same verbal projection may take scope in between, as we can again see with the two possible scopings of the two quantifiers in (44), in particular in (44b), where the subject quantifier intervenes between the verb and the object quantifier.

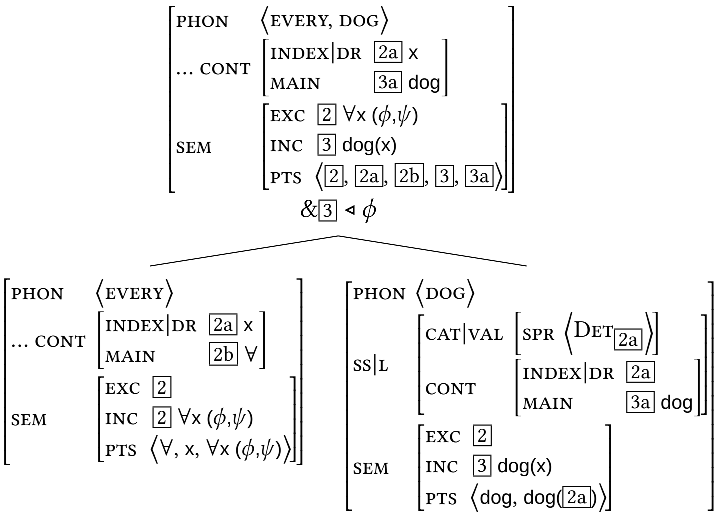
The local semantics of signs is split from the combinatorial *lrs* structures in parallel to the separation of local syntactic structure from the syntactic tree structure. The local semantics remains under the traditional CONTENT attribute, where it is available for lexical selection by the valence attributes. The LOCAL value of the noun *dog* illustrates the relevant structure:

(47) $\begin{bmatrix} \textit{local} & \\ \text{CAT}|\text{VAL} & \begin{bmatrix} \text{SPR} & \left\langle \text{DET}_{\boxed{1}} \right\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{INDEX}|\text{DR} & \boxed{1}\ \mathsf{x} \\ \text{MAIN} & \mathsf{dog} \end{bmatrix} \end{bmatrix}$

The attribute DISCOURSE-REFERENT (DR) contains the variable that will be the argument of the unary predicate dog, which is the MAIN semantic contribution of the lexeme. The variable, x, does not come from the noun but is available to the noun by selection of the determiner by the valence attribute SPR. The subscripted tag $\boxed{1}$ on the SPR list indicates the identity of DR values of the determiner and the nominal head *dog*. A principle of local semantics says that MAIN values and DR values are inherited along the syntactic head path.

The semantics of phrases follows from the interaction of the (lexical) selection of local semantic structures and the semantic combinatorics that results from the principles in (46) and the clauses of the Semantics Principle:

(48)



For ease of readability, the notation above omits the lambda abstractions from the generalized quantifier and chooses a notation from first order logic and not all structure sharings between pieces of the logical representation are made explicit in (48). The head noun *dog* contributes (on PARTS, PTS), the predicate dog and the application of the predicate to a lexically unknown argument, $\boxed{2a}$, identical with the DR value of *dog*. As shown in (47), the DR value of the noun is shared with the DR value of the selected determiner, which is the item contributing the variable x to the representation. In addition, *every* contributes the quantifier and the application of the quantifier to its arguments. The clause of the Semantics Principle which restricts the combination of quantificational determiners with nominal projections identifies the INC of *every* with the EXC of *dog*, and requires that the INC of dog ($\boxed{3}$) be a subterm of the restrictor of the quantifier, ϕ (notated as '$\boxed{3} \triangleleft \phi$', conjoined to the AVM describing the phrase). The identification of the EXC and INC of *every* follows from (46b-c). According to this analysis, the semantic representation of the phrase *every dog* is a universal quantification with dog(x) in the restrictor and unknown scope (ψ). The scope will be determined when the noun phrase combines with a verb phrase. For example, such a verb phrase could be *barks*. If its semantics is represented as a unary predicate bark, the predicate and its application to a single argument are contributed by the verb phrase, and local syntactic selection of the subject *every dog* by the verb *barks* identifies this argument as variable x, parallel to the selection of the quantifier's variable by *dog* above. The relevant clause of the Semantics Principle requires that bark(x) be a subterm of ψ, and the EXC, $\boxed{2}$, of the complete sentence receives the value $\forall$x (dog(x), bark(x)) as the only available reading in accordance with the EXCONT Principle:

(49)

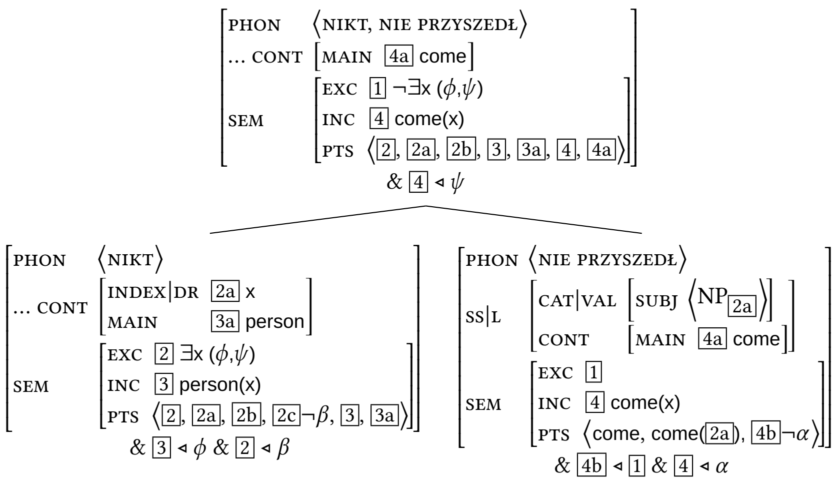
$$\left[\begin{array}{ll} \textsc{phon} & \langle \textsc{every, dog, barks} \rangle \\ \ldots\ \textsc{cont} & \left[\textsc{main}\ \boxed{4a}\ \text{bark}\right] \\ \textsc{sem} & \left[\begin{array}{ll} \textsc{exc} & \boxed{2}\ \forall \text{x}\,(\phi,\psi) \\ \textsc{inc} & \boxed{4}\ \text{bark(x)} \\ \textsc{pts} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{3}, \boxed{3a}, \boxed{4}, \boxed{4a} \rangle \end{array}\right] \end{array}\right] \\ \&\ \boxed{4} \triangleleft \psi$$

$$\left[\begin{array}{ll} \textsc{phon} & \langle \textsc{every, dog} \rangle \\ \ldots\ \textsc{cont} & \left[\begin{array}{ll} \textsc{index}|\textsc{dr} & \boxed{2a}\ \text{x} \\ \textsc{main} & \boxed{3a}\ \text{dog} \end{array}\right] \\ \textsc{sem} & \left[\begin{array}{ll} \textsc{exc} & \boxed{2}\ \forall \text{x}\,(\phi,\psi) \\ \textsc{inc} & \boxed{3}\ \text{dog(x)} \\ \textsc{pts} & \langle \boxed{2}, \boxed{2a}, \boxed{2b}, \boxed{3}, \boxed{3a} \rangle \end{array}\right] \end{array}\right] \\ \&\ \boxed{3} \triangleleft \phi \qquad \left[\begin{array}{ll} \textsc{phon} & \langle \textsc{barks} \rangle \\ \textsc{ss}|\textsc{l} & \left[\begin{array}{ll} \textsc{cat}|\textsc{val} & \left[\textsc{subj}\ \langle \text{NP}_{\boxed{2a}} \rangle\right] \\ \textsc{cont} & \left[\textsc{main}\ \boxed{4a}\ \text{bark}\right] \end{array}\right] \\ \textsc{sem} & \left[\begin{array}{ll} \textsc{exc} & \boxed{2} \\ \textsc{inc} & \boxed{4}\ \text{bark(x)} \\ \textsc{pts} & \langle \text{bark, bark(}\boxed{2a}\text{)} \rangle \end{array}\right] \end{array}\right]$$

The identity of the restrictor ϕ of the universal quantifier with $\boxed{3}$ dog(x) and of its scope ψ with $\boxed{4}$ bark(x) are determined at utterance level by the lack of other material that could be added to the two arguments of the quantifier. For example, an extraposed relative clause which belongs to *every dog* could consistently contribute its meaning representation to the restrictor, and only the absence of such additional semantic material leads to the inferred identity of $\boxed{3}$ with ϕ.

Underspecification of the structure of meaning representations in the clauses of the Semantics Principle and in lexical entries interacts with the possibility of structure sharing. If two pieces of meaning representation have the same shape and obey compatible structural conditions (as determined by relevant subterm contraints) they can be identical. Even stronger, in certain grammatical constellations, principles of grammar may even strictly require their identity. Lexical underspecification of meaning contributions moreover permits the shared construction of functors such as the construction of a polyadic quantifier from several lexical items in a sentence. These two applications of LRS lead to new possibilities of semantic composition compared to standard compositional semantics in generative grammar, because functors can be composed in (logical) syntax which cannot be semantically decomposed or cannot be decomposed within the structural limits of a surface-oriented syntax, i.e. a syntactic structure which only reflects syntactic but not semantic composition.

(50)



Consider the semantic representation of the Polish sentence *nikt nie przyszed* ‘nobody came’ in (50). Negated finite verbs in Polish contribute a negation that must be realized within the verbs EXCONT ($\boxed{4b}$ ◃ $\boxed{1}$) and outscopes the INCONT of the verb ($\boxed{4}$ ◃ α). Similarly, the existential quantifier of the n-word *nikt* ‘nobody’ is outscoped by negation ($\boxed{2}$ ◃ β). However, in addition to the familiar restriction when the quantificational subject combines with the finite verb, Polish as a strict negative concord language requires that a negated finite verb be in the scope of at most one negation in its EXCONT, entailing identity of the two negations, $\boxed{2c} = \boxed{4b}$, and the single negation reading *nobody came* as the only admissible reading of the sentence shown in (50). To capture obligatory negation marking at finite words in Polish, a second principle of negative concord rules that if a finite verb is in the scope of negation in its EXCONT, it must itself be a contributor of negation (Richter & Sailer 2001).

The idea of identifying contributions form different constituents in an utterance is even more pronounced in cases of unreducible polyadic quantification. The reading of (51a) in which each unicorn from a collection of unicorns has a set of favorite meadows that is not the same as the set of favorite meadows of any other unicorn is known to be expressible by a polyadic quantifier taking two sets and a binary relation as arguments (51d) but it cannot be expressed by two independent monadic quantifiers (Keenan 1992).

(51) a. Every unicorn prefers different meadows.

b. different meadows: $(\gamma', \Delta)(\sigma_1, \lambda y.meadow(y), \lambda v_1 \lambda y.\rho')$

c. every unicorn: $(\forall, \gamma)(\lambda x.unicorn(x), \sigma_2, \lambda x \lambda v_2.\rho)$

d. $(\forall, \Delta)(\lambda x.unicorn(x), \lambda y.meadow(y), \lambda x \lambda y.prefer(x, y))$

(51) sketches the LRS solution to this puzzle in Richter (2016). The adjective *different* contributes an incomplete polyadic quantifier of appropriate type which integrates the representation of the nominal head of its NP into the second restrictor but leaves open a slot in the representation of its functor for another quantifier it must still combine with (51b). The determiner *every* underspecifies the realization of its quantifier in such a way that one of the possible representations yields (51c) for *every unicorn*, which is exactly of the right shape to be identified with the representation of *different meadows*, leading to the expression in (51d) for (51a). Lahm (2016) presents an alternative account of such readings with *different* using Skolem functions which also hinges on LRS-specific techniques. Iordăchioaia & Richter (2015) study Romanian negative concord constructions and represent their readings using polyadic negative quantifiers; Lahm (2018) develops a lexicalized theory of plural semantics.

6.2.2 Representation languages and notational conventions

Any LRS grammar relies on an encoding of the syntax of an appropriate semantic representation language in the feature logic. In principle, any finitary logical language can be encoded in RSRL, which covers every language that has been proposed for meaning representations in linguistics. Work in LRS has so far been couched mostly in variants of Two-sorted Type Theory (Ty2, Gallin 1975) as one of the standard languages of formal semantics, or in Montague's Intensional Logic. The type system of these logical languages are useful for underspecified descriptions in semantic principles, since relevant groups of expressions can be generalized over by their type without reference to their internal structure. For example, a clause of the Semantics Principle can use the type of generalized quantifiers to distinguish quantificational complement daughters of verbal projections and state the necessary restrictions on how they are integrated with the semantics of the verbal head daughter, while other types of complement daughters are treated differently and may even not at all be restricted by a clause in the Semantics Principle in how they integrate with the verbal semantics. The latter is often the case with proper names and definite descriptions, which can be directly integrated with the semantics of the verb by lexical argument selection.

Encodings of semantic representations in feature logic are usually assumed as given by the background LRS theory. Examples of encodings can be found in Sailer (2000) and Richter (2004). Sailer (2000) offers a correspondence proof of the encoded structures with a standard syntax of languages of Ty2. As descriptions of logical terms in literal feature logic are very cumbersome to read and write, and offer no practical advantage or theoretical insight, all publications use notational shortcuts and employ logical expressions with metavariables for their descriptions instead. As nothing depends on feature logical notation, the gain in readability outweighs any concerns about notational precision.

References

Asudeh, Ash & Richard Crouch. 2002. Glue Semantics for HPSG. In Beerman Dorothee van Eynde Frank Hellan Lars (ed.), *Proceedings of the 8th International HPSG Conference*, 1–19. Stanford: CSLI Publications.

Bach, Emmon. 1976. An extension of classical transformational grammar. In *Problems of linguistic metatheory (proceedings of the 1976 conference)*, 183–224. Michigan State University.

Baker, C. L. 1970. *Indirect questions in english*. University of Illinois at Urbana-Champaign dissertation.

Barwise, John & Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4. 159–219.

Cooper, Robin. 1975. *Montague's semantic theory and transformational syntax*. University of Massachusetts, Amherst dissertation.

Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel.

Copestake, Ann. 2007. Semantic composition with (robust) minimal recursion semantics. In *DeepLP '07 proceedings of the workshop on deep linguistic processing*, 73–80. Prague.

Copestake, Ann. 2009. Slacker semantics: why superficiality, dependency and avoidance of commitment can be the right way to go. In *Proceedings of the 12th conference of the european chapter of the ACL*, 1–9. Athens: Association for Computational Linguistics.

Copestake, Ann, Dan Flickinger, Rob Malouf, Susanne Riehemann & Ivan Sag. 1995. Translation using minimal recursion semantics. In *Proceedings of the sixth international conference on theoretical and methodological issues in machine translation (TMI95)*. Leuven, Belgium.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3. 281–332. DOI:10.1007/s11168-006-6327-9

Copestake, Ann, Alex Lascarides & Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th ACL meeting*. Toulouse, France.

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics and Information-based Morphology. *Journal of Linguistics* 52(2). 311–374.

Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxfor: Oxford University Press.

Egg, Markus. 2004. Flexible semantic construction: the case of reinterpretation. Saarbrücken.

Fillmore, Charles. 1999. Inversion and constructional inheritance. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 113–128. Stanford: CSLI Publications.

Gallin, Daniel. 1975. *Intensional and higher-order modal logic*. Amsterdam: North Holland.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Groenendijk, Jeroen & Martin Stokhof. 1997. Questions. In J. Van Benthem & A Ter Meulen (eds.), *Handbook of logic and language*, 1055–1124. Cambridge, MA: MIT Press.

Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Wiley-Blackwell.

Hornstein, Norbert. 1995. *Logical form, from GB to minimalism*. Oxford: Blackwell.

Iordăchioaia, Gianina & Frank Richter. 2015. Negative concord with polyadic quantifiers. The case of Romanian. *Natural Language and Linguistic Theory* 33(2). 607–658.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer.

Kasper, Robert. 1997. The semantics of recursive modification. Ohio State University.

Kay, Martin. 1979. Functional grammar. In Christine et al. Chiarello (ed.), *Proceedings of the fifth meeting of the berkeley linguistics society*, 142–158. Berkeley: Berkeley Linguistics Society.

Kay, Martin, Mark Gawron & Peter Norvig. 1992. *Verbmobil: a translation system for face-to-face dialog*. Stanford: CSLI Publications.

Keenan, Edward. 1992. Anaphora, language variation, and language universals. In J. Guenter, B. Kaiser & C. Zoll (eds.), *Proceedings of the 19th annual meeting of the berkeley linguistics society*, 503–516. Berkeley: Berkeley Linguistics Society.

Keenan, Edward & Leonard Faltz. 1985. *Boolean semantics for natural language*. Dordrect: Reidel.

Klein, Ewan & Ivan Sag. 1985. Type-driven translation. *Linguistics and Philosophy* 8. 163–201.

Koenig, Jean-Pierre & Anthony Davis. 2003. Semantically transparent linking in HPSG. In Stefan Müller (ed.), *Proceedings of the HPSG 2003 conference*, 222–235. Stanford: CSLI Publications.

Ladusaw, William. 1988. A proposed distinction between *Levels* and *Strata*. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm 2*. Seoul: Hanshin Publishing Co.

Lahm, David. 2016. ‘‘different” as a restriction on Skolem functions. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burdorf (eds.), *Proceedings of the 26th semantics and linguistic theory conference (SALT 26)*, 546–565. LSA. DOI:10.3765/salt.v26i0.3793

Lahm, David. 2018. Plural in Lexical Resource Semantics. In Stefan Müller & Frank Richter (eds.), *Proceedings of the 25th International Conference on Head-Driven Phrase Structure Grammar, University of Tokyo*, 89–109. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2018/hpsg2018-lahm.pdf.

May, Robert. 1985. *Logical form: its structure and derivation*. Cambridge, MA: MIT Press.

Montague, Richard. 1974. *Formal philosophy*. New Haven: Yale University Press.

Partee, Barbara. 1984. Compositionality. In Fred Landman & Frank Veltman (eds.), *Varieties of formal semantics*, 281–311. Dordrecht: Foris.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Pollard, Carl & Ivan Sag. 1987. *Information-based syntax and semantics*. Stanford: CSLI Publications.

Pollard, Carl & Eun Jung Yoo. 1998. A unified theory of scope for quantifiers and *wh*-phrases. *Journal of Linguistics* 34. 415–445.

Przepiórkowski, Adam. 1998. 'A Unified Theory of Scope' revisited: Quantifier retrieval without spurious ambiguities. In Gosse Bouma, Geert-Jan Kruijff & Richard T. Oehrle (eds.), *Procedings of the FHCG98. Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar and Categorial Grammar*, 185–195. Saarbrücken. http://nlp.ipipan.waw.pl/~adamp/Papers/1998-fhcg/.

Richter, Frank. 2004. *A mathematical formalism for linguistic theories with an application in Head-Driven Phrase Structure Grammar*. Universität Tübingen Phil. Dissertation (2000). https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46230, accessed 2018-2-25.

Richter, Frank. 2016. Categorematic unreducible polyadic quantifiers in Lexical Resource Semantics. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 599–619. Stanford, CA: CSLI Publications.

Richter, Frank & Manfred Sailer. 2001. Polish negation and lexical resource semantics. In Geert-Jan Kruijff, Lawrence S. Moss & Richard T. Oehrle (eds.), *Proceedings FG-MOL 2001* (Electronic Notes in Theoretical Computer Science 53). Elsevier Science.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In *ESSLLI 2003 course material*, vol. 5 (Collegium Logicum). Kurt Gödel Society Wien.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sailer, Manfred. 2000. *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar*. Eberhard-Karls-Universität Tübingen Dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46191, accessed 2018-2-25.

Searle, John R. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge, UK: Cambridge University Press.

Zeevat, Henk. 1989. A compositional approach to discourse representation theory. *Linguistics and Philosophy*. 1st ser. 12. 95–131.

Chapter D

Information structure

Kordula De Kuthy

Universität Tübingen

Information structure as the hinge between sentence and discourse has been at the center of interest for linguists working in different areas such as semantics, syntax or prosody for several decades. A constraint-based grammar formalism such as HPSG that encodes multiple levels of linguistic representation within the architecture of signs opens up the possibility to elegantly integrate such information about discourse properties. In this chapter, I discuss a number of approaches that have explored how to best integrate information structure as a separate level into the representation of signs. I discuss which lexical and phrasal principles have been implemented in these approaches and how they constrain the distribution of the various information structural features. Finally, I discuss how the various approaches are used to formulate theories about the interaction of syntax, prosody and information structure. In particular, we will see several cases where (word order) principles that used to be stipulated in syntax can now be formulated as an interaction of syntax and discourse properties.

1 Introduction

The *information structure* of a sentence captures how the meaning expressed by the sentence is integrated into the discourse. The *information structure* thus encodes which part of an utterance is informative in which way, in relation to a particular context. A wide range of approaches exists with respect to the question of what should be regarded as the primitives of the information structure.

It is now commonly assumed that there are three basic dimensions of information structure[1] that are encoded in natural languages and that have been assumed as the basic primitives: (i) a distinction between what is new information

[1] For a comprehensive overview of the different research strands with respect to the information structural dimension, see Kruijff-Korbayová & Steedman (2003).

Kordula De Kuthy. 2019. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 709–743. Berlin: Language Science Press. DOI:??

advancing the discourse (*focus*) and what is known, i.e., anchoring the sentence in existing (or presupposed) knowledge or discourse (*background*), (ii) a distinction between what the utterance is about (*topic*, *theme*) and what the speaker has to say about it (*comment*, *rheme*), and (iii) a dimension referred to as *information status* where entities that have already been mentioned in the discourse (*given*) are distinguished from those that have not been mentioned (*new*). For all three ways of partitioning the information structure, we find approaches within the HPSG framework. Example (1) illustrates how one utterance in the context of a question can be structured according to different partitionings of information structure.

(1) Q: What does Sarah drink?

	background		*focus*
A:	Sarah	drinks	TEA.
	topic	*comment*	

The focus/background division with focus as the part of an utterance that is informative with respect to the discourse is one of the most commonly adopted partitionings when studying information structure, and thus many approaches within the HPSG architecture as well assume a division into focus and background, such as the ones that will be discussed in this article: Engdahl & Vallduví (1996), De Kuthy (2002), Webelhuth (2007), Paggio (2009), Bildhauer (2008), Song & Bender (2012) and Song (2017). Less common within the HPSG framework are approaches that take topic, i.e. the material that an utterance is about, as the central notion and assume topic and comment (or theme and rheme) as the primitives of the information structure. Most approaches discussed here assume that the background has one designated (mostly referential) element functioning as the topic (or link), among them Engdahl & Vallduví (1996), De Kuthy (2002), Paggio (2009) and Song (2017).

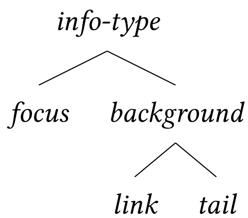
With respect to information status (including primitives such as new and given mentioned above), the discourse status of referential elements is of interest, i.e. whether they can be linked to previously mentioned items, i.e. whether they are (discourse) old or given, or whether they haven't been mentioned before and are thus (discourse) new. The representation of information status has received comparatively little interest within the HPSG community; the approach by De Kuthy & Meurers (2011) is one of the few that explicitly integrate this dimension into their information structural architecture.

The need to represent discourse properties in a grammar architecture of signs results from the insight that in many, if not all, languages, the way utterances are realized via their syntactic structure, morphological patterns and prosody very often interacts with discourse requirements of these utterances. In other words, approaches dealing with constraints on word order in a particular construction need to encode that this particular word order is only grammatical given a particular context, or a particular accent pattern has to be connected to a particular discourse status of the accented elements.[2] Most of the approaches discussed here deal with such interface questions, and I therefore discuss the particular word order and phonetic theories that have been implemented in Sections 6 and 7 in detail. As a starting point, however, I will first discuss the various architectural designs that have been implemented in order to be able to formulate the specific theories integrating discourse constraints.

2 Information structure in the architecture of signs

Several ways of representing information structure within the architecture of signs have been pursued as part of the HPSG framework: one of the earliest approaches, which is similar to the idea of F-marking as pursued in many syntax-based approaches to information structure in Generative Grammar (such as Jackendoff 1972; Selkirk 1982), has been proposed by Manandhar (1994b). He assumes that all signs have an additional appropriate feature INFO-STRUC which takes as its value objects of the sort *info-type*. A sign can then have one of the subtypes of *info-type* shown in (2) as its informational marking.

(2) Type hierarchy under *info-type* of Manandhar (1994b):



The distribution of the INFO-STRUC values in a sign is determined by the *Focus Inheritance Principle*, which enforces that in every phrase, the INFO-STRUC value of the mother subsumes the values of the INFO-STRUC of all of its daughters. The consequence of this principle is that if one daughter in a phrase is in the focus

[2]For some examples in the literature where this has been explored for word order phenomena, see for example Ambridge & Goldberg (2008), De Kuthy & Konietzko (2019) and Culicover & Winkler (2019).

and the other one in the background, then the mother's INFO-STRUC value is the smallest common supertype of both, namely *info-type*.

There are two problematic aspects of such an architecture. Firstly, it leads to a proliferation of syntactic markup of non-syntactic properties, in particular once one considers the full range of information structural notions, such as focus and focus projection, multiple foci and the marking of other discourse functions such as topic. And secondly, the perspective of information structure as resulting from an independent interpretation process of syntactic markup does not support a view of syntax, information structure and intonation as directly interacting modules, a view that can be nicely implemented in a multi-layer framework such as HPSG. More common are thus approaches that encode the information structure as a separate layer, i.e., a feature with its own structural representation.

In the original setup of signs introduced in Pollard & Sag (1994), the feature CONTEXT is introduced as part of *local* objects as a place to encode information relating to the pragmatic context (and other pragmatic properties) of utterances. In Engdahl & Vallduví (1996) it is argued that it would be most natural to also represent information structural information as part of this CONTEXT feature. Engdahl & Vallduví (1996) thus introduce the feature INFO-STRUC as part of the CONTEXT and since they couch their approach into Vallduví's (1992) information packaging account, INFO-STRUC is further divided into FOCUS and GROUND. All INFO-STRUC features take entire signs as their values. The complete specification is shown in (3).

(3) Information structure in Engdahl & Vallduví (1996: 56)

$$\begin{bmatrix} \textit{sign} \\ \text{SYNSEM}|\text{LOCAL}|\text{CONTEXT}|\text{INFO-STRUC} \begin{bmatrix} \text{FOCUS} & \textit{sign} \\ \text{GROUND} & \begin{bmatrix} \text{LINK} & \textit{sign} \\ \text{TAIL} & \textit{sign} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Another approach locating the representation of information structure within the CONTEXT feature is the one by Paggio (2009) as part of a grammar of Danish. The INFOSTR features TOPIC, FOCUS and BG take as their values lists of indices. Since Paggio (2009) includes Minimal Recursion Semantics (MRS, Copestake et al. 2005) as the semantic representation,[3] these indices can be structure-shared with the argument indices of the semantic relations collected on the RELS list of the content of a sign. The basic setup is illustrated in (4).

(4) Information structure in Paggio (2009: 149):

[3] A detailed discussion of the properties and principles of MRS as implemented in HPSG can be found in Koenig & Richter (2019), Chapter C of this volume.

$$\begin{bmatrix} \textit{sign} & \\ \textsc{synsem}|\textsc{local}|\textsc{context}|\textsc{infostr} & \begin{bmatrix} \textsc{focus} & \textit{list-of-indices} \\ \textsc{topic} & \textit{list-of-indices} \\ \textsc{bg} & \textit{list-of-indices} \end{bmatrix} \end{bmatrix}$$

Several approaches encode information structure as part of the CONTENT, such as Song (2017) and Song & Bender (2012). Since they also use MRS as the semantic representation language, they enrich the architecture of *mrs* structures. The information structure itself is encoded via a feature ICONS (individual constraints) that is introduced parallel to HCONS (handle constraints) as part of the CONTENT, as shown in (5). Song (2017) and Song & Bender (2012) use *diff-list* as values for the features RELS, HCONS and ICONS (expressed by the "!" at the beginning and the end of the list). This type of list includes an explicit pointer to the last element of the list.

(5) Information structure in Song & Bender (2012) and Song (2017: 116):

$$\begin{bmatrix} \textit{sign} & \\ \textsc{synsem}|\textsc{local}|\textsc{content} & \begin{bmatrix} \textit{mrs} & \\ \textsc{hook} & \begin{bmatrix} \textit{hook} & \\ \textsc{icons-key} & \textit{info-str} \\ \textsc{clause-key} & \textit{event} \end{bmatrix} \\ \textsc{rels} & \textit{diff-list} \\ \textsc{hcons} & \textit{diff-list} \\ \textsc{icons} & \left\langle !\ \ldots, \begin{bmatrix} \textit{info-str} & \\ \textsc{clause} & \textit{individual} \\ \textsc{target} & \textit{individual} \end{bmatrix}, \ldots\ ! \right\rangle \end{bmatrix} \end{bmatrix}$$

The type *info-str* used as the value for elements on the ICONS list is divided into an elaborate hierarchy with several subtypes, such as *semantic-focus, constrast-focus, focus-or-topic, non-focus*, etc. (cf. Song 2017: 114). The elements of type *info-str* on the ICONS list have two appropriate features CLAUSE and TARGET. TARGET is always structure-shared with the respective sign's ARG0 value, and the value of CLAUSE is always structure-shared with the INDEX value of the predicate that is the semantic head of the clause.

As pointed out by De Kuthy (2002), assuming that the information structure is part of *local* objects (which it is if it is part of the CONTENT in HPSG as proposed by Engdahl & Vallduví 1996 or part of the CONTENT) is problematic in connection with a trace-based account of unbounded dependencies. Traces should not contribute anything to the information structure of a sentence. If one wants to develop an information structure approach which is independent of the decision of which kind of UDC theory one assumes, the only options for placing the information structure attribute are under *synsem* objects or at the top level.

Information structure as part of *synsem* objects would suggest that it plays a role in syntactic selection. This possibility is assumed in Bildhauer & Cook (2011), and they thus represent INFO-STRUC as a feature appropriate for *synsem* objects (their account will be discussed in more detail in Section 6). A third possibility is argued for in De Kuthy (2002) and Bildhauer (2008), namely that information structure should not be part of *synsem* objects. As a result, they encode information structure again as an additional feature of signs (similar to the approach by Manandhar 1994b discussed above), but it is argued that the appropriate values should be semantic representations. Using indices as the value of information structure-related features (as in the approaches by Paggio 2009, Song & Bender 2012 and Song 2017) is again problematic whenever two constituents share their index value, but only one of them is assigned a particular information structural function. For example, under the assumption that in a head-adjunct phrase the index is structure-shared between an intersective adjective and the nominal head (as in *red car*), there is no way to relate a particular information structure function (e. g., contrast) to the adjective alone (as in *RED car*).

In De Kuthy (2002), a tripartite partition of information structure into focus, topic and background is introduced. As to the question of what kinds of objects should be defined as the values of these features, De Kuthy proposes the values of the INFO-STRUC features to be chunks of semantic information. It is argued that the semantic representation proposed in Pollard & Sag (1994) is not appropriate for her purpose, because the semantic composition is not done in parallel with the syntactic build-up of a phrase. Instead, the Montague-style (cf. Dowty et al. 1981) semantic representation for HPSG proposed in Sailer (2000) is adopted, in which CONTENT values are regarded as representations of a symbolic language with a model-theoretic interpretation. As the semantic object language under CONTENT the language Ty2 (cf. Gallin 1975) of two-sorted type theory is chosen. The resulting feature architecture is shown in (6).

(6) The structure of INFO-STRUC in De Kuthy (2002: 165):

$$\begin{bmatrix} \textit{sign} & \\ \text{PHON} & \textit{list} \\ \text{SYNSEM} & \textit{synsem} \\ \text{INFO-STRUC} & \begin{bmatrix} \textit{info-struc} & \\ \text{FOCUS} & \textit{list-of-mes} \\ \text{TOPIC} & \textit{list-of-mes} \end{bmatrix} \end{bmatrix}$$

The information structure is encoded in the attribute INFO-STRUC that is appropriate for signs and has the appropriate features FOCUS and TOPIC, with lists of so-called meaningful expressions (semantic terms, cf. Sailer 2000) as values. These meaningful expressions (that are also used as the representation of logical forms as the CONT values) are lambda terms formulated in a predicate logic language as discussed in more detail in Section 3.2.2 in (13).

3 Information structure principles

The approaches sketched above all assume that signs contain some kind of structural representation of information structure, with the consequence that they need to introduce principles that constrain the values of the information structural features. Most approaches thus formulate two types of principles as part of their grammar fragment: one set of principles on the lexical level tying information structure to word level properties such as accents, and another set of principles on the phrasal level determining the distribution of information structure values between mother and daughters in a phrase.

3.1 Instantiating information structure on the word level

In the approach of Engdahl & Vallduví (1996), prosodic properties of English, in particular accent placement, are tied to specific information structural properties of words and phrases. On the word level, they thus introduce two principles that instantiate the information structure FOCUS and LINK when the word has a particular accent. The two principles are shown in (7).

(7) Information structure of words (Engdahl & Vallduví 1996: 56):

$$\textit{word} \longrightarrow \boxed{1}\begin{bmatrix}\textsc{phon}|\textsc{accent} & A \\ \textsc{info-struc}|\textsc{focus} & \boxed{1}\end{bmatrix}$$

$$\textit{word} \longrightarrow \boxed{1}\begin{bmatrix}\textsc{phon}|\textsc{accent} & B \\ \textsc{info-struc}|\textsc{ground}|\textsc{link} & \boxed{1}\end{bmatrix}$$

Words with an A accent always contribute focal information, i.e. the entire sign is structure-shared with the INFO-STRUC|FOCUS value; words carrying a B accent contribute link information, i.e. the entire sign is structure-shared with the INFO-STRUC|GROUND|LINK value.[4]

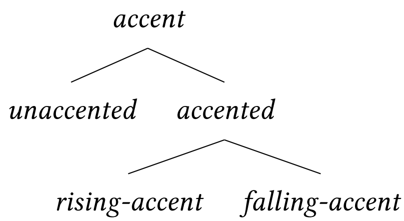
A similar set of word level principles is introduced in the approach of De Kuthy (2002), where the information structure of utterances in German is also tied to

[4]The usage of the terms “A accent” and “B accent” goes back to Jackendoff (1972).

words carrying particular accent patterns. The phonology of signs is altered as shown in Figure 8 to include an ACCENT attribute to encode whether a word receives an accent or not, and whether it is a rising or falling accent, should it receive one.

(8) Representing pitch accents and accent type hierarchy according to De Kuthy (2002: 166):

$$\begin{bmatrix} \textit{sign} & \\ \text{PHON} & \begin{bmatrix} \text{PHON-STRING} & \textit{list} \\ \text{ACCENT} & \textit{accent} \end{bmatrix} \end{bmatrix}$$



The information structure of words is defined through the principle shown in Figure 9 which assigns the semantic contribution of the word to the focus or topic specification in the information structure representation of that word, depending on the type of accent the word receives.

(9) Principle assigning information structure to words (De Kuthy 2002: 167):

$$\textit{word} \longrightarrow \begin{bmatrix} \text{PHON|ACCENT} & \textit{falling-accent} \\ \text{SS|LOC|CONT|LF} & \boxed{1} \\ \text{INFO-STRUC} & \begin{bmatrix} \text{FOCUS} & \langle \boxed{1} \rangle \\ \text{TOPIC} & \langle \rangle \end{bmatrix} \end{bmatrix} \vee$$

$$\begin{bmatrix} \text{PHON|ACCENT} & \textit{unaccented} \\ \text{INFO-STRUC} & \begin{bmatrix} \text{FOCUS} & \langle \rangle \\ \text{TOPIC} & \langle \rangle \end{bmatrix} \end{bmatrix} \vee \ldots$$

Here only two cases are spelled out, one for *falling-accent* signalling focus, and one for unaccented words not contributing anything to the information structure. Other possible cases could for example be a specific accent (like a fall-rise) signalling topic, i.e. a non-empty TOPIC list.

In the approach of Song (2017), lexical items are subtypes of four different *icons-lex-item* types, which specify whether lexical items can contribute any information structural information to the *icons* list, and if yes, how many items can do this. These four lexical subtypes are shown in (10).

(10) Lexical types specifying ICONS values (Song 2017: 137)

a. $\begin{bmatrix} \textit{no-icons-lex-item} \\ \text{MKG} & \begin{bmatrix} \text{FC} & \textit{na} \\ \text{TP} & \textit{na} \end{bmatrix} \\ \text{ICONS} & \langle\, !\, !\, \rangle \end{bmatrix}$

b. $\begin{bmatrix} \textit{basic-icons-lex-item} \\ \text{ICONS} & \langle\, !\, !\, \rangle \end{bmatrix}$

c. $\begin{bmatrix} \textit{one-icons-lex-item} \\ \text{ICONS} & \langle\, ![\,]!\, \rangle \end{bmatrix}$

d. $\begin{bmatrix} \textit{two-icons-lex-item} \\ \text{ICONS} & \langle\, ![\,],[\,]!\, \rangle \end{bmatrix}$

Lexical entries for elements that cannot be marked with respect to information structure are of type *no-icons-lex-items*, such as relative pronouns or expletives in English. Nominal items, such as common nouns, proper nouns and pronouns, have lexical entries of type *basic-icons-lex-item*. These types of words can have an information structural marking, but don't have to. The two other lexical subtypes are used for verbs with one clausal argument (*one-icons-lex-item*) or two clausal arguments (*two-icons-lex-items*). The information structural contribution of these clausal arguments then has to be part of the verb's ICONS list. All other verbs are not required to have any elements on their ICONS list and can thus also be of type *basic-icons-lex-item*.

To capture further constraints on the information structure properties on the word level, such as accent patterns triggering focus or topic, lexical rules are formulated in Song (2017) that derive lexical entries with the respective specifications. One such set of lexical rules for A and B accents in English is discussed in Section 7.

3.2 Information structure principles on the phrasal level

3.2.1 Information packaging (Engdahl & Vallduví 1996)

One of the first approaches integrating an explicit representation of information structure into the HPSG architecture, Engdahl & Vallduví (1996) encode the information structure as part of the CONTEXT of signs with the help of an additional feature INFO-STRUC. As discussed above, on the lexical level the instantiation of these features can be triggered by phonetic properties, such as certain accents, for intonation languages like English. Phrasal signs must then satisfy the INFO-STRUC instantiation constraints in (11).[5]

[5]Engdahl and Vallduví's formulation of the principle is incompatible with the model theoretic view of HPSG in Pollard & Sag (1994). Feature structures are complete models of objects, thus

(11) INFO-STRUC *instantiation principles for English:*

Either (i) if a DAUGHTER's INFO-STRUC is instantiated, then the mother inherits this instantiation (for narrow foci, links and tails),

or (ii) if the most oblique DAUGHTER's FOCUS is instantiated, then the FOCUS of the mother is the sign itself (wide focus).

An example including a wide VP focus licensed by the principle in (11) with the relevant INFO-STRUC values is shown in Figure 1.

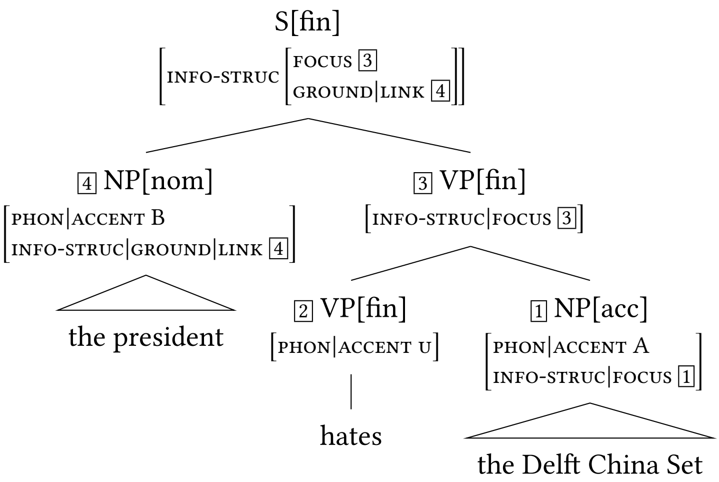


Figure 1: An example for VP focus in Engdahl & Vallduví (1996: 59)

In this example, the rightmost NP daughter *the Delft China Set* carries an A accent. According to the principle in (7) shown earlier, the entire sign is thus structure-shared with the focus value (or, in Engdahl & Valluvi's terms, the focus value "is instantiated"). As a consequence, the second clause of the principle in (11) applies and the focus value of the VP mother is the sign itself, which is then inherited by the sentence. Several aspects of the licensing of the structure in Figure 1 are not properly spelled out in Engdahl & Vallduvi's approach. For example, the analysis seems to presuppose a set of additional principles for focus inheritance in nominal phrases which do not straightforwardly follow from the principles formulated in (11).

there is no way in which a value can not be instantiated in a feature structure. Only descriptions of feature structures can be underspecified, but not the feature structures themselves.

3.2.2 Information structure as structured meanings (De Kuthy 2002)

The so-called structured meaning approach (von Stechow 1981; Jacobs 1983; Krifka 1992) to information structure provides a compositional semantic mechanism based on separate representations of the semantic contribution of the focus and that of the background. De Kuthy (2002), De Kuthy & Meurers (2003) and Webelhuth (2007) worked out how such a structured meaning approach can be integrated into the HPSG architecture.

As discussed above, in De Kuthy (2002), the information structure is encoded in the attribute INFO-STRUC that is appropriate for signs and has the appropriate features FOCUS and TOPIC, with lists of so-called meaningful expressions as values. The background of a sentence in De Kuthy's approach is then defined to be that part of the logical form of the sentence which is neither in focus nor in topic. This characterization of background closely resembles the definition of background employed by the *structured meaning* approach to focus (cf. Krifka 1992). The INFO-STRUC value of a simple sentence with the focus as indicated in (12) is thus structured as shown in (13).

(12) Peter [[liest ein BUCH]]$_F$.
Peter reads a book

(13) A sign representation including information structure (De Kuthy 2002: 163):

$$\begin{bmatrix} \textsc{s}|\textsc{loc}|\textsc{cont}|\textsc{lf} & \exists x[\mathit{book}'(x) \wedge \mathit{read}'(p,x)] \\ \textsc{info-struc} & \begin{bmatrix} \textsc{focus} & \langle \lambda y \exists x[\mathit{book}'(x) \wedge \mathit{read}'(y,x)]\rangle \\ \textsc{topic} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

The information-structure values of phrases are constrained by principles such as the one in (14). The original principle formulated in De Kuthy (2002: 169) only contains the first two disjuncts shown in (14). The third disjunct is added in De Kuthy & Meurers (2003). Sentences where the focus or the topic does not project represent the most basic case: only those words bearing an accent are in the topic or in the focus of an utterance.

(14) Principle 1: Extended focus projection principle (De Kuthy & Meurers 2003):

$$\textit{phrase} \rightarrow \begin{bmatrix} \text{INFO-STR}|\text{FOCUS } \boxed{1} \oplus \textit{collect-focus}\left(\boxed{2}\right) \\ \text{HEAD-DTR}|\text{INFO-STR}|\text{FOCUS } \boxed{1} \\ \text{NON-HEAD-DTRS } \boxed{2} \end{bmatrix} \lor$$

$$\begin{bmatrix} \text{PHON}|\text{PHON-STR } \textit{list} \oplus \boxed{2} \\ \text{SS}|\text{LOC} \begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{noun} \lor \textit{prep} \\ \text{CONT}|\text{LF} & \boxed{3} \end{bmatrix} \\ \text{INFO-STR}|\text{FOCUS } \langle \boxed{3} \rangle \\ \textit{any-dtr}\left(\begin{bmatrix} \text{PHON}|\text{PHON-STR} & \boxed{2} \\ \text{SS}|\text{L}|\text{CONT}|\text{LF} & \boxed{4} \\ \text{INFO-STR}|\text{FOCUS} & \langle \boxed{4} \rangle \end{bmatrix}\right) \end{bmatrix} \lor$$

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC} \begin{bmatrix} \text{CAT}|\text{HEAD} & \textit{verb} \\ \text{CONT}|\text{LF} & \boxed{3} \end{bmatrix} \\ \text{INFO-STR}|\text{FOCUS } \langle \boxed{3} \rangle \\ \text{NON-HEAD-DTRS } \left\langle \ldots, \begin{bmatrix} \text{SYNSEM} \begin{bmatrix} \text{FPP } \textit{plus} \\ \text{LOC}|\text{CONT}|\text{LF } \boxed{4} \end{bmatrix} \\ \text{INFO-STR}|\text{FOCUS } \langle \boxed{4} \rangle \end{bmatrix}, \ldots \right\rangle \end{bmatrix} \lor \ \ldots$$

In this case, the mother of a phrase just collects the focus values of all her daughters as ensured by the principle in (14).[6] The relation *collect-focus* ensures that from the list of non-head daughters, the FOCUS value of every non-head daughter is added to the list of FOCUS values of the entire phrase. A similar principle is needed to determine the TOPIC value of phrases.

For cases of so-called focus projection[7] in NPs and PPs, it is assumed in De Kuthy (2002: 169) that it is sufficient to express that the entire NP (or PP) can be focused if the rightmost constituent in that NP (or PP) is focused, as expressed by the second disjunct of the principle in (14). If focus projection is possible in a certain configuration then this is always optional, therefore the focus projection principle for nouns and prepositions is formulated as a disjunct. The second

[6] The presentation differs from that in De Kuthy (2002); it is the one from De Kuthy & Meurers (2003). Definitions of the auxiliary relations:

$$\textit{any-dtr}\left(\boxed{1}\right) := \left[\text{HEAD-DTR } \boxed{1}\right].$$
$$\textit{any-dtr}\left(\boxed{1}\right) := \left[\text{NON-HEAD-DTRS } \text{element}\left(\boxed{1}\right)\right].$$

$$\textit{collect-focus}\left(\langle\rangle\right) := \langle\rangle.$$
$$\textit{collect-focus}\left(\left\langle \left[\text{INFO-STRUC}|\text{FOCUS } \langle\boxed{1}\rangle\right] \mid \boxed{2} \right\rangle\right) := \left\langle \boxed{1} \mid \textit{collect-focus}\left(\boxed{2}\right)\right\rangle.$$

[7] Focus projection is a term commonly used to describe the fact that in an utterance with prosodic marking of focus on a word, this marking can lead to ambiguity, in that different constituents containing the word can be interpreted as focused (cf. Gussenhoven 1983; Selkirk 1995).

disjunct of the principle in (14) ensures that a phrase headed by a noun or a preposition can only be in the focus (i.e., its entire logical form is token identical to its FOCUS value) if the daughter that contributes the rightmost part of the phonology of the phrase is entirely focused itself. The relation *any-dtr* is a description of a sign with a head daughter or a list of non-head daughters and thereby ensures that it can be either the head (i.e. head daughter) of the phrase itself, or any non-head daughter that meets the condition of being focused. Again, a similar principle needs to be provided for the TOPIC value of nominal and prepositional phrases.

For the verbal domain, the regularities are known to be influenced by a variety of factors, such as the word order and lexical properties of the verbal head (cf., e.g., von Stechow & Uhmann 1986). Since verbs need to be able to lexically mark which of their arguments can project focus when they are accented, De Kuthy & Meurers (2003) introduce the boolean-valued feature FOCUS-PROJECTION-POTENTIAL (FPP) for objects of type *synsem*. Figure 15 shows the relevant part of the lexical entry of the verb *lieben* 'love' which allows projection from the object but not from the subject:

(15) The focus projection potential of *lieben* (De Kuthy & Meurers 2003):

$$\begin{bmatrix} \text{PHON|PHON-STR} \ \langle \textit{lieben} \rangle \\ \text{ARG-S} \ \left\langle \begin{bmatrix} \text{LOC|CAT|HEAD} \begin{bmatrix} \textit{noun} \\ \text{CASE} \ \ \textit{nom} \end{bmatrix} \\ \text{FPP} \ \textit{minus} \end{bmatrix}, \begin{bmatrix} \text{LOC|CAT|HEAD} \begin{bmatrix} \textit{noun} \\ \text{CASE} \ \ \textit{acc} \end{bmatrix} \\ \text{FPP} \ \textit{plus} \end{bmatrix} \right\rangle \end{bmatrix}$$

The third disjunct of the principle in (14) then specifies under which circumstances focus can project in the verbal domain: a phrase headed by a verb can only be in the focus (i.e., its entire logical form is token identical to an element of its FOCUS value) if the daughter that has the focus projection potential (FPP *plus*) is entirely focused itself.

3.2.3 Information structure principles in MRS

As introduced above, in the MRS based approach of Paggio (2009), the information structure is part of the CONTEXT, consisting of FOCUS, TOPIC and BACKGROUND features which are structure-shared with the respective INDEX values of the semantic representation of a phrase. Paggio (2009) connects the distribution of information structure values to particular clausal types and introduces new phrasal subtypes which constrain the distribution of information structure in the respective phrase. One such new phrasal subtype is the type *focus-inheritance* as defined in (16), which then has to be cross-classified with every basic phrasal subtype (such as *hd-comp, hd-spec, hd-adj*, etc.) in order to constrain the distribution of focus values across all phrasal subtypes.

(16) Principle for focus-inheritance (Paggio 2009: 155):

$$\textit{focus-inheritance} \longrightarrow \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{context} & \begin{bmatrix} \textsc{focus} & \langle \boxed{2}, \boxed{1} \rangle \\ \textsc{bg} & \boxed{3} \end{bmatrix} \\ \textsc{hd}|\textsc{synsem}|\textsc{loc}|\textsc{context} & \begin{bmatrix} \textsc{focus} & \boxed{1} \\ \textsc{bg} & \boxed{3} \end{bmatrix} \\ \textsc{non-hd} & \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{context}|\textsc{focus} & \langle \boxed{2} \rangle \\ \textsc{accent} & \textit{true} \end{bmatrix} \end{bmatrix}$$

The principle in (16) ensures that for signs of type *focus-inheritance*, the list of focus values of the mother is the list of focus values of the head daughter[8] plus the focus value of the non-head daughter, in case it is accented. Similar principles are defined for the inheritance of background values, also depending on the accent status of the non-head daughter. Paggio further assumes that each phrasal subtype has further subtypes connecting it to one of the information structure inheritance phrasal types. For example, she assumes that there is a phrasal subtype *focus-hd-adj* that is a subtype both of *hd-adj* and of *focus-inheritance*. Finally, clausal types are introduced that account for the information structure values at the top level of a clause. For example, the specification for *decl-main-all-focus* as shown in (17) is a clause in which both the background and the topic values are empty and the mother collects the focus values from the head and the non-head daughters.[9]

(17) Declarative all-focus construction (Paggio 2009: 160):

$$\begin{bmatrix} \textit{decl-main-all-focus} & \\ \textsc{ctxt}|\ldots & \begin{bmatrix} \textit{all-focus} & \\ \textsc{topic} & \langle\rangle \\ \textsc{focus} & \langle \boxed{2}, \boxed{1} \rangle \\ \textsc{bg} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textsc{ctxt}|\ldots|\textsc{focus} & \langle \boxed{1} \rangle \end{bmatrix} \qquad \begin{bmatrix} \textsc{ctxt}|\ldots|\textsc{focus} & \boxed{2} \end{bmatrix}$$

[8] This is not correctly specified in the original principle as formulated by Paggio (2009). If the head daughter can have a list with more than one element as its FOCUS value, then this entire list would have to be added to the list of FOCUS values of the mother, and not just be one element of that list.

[9] Again, the list specifications as formulated by Paggio (2009) are not entirely correct: if the head daughter's FOCUS value $\boxed{2}$ is a list with more than one element, this list has to be added to the list of FOCUS values of the mother.

Different from Paggio's approach, Song & Bender (2012) and Song (2017) locate the representation of information structure within the MRS-based CONTENT value of signs. The list elements of information structural values that are built up for a phrase consist of focus, background or topic elements co-indexed with the semantic INDEX values of the daughters of that phrase. The main point of their approach is that they want to be able to represent underspecified information structural values, since very often a phrase, for example with a certain accent pattern, is ambiguous with respect to the context in which it can occur and thus is ambiguous with respect to its information structure values. An example they discuss is the one in (18), where the first sentence could be an answer to the question *What barks?* and thus signal narrow focus, whereas the second utterance could be an answer to the question *What happened?* and signal broad focus.

(18) a. [[The DOG]]$_F$ barks.
   b. [[The DOG barks]]$_F$.

The approach pursued in Song & Bender (2012) thus assumes that the two possible readings in (18) are further specializations of one MRS which is associated with one syntactic structure and includes underspecified values, in particular the type of the ICONS element for the constituent *barks*, leaving it open whether that is part of the focus or not.

In Song (2017), this approach is further spelled out and lexical rules allowing transitive and ditransitive verbs to be a possible source for focus projection. In an example such as (19), Song (2017) assumes that focus can only project if the last argument is accented as in (19b) (here accent is shown on the noun *book* in small caps), but not if some other argument is accented, as in (19a), where the proper noun *Lee* is accented.

(19) a. Kim sent LEE the book.
   b. Kim sent Lee the BOOK.

Accordingly, there are two lexical entries for the verb *send*, which are derived by the lexical rules shown in (20).

(20) Focus projection lexical rules (Song 2017: 227):
   a. *no-focus-projection-rule* $\longrightarrow$

$$\begin{bmatrix} \textsc{index} & \boxed{1} \\ \textsc{icons-key} & \boxed{2} \\ \textsc{val} & \begin{bmatrix} \textsc{subj} & \left\langle \begin{bmatrix} \textsc{icons-key non-focus} \end{bmatrix} \right\rangle \\ \textsc{comps} & \left\langle \begin{bmatrix} \textsc{mkg}|\textsc{fc} + \end{bmatrix} \begin{bmatrix} \textsc{mkg}|\textsc{fc} - \\ \textsc{icons} \langle ! \, ! \rangle \end{bmatrix} \right\rangle \end{bmatrix} \\ \textsc{c-cont}|\textsc{icons} & \langle \, ! \, \boxed{2} \begin{bmatrix} \textit{non-focus} \\ \textsc{target} \; \boxed{1} \end{bmatrix} ! \rangle \\ \textsc{dtr} \; \textit{lex-rule-infl-affixed} & \end{bmatrix}$$

b. *focus-projection-rule* $\longrightarrow$

$$\begin{bmatrix} \textsc{clause-key} & \boxed{1} \\ \textsc{val}|\textsc{comps} & \left\langle \begin{bmatrix} \textsc{mkg}|\textsc{fc} - \\ \textsc{index} \quad \boxed{2} \end{bmatrix} \begin{bmatrix} \textsc{mkg}|\textsc{fc} + \\ \textsc{icons} \langle \, ! \begin{bmatrix} \textit{semantic-focus} \end{bmatrix} ! \rangle \end{bmatrix} \right\rangle \\ \textsc{c-cont}|\textsc{icons} & \langle \, ! \begin{bmatrix} \textit{non-focus} \\ \textsc{target} \; \boxed{2} \\ \textsc{clause} \; \boxed{1} \end{bmatrix} ! \rangle \\ \textsc{dtr} \; \textit{lex-rule-infl-affixed} & \end{bmatrix}$$

The lexical rule *no-focus-projection-rule* requires lexical entries to have a non-focus-marked element as the last element on the COMPS list, and in addition the word itself has an ICONS-KEY of type *non-focus* preventing the word itself from being focused. The lexical rule *focus-projection-rule* has a focus-marked element as the last element in the COMPS list. It is not further specified whether only that focussed complement or also the word itself contributes anything to the ICONS value. In the example (19b), if the verb *sent* is licensed by the rule *focus-projection-rule*, either only *the book*, or the entire VP *sent Lee the book*, or even the entire sentence *Kim sent Lee the book* could be focused.

Since the approach of Song (2017) is part of a larger grammar fragment (the LinGO Grammar Matrix; Bender et al. 2010) with the aim of parsing and generating sentences from a large number of different languages, it contains a multitude of lexical and phrasal types and principles. Some of these specifications are introduced to capture very language-specific information structure properties (such as morphological markings, word order constraints, etc.), while others are necessary for the specific way in which grammar fragments in the LinGO Grammar Matrix are implemented and processed. It would be far beyond the scope of this article to discuss all these principles and specifications in detail and I therefore only included the most essential aspects of Song's approach in my discussion here.

4 Topics

Most HPSG approaches are based on a focus/background division of the information structure. To capture aspects of a topic vs. comment distinction, or to be able to specify topics as a special element in the background, they include an additional feature or substructure for topics. Engdahl & Vallduví (1996), for example, divide the GROUND into LINK and TAIL, where the link is a special element of the background linking it to the previous discourse, just like topics. In the approaches of De Kuthy (2002) and Paggio (2009), an additional feature TOPIC is introduced, parallel to FOCUS and BACKGROUND, in order to distinguish discourse referents as topics from the rest of the background.

Most approaches don't introduce separate mechanisms for the distribution of TOPIC values, but rather assume that similar principles as the ones introduced for focus can constrain topic values, as mentioned above for the approach of De Kuthy (2002). A more specific example can be found in Paggio (2009), where a constraint on topicalization constructions including a topic-comment partitioning is formulated, as illustrated in Figure 2. This *inv-topic-comment* phrasal type

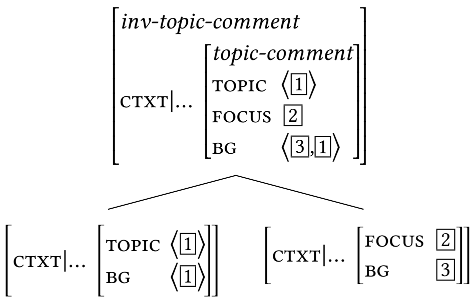


Figure 2: Topicalization construction with extracted topic

constrains the information structure values of topicalization constructions in Danish characterized by subject verb inversion,[10] where the topic corresponds to the topicalized complement, as illustrated by the example in (21) from Paggio (2009).

[10] Although Danish is generally considered to be a V2 language, where any kind of constituent (not only the subject) can occur in the position before the finite verb, Paggio (2009) seems to assume that clauses in which a dependent different from the subject, i.e. an object or some adjunct phrase, occurs before the finite verb have a different structure than those where the subject occurs in the sentence-initial position.

(21) og [i det nederste vindue]$_T$ [tager man og saetter urtepotten]$_F$
and in the lowest window takes one and puts flowerpot.DEF
'And in the lowest window you take and put the flowerpot.'

In Song (2017), a number of lexical and phrasal principles are provided with the purpose of licensing topic-comment structures. The principles and lexical entry in (22) are spelled out in order to license *wa* constructions in Japanese with a left dislocated topic phrase.

(22) Licensing topic-comment structures in Song (2017):

a. $$\textit{topic-comment} \longrightarrow \begin{bmatrix} \textsc{l-periph} & + \\ \textsc{mkg} & \textit{tp} \\ \textsc{hd}|\textsc{mkg}|\textsc{tp} & - \\ \textsc{nhd} & \begin{bmatrix} \textsc{mkg} & \textit{tp} \\ \textsc{l-periph} & + \end{bmatrix} \end{bmatrix}$$

b. $$\textit{top-scr-comp-head} \longrightarrow \begin{bmatrix} \textsc{ha}|\textsc{val}|\textsc{comps} & \langle\rangle \\ \textsc{nhd}|\textsc{icons-key} & \textit{constrast-topic} \end{bmatrix}$$

c. $$\textit{wa-marker} \longrightarrow \begin{bmatrix} \textsc{stem} & \langle \textit{wa} \rangle \\ \textsc{incons-key} & \boxed{2} \\ \textsc{mkg} & \textit{tp} \\ \textsc{comps} & \langle \begin{bmatrix} \textsc{index} & \boxed{1} \end{bmatrix} \rangle \\ \textsc{incons} & \langle ! \; \boxed{2} \begin{bmatrix} \textit{contrast-or-topic} \\ \textsc{target} \; \boxed{1} \end{bmatrix} ! \rangle \end{bmatrix}$$

The constraint in (22a) on the phrasal subtype *topic-comment* ensures that only the non-head daughter is marked as a topic, whereas the head daughter functions as the comment (and presumably contains some focused material). The specification [L-PERIPH +] indicates that a constituent with this feature value cannot be combined with another constituent leftward.

A Japanese topic-comment structure, such as the one in (23) (Song 2017: 198), is licensed by the phrasal subtype *top-scr-comp-head*, i.e. it is assumed that the fronted complement, the *wa*-marked NP *sono hon wa* 'the book' is scrambled to the left peripheral position and is interpreted as a contrastive topic phrase.

(23) sono hon wa Kim ga yomu.
Det book wa Kim nom read
'This book, Kim read.'

The topic marker *wa* in Japanese is treated as an adposition with the lexical specifications shown in (22c). The entire sentence is thus licensed as a head complement structure, where the object NP is scrambled to the sentence initial position and functions as a contrastive topic. The *tp* marking of the entire *topic-comment* phrase ensures that this phrase cannot be embedded as the comment in another *topic-comment* phrase.

5 Givenness

In De Kuthy & Meurers (2011), it is shown how the HPSG approach to information structure of De Kuthy (2002) and colleagues can be extended to capture givenness and to make the right predictions for so-called *deaccenting*, which has been shown to be widespread (Büring 2006). In contrast to Schwarzschild (1999), who spells out his approach in the framework of alternative semantics (Rooth 1992), they show how the notion of givenness can be couched in a standard structured meaning approach – thereby preserving the explicit, compositional representations of focus.

The example in (24) illustrates the necessity to include information about givenness into the information structural setup.

(24) The conference participants are renting all kind of vehicles. Yesterday, Bill came to the conference driving a red convertible and today he's arrived with a blue one.

  a. What did John rent?

  b. He (only) rented [[a GREEN convertible]]$_F$.

The context in (24) introduces some conference participants, Bill, the rental of vehicles and red and blue convertibles into the discourse. Based on this context, when considering the question (24a) asking for the object that John is renting as the focus, one can answer this question with sentence (24b), where *a green convertible* is the focus: out of all the things John could have rented, he picked a green convertible. In this focus, only *green* is new to the discourse, whereas convertibles were already given in the context, and still the entire NP is in the focus.

To capture such cases of focus projection, an additional feature GIVEN is introduced as part of the setup of De Kuthy (2002), as already discussed in Section 3.2.2. The relation between pitch accents and the information structure of words is still defined by the principle shown in (25), depending on the type of accent the word receives.

(25) Relating intonation and information structure for words (De Kuthy & Meurers 2011):

$$\textit{word} \longrightarrow \begin{bmatrix} \textsc{phon|accent} & \textit{accented} \\ \textsc{ss|loc|cont|lf} & \boxed{1} \\ \textsc{struc-meaning} & \begin{bmatrix} \textsc{focus} & \langle \boxed{1} \rangle \\ \textsc{given} & \langle \rangle \end{bmatrix} \end{bmatrix} \lor \begin{bmatrix} \textsc{phon|accent} & \textit{unaccented} \\ \textsc{struc-meaning} & \begin{bmatrix} \textsc{focus} & \langle \rangle \\ \textsc{given} & \langle \rangle \end{bmatrix} \end{bmatrix} \lor \ldots$$

In addition, the Focus Projection Principle originally introduced in De Kuthy (2002) and then extended in De Kuthy & Meurers (2003) is extended with a disjunct capturing focus projection in the presence of givenness (De Kuthy & Meurers 2011). Figure 26 shows the resulting principle.

(26) Extended Focus Projection Principle including Givenness (De Kuthy & Meurers 2011):

$$\textit{phrase} \longrightarrow \begin{bmatrix} \textsc{struc-meaning|focus}\ \boxed{1} \oplus \textit{collect-focus}\big(\boxed{2}\big) \\ \textsc{head-dtr|info-str|focus}\ \boxed{1} \\ \textsc{non-head-dtrs}\ \boxed{2} \end{bmatrix} \lor$$

$$\begin{bmatrix} \textsc{phon|phon-str}\ \textit{list} \oplus \boxed{2} \\ \textsc{ss|loc} \begin{bmatrix} \textsc{cat|head} & \textit{noun} \lor \textit{prep} \\ \textsc{cont|lf} & \boxed{3} \end{bmatrix} \\ \textsc{struc-meaning|focus}\ \langle \boxed{3} \rangle \\ \textit{any-dtr} \left(\begin{bmatrix} \textsc{phon|phon-str} & \boxed{2} \\ \textsc{ss|l|cont|lf} & \boxed{4} \\ \textsc{struc-meaning|focus} & \langle \boxed{4} \rangle \end{bmatrix} \right) \end{bmatrix} \lor$$

$$\begin{bmatrix} \textsc{synsem|loc} \begin{bmatrix} \textsc{cat|head} & \textit{verb} \\ \textsc{cont|lf} & \boxed{3} \end{bmatrix} \\ \textsc{struc-meaning|focus}\ \langle \boxed{3} \rangle \\ \textsc{non-head-dtrs}\ \left\langle \ldots, \begin{bmatrix} \textsc{synsem} \begin{bmatrix} \textsc{fpp}\ \textit{plus} \\ \textsc{loc|cont|lf}\ \boxed{4} \end{bmatrix} \\ \textsc{struc-meaning|focus}\ \langle \boxed{4} \rangle \end{bmatrix}, \ldots \right\rangle \end{bmatrix} \lor$$

$$\begin{bmatrix} \textsc{ss|loc|cont|lf}\ \boxed{3} \\ \textsc{struc-meaning|focus}\ \langle \boxed{3} \rangle \\ \textsc{dtrs-list} \left(\textit{given-sign-list} \bigcirc \left\langle \begin{bmatrix} \textsc{ss|l|cont|lf} & \boxed{4} \\ \textsc{struc-meaning|focus} & \langle \boxed{4} \rangle \end{bmatrix} \right\rangle \right) \end{bmatrix} \lor \ldots$$

The new fourth disjunct of the Extended Focus Projection Principle[11] captures the cases previously unaccounted for where given material in a focused phrase is deaccented. Focus in those examples can project from a focused daughter in a position which normally does not allow focus projection. This only is an option if all other daughters in that focused phrase are *given*. Spelling this out, the fourth

[11] The auxiliary relations are defined as:

disjunct of the principle in (26) specifies that the mother of a phrase can be in the focus (i.e., the entire LF value of the mother's CONTENT is token identical to an element on the mother's FOCUS list) if it is the case that the list of all daughters (provided by *dtrs-list*, a relational description of a list containing signs that are given) consists of *given* signs into which a single *focused* sign is shuffled (◯).[12] [13] As before, a sign is focused if its LF value is token identical to an element of its FOCUS value; and a sign is given if its LF value is token identical to an element of its GIVEN value.

The pitch accent in this example is on the adjective *green* so that the principle in (9) on p. 716 licenses structure sharing of the adjective's content with its FOCUS value. In the context of the question (24a), the entire NP *a green convertible* from example (24b) is in the focus. In the phrase *green convertible*, the clause licensing focus projection in NPs does not apply, since the adjective *green*, from which the focus has to project in this case, is not the rightmost element of the phrase. What does apply is the fourth disjunct of the principle licensing focus projection in connection with givenness. Since the noun *convertible* is given, the adjective *green* is the only daughter in the phrase that is not given and focus is allowed to project to the mother of the phrase. In the phrase *a green convertible*, focus projection is again licensed via the clause for focus projection in noun phrases, since the focused phrase *green convertible* is the rightmost daughter in that noun phrase.

6 Information structure and word order

The explicit representation of information structure as part of signs in HPSG opens up the possibility of providing explanations for constraints previously stipulated in syntax, such as word order constraints, by deriving the constraints from the nature of the integration of a sentence into the discourse. Many of the ap-

$$\textit{dtrs-list}\left(\left\langle \boxed{1} | \boxed{2} \right\rangle\right) := \begin{bmatrix} \text{HEAD-DTR} & \boxed{1} \\ \text{NON-HD-DTRS} & \boxed{2} \end{bmatrix}$$

$$\textit{given-sign-list} := \langle\rangle.$$

$$\textit{given-sign-list} := \left\langle \begin{bmatrix} \text{SS}|\text{L}|\text{CONT}|\text{LF} & \boxed{1} \\ \text{STRUC-MEANING} & \begin{bmatrix} \text{GIVEN} & \left\langle \boxed{1} \right\rangle \end{bmatrix} \end{bmatrix} \middle| \ \textit{given-sign-list} \right\rangle.$$

[12] The relation "shuffle" ◯ is used as originally introduced in Reape (1994): the result is a list that contains all element from the two input lists and the order of elements from the original lists is preserved.

[13] If only binary structures are assumed, as in the examples in this chapter, the principle can be simplified. Here, I kept the general version with recursive relations following De Kuthy & Meurers (2003), which also supports flatter structures.

proaches discussed in the previous section employ the information structural architecture exactly in this way and formulate principles linking word order to discourse properties.

One first such approach is presented in Engdahl & Vallduví (1996), where word order constraints for Catalan are couched into the information structure setup discussed in Section 3.2. The basic observation is that in Catalan, the word order within the sentential core is and that every constituent within this sentential core is interpreted as focal. If an argument of the main verb of a sentence is to be interpreted as non-focal, it must be clitic-dislocated. The example in (27) from Engdahl & Vallduví (1996) illustrates the two possible cases: the argument *a Barcelona* 'to Barcelona' can be topicalized as in (27b) or positioned at the end of the sentence as in (27c) in order to be interpreted as non-focal.

(27) a. Ahir [[va tornar a Barcelona el PRESIDENT]]$_F$.
yesterday 3s-past-return to Barcelona the president

b. A Barcelona$_1$ [[hi$_1$ va tornar el PRESIDENT]]$_F$.
to Barcelona there returned the president

c. [[Hi$_1$ va tornar el PRESIDENT]]$_F$ a Barcelona$_1$.
there returned the president to Barcelona
'Yesterday, the president returned to Barcelona.'

With respect to modeling this within the HPSG account, they assume that phrases associated with a LINK interpretation should be constrained to be left dislocated, whereas phrases associated with a TAIL interpretation should be right attached. They thus introduce the following ID schema for Catalan:

(28) *Head-Dislocation Schema for Catalan:*
The DTRS value is an object of sort *head-disloc-struc* whose HEAD-DTR|SYNSEM|LOCAL|CATEGORY value satisfies the description $\left[\text{HEAD } \textit{verb}\left[\text{VFORM } \textit{finite}\right]\text{, SUBCAT } \langle\rangle\right]$, and whose DISLOC-DTRS|CONTEXT|INFO-STRUC value is instantiated and for each DISLOC-DTR, the HEAD-DTR|SYNSEM|LOCAL|CONTENT value contains an element which stands in a *binding* relation to that DISLOC-DTR.

The principle requires that the information structure value of dislocated daughters of a finite sentence has to be GROUND. An additional LP statement then captures the relation between the directionality of the dislocation and a further restriction of the GROUND value, as illustrated in (29).

(29) LP constraint on information structure in Catalan (Engdahl & Vallduví 1996):

LINK > FOCUS > TAIL

This LP statement is meant to ensure that link material must precede focus material and focus material must precede tails. By this, Engdahl & Vallduví (1996) want to ensure that left-detached constituents are always interpreted as links and right-detached constituents as tails.

The insights from Engdahl & Vallduvi's approach are the basis for an approach accounting for clitic left dislocation in Greek presented by Alexopoulou & Kolliakou (2002). The representation of information structure with the features FOCUS and GROUND (further divided into LINK and TAIL) is taken over as well as the phonological constraints on words and the information structure instantiation principle. In order to account for clitic left dislocation, as illustrated in (30), an additional feature CLITIC is introduced as appropriate for *nonlocal* objects.

(30) a. Pii simetehoun s' afti tin paragogi?
who take part in that the production
'Who contributed to this production?'

b. Tin parastasi *ti* skinothetise o Karolos KOUN …
the performance FEM.3SG.ACC directed the Karolos Koun
'Karolos Koun directed the performance …'

The Linkhood Constraint shown in (31) ensures that links (i.e. elements whose INFO-STRUC|LINK value is instantiated) can only be fillers that are "duplicated" in the morphology by a pronominal affix, i.e. it is required that there is an element $\boxed{1}$ on the CLITIC list of the head daughter that is structure-shared with the filler's HEAD value. The use of the disjoint union relation $\uplus$[14] ensures that the singleton element $\boxed{1}$ representing the doubled clitic is the only element on the phrase's clitic list with these specifications. In addition, it is required that the filler-daughter $\boxed{2}$ is structure-shared with the LINK attribute in the information structure of the mother.

(31) The Linkhood Constraint for clitic left dislocation phrases (Alexopoulou & Kolliakou 2002: 238):

$$\begin{bmatrix} \textit{clitic-left-disloc-phrase} & \\ \text{INFO-STRUC|LINK} & \{\boxed{2}\} \\ \text{CLITIC} & \boxed{\Sigma_2} \end{bmatrix} \rightarrow \boxed{2}\begin{bmatrix}\text{PHON|ACCENT} & u \\ \text{HEAD} & \boxed{1}\end{bmatrix}, \textbf{H}\begin{bmatrix} \textit{phrase} & \\ \text{HEAD} & \textit{verb} \\ \text{CLITIC} & \{\boxed{1}\} \uplus \boxed{\Sigma_2}\end{bmatrix}$$

14 Alexopoulou & Kolliakou (2002) provide no exact definition for the use of the symbol $\uplus$ (disjoint union), but a definition that is often used within HPSG approaches can be found in Manandhar (1994a).

The Linkhood Constraint thus has two purposes: it ensures clitic doubling and it connects the particular word order of left dislocated phrase to discourse properties by requiring the filler daughter to be the link of the entire clause.

Other approaches dealing with left dislocated phrases are the ones by De Kuthy (2002) and De Kuthy & Meurers (2003); the latter relates the occurrence of discontinuous NPs in German to specific information structural contexts, while De Kuthy & Meurers (2003) show that the realization of subjects as part of fronted non-finite constituents and its constraints can be accounted for based on independent information structure conditions.

Based on the setup discussed in Section 3.2.2 above, constraints are formulated that restrict the occurrence of discontinuous NPs and fronted VPs based on their information structure properties. The type of discontinuous NPs at the center of De Kuthy's approach are so-called NP-PP split constructions, in which a PP occurs separate from its nominal head, as exemplified in (32).

(32) a. *Über Syntax* hat Max sich [ein Buch] ausgeliehen.
about syntax has Max self a book borrowed
'Max borrowed a book on syntax.'

b. [Ein Buch] hat Max sich *über Syntax* ausgeliehen.
a book has Max self about syntax borrowed

The information structure properties of discontinuous noun phrases are summarized in De Kuthy (2002: 176) in the following principle:

> In an utterance, in which a PP occurs separate from an NP, either the PP or the NP must be in the focus or in the topic of the utterance, but they cannot both be part of the topic or the same focus projection. (De Kuthy 2002: 176)

The last restriction can be formalized as: the PP's or NP's CONTENT values cannot be part of the same meaningful-expression on the FOCUS list or the TOPIC list of the INFO-STRUC value of the utterance.

As discussed in De Kuthy & Meurers (2003), it has been observed that in German it is possible for unergative and unaccusative verbs to realize a subject as part of a fronted non-finite verbal constituent (Haider 1990). This is exemplified in (33) with examples from Haider (1990: 94):

(33) a. [Ein Fehler unterlaufen] ist meinem Lehrer noch nie.
an.NOM error crept.in is my.DAT teacher still never
'So far my teacher has never made a mistake.'

b. [Haare wachsen] können ihm nicht mehr.
hair.NOM grow can him.DAT not anymore
'His hair cannot grow anymore.'

c. [Ein Außenseiter gewonnen] hat hier noch nie.
an.NOM outsider won has hier still never
'An outsider has never won here yet.'

In order to account for the context-sensitive occurrence of such fronted verbal constituents, specific information structure properties of fronted verb phrases need to be expressed in a principle expressing what De Kuthy & Meurers refer to as Webelhuth's generalization (Webelhuth 1990: 53): In an utterance in which a verb phrase occurs as a fronted constituent (i.e., the filler of a head-filler phrase) this entire verb phrase must be in the focus of the utterance (i.e., the FOCUS value of the fronted constituent must be identical to its semantic representation). The formalization of this principle provided by (De Kuthy & Meurers 2003) is shown in (34).

(34) Webelhuth's generalization (De Kuthy & Meurers 2003):

$$\begin{bmatrix} \textit{head-filler-phrase} \\ \text{NON-HEAD-DTR}|\text{SYNSEM}|\text{LOC}|\text{CAT}|\text{HEAD}\ \ \textit{verb} \end{bmatrix} \rightarrow \begin{bmatrix} \text{INFO-STRUC}|\text{FOCUS}\ \langle\boxed{1}\rangle \\ \text{NON-HEAD-DTR}\ \begin{bmatrix} \text{INFO-STRUC}|\text{FOCUS} & \langle\boxed{1}\rangle \\ \text{SYNSEM}|\text{LOC}|\text{CONT}|\text{LF} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

Combining the new lexical specifications, the focus projection rule for the verbal domain and the partial fronting focus requirement with the basic setup of De Kuthy (2002), one obtains a theory which predicts that subjects can only be part of a fronted verb phrase if they can be the focus exponent.[15] The sketch of an analysis for an example such as (33c) is illustrated in Figure 3. The entry of *gewinnen* 'to win' (the base form of the verb *gewonnen*) in (33c) in Figure 4 encodes the lexical property that the subject of this intransitive verb has focus projection potential.

Under the assumption that in (33c) the noun *Außenseiter* carries a pitch accent, the information structure principle for words in (9) on p. 716 ensures that the noun contributes its LOGICAL-FORM value to its FOCUS value. The focus projection principle in (14) on p. 720 ensures that the focus can project over the

[15]Not every element in a syntactic phrase corresponding to the focus is prosodically prominent. Generally only one elementis: the so-called *focus exponent* (cf. Selkirk 1995).

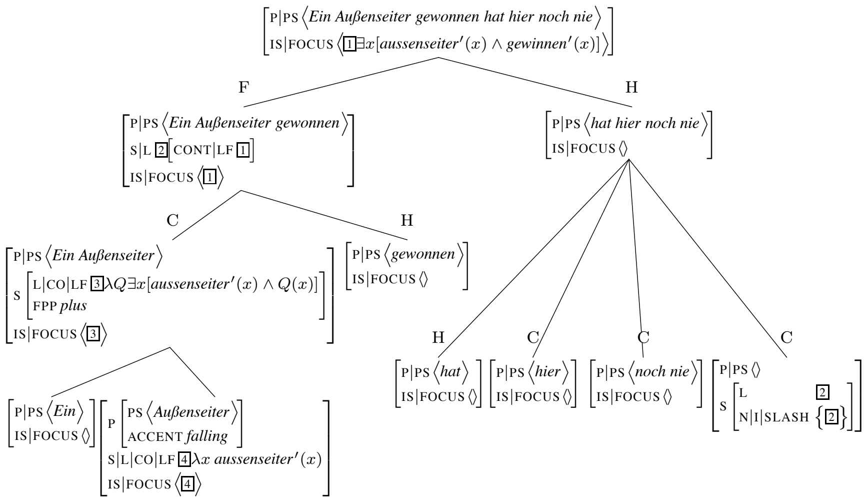


Figure 3: Partial VP fronting in De Kuthy & Meurers (2003)

$$\begin{bmatrix} \textsc{phon} & \langle \textit{gewinnen} \rangle \\ \textsc{arg-s} & \left\langle \begin{bmatrix} \textsc{fpp} \ \textit{plus} \\ \textsc{loc}|\textsc{cat}|\textsc{head}|\textsc{case} \ \textit{nom} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 4: The lexical entry of *gewinnen* ‘to win’

entire NP *ein Außenseiter*, i.e., its FOCUS element is identical to its LF value. Since *ein Außenseiter* as the subject of *gewonnen* in the tree in Figure 3 is lexically marked as FPP *plus*, the principle governing focus projection in the verbal domain in (14) licenses the focus to project over the entire fronted verb phrase *ein Außenseiter gewonnen*. The fronted constituent thus contributes its LF value to its FOCUS value. In this example, the focus does not project further, so that in the head-filler phrase the focus values of the two daughters are simply collected as licensed by the first disjunct of the focus principle in (14) discussed earlier in Section 3.2.2. As a result, the FOCUS value of the fronted verb phrase is the FOCUS value of the entire sentence. Finally, note that the example satisfies Webelhuth's generalization, which requires a fronted verb phrase to be the focus of the utterance as formalized in the principle in (34).

In the same spirit, Bildhauer & Cook (2010) show that sentences in which multiple elements have been fronted are directly linked to specific types of information structure. In German, a V2 language, normally exactly one constituent occurs in the position before the finite verb in declarative sentences. But so-called multiple fronting examples with more than one constituent occurring before the finite verb have been well attested in naturally-occurring data (Müller 2003). Two examples from Bildhauer & Cook (2010) are shown in (35).[16]

(35) a. [Dem Saft] [eine kräftigere Farbe] geben Blutorangen.
to.the juice a more.vivid colour give blood.oranges
'What give the jiuce a more vivid colour is blood oranges.'

b. [Stets] [einen Lacher] [auf ihrer Seite] hatte die Bubi Ernesto Family.
always a laugh on their side had the Bubi Ernesto Family
'Always good for a laugh was the Bubi Ernesto Family.'

But, as discussed by Bildhauer & Cook, such multiple fronting examples seem to require very special discourse conditions in order to be acceptable. Just like the fronted verb phrases discussed in De Kuthy & Meurers (2003) above, Bildhauer & Cook (2010) propose analyzing multiple fronting constructions in German as head-filler phrases which in this case introduce a topic shift. Following the approach by Müller (2005), multiple fronting configurations can be identified via the filler daughter which must have a HEAD|DSL (double slash) value of type *local*.[17] As introduced above, Bildhauer & Cook (2010) assume that an information structure attribute is specified in *synsem* objects, with the features FOCUS and TOPIC taking lists of *elementary predications* as their values. In general, multiple fronting *head-filler* phrases are restricted by the constraint in (36).

(36) Relating multiple fronting to focus (Bildhauer & Cook 2010: 75):

$$\begin{bmatrix} \textit{head-filler-phrase} \\ \textsc{non-head-dtrs}\ \ \langle [\textsc{loc|cat|head|dsl}\ \ \textit{local}] \rangle \end{bmatrix} \rightarrow [\textsc{is}\ \ \textit{pres} \vee \textit{a-top-com} \vee \ldots]$$

$$\begin{bmatrix} \textit{head-filler-phrase} \\ \textsc{is}\ \ \textit{pres} \end{bmatrix} \rightarrow \begin{bmatrix} \textsc{ss|loc|cat|head|dt} & \langle [\textsc{loc|cont|rels}\ \boxed{1}] \rangle \\ \textsc{hd-dtr|ss|is|focus} & \langle \boxed{1} \rangle \end{bmatrix}$$

[16] The examples are corpus examples that were extracted by Bildhauer & Cook (2010) from Deutsches Referenzkorpus (DeReKo), hosted at the Institut für Deutsche Sprache, Mannheim: http://www.ids-mannheim.de/kl/projekte/korpora

[17] In Müller's (2005) formalization, filler daughters in multiple fronting configurations (and only in these) have a HEAD|DSL value of type *local*, i.e., they contain information about an empty verbal head. The DSL ('double slash') feature is needed to model the HPSG equivalent of verb movement from the sentence-final position to initial position.

The first constraint ensures that *head-filler* phrases that are instances of multiple frontings are restricted to have an IS-value of an appropriate type.[18] The second constraint then ensures that in presentational multiple frontings, the designated topic of the head daughter (i.e. the verbal head of the *head-filler-phrase*) must be focused. The feature DT (designated topic) lexically specifies which element, if any, is normally realized as the topic of a particular verb. This constraint thus encodes what Bildhauer & Cook (2010) call "topic shift": the non-fronted element in a multiple fronting construction that would preferably be the topic is realized as a focus. A similar constraint is introduced for another instance of multiple frontings, which is called *propositional assessment* multiple fronting. Here it has to be ensured that the designated topic must be realized as the topic somewhere in the head daughter and the head daughter must also contain a focused element.

Webelhuth (2007) provides another account of the special information structural requirements of fronted constituents, in this case of predicate fronting in English that is based on the interaction of word order and information structural constraints.

(37) I was sure that Fido would bark and *bark he did.*

The principles part of Webelhuth's account require that in such cases of predicate fronting, the auxiliary is focused and the remainder of the sentence is in the background. The two principles needed for this interaction are shown in (38).

(38) Predicate preposing phrases (Webelhuth 2007):

$$\begin{bmatrix} \textit{aux-wd} & \\ \text{ARG-S} & \langle \text{NP, gap-ss} \rangle \end{bmatrix} \rightarrow \begin{bmatrix} \text{SS}|\text{STATUS} & \textit{foc} \\ \text{ARG-S} & \langle \begin{bmatrix} \text{STATUS} & \textit{bg} \end{bmatrix}, \textit{gap-ss} \rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{PRED-PREPOS-PH} \end{bmatrix} \rightarrow \begin{bmatrix} \textit{HD-FILL-PH} & \\ \text{NON-HD-DTR} & \begin{bmatrix} \text{SS}|\text{STATUS} & \textit{bg} \end{bmatrix} \end{bmatrix}$$

The first constraint ensures that auxiliary words whose predicate complement has the potential to be preposed (i.e. is of type *gap-ss*) have the information status *focus*, whereas the status of the first argument (the subject) is *background*. Additional constraints then ensure that auxiliary words with a gapped second argument can only occur in predicate preposing phrases, and vice versa, that predicate preposing phrases contain the right kind of auxiliary.

[18] Bildhauer & Cook (2010: 75) assume that the type *is* as the appropriate value for IS has several subtypes specifying specific combinations of TOPIC and FOCUS values, such as *pres* for presentational focus or *a-top-com* for assessed-topic-comment.

7 Information structure and prosody

A lot of languages mark information structure prosodically, like for example English and German, where pitch accents of various shapes are used to mark focus. Accordingly, several of the approaches discussed above include a component which enriches the phonology representation of signs such that it allows the integration of the necessary prosodic aspects like accents.

Engdahl & Vallduví (1996) assume that signs can be marked for particular accents signaling focus or links in English, so-called A and B accents. In a similar way, De Kuthy (2002) extends the value of PHON such that it includes a feature ACCENT, in order to formulate constraints on the connection between accents and information structure markings. Most of approaches discussed above do not include a detailed analysis of the prosodic properties of the respective language that is being investigated with respect to discourse properties. As a result, most approaches do not go beyond the postulation of one or two particular accents, which are then somehow encoded as part of the PHON value. These accents more or less serve as an illustration of how lexical principles can be formulated within a particular theory that constrain the distribution of information structural values on the lexical level. The more articulate such a representation of PHON values including accent pattern, intonation contours, boundary tone, etc. is, the more detailed the principles could be that are needed to connect information structure to prosodic patterns in languages that signal discourse properties via intonation contours.

In Bildhauer (2008), one such detailed account of the prosodic properties of Spanish is developed together with a proposal for how to integrate prosodic aspects into the PHON value, also allowing a direct linking of the interaction of prosody and information structure. In his account, the representation of PHON values in HPSG is enriched to include four levels of prosodic constituency: phonological utterance, intonational phrases, phonological phrases and prosodic words. The lowest level, prosodic words of type *pwrd*, include the feature SEGS, which corresponds to the original PHON value assumed in HPSG, and additional features such as PA for pitch accents or BD for boundary tones, which encodes whether a boundary tone is realized on that word. The additional features UT (phonological utterance), IP (intonational phrase) and PHP (phonological phrase) encode via the typ *epr* (edges and prominence) which role a prosodic word plays in higher level constituents. For example, the feature DTE (designated terminal element) specifies whether the word is the most prominent one in a phonological phrase. A sign's PHON list then contains all *pwrd* objects, and relational constraints define the role each prosodic word plays in the higher prosodic constituents. This flat representation of prosodic constituency still makes it possible to express constraints about intonational contours associated with certain utterance types. One example discussed in Bildhauer's work is the contour associated with broad focus declaratives in Spanish, which can be decomposed into a sequence of late-rise (L*H) prenuclear accents, followed by an early-rise nuclear accent (LH*), followed by a low boundary tone (L%). The constraint introduced to model this contour for declarative utterances thus instantiates the BD value (boundary tone) of the last *pwrd* (prosodic word) in the PHON list to *low*, instantiates a nuclear pitch accent *low-high-star* on this rightmost prosodic word and ensures that a prenuclear pitch accent *low-star-high* is instantiated on every preceding compatible prosodic word. The constraint encoding this is shown in (39).

(39) Intonational contour of Spanish declarative utterances (Bildhauer 2008: 142):

$$\begin{aligned}\textit{decl-tune}\left(\boxed{1}\right) \leftrightarrow\ & \boxed{1} = \boxed{2} \oplus \left\langle \begin{bmatrix}\text{PA} & \textit{low-high-star}\\ \text{BD} & \textit{low}\end{bmatrix} \right\rangle \wedge \\ & \boxed{2} = \textit{list}\left(\begin{bmatrix}\text{BD} & \textit{none}\end{bmatrix}\right) \wedge \\ & \boxed{2} = \textit{list}\left(\begin{bmatrix}\text{PA} & \textit{none}\end{bmatrix}\right) \bigcirc\ \textit{list}\left(\begin{bmatrix}\text{PA} & \textit{low-star-high}\end{bmatrix}\right)\end{aligned}$$

$$\begin{bmatrix}\textit{sign} & \\ \text{EMBED} & -\end{bmatrix} \longrightarrow \begin{bmatrix}\text{PHON} & \boxed{1}\end{bmatrix} \wedge \textit{decl-tune}\left(\boxed{1}\right)$$

The second constraint in (39) ensures that only unembedded utterances can be constrained to the declarative prosody described above. That this specific contour is then compatible with a broad focus reading is ensured by an additional principle expressing a general focus prominence constraint for Spanish, namely that focus prominence has to fall on the last prosodic word in the phonological focus domain, which, in the case of a broad focus, can be the entire utterance. The principle formulated in Bildhauer's account is shown in (40).

(40) Focus prominence in Spanish (Bildhauer 2008: 146):

$$\begin{bmatrix}\textit{sign} & \\ \text{CONT} & \boxed{1}\\ \text{FOC} & \boxed{1}\end{bmatrix} \rightarrow \begin{bmatrix}\text{PHON} & \textit{list} \oplus \left\langle \begin{bmatrix}\text{UT}|\text{DTE} & +\end{bmatrix} \right\rangle\end{bmatrix}$$

Since only words that are the designated terminal element (DTE) can bear a pitch accent, the interplay of the two principles above ensures that in utterances with a declarative contour the entire phrase can be in the focus. These principles thus illustrate nicely not only how lexical elements can contribute to the information structure via their prosodic properties, but also how entire phrases with specific prosodic properties can be constrained to have specific information structural properties.

The approach of Song (2017) also includes a component that captures the interaction between prosodic properties of utterances and the effect on the information structure. In order to include information structural constraints of the so-called A and B accents in English, several components of Bildhauer's (2007) phonological architecture are adapted for the information structural setup in Song (2017). Among them is the idea that in a phonological phrase (encoded in the feature phonological utterance UT), focus prominence is related to the most prominent word in that phrase, which is encoded via the constraint in (41).

(41) Prosodic marking of focus (Song 2017: 159):

$$\textit{lex-rule} \longrightarrow \begin{bmatrix} \text{UT}|\text{DTE} & \boxed{1} \\ \text{MKG}|\text{FC} & \boxed{1} \end{bmatrix}$$

Specific lexical principles for the A and B accents then ensure the correct information structural marking and specifiy which type of element has to be present on the ICONS list. The specification necessary for English A accents that signal focus (here characterized as *high-star*) are shown in (42).

(42) Focus marking of A accents in English (Song 2017: 160):

$$\textit{fc-lex-rule} \longrightarrow \begin{bmatrix} \text{UT}|\text{DTE} & + \\ \text{PA} & \textit{high-star} \\ \text{MKG} & \textit{fc-only} \\ \text{INDEX} & \boxed{1} \\ \text{INCONS-KEY} & \boxed{2} \\ \text{C-CONT}|\text{ICONS} & \langle\ !\ \boxed{2}\begin{bmatrix} \textit{semantic-focus} \\ \text{TARGET}\ \boxed{1} \end{bmatrix} !\ \rangle \\ \text{DTR}|\text{HEAD} & \textit{+nv} \end{bmatrix}$$

8 Conclusion

I have discussed various possibilities for how to represent information structure within HPSG's sign-based architecture. Several approaches from the HPSG literature were presented which all have in common that they introduce a separate feature INFO-STRUC into the HPSG setup, but they differ in (i) where they locate such a feature, (ii) what the appropriate values are for the representation of information structure and (iii) how they encode principles constraining the distribution and interaction of information structure with other levels of the grammatical architecture. Finally, I discussed a number of theories in which phenomena such as word order are constrained to only be well-formed when they exhibit specific information structural properties.

References

Alexopoulou, Theodora & Dimitra Kolliakou. 2002. On linkhood, topicalization and clitic left dislocation. *Journal of Linguistics* 38(2). 193–245.

Ambridge, Ben & Adele Goldberg. 2008. The island status of clausal complements: evidence in favor of an information structure explanation. *Cognitive Linguistics* 19(3). 357–389.

Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson & Safiyyah Saleem. 2010. Grammar customization. *Research on Language & Computation* 8(1). 23–72. http://dx.doi.org/10.1007/s11168-010-9070-1. 10.1007/s11168-010-9070-1.

Bildhauer, Felix. 2008. *Representing information structure in an HPSG grammar of Spanish*. Universität Bremen Dissertation.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topic-hood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Bildhauer, Felix & Philippa Cook. 2011. Multiple fronting vs. vp fronting in german. In Amir Zeldes & Anke Lüdeling (eds.), *Proceedings of Quantitative Investigations in Theoretical Linguistics 4*, 18–21. https://korpling.german.hu-berlin.de/qitl4/QITL4_Proceedings.pdf, accessed 2016-9-22.

Büring, Daniel. 2006. Focus projection and default prominence. In Valéria Molnár & Susanne Winkler (eds.), *The Architecture of Focus* (Studies in Generative Grammar 82), 321–346. Berlin: Mouton De Gruyter.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Culicover, Peter & Susanne Winkler. 2019. Why topicalize VP? In Verner Egerland, Valéria Molnár & Susanne Winkler (eds.), *Architecture of Topic* (Studies in Generative Grammar 136), 173–202. Berlin/New York: Mouton De Gruyter.

De Kuthy, Kordula. 2002. *Discontinuous NPs in German* (Studies in Constraint-Based Lexicalism 14). Stanford, CA: CSLI Publications.

De Kuthy, Kordula & Andreas Konietzko. 2019. Information structural constraints on PP topicalization from NPs. In Verner Egerland, Valéria Molnár & Susanne Winkler (eds.), *Architecture of Topic* (Studies in Generative Grammar 136), 203–222. Berlin/New York: Mouton De Gruyter.

De Kuthy, Kordula & Walt Detmar Meurers. 2003. The secret life of focus exponents, and what it tells us about fronted verbal projections. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 97–110. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-2-25.

De Kuthy, Kordula & Walt Detmar Meurers. 2011. Integrating GIVENness into a structured meaning approach in HPSG. In *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar*, 289–301. Stanford: CSLI Publications.

Dowty, David R., R. E. Wall & S. Peters. 1981. *Introduction to Montague Semantics*. Dordrecht: D. Reidel Publishing Company.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover & Enric Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science, vol. 12: Studies in HPSG*, chap. 1, 1–32. Edinburgh: Centre for Cognitive Science, University of Edinburgh. ftp://ftp.cogsci.ed.ac.uk/pub/CCS-WPs/wp-12.ps.gz, accessed 2018-3-20.

Gallin, Daniel. 1975. *Intensional and higher-order modal logic*. Amsterdam: North Holland.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19. 377–419.

Haider, Hubert. 1990. Topicalization and other puzzles of German syntax. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers* (Linguistik Aktuell/Linguistics Today 5), 93–112. Amsterdam: John Benjamins Publishing Co.

Jackendoff, Ray S. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jacobs, Joachim. 1983. *Fokus und Skalen: Zur Syntax und Semantik der Gradpartikeln im Deutschen* (Linguistische Arbeiten 138). Tübingen: Max Niemeyer Verlag.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Krifka, Manfred. 1992. A compositional semantics for multiple focus constructions. In Joachim Jacobs (ed.), *Informationsstruktur und Grammatik* (Linguistische Berichte Sonderheft 4 / 1991–92). Opladen/Wiesbaden: Westdeutscher Verlag.

Kruijff-Korbayová, Ivana & Mark J. Steedman. 2003. Discourse and information structure. *Journal of Logic, Language and Information: Special Issue on Discourse and Information Structure* 12(3). 249–259.

Manandhar, Suresh. 1994a. An attributive logic of set descriptions and set operations. In James Pustejovsky (ed.), *32th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 255–262. Las Cruses: Association for Computational Linguistics.

Manandhar, Suresh. 1994b. Encoding information packaging in HPSG. In Elisabet Engdahl (ed.), *Integrating information structur into constraint-based and categorial approaches*. Amsterdam: ILLC. DYANA-2 Report R.1.3.B.

Müller, Stefan. 2003. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330.

Paggio, Patrizia. 2009. The information structure of Danish grammar constructions. *Nordic Journal of Linguistics* 32(1). 137–164.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Reape, Michael. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-driven Phrase Structure Grammar*, 151–197. Stanford: CSLI Publications.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Sailer, Manfred. 2000. *Combinatorial semantics and idiomatic expressions in Head-Driven Phrase Structure Grammar*. Eberhard-Karls-Universität Tübingen Dissertation. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/46191, accessed 2018-2-25.

Schwarzschild, Roger. 1999. Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7(2). 141–177.

Selkirk, Elizabeth. 1982. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Selkirk, Elizabeth. 1995. Sentence prosody: intonation, stress, and phrasing. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 550–569. London: Basil Blackwell.

Song, Sanghoun. 2017. *Modeling information structure in a cross-linguistic perspective* (Topics at the Grammar-Discourse Interface 1). Berlin: Language Science Press.

Song, Sanghoun & Emily M. Bender. 2012. Individual constraints for information structure. In Stefan Müller (ed.), *Proceedings of the 19th International Confer-*

ence on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon, 330–348. http://cslipublications.stanford.edu/HPSG/2012/song-bender.pdf.

Vallduví, Enric. 1992. *The informational component*. New York, NY: Garland.

von Stechow, Arnim. 1981. Topic, focus, and local relevance. In Wolfgang Klein & Willem Levelt (eds.) (Synthese Language Library 13), 95–130. Dordrecht: D. Reidel Publishing Company.

von Stechow, Arnim & Susanne Uhmann. 1986. Some remarks on focus projection. In Werner Abraham & S. de Meij (eds.), *Topic, focus, and configurationality: Papers from the 6th Groningen Grammar Talks, Groningen, 1984* (Linguistik Aktuell/Linguistics Today 4), 295–320. Amsterdam: John Benjamins Publishing Co.

Webelhuth, Gert. 1990. Diagnostics for structure. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers* (Linguistik Aktuell/Linguistics Today 5), 41–75. Amsterdam: John Benjamins Publishing Co.

Webelhuth, Gert. 2007. Complex topic-comment structures in HPSG. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 306–322. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/, accessed 2018-2-25.

Part IV

Other areas of linguistics

Chapter E

Diachronic syntax

Ulrike Demske

Universität Potsdam

Basic questions of language change concern the what, how, why of a change. The present paper focusses on syntactic change and addresses the question how a representational framework as HPSG can be used to model syntactic change. Taking the history of German as a showcase, different types of morpho-syntactic changes are considered including changes within the verb phrase as well as the noun phrase.

1 Dimensions of syntactic change

Syntactic change links the language of two speech communities which are chronologically related such as Old and Modern Norwegian. Only the grammar of the first speech community allows null subjects in restricted contexts, while referential subjects are obligatory in Modern Norwegian. As shown by Kinn (2015), Norwegian changes from a partial null-subject language to a language disallowing null subjects.

(1) a. Sægir hann þat at ægi man *pro* satt vera. (Old Norwegian)
says he that that not can [it] true be
'He says that it cannot be true.'

b. Han sier at ∗(han) ikke kan komme. (Modern Norwegian)
he says that he not can come
'He says that he cannot come.'

The reinterpretation of non-subjects as subjects in the history of English provides a second example (Denison 1993): A verb like *langian* 'long for' selects an accusative but no nominative noun phrase in Old English as exemplified in (2).

Ulrike Demske. 2019. Diachronic syntax. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 747–772. Berlin: Language Science Press. DOI:??

As shown by the translation, the experiencer argument of the verb *long* has been promoted from object to subject in Modern English.

(2) þa ongan hine eft langian on his cyþþe (Old English)
then began him.ACC again long for his native land
'Then he began again to long for his native land.'

Linking (1a) and (1b) as well as the English versions of (2) in terms of syntactic change would amount to the linking of surface manifestations instead of the underlying grammars. And it is in fact the underlying grammar that is affected by syntactic change triggered by internal and external factors. Internal factors belong to the linguistic system and may be part either of the syntax or any other module of the grammar as the morphology or the lexicon. As regards the syntactic change in the history of English, illustrated above, it may be taken as lexically driven, modifying the argument structure of verbs with experiencer objects, as suggested by (Denison 1993). Lightfoot (1979) on the other hand considers the change to be syntactically driven, i.e. as an effect of the word order change from OV to VO orders in the history of English.

The impact of morphological changes on the syntactic component of the grammar has always been of particular importance in diachronic syntax. Strong correlations are supposed to hold between a rich morphology on the one hand and the possibility of for instance word order variation or null subjects. Thus an attested collapse of case distinctions may be made up by a less variable word order with structural positions conveying the information formerly provided by case markers. The loss of morphology may also be compensated by replacing case marked noun phrases by prepositional phrases with prepositions now providing the relevant information. A pertinent example from English concerns the loss of adnominal genitives as a consequence of the loss of the genitive as a morphological case. While the postnominal genitive phrase is superseded by the *of*-phrase, the prenominal genitive marker is reanalyzed as a clitic (Allen 2006). In Old English, adnominal genitives are attested in pre- and post-head position.

(3) a. & þær wæs Kola ðæs cyning-es heahgereafa (Old English)
and there was Kola the.GEN king-GEN high reeve
'And there was Kola, the king's high reeve.'

b. þæt wæron þa ærestan scipu deniscra monna þe … (OE)
that were the first ships Danish.GEN men.GEN which …
'those were the first ships of Danish men which …'

Syntactic change may also be driven by external factors such as language processing, information packaging or language contact: According to Hawkins (2004), grammars conventionalize syntactic structures depending on their degree of preference in performance. And the rise of VO patterns in the German variety of Cimbrian is due to its contact with the Italian VO order in the North Eastern part of Italy (Grewendorf & Poletto 2005).

(4) a. Haütte die Mome hat gebläscht di Piattn. (Cimbrian)
today the mother has washed the plates
'Today, mother has washed the plates.'

b. Heute hat die Mutter die Teller gespült. (Present-day German)
today has the mother the plates washed

Derivational approaches such as Government and Binding and Minimalism assume that language change happens in the course of first language acquisition. Individual language learners reanalyze the linguistic input they get during the acquisition process, resulting in the resetting of a parameter value (Lightfoot 1979). The nature of parameters has substantially changed from GB to Minimalism: Resetting of a parameter value in terms of GB meant for the loss of null subjects, as illustrated in (1a) and (1b), for instance that the value from the null-subject parameter changed from [+] to [-] in the history of Norwegian. The minimalist view of parameters restricts parametric variation to lexical items in general and to the formal features of functional heads in particular (Borer-Chomsky Conjecture according to Baker (2008)). The null-subject parameter will then be phrased in terms of a feature bundle associated with a functional verbal head with the loss of null-subjects triggered by the absence of particular features. Depending on the range of specific formal features, Biberauer & Roberts (2017) distinguish four types of parameters according to their size: macro-, meso-, micro-, and nanoparameters, with the first type being the most stable one in the development of a language (e.g. rigid head-final or head-initial order). They further claim that formal features are not pre-specified in Universal Grammar but emerge from the interaction of Universal Grammar, primary linguistic data, and general cognitive optimization strategies in the sense of Chomsky (2005). Two principles in particular are suggested by Biberauer & Roberts (2017): (i) feature economy, restricting the acquisition of features to those with robust evidence in the input data, thereby minimizing computation; and (ii) input generalization, requiring to make maximal use of the acquired features. A case in point for the latter strategy are languages with harmonic head-final or head-initial word orders triggered by the generalization of the head parameter. The resetting of parameters in a Minimalist framework is therefore no longer restricted to the linguistic system, but includes external factors such as acquisition strategies. Since robust evidence in the primary linguistic data is required to allow a particular formal feature to be present in the underlying grammar, its frequency in the linguistic input will play a prominent role. And here further external factors come into play: Processing ease as well as language contact might affect the frequency of individual variants in the lifespan of a speaker, thus modifying her output and consequently the input of language learners. A case in point is the continuous form in Pennsylvania German (Louden 1988): While the pattern is still restricted in Common Pennsylvania German (1850–1950), it is widely spread in Plain Pennsylvania German (since 1950) motivated by the contact with its English counterpart. Another instance is the contact between Low German and Swedish. As shown by Petzell (2016), Swedish scribes of Low German frequently use a VO instead of an OV order, motivated by their L1 language. Such changes might result in changes of the underlying grammar of subsequent generations.

(5) a. er lasst de Hund los (Common Pennsylvania German)
he lets the dog loose

b. er is de Hund an los lasse (Plain Pennsylvania German)
he is the dog PREP loose let
'He is letting the dog loose.'

Under the assumption that syntactic change is triggered by changes affecting the lexicon, one might conclude that there is no such thing as syntactically triggered change, as Biberauer & Walkden (2015) actually suggest. We would then expect that lexicalist approaches to syntactic change such as HPSG or LFG are at least equally well suited as derivational approaches to model syntactic change, cf. Vincent (2001) and Börjars & Vincent (2017) for LFG. The remainder of the paper will hence provide some case studies to illustrate how HPSG can be used to model the way structure can change over time. All case studies are taken from the history of German.

The outline is as follows: Grammaticalization processes are exemplified by the rise auf auxiliary verbs in section 2.1. Changes with respect to word order are addressed in sections 2.2 regarding verb clusters and 2.3 with respect to word order changes within the noun phrase. Overall, section 2.3 is devoted to various changes affecting the left periphery of noun phrases and also touches upon the issue why particular changes happen.

2 Case studies

The case studies presented in this section are supposed to exemplify different types of syntactic change. Grammaticalization processes typically give rise to grammatical markers with the development of auxiliary verbs from main verbs figuring as a prominent example. The next section will therefore deal with the emergence of the passive auxiliaries *kriegen*, *bekommen* 'get' and *geben* 'give' in the history of German. A second type of change are word order changes which are the topic of section 2.2, focussing on the order of verbs at the right clausal periphery in German, a syntactic change which has attracted a lot of interest from a descriptive as well as a more formal perspective. The third case study addresses several changes affecting the left periphery of noun phrases, including the grammaticalization of the definite determiner as well as the word order change of adnominal genitives.

2.1 Grammaticalization: rise of auxiliary verbs

Grammaticalization processes have an impact on the way grammatical information is marked in a language. Grammatical information such as verbal mood may be expressed either by morphological means as in (6a) or by syntactic means as in (6b) which uses the auxiliary verb *werden* to convey a modal meaning. With the grammatical meaning being alike, a morphological marker such as *-e* in (6) is taken to be further down on the grammaticalization cline as the corresponding auxiliary verb because of its morphological boundedness.

(6) a. Fred ging-e ins Kino. (Present-day German)
Fred went-SBJV to.the movies

b. Fred würde ins Kino gehen. (Present-day German)
Fred SBJV to.the movies go
'Fred would go to the movies.'

Further prominent examples of grammaticalization processes involve the rise of negation markers in French such as *pas* from the Latin noun *passus* 'step' or *personne* 'no-one' which both were originally restricted to a positive meaning. A well-known German example for a grammaticalization process is the subordinating conjunction *weil* 'because' evolving from the noun *Weile* 'while'. Grammaticalization typically consists of a reduction in meaning (lexical » grammatical) and in form (syntactic marker » morphological marker).[1]

[1] Cf. Lehmann (2015) for his seminal work on issues of grammaticalization.

As regards the class of auxiliary verbs in Present-day German, the verbs *kriegen* and *bekommen*, both meaning 'get', are fairly recent members of this class (Reis 1976). They pattern with passive auxiliaries such as *werden* 'get' and *sein* 'be' in German, triggering classic diagnostics for passive constructions: An object of the active counterpart figures as subject of the finite verb in the passive clause, and the subject of the active clause is optionally realized as *von*-PP in its passive equal. A crucial difference to the canonical passive construction concerns the status of the object: It is not the direct but the indirect object which becomes the subject of the passive clause, hence the term *dative passive*.[2] In (7), the indirect object of *sagen* 'tell' was promoted to grammatical subject of *bekommen*,favored by the fact that both verbs select for experciencer arguments although in different slots of their argument structure. The *dass* 'that'-clause can only be understood as an argument of *sagen* 'tell', providing evidence for an analysis of *bekommen* 'get' as auxiliary.

(7) Ich habe nur gesagt bekommen, dass er Probleme mit den Hinterreifen hat. (BRZ13/APR.07807)
I have only told got that he problems with the rear tires has
'I was just told that he has problems with his rear.tires.'

Both verbs *kriegen* 'get' as well as *bekommen* 'get' have developed a grammatical meaning alongside their lexical meaning in the recent history of German (Glaser 2005; Lenz 2012). Early examples for their use as auxiliary verbs come from the 16th and the 17th century with both auxiliaries typically combining with past participles of lexical verbs which select for experiencer arguments in object position (*schenken* 'grant', *schicken* 'send'), which may be interpreted as experiencer arguments of the governing predicate. The same holds for the theme argument which may be selected either by the lexical verb or the emerging auxiliary. Examples as in (8) therefore represent the beginning of the grammaticalization process. Diagnostics for the passive construction include the *von*-phrase in (8a) as well as the valency alternation turning the indirect objects of the corresponding active clause into subjects of the passive clause.

[2] The focus here is on the passive meaning of the verbs in question. Resultative uses of *bekommen, kriegen* 'get' are disregarded in the present context. *Selma kriegt die Gleichung gelöst* 'Selma manages to solve the equation', means that *Selma* is an agent argument reaching a particular goal with some effort with the auxiliary conveying an active meaning. Cf. Reis (1985) for a discussion of both types of constructions.

(8) a. da hatte ich eben ein paar Ducaten vom Herrn geschenckt
EXPL had I just a few ducats by the master granted
kriegt (1672: Weise)
got
'I just got a few ducats by the master.'

b. Es verdrüst sie/ daß du Wein hast geschickt bekommen/ und sie
EXPL annoys her that you wine have send got and she
keinen. (1695: Reuter)
none
'She is annoyed, because you got wine, but she has not.'

Examples like (8) are taken to represent stage 1 in the grammaticalization process, while examples like (7) are instantiations of stage 2, because the direct object can only be interpreted as argument of the non-finite verb.[3] Stage 3 on the grammaticalization cline includes verbs with a privative semantics, indicating that the emerging auxiliary is grammaticalized to such a degree that indirect objects are no longer restricted to recipient arguments. The use of intransitive verbs like *helfen* 'help' are considered to represent stage 4 on the grammaticalization path (Ebert 1978).

(9) a. Aber nach einer Woche hatte sie noch nicht einmal die Fäden
but after one week had she.NOM not again the stitches
gezogen bekommen. (RHZ03/DEZ.09011 RZ)
removed got
'After one week, she had not again removed the stitches.'

b. Sie wollen konkret geholfen bekommen. (PHE/W18.00094)
sie.NOM want definitely helped get
'They definitely want to get help.'

The use of *kriegen, bekommen* 'get' as passive auxiliaries is attested in all German varieties with stilistic differences between the two verbs.[4] This does not hold for the verb *geben* 'give' which may be used as a passive auxiliary only in certain dialects. Like *kriegen* and *bekommen*, the verb *geben* developed into a passive auxiliary in the recent history of German (Lenz 2007). As (10) shows, the experiencer argument of *mitnehmen* 'give a lift' appears as subject of the auxiliary

[3] These examples provide evidence against the claim that the direct object is assigned a thematic role by both the finite and the nonfinite verb, cf. Haider (1986).

[4] While *bekommen* is used in the standard variety of German, *kriegen* is confined to less formal registers. The verb *erhalten* 'get' is only rarely used as an auxiliary with a passive meaning.

geben 'give', while the agent argument is realized by a *von*-phrase as expected in a passive construction.[5]

(10) De Tobi gebbt vom Yannick mitgehol. (West Central German)
the Tobi gives by the Yannick given a lift
'Tobi is given a lift by Yannik.'

In a minimalistic framework, grammaticalization processes are modeled as a categorial reanalysis of lexical categories as functional categories or of hierarchically lower functional categories as higher ones in line with the Borer-Chomsky Conjecture (Roberts & Roussou 2003). In many instances, grammaticalization means the loss of features, as predicted by a cognitive optimization strategy as feature economy. How do declarative frameworks such as LFG or HPSG account for this type of change?

An HPSG analysis of the dative passive in Present-day German includes a lexical entry for the passive auxiliary *bekommen* and a lexical rule deriving the lexical entry for a passive verb from its active counterpart. The feature description for the passive auxiliary in (11) captures the fact that auxiliaries are underspecified for their argument structure, i.e. *become* takes as its complement the passive participle and all the complements the passive verb requires. In contrast to the passive participle in the standard passive, participles in dative passive constructions may assign accusative case when an appropriate argument is present. Examples as (9b) have shown that the dative passive is not restricted to ditransitive verbs in Present-day German as has been claimed by Kordoni & van Noord (2009).

(11) *bekommt*-AUX 'gets':

$$\begin{bmatrix} \textsc{phon}\ \langle\ \textit{bekommt}\ \rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{vform} & \textit{fin} \\ \textsc{aux} & + \end{bmatrix} \\ \textsc{arg-st} & \left\langle \begin{bmatrix} \textsc{head} & \textit{passive-part} \\ \textsc{arg-st} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}, (\text{NP}[\textit{acc}]_{\boxed{2}}) \right\rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The passive lexical rule derives the passive participle subcategorized by the auxiliary *bekommen* 'get', illustrated by way of its input and output feature description

[5] Thanks go to Christian Ramelli for his native speaker judgment as regards the West Central German variety spoken in the Eastern part of the Saarland.

in (15) and (14), cf. Müller (2018: 285) for the standard passive construction in German. The dative passive rule suppresses the most prominent argument, the agent, of the active verb and promotes the argument marked for dative case by the active verb, the beneficiary, to subject of the passive verb *verordnet* 'prescribed'. The theme argument is not affected by the dative passive rule.[6]

(12) Er bekommt strenge Bettruhe verordnet.
he.NOM gets strict bed rest.NOM prescribed
' Strict bed rest was prescribed to him.'

(13) Input *verordn-* 'prescribe'

$$\begin{bmatrix} \text{PHON} & \langle \textit{ verordn } \rangle \\ \text{SYNSEM|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}, \text{NP}[\textit{dat}]_{\boxed{2}}, \text{NP}[\textit{acc}]_{\boxed{3}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{4}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \textit{verordnen} & \\ \text{EVENT} & \boxed{4} \\ \text{AGENT} & \boxed{1} \\ \text{BENEFACTIVE} & \boxed{2} \\ \text{THEME} & \boxed{3} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

[6] Assuming a lexical rule to derive a passive participle that fits the requirements of the passive auxiliary *bekommen*, results in three different participles, i.e. perfect participle, passive participle for standard passive constructions and passive participle for dative passive constructions. Müller (2018) argues that only one lexical rule is needed to derive the past participle. According to him, the respective auxiliary determines which arguments are realized in a perfect or a passive context. Whether his analysis is better suited to fit the participle data in German does not touch upon the issue of the grammaticalization process illustrated here.

(14) Output *verordnet* 'prescribed'

$$\begin{bmatrix} \textsc{phon} & \langle \textit{verordnet} \rangle \\ \textsc{synsem|loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} & \\ \textsc{vform} & \textit{passive-part} \end{bmatrix} \\ \textsc{arg-st} & \left\langle \text{NP}[\textit{nom}]_{\boxed{2}}, \text{NP}[\textit{acc}]_{\boxed{3}} \right\rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{index} & \boxed{4}\ \textit{event} \\ \textsc{rels} & \left\langle \begin{bmatrix} \textit{verordnen} & \\ \textsc{event} & \boxed{4} \\ \textsc{agent} & \boxed{1} \\ \textsc{benefactive} & \boxed{2} \\ \textsc{theme} & \boxed{3} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

In view of the analysis of the dative passive in Present-day German, its rise in the history of German may be modeled in terms of a change modifying the CONTENT feature of the verb *bekommen*. Semantic bleaching of the main verb in the course of its grammaticalization process affects above all its argument structure: The lexical verb *bekommen* assigns the thematic role beneficiary to its subject argument and a theme role to its direct object, cf. (15). The auxiliary *bekommen* on the other hand has no lexical meaning and does not assign thematic roles. The rise of the passive auxiliary *bekommen* can be thought of as taking place in two steps: Starting with the lexical verb *bekommen*, the first step includes its semantic bleaching such that the verb admits verbal complements headed by ditransitive verbs denoting a change of possession. The second step is best characterized by the further bleaching of the auxiliary-to-be: Its distribution is no longer governed by a semantic restriction to verbs denoting a change of possession, but only by a syntactic one, because it requires verbs with a dative object in its active counterpart, cf. (9a). Alongside its bleached version, the verb *bekommen* retains its original status as a lexical verb. The diachronic scenario sketched for *bekommen* 'get' is supposed to hold for the other passive auxiliaries as well.

(15) *bekommt* 'gets':

$$\begin{bmatrix} \text{PHON} & \langle \textit{gets} \rangle \\ \text{SYNSEM}|\text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \left\langle \text{NP}[\textit{nom}]_{\boxed{1}}, \text{NP}[\textit{acc}]_{\boxed{2}} \right\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{INDEX} & \boxed{3}\ \textit{event} \\ \text{RELS} & \left\langle \begin{bmatrix} \textit{bekommen} & \\ \text{EVENT} & \boxed{3} \\ \text{BENEFACTIVE} & \boxed{1} \\ \text{THEME} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Other studies on the rise of auxiliary verbs in representational theories of grammar include Schwarze (2001) who models this change within the framework of lexical functional grammar. He suggests to capture the rise of auxiliary verbs in Romance by a difference with respect to the f-structure features in the lexical entries between lexical and auxiliary verbs. In actual Romance, the f-structure of a passive clause is headed by the non-finite verb, while the modern Romance counterparts of *esse* 'be' may no longer function as heads of an f-structure with their transition to auxiliary status.

2.2 Word Order Changes in the Verbal Complex

Present-day German is an OV language with the finite verb occuring at the right edge of subordinate clauses. In case more than one verb appears in final position, the canoncial order is descending, i.e. $V_3V_2V_1$ with the governing verb following the governed verb as illustrated by the three-place verb cluster in (16). The occurrence of non-canonical orders like $V_1V_3V_2$ is restricted in Standard German with respect to the number of verbs in the cluster (at least three verbs) and the type of auxiliary. While the auxiliary verb *haben* 'have' requires the non-canonical word order, cf. (17a), the tense auxiliary *werden* 'will' and the modal verbs may occur with both orders. As regards the auxiliary *sein* 'be', the non-canonical order is not available (17b).

(16) Auch wenn Selma die Noten gefunden$_3$ haben$_2$ wird$_1$,
even if Selma the notes found have will
'Even if Selma will have found the music'

(17) a. Auch wenn Fred das hätte$_1$ wissen$_3$ müssen$_2$,
even if Fred this had know must
'even if Fred should have known this'

b. ∗Auch wenn er gestern in der Vorlesung ist$_1$ gesehen$_3$ worden$_2$,
even if he yesterday in the lecture is seen been
'Even if he has been seen in the lecture yesterday '

Neither of these restrictions holds for non-standard varieties of German: Upper German dialects provide ample evidence for both descending and ascending verb orders, even if the verb cluster includes only two verbs (Dubenion-Smith 2010). Likewise historical stages of German witness a wide variety of word orders regardless of the number of verbs appearing in final position and independent of the nature of the auxiliary (Ebert 1981; Härd 1981; Sapp 2011). The Early New High German examples in (18) render attestations for a two-place verbal complex with *haben* preceding a past participle and a three-place verbal complex with the auxiliary verb *sein* preceding two past participles. Both patterns are ruled out in the standard varieties of Present-day German.

(18) a. uns ist ein Abentüer widerfaren underwegen, daz uns ein Wolff
us is an adventure happened on the way that us a wolf
vil Leids hat$_1$ gethon$_2$ (Ulenspiegel 226.4)
much harm has done
'An adventure happened to us on the way: a wolf has done much harm to us.'

b. so schreibt man auch aus Holl. das newlich in Frießlandt ein
so writes one also from Holland that recently in Friesland a
fewriger fliegender Trach sey$_1$ gesehen$_3$ worden$_2$ (Aviso 35.14)
fiery flying dragon were seen been
'News come from the Netherlands that a fiery flying dragon has been seen in Friesland.'

The restrictions effective in Present-day German arise in a two-step process: The order V_2V_1 becomes fixed with two-place verbal complexes troughout the 16th century, while it took about a hundred more years for the order $V_3V_2V_1$ to become the canonical order for three-place verbal complexes (Ebert 1981; Härd 1981; Sapp 2011).

How is this change modeled in a representational framework such as HPSG? Auxiliary verbs and their verbal complements as given in (16) through (18) are supposed to build verb clusters with the arguments of the respective verbal complement being attracted by the auxiliary (Hinrichs & Nakazawa 1994; Pollard 1994; Kathol 2000; Müller 2002; 2008).[7] Accordingly, the structure of a verb cluster exemplifying the canonical descending order can be represented as in the passive verb cluster *dass sie die Fäden gezogen bekam* 'that she had removed the stitches':

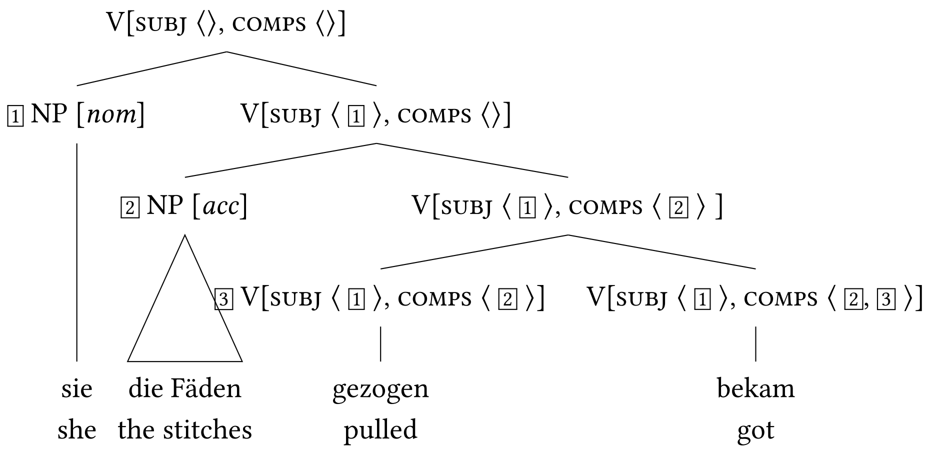


Figure 1: Passive verb cluster in *dass sie die Fäden gezogen*$_2$ *bekam*$_1$

The variation regarding the order of auxiliary and lexical verb in verb clusters is currently addressed from two perspectives: (i) non-canonical patterns of three-place verb clusters in Present-day German as in (17a) which figure under the notion of *Oberfeldbildung* since Bech (1955), and (ii) the canonical order in Dutch verb clusters which is ascending instead of descending, i.e. $V_1V_2V_3$. Building on previous proposals, one way to account for the word order change affecting the verbal complex in the history of German would include the assumption of an appropriate head feature as advocated by Hinrichs & Nakazawa (1994). They emphasize that a lexical approach to the word order within the verb cluster would account also for the variation on the level of individual speakers.

Recent work suggests the head feature GOVR which indicates the direction of government of non-finite verbs and was proposed to capture synchronic variation in German and Dutch (Bouma & van Noord 1996; Kathol 2000; Augustinus

[7] Abeillé & Godard (2002) argue that verbal complements of auxiliary verbs in French are part of a flat VP when the auxiliary conveys tense information, while the passive auxiliary *être* 'be' takes a VP complement.

2015).[8] From a diachronic perspective, one would either assume that infinitival as well as participial complements carry a feature GOVR which is underspecified in earlier stages of German (= GOVR *dir*) allowing for the attested variation of word orders, cf. figure 2, or one would go on the assumption that the head feature GOVR arises only later in the history of German. In the context of example (18b), the value of GOVR may be determined as follows: The participle *gesehen* 'seen' is governed by the auxiliary *worden* 'been' appearing on its right side and the GOVR value of the verb cluster *gesehen worden* is feature shared with its head daughter *worden* which is governed by the auxiliary *worden*.

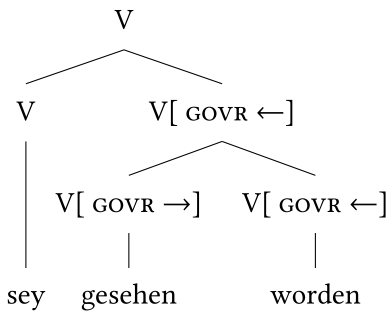


Figure 2: Three-place verb cluster in Early New High German

In Present-day German, a three-place verb cluster including the auxiliary *sein* 'be' exhibits the canonical word order $V_3V_2V_1$. In contrast to Early New High German, the feature value is provided by the lexical entry of the respective verb. A partial lexical description of the passive participle *worden* 'been' is illustrated below, indicating that its governing auxiliary has to appear on the right side.

(19) *worden*-AUX 'been':

$$\begin{bmatrix} \text{PHON} \ \langle \ \textit{worden} \ \rangle \\ \text{SYNSEM|LOC|CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{verb} & \\ \text{VFORM} & \textit{passive-part} \\ \text{AUX} & + \\ \text{GOVR} & \rightarrow \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

As has been suggested in work on synchronic variation as regards verb cluster in German and Dutch, diachronic variation can be modeled in a straightforward

[8] The head feature suggested by Hinrichs & Nakazawa (1994) is FLIP which indicates government to the right when exhibiting a positive value as in example (17a) above.

way by building on lexical entries of verbs. The analysis of the word order change sketched above makes use of the possibility that a lexical feature may be underspecified in a particular variety of a language.

2.3 Left Periphery of Noun Phrases

The third case study presents a bundle of changes affecting the left periphery of noun phrases in the history of German. All changes might be due to a single change as regards the relationship between nominal and determiner (Demske 2001).

The historical record indicates that the distribution of adjectival inflection types is semantically governed in Old High German, i.e. definite determiners trigger weak adjectival inflection, whereas indefinite determiners call for strong adjectival inflection. In (20), the weak declension type is triggered by the demonstrative and the possessive determiner, respectively, while the strong declension type in (21) is motivated by the indefinite pronoun. In contrast to Present-day German the strong declension type is used irrespective of the morphology of the indefinite determiner (cf. *ein* vs. *einemo*). Old High German behaves in this respect as Modern Icelandic.[9]

(20) a. thiz irdisg-a dal (OHG)
this worldly-WEAK valley
'this worldly valley'

b. min liob-o sun (OHG)
my good-WEAK sun
'my good sun'

(21) a. ein arm-az uuîb (OHG)
a poor-STRONG woman
'a poor woman'

b. einemo diur-emo merigrioze (OHG)
a valuable-STRONG pearl
'a valuable pearl'

In Present-Day German the distribution of adjectival inflection types is morphologically governed: If grammatical features are overtly marked by the determiner,

[9] In Modern Icelandic, a definite noun phrase requires a weakly inflected adjective and an indefinite noun phrase a strongly inflected one: *þessi raud-i hestur* 'this red horse' vs. *raud-ur hestur* ' a red horse'.

the following adjective instantiates the weak inflection type, otherwise the adjective exhibits strong inflection (cf. *ein* vs. *einem*).

(22) a. ein herausragend-er Cellist (PDG)
an outstanding-STRONG cellist

b. einem herausragend-en Cellisten (PDG)
an outstanding-WEAK cellist
'an outstanding cellist'

The changing nature of the relationship between determiner and nominal may be captured by a change in the feature structure of the determiner: In Old High German, it is the CONTENT feature of the determiner that drives the distribution of adjectival declension types, in Present-day German on the other hand, the distribution is driven by its CAT feature. The feature description of the indefinite determiner Old High German includes as AGRfeatures CASE, NUMBER and GENDER[10]. The feature SPR indicates that the determiner requires a nominal expression lacking a specifier, i.e. NOM(Sag et al. 2003). The index $\boxed{1}$ signals the structure sharing of the determiner's CONTENT feature with the DECLfeature of the nominal expression. Determiner and nominal agree in the following example with respect to indefiniteness.

(23)
$$\begin{bmatrix} \text{PHON} & \langle\, \textit{einemo} \,\rangle \\ \text{SYNSEM|LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{det} & \\ \text{AGR} & \begin{bmatrix} \text{CASE} & \textit{dat} \\ \text{NUM} & \textit{sg} \\ \text{GEND} & \textit{masc} \end{bmatrix} \end{bmatrix} \\ \text{SPR} & \langle\, \text{NOM}\,[\text{DECL}\ \boxed{1}\,]\,\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{DET} & \textit{exists} \\ \text{RESTIND} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

In Present-day German, the DECLvalue of NOMno longer conveys information about its definiteness. The determiner selects for NOMaccording to categorial features: In case the determiner provides information on the AGRvalue of the noun phrase, it asks for a weakly inflected NOMas in (24).[11] Otherwise the information

[10] Disregarding the context in (21), the determiner *einemo* may also agree with a neuter noun.
[11] In my view, DECLis not a HEAD feature of the determiner, since declension type is an inherent feature of determiners. Cf. however Kiss (1995: 72).

in question has to be provided by a strongly inflected NOM.[12]

(24)
$$\begin{bmatrix} \textsc{phon} & \langle \textit{einem} \rangle \\ \textsc{synsem|loc|cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{det} & \\ \textsc{agr} & \begin{bmatrix} \textsc{case} & \textit{dat} \\ \textsc{num} & \textit{sg} \\ \textsc{gend} & \textit{masc} \end{bmatrix} \end{bmatrix} \\ \textsc{spr} & \langle \textsc{nom} \, [\textsc{decl} \; \textit{weak}] \rangle \end{bmatrix} \end{bmatrix}$$

The changing nature of the relationship between determiner and NOMcan be modeled by modifying the feature structure of SPR in the lexical description of the determiner. A possible motivation for this change comes from the increasing grammaticalization of the definite determiner: While the determiner is attested above all with sortal concepts in Old High German testifying to its use as a demonstrative, it lacks in cases where the head noun refers to functional concepts that are inherently unambiguous (Demske 2001). The examples in (25) illustrate this distribution: A head noun like *figboum* 'figtree' has a sortal meaning, i.e. its unique referent is specified by the context in (25a), while the noun *erda* 'earth' denotes a functional concept which refers unambiguously irrespective of particular situations (25b). A demonstrative determiner as *ther* 'this one' is hence excluded in such a context.

(25) a. Inti quad Imo, niomer fon thir uuahsmo arboran uuerde zi éiuuidu thô sâr sliumo arthorr&a **ther figboum**. (OHG)

'and he saith unto it, Let there be no fruit from thee henceforward for ever. And immediately the fig tree withered away.'

b. Inti **erda** giruorit uuas Inti steina gislizane uuarun (OHG)

'And the earth shook and the rocks were split.'

The rise of weak definites in the course of Early New High German represents a further step in the grammaticalization process of the definite determiner: Relational nouns with their argument in postnominal position may lack any determiner (26a) or they exhibit either the indefinite or the definite determiner (26b, 26c). The third pattern is the default pattern in Present-Day German (Demske 2019).

[12] Pollard & Sag (1994: 373) point out that nouns like *Verwandter* 'relative' support the assumption that DECLis a feature not only of adjectives but also of nouns, because the declension class of these nouns is governed by the respective determiner.

(26) a. Vnd wie wol ich bin **sone eins konigs** (ENHG)

'And though I am the son of a king.'

b. Sie namen **einen Zahn eines Thiers** / welches so groß ist wie eine Ratte (ENHG)

'They took the tooth of an animal which was as big as a rat.'

c. ich verstunde gleich aus ihrem Diskurs (…) daß ihr Mann beim Senat wäre, und ohngezweifelte Hoffnung hätte, denselben Tag **die Stell eines Landvogts oder Landamtmanns** zu bekommen (ENHG)

'From her conversation I came to discover that her husband was in the senate (…) He was also supposed to have had good expectations of receiving the position of a district governor or a bailiff that very day.'

In Present-Day German, definite determiners are used with sortal, functional and relational nouns, cf. (27), suggesting that the definite determiner is no longer licensed on semantic grounds as in Old High German, but on morphological grounds. The changing distribution of the definite determiner fits nicely the assumption that the specifier relation originally is semantically based and then turns into a morphologically based relationship.

(27) a. Der Mond ist aufgegangen. (PDG)

'The moon has risen.'

b. Sie ist die Tochter eines Unternehmers. (PDG)

'She is the daughter of an entrepreneur'

The increasing grammaticalization of the nominal left periphery has further effects on pre-head constituents. The demonstrative pronoun *solch* 'such' has either a sortal or an individual reading in older stages of German allowing for singular count nouns to occur without another determiner (28). In Present-Day German, *solch* is restricted to a sortal interpretation and exhibits classical diagnostics for adjectivehood including the requirement of a determiner with singular count nouns (Demske 2005). The developing restrictions governing the use of the demonstrative pronoun *solch* 'such' are easily modeled within a lexical approach, going on the assumption that there are either two lexical entries for the demonstrative in earlier stages of German or that the HEAD feature allows both a SPR as well as a MOD relation between the demonstrative and the nominal head. With the development of the demonstrative *ther* 'this one' into the definite determiner, the demonstrative *solch* conventionalizes its use as a demonstrative adjective while losing the one as a determiner.

(28) derselbe Landherr hatt an jetzo **solche Türckin** / weil jhr man nicht zur hand / sondern verreist gewest / wider vnder die Türcken vmb etliche Türckische Teppich vnd andere sachen verkaufft / (ENHG)

'this overlord has sold now this Turkish woman to other Turkish people for several Turkish tapestry and other things, because her husband has been abroad.'

A significant role in the history of the left periphery is played by the adnominal genitive. Three stages have to be distinguished in its development: In the first stage, genitive noun phrases systematically appear in prenominal position as attested in OHG sources (29). The genitive is a prenominal complement at this stage indicated by the preceding determiner and adjective as illustrated by the second example. Note that the determiner and the adjective are marked for dative case as required by the initial preposition *in* 'in'.

(29) a. scouuot thes accares lilia uuvo sie uuahsen (OHG)
observe the field's lilies how they grow
'Observe how the lilies of the field grow.'

b. In dhemu heilegin daniheles chiscribe (OHG)
in the holy Daniel's scripture
'in the holy scripture of Daniel'

The adnominal genitive of stage two also occurs in prenominal position, provided that it denotes humans or animals, cf. (30). All adnominal genitives marked [-animate] are now limited to postnominal position as testified by historical data from the Early New High German period (Ebert 1988). The prenominal genitive is still a full noun phrase in stage two, allowing not only for pre-head, but also for post-head dependents as shown by the postnominal modifiers in (30), which can be phrasal or sentential.

(30) a. **Der Frawen zu vnseren zeiten** kunst weyßheit vnd tugende ist nit not zu erzelen (ENHG)

'The art, wisdom and virtue of women in our days does not need to be recounted'

b. Dieser Tagen seyn allhie **der Evangel. Fürsten vnnd Südt/ so zu Schwäbschen Hall jüngst beysamen gewest/** Abgesandte alher komen (ENHG)

'These days envoys of the Protestant sovereigns and cities have come here, after they have met at Schwäbisch Hall recently.'

The final stage in the development of the adnominal genitive is represented by Present-day German: The prenominal position is restricted to proper names and kinship terms disallowing any pre-head or post-head modifier. Genitive phrases headed by individual nouns appear in postnominal position irrespective of the feature [±animate].

(31) a. Selmas/Vaters altes Fahrrad
Selma's/Daddy's old bike

b. *der Menschheit ältester Traum
the mankind's oldest dream

(32) a. das Gartenhaus des alten Goethe
the summer house of the old Goethe

b. *des alten Goethe Gartenhaus
the old Goethe summer house
'the summer house of late Goethe'

The historical scenario sketched for the development of the adnominal genitive fits in well with other changes affecting the left periphery of noun phrases: While the adnominal genitive is a full noun phrase in Old High German functioning as a complement of the head noun, the adnominal genitive is used as a possessive determiner in Present-day German. Again the change can be modeled as a change affecting the relation between a prenominal constituent and the nominal head. The Old High German demonstrative pronouns *ther* 'this one' and *sulîh* 'such' conventionalize either an individual or a sortal reading in the history of German. Patterning with the former, the prenominal genitive conventionalizes a determiner relation to the nominal, provided that it can contribute a possessive meaning. Adnominal genitives with the feature [-animate] are consequently postponed and they retain their grammatical function as complements of the head noun. Prenominal genitives on the other hand become possessive determiners establishing a SPR relation to nominals with strong inflection (Pollard & Sag 1994: 54):

(33) Lexical item for *Selmas*:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{Selmas} \rangle \\ \textsc{synsem}|\textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \textit{det} \\ \textsc{spr} & \langle\ \textsc{nom}\ [\textsc{decl}\ \textit{strong},\ \textsc{index}\ \boxed{1},\ \textsc{restr}\ \boxed{2}\,]\ \rangle \end{bmatrix} \\ \textsc{content} & \begin{bmatrix} \textsc{det} & \textit{the} \\ \textsc{restind} & \begin{bmatrix} \textsc{index} & \boxed{1} \\ \textsc{restr} & \left\langle \begin{bmatrix} \textsc{reln} & \textit{poss} \\ \textsc{possessor} & \boxed{3} \\ \textsc{possessed} & \boxed{1} \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix} \end{bmatrix} \\ \textsc{context}|\textsc{backgr} & \left\langle \begin{bmatrix} \textsc{reln} & \textit{naming} \\ \textsc{bearer} & \boxed{3} \\ \textsc{name} & \textit{Selma} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The reanalysis of the relation between prenominal possessive and nominal in the history of German not only affects the pre-head genitive but also the possessive pronoun. In Middle High German and still in Early New High German, the possessive pronoun patterns with adjectives considering its co-occurrence patterns: It may follow a definite determiner and may even agree with another adjective as regards its declension type as in (34a). Here possessive pronoun and adjective both exhibit weak declension triggered by the preceding definite determiner. At this stage in the history of German, the possessive pronoun ist still a constituent of the nominal. The Early New High German example in (34b) shows that a possessive pronoun may also follow a prenominal adjective suggesting that it functions as a modifier itself (note also the agreement with respect to declencion type between adjective and possessive pronoun). This word order is excluded in Present-day German.

(34) a. die iuwer-n scoen-en tohter (MHG)
the your-WEAK beautiful-WEAK daughter
'your beautiful daughter'

b. mit gross-em jhr-em Rhum vnd Lob (ENHG)
with big-STRONG their-STRONG glory and praise
'with their big glory and praise'

Once again, it is the relation between prenominal element and head noun which is subject to change: The possessive pronoun behaves as an adjectival modifier in earlier stages of German, before it becomes a possessive determiner (i.e. MOD relation develops into SPR relation). According to Pollard & Sag (1994: 54), its lexical description looks very much like the description of the prenominal genitive given above (disregarding the CONTENT and the CONTEXT value).

Changes at the left periphery of noun phrases start with the grammaticalization of the definite determiner throughout the period of Old High German, testified by a widening of its distribution. In Present-day German, not only sortal, but also functional and relational nouns combine with the definite determiner. The steady increase in the use of the determiner triggers a reanalysis of the relation between determiner and nominal: A semantically driven relation turns into a relation that is also morphologically based as evidenced by the changing trigger for the adjectival declension type. The determiner is consequently licensed by the CAT feature in the noun's feature structure as shown for the functional noun *Mond* 'moon' which subcategorizes for its determiner in Present-day German but has not done so in Old High German.

(35)
$$\begin{bmatrix} \text{PHON} & \langle \textit{Mond} \rangle \\ \text{SYNSEM|LOC|CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{SPR} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{det} & \\ \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{ARG-ST} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

The reanalysis of the relation between determiner and nominal has consequences for the interpretation of other prenominal constituents: Possessive pronouns as well as possessor phrases are likewise taken to instantiate a specifier relation to the nominal in question, thus augmenting the class of determiners in German. In addition, pre-head constituents precluding a specifier interpretation are postponed, cf. genitive complements with [-animate], and pre-head constituents ambiguous between a specifier and a modifier reading are limited to one interpretation, cf. *solch* 'such' behaving as a demonstrative adjective in Present-day German. All changes can be modeled in a straightforward way as lexical changes.

3 Summary

Recent years witnessed a growing consensus that syntactic change is best accounted for in the lexicon of a language. The consensus holds across frameworks: Biberauer & Walkden (2015) highlight the role of the lexicon in Minimalism, the volume by Butt & King (2001) exhibits case studies of syntactic change in the representational framework of LFG. The present contribution set out to show how the typed feature structures of HPSG can be used to model the way syntactic structures change over time. Different types of morpho-syntactic change have been considered in the history of German: the grammaticalization of auxiliary verbs and of demonstrative pronouns, word order changes affecting verb and noun phrases and changing relations between prenominal constituents and the respective nominal. In all cases, the change in question can be modeled in terms of feature structures in the lexicon of a language.

Abbreviations

Acknowledgements

Sources

Include a list the historical sources used in the paper?

References

Abeillé, Anne & Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3). 404–452.

Allen, Cynthia. 2006. Case syncretism and word order change. In Ans van Kemenade & Bettelou Los (eds.), *The Handbook of the History of English*, 201–223. Oxford: Blackwell Publishing.

Augustinus, Liesbeth. 2015. *Complement Raising and Cluster Formation in Dutch: a Treebank-Supported Investigation*. Utrecht: LOT dissertation.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation* (Linguistics Today 132), 351–373. Benjamins Publishing.

Bech, Gunnar. 1955. *Studien über das deutsche verbum infinitum*. Copenhagen: Munksgaard.

Biberauer, Theresa & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge Handbook of Historical Syntax*, 134–162. Cambridge: Cambridge University Press.

Biberauer, Theresa & George Walkden. 2015. Introduction: changing views of syntactic change. In Theresa Biberauer & George Walkden (eds.), *Syntax over Time: Lexical, Morphological, and Information-Structural Interactions* (Oxford Studies in Diachronic and Historical Linguistics 15), 1–13. Oxford: Oxford University Press.

Börjars, Kersti & Nigel Vincent. 2017. Lexical-Functional Grammar. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge Handbook of Historical Syntax*, 642–663. Cambridge: Cambridge University Press.

Bouma, Gosse & Gertjan van Noord. 1996. Word order constraints on verb clusters in German and Dutch. In Erhard Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex Predicates in Nonderivational Syntax* (Syntax and Semantics 30), 43–72. Academic Press.

Butt, Miriam & Tracy Holloway King (eds.). 2001. *Time over Matter: Diachronic Perspectives on Morphosyntax* (Studies in Constraint-Based Lexicalism). Stanford, CA: CSLI Publications.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Demske, Ulrike. 2001. *Merkmale und Relationen: Diachrone Studien zur Nominalphrase des Deutschen* (Studia Linguistica Germanica 56). Berlin: De Gruyter.

Demske, Ulrike. 2005. Weshalb Demonstrativpronomina nicht immer Determinantien sind. In Franz-Josef d'Avis (ed.), *Deutsche Syntax. Empirie und Theorie. Symposium in Göteborg 13.-15. Mai 2004* (Göteborger Germanistische Forschungen 46), 53–80. Göteborg: Acta Universitatis Gothoburgensis.

Demske, Ulrike. 2019. The Grammaticalization of the Definite Article in German: From Demonstratives to Weak Definites. In Johanna Flick & Renata Szczepaniak (eds.), *Walking on the Grammaticalization Path of the Definite Article in German: Functional Main and Side Roads.* Amsterdam: Benjamins.

Denison, David. 1993. *English Historical Syntax.* London: Longman.

Dubenion-Smith, Shannon A. 2010. Verbal Complex Phenomena in West Central German: Empirical Domain and Multi-Causal Account. *Journal of Germanic Linguistics* 22(2). 99–191.

Ebert, Robert Peter. 1978. *Historische Syntax des Deutschen.* Stuttgart: Metzler.

Ebert, Robert Peter. 1981. Social and stylistic variation in the order of auxiliary and nonfinite verbs in dependent clauses in Early New High German. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 103. 204–237.

Ebert, Robert Peter. 1988. Variation in the position of the attributive genitive in the sixteenth century German. *Monatshefte* 80. 32–49.

Glaser, Elvira. 2005. 'Krieg' und 'kriegen': zur Arealität der 'bekommen'-Periphrasen. In Annelies Häcki Buhofer, Ulla Kleinberger Günther & Elisabeth Piirainen (eds.), *Krieg und Frieden. Auseinandersetzung und Versöhnung in Diskursen*, 43–64. Tübingen: Niemeyer.

Grewendorf, Günther & Cecilia Poletto. 2005. Von OV zu VO: ein Vergleich zwischen Zimbrisch und Plodarisch. In Ermenegildo Bidese, James R. Dow & Thomas Stolz (eds.), *Das Zimbrische zwischen Germanisch und Romanisch*, 114–128. Bochum: Brockmeyer.

Haider, Hubert. 1986. Fehlende Argumente: vom Passiv zu kohärenten Infinitiven. *Linguistische Berichte* 101. 3–33.

Härd, John Evert. 1981. *Studien zur Struktur mehrgliedriger deutscher Nebensatzprädikate. Diachronie und Synchronie*. Gothenburg: Acta Universitatis Gothoburgensis.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Hinrichs, Erhard & Tsuneko Nakazawa. 1994. Linearizing AUXs in German Verbal Complexes. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–37. Stanford, CA: CSLI Publications.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University Press.

Kinn, Kari. 2015. *Null Subjects in the History of Norwegian*. Oslo: University of Oslo dissertation.

Kiss, Tibor. 1995. *Merkmale und Repräsentationen. Eine Einführung in die deklarative Grammatikanalyse*. Opladen: Westdeutscher Verlag.

Kordoni, Valia & Gertjan van Noord. 2009. Passives in the Germanic languages: the case of Dutch and German. *Groninger Arbeiten zur Germanistischen Linguistik* 49. 77–96.

Lehmann, Christian. 2015. *Thoughts on grammaticalization*. 3rd edn. Berlin: Language Science Press.

Lenz, Alexandra. 2007. Zur Grammatikalisierung von 'geben' im Deutschen und Lëtzebuergeschen. *Zeitschrift für germanistische Linguistik* 35. 52–82.

Lenz, Alexandra. 2012. On the genesis of the German recipient passive – two competing hypotheses in the light of current dialect data. In Gunther de Vogelaer & Guido Seiler (eds.), *The dialect laboratory. Dialects as a testing ground for theories of language change* (SLCS 128), 121–138. Amsterdam: Benjamins.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Louden, Mark. 1988. *Bilingualism and Syntactic Change in Pennsylvania German*. Cornell University dissertation.

Müller, Stefan. 2002. *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2008. *Head-driven phrase structure grammar*. 2nd edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg.

Müller, Stefan. 2018. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 2nd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.5281/zenodo.1193241

Petzell, Erik M. 2016. Low German with a Swedish twist – Contact-induced word order transfer in the 15th century. *Ampersand* 3. 143–150.

Pollard, Carl J. 1994. Toward a Unified Account of Passive in German. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 273–296. Stanford, CA: CSLI Publications.

Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Reis, Marga. 1976. Zum grammatischen Status der Hilfsverben. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 98. 64–82.

Reis, Marga. 1985. Mona Lisa kriegt zuviel – Vom sogenannten 'Rezipientenpassiv' im Deutschen. *Linguistische Berichte* 96. 140–155.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change*. Cambridge: Cambridge University Press.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. 2 edition. Stanford, CA: CSLI Publications.

Sapp, Christopher D. 2011. *The Verbal Complex in Subordinate Clauses from Medieval to Modern German* (Linguistics Today 173). Amsterdam: John Benjamins.

Schwarze, Christioph. 2001. Representation and variation: on the development of Romance auxiliary syntax. In Miriam Butt & Tracy Holloway King (eds.), *Time over Matter: Diachronic Perspectives on Morphosyntax*, 143–172. Stanford, CA: CSLI Publications.

Vincent, Nigel. 2001. LFG as a model of syntactic change. In Miriam Butt & Tracy Holloway King (eds.), *Time over Matter: Diachronic Perspectives on Morphosyntax*, 1–42. Stanford, CA: CSLI Publications.

Chapter F

Acquisition

Jonathan Ginzburg
Université Paris Diderot

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse

Jonathan Ginzburg. 2019. Acquisition. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 773–774. Berlin: Language Science Press. DOI:??

rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter G

Processing

Tom Wasow

Although not much psycholinguistic research has been carried out in the framework of HPSG, the architecture of the theory fits well with what is known about human language processing. This chapter enumerates aspects of this fit. It then discusses two phenomena, island constraints and relative clauses, in which the fit between experimental evidence on processing and HPSG analyses seems particularly good.

1 Introduction

Little psycholinguistic research has been guided by ideas from HPSG (but see Konieczny 1996 for a notable exception). This is not so much a reflection on HPSG as on the state of current knowledge of the relationship between language structure and the unconscious processes that underlie language production and comprehension. Other theories of grammar have likewise not figured prominently in theories of language processing, at least in recent decades.[1] The focus of this chapter, then, will be on how well the architecture of HPSG comports with available evidence about language production and comprehension.

My argument is much the same as that put forward by Sag et al. (2003: Chapter 9), and Sag & Wasow (2011; 2015), but with some additional observations about the relationship between competence and performance. I presuppose the "competence hypothesis" (see Chomsky 1965: Chapter 1), that is, that a theory of language use (performance) should incorporate a grammar representing the

[1]Half a century ago, the Derivational Theory of Complexity (DTC) was an attempt to use psycholinguistic experiments to test aspects of the grammatical theory that was dominant at the time. The DTC was discredited in the 1970s, and the theory it purported to support has long-since been superseded. See Fodor et al. (1974) for discussion.

Tom Wasow. 2019. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 775–796. Berlin: Language Science Press. DOI:??

knowledge of language (competence) that is drawn on in everyday comprehension and production, as well as in other linguistic activities, such as language games and the (often artificial) tasks employed in psycholinguistic experiments.

The primary reason for adopting the competence hypothesis is parsimony: a theory of language use is simpler if it does not have to repeat much the same information about the language in both its production and comprehension components. This information would include things like the vocabulary, the preferred word orders, and most of the rest of what linguists encode in their grammars. A performance theory that incorporates a grammar only needs to include such information once.[2] Moreover, to the extent that the theoretical constructs of the grammar play a role in modeling both production and comprehension, the overall theory is simpler.

There is also, however, an empirical reason for preferring a model with a good fit between competence and performance. As noted by Bresnan et al. (2001), preferences that are only statistical tendencies in some languages can show up in others as categorical requirements. The example they discuss in detail is the avoidance of clauses with third-person subjects but first- or second-person objects or obliques. In English, this is a powerful statistical tendency, which they document by showing that the passivization rate in the Switchboard corpus is very significantly lower when the agent is first- or second-person than when it is third-person. In Lummi (a Salish language of British Columbia), this preference is categorical: clauses with third-person subjects but first- or second-person objects or obliques are simply unacceptable. Hawkins (2004; 2014) argues that such examples are by no means exceptional, and formulates the following "Performance-Grammar Correspondence Hypothesis" (PGCH):

> Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by frequency of use and

[2] There are of course some discrepancies between production and comprehension that need to be accounted for in a full theory of language use. For example, most people can understand some expressions that they never use, including such things as dialect-specific words or accents. But these discrepancies are on the margins of speakers' knowledge of their languages. The vast majority of the words and structures that speakers know are used in both production and comprehension. Further, it seems to be generally true that what speakers can produce is a proper subset of what they can comprehend. Hence, the discrepancies can plausibly be attributed to performance factors such as memory or motor habits. See Gollan et al. (2011) for evidence of differences between lexical access in production and comprehension. See Momma & Phillips (2018) for arguments that the structure-building mechanisms in production and comprehension are the same. For a thoughtful discussion of the relationship between production and comprehension, see MacDonald (2013) and the commentaries published with it.

ease of processing.[3]

There are two ways in which a processing model incorporating a grammar might capture this generalization. One is to give up the widespread assumption that grammars provide categorical descriptions, and that any quantitative generalizations must be extra-grammatical; see Francis (2019) for arguments supporting this option, and thoughtful discussion of literature on how to differentiate processing effects from grammar. For example, some HPSG feature structures might allow multiple values for the same feature, but with probabilities (adding up to 1) attached to each value.[4] I hasten to add that fleshing out this idea into a full-fledged probabilistic version of HPSG would be a large undertaking, well beyond the scope of this chapter; see Linadarki (2006) and Miyao & Tsujii (2008) for work along these lines. But the idea is fairly straightforward, and would allow, for example, English to have *in its grammar* a non-categorical constraint against clauses with third-person subjects and first- or second-person objects or obliques.

The second way for a theory adopting the competence hypothesis to represent Hawkins's PGCH would be to allow certain generalizations to be stated either as grammatical constraints (when they are categorical) or as probabilistic performance constraints. This requires a fit between the grammar and the other components of the performance model that is close enough to permit what is essentially the same generalization to be expressed in the grammar or elsewhere. In the case discussed by Bresnan et al., for example, treating the constraint in question as part of the grammar of Lummi but a matter of performance in English would require that both the theory of grammar and models of production would include, minimally, the distinction between third-person and other persons, and the distinction between subjects and non-subjects. Since virtually all theories of grammar make these distinctions, this observation is not very useful in choosing among theories of grammar. I will return later to phenomena

[3]In the Bresnan et al. example, I know of no experimental evidence that clauses with third-person subjects and first- or second-person objects are difficult to process. But a plausible case can be made that the high salience of speaker and addressee makes the pronouns referring to them more accessible in both production and comprehension than expressions referring to other entities. In any event, clauses with first- or second-person subjects and third-person objects are far more frequent than clauses with the reverse pattern in languages where this has been checked. Thus, the Bresnan et al. example falls under the PGCH, at least with respect to "frequency of use" .

[4]I discussed this idea many times with the late Ivan Sag. He made it clear that he believed grammatical generalizations should be categorical. In part for that reason, this idea was not included in our joint publications on processing and HPSG.

that bear on the choice among grammatical theories, at least if one accepts the competence hypothesis.

Since its earliest days, HPSG research has been motivated in part by considerations of computational tractability (see Flickinger, Pollard & Wasow 2019, Chapter 2 of this volume, for discussion). Some of the design features of the theory can be traced back to the need to build a system that could run on the computers of the 1980s. Despite the obvious differences between human and machine information processing, some aspects of HPSG's architecture that were initially motivated on computational grounds have turned out to fit well with what is known about human language processing. A prime example of that is the computational analogue to the competence hypothesis, namely the fact that the same grammar is used for parsing and generation. In Section 3, I will discuss a number of other high-level design properties of HPSG, arguing that they fit well with what is known about human language processing, which I summarize in Section 2. In Section 4, I will briefly discuss two phenomena that have been the locus of much discussion about the relationship between grammar and processing, namely island constraints and differences between subject and object relative clauses.

2 Key facts about human language processing

In this section I review a number of well-known general properties of human language processing. Most of them seem evident from subjective experience of language use, but there is supporting experimental evidence for all of them.

2.1 Incrementality

Both language production and comprehension proceed incrementally, from the beginning to the end of an utterance. In the case of production, this is evident from the fact that utterances unfold over time. Moreover, speakers very often begin their utterances without having fully planned them out, as is evident from the prevalence of disfluencies. On the comprehension side, there is considerable evidence that listeners (and readers) begin analyzing input right away, without waiting for utterances to be complete. A grammatical framework that assigns structure and meaning to initial substrings of sentences will fit more naturally than one that doesn't into a processing model that exhibits this incrementality we see in human language use.

I hasten to add that there is also good evidence that both production and comprehension involve anticipation of later parts of sentences. While speakers may not have their sentences fully planned before they begin speaking, some planning of downstream words must take place. This is perhaps most evident from instances of nouns exhibiting quirky cases determined by verbs that occur later in the clause. For example, objects of German *helfen*, 'help', take the dative case, rather than the default accusative for direct objects. But in a sentence like (1), the speaker must know that the verb will be one taking a dative object at the time the dative case article *dem* is uttered.

(1) Wir werden dem Kind bald helfen. (German)
we will the.DAT child soon help
'We will help the child soon.'

Likewise, in comprehension there is ample evidence that listeners and readers anticipate what is to come. This has been demonstrated using a variety experimental paradigms. Eye-tracking studies (see Tanenhaus, Spivey-Knowlton & Sedivy 1995, Altmann & Kamide 1999, Arnold et al. 2007, among many others) have shown that listeners use semantic information and world knowledge to predict what speakers will refer to next.

Thus, a theory of grammar that fits comfortably into a model of language use should provide representations of initial substrings of utterances that can be assigned (partial) meanings and be used in predicting later parts of those utterances.

2.2 Non-modularity

Psycholinguistic research over the past four decades has established that language processing involves integrating a wide range of types of information on an as-needed basis. That is, the various components of the language faculty interact throughout their operation. A model of language use should therefore *not* be modular, in the sense of Jerry Fodor's influential (1983b) book, *The Modularity of Mind*.[5]

[5] Much of the psycholinguistic research of the 1980s was devoted to exploring modularity – that is, the idea that the human linguistic faculty consists of a number of distinct "informationally encapsulated" modules. While Fodor's book was mostly devoted to arguing for modularity at a higher level, where the linguistic faculty was one module, many researchers at the time extended the idea to the internal organization of the linguistic faculty, positing largely autonomous mechanisms for phonology, syntax, semantics, and pragmatics, with the operations of each of these sub-modules unaffected by the operations of the others. The outcome of years of experimental studies on the linguistic modularity idea was that it was abandoned by most psycholinguists. For an early direct response to Fodor, see Marslen-Wilson & Tyler (1987).

Some casual observations argue against modular language processing. For example, the famously ambiguous sentences (2a) and (2b) can be disambiguated in speech by the stress patterns.

(2) a. I forgot how good beer tastes.
b. Dogs must be carried.

The two meanings of (2a) correspond to two different parses (one with *good* as part of the noun phrase *good beer* and the other with *how good* as a verb phrase modifier). The two meanings of (2b) have the same syntactic structure, but differ in whether the requirement is that all dogs be carried, or that everyone carry a dog. This interaction of prosody with syntax (in the case of (2a)) and with semantics (in the case of (2b)) is produced and perceived before the end of the utterance, suggesting that phonological information is available in the course of syntactic and semantic processing.

Moreover, non-linguistic knowledge influences the disambiguation in both of these cases. If (2a) is preceded by "I just finished three weeks without alcohol", the natural interpretation of *good* is as a modifier of *tastes*; but following "I just finished three weeks drinking only Bud Light", *good* is more naturally interpreted as a modifier of *beer*. In the case of (2b), only one interpretation (that anyone with a dog must carry it) is plausible, given our knowledge of the world. Indeed, most non-linguists fail to see the ambiguity of (2b) without a lengthy explanation.

More rigorous evidence of the non-modular character of language processing has been provided by a variety of types of experiments. The work of Michael Tanenhaus and his associates, using eye-tracking to investigate the time-course of sentence comprehension, played an important role in convincing most psycholinguists that human language understanding is non-modular. See, for example, Eberhard et al. (1995), McMurray et al. (2008), Tanenhaus, Spivey-Knowlton, Eberhard, et al. (1995), Tanenhaus et al. (1996), and Tanenhaus & Trueswell (1995). A recent survey of work arguing against modularity in language processing is provided by Spevack et al. (2018).

2.3 Importance of words

The individual properties of words play a central role in how people process phrases and sentences. Consider, for example, what is probably the most famous sentence in psycholinguistics, (3), due originally to (Bever 1970).

(3) The horse raced past the barn fell.

The extreme difficulty that people who have not previously been exposed to (3) have comprehending it depends heavily on the choice of words. A sentence like (4), with the same syntactic structure, is far easier to parse.

(4) The applicant interviewed in the morning left.

Numerous studies (e.g. Ford et al. 1982; Trueswell et al. 1993; MacDonald et al. 1994; Bresnan et al. 2007; Wasow et al. 2011) have shown that such properties of individual words as subcategorization preferences, semantic categories (e.g. animacy), and frequency of use can influence the processing of utterances.

2.4 Influence of context

Much of the evidence against modularity of the language faculty is based on the influences of non-linguistic context and world knowledge on language processing. The well-known McGurk effect (McGurk & MacDonald 1976) and the Stroop effect (Stroop 1935) demonstrate that, even at the word level, visual context can influence linguistic comprehension and production.

Linguistic context also clearly influences processing, as the discussion of examples (2a) and (2b) above illustrates. The same conclusion is supported by numerous controlled studies, including, among many others, those described by Crain & Steedman (1985), Altmann & Steedman (1988), Branigan (2007), Traxler & Tooley (2007), Matsuki et al. (2011), and Spevack et al. (2018). The last of these references concludes (p. 11), "when humans and their brains are processing language with each other, there is no format of linguistic information (e.g., lexical, syntactic, semantic, and pragmatic) that cannot be rapidly influenced by context."

2.5 Speed and accuracy of processing

A good deal of psycholinguistic literature is devoted to exploring situations in which language processing encounters difficulties, notably work on garden paths (in comprehension) and disfluencies (in production). Much more striking than the existence of these phenomena, however, is how little they matter in everyday language use. While ambiguities abound in normal sentences (see Wasow 2015), comprehenders very rarely experience noticeable garden paths. Similarly, disfluencies in spontaneous speech occur in nearly every sentence but rarely disrupt communication.

People are able to use speech to exchange information remarkably efficiently. A successful account of human language processing must explain why it works as well as it does.

3 Features of HPSG that fit well with processing facts

In this section, I review some basic design features of HPSG, pointing out ways in which they comport well with the properties of language processing listed in the previous section.

3.1 Constraint-based

Well-formedness of HPSG representations is defined by the simultaneous satisfaction of a set of constraints that constitutes the grammar. This lack of directionality allows the same grammar to be used in modeling production and comprehension.

check for crossref-erenc-ing other chapters once they are in

Consider, for instance, the example of quirky case assignment illustrated in (1) above. A speaker uttering (1) would need to have planned to use the verb *helfen* before beginning to utter the object NP. But a listener hearing (1) would encounter the dative case on the article *dem* before hearing the verb and could infer only that a verb taking a dative object was likely to occur at the end of the clause. Hence, the partial mental representations built up by the two interlocutors during the course of the utterance would be quite different. But the grammatical mechanism licensing the combination of a dative object with this particular verb is the same for speaker and hearer.

In contrast, theories of grammar that utilize sequential operations to derive sentences impose a directionality on their grammars. If such a grammar is then to be employed as a component in a model of language use (as the competence hypothesis stipulates), its inherent directionality becomes part of the models of both production and comprehension. But production involves mapping meaning onto sound, whereas comprehension involves the reverse mapping. Hence, a directional grammar cannot fit the direction of processing for both production and comprehension.[6]

Branigan & Pickering (2017) argue at length that "structural priming provides an implicit method of investigating linguistic representations".[7] They go on to conclude (p. 14) that the evidence from priming supports "frameworks that …

[6] This was an issue for early work in computational linguistics that built parsers based on the transformational grammars of the time, which generated sentences using derivations whose direction went from an underlying structure largely motivated by semantic considerations to the observable surface structure. See, for example, Hobbs & Grishman (1975).

[7] Priming is the tendency for speakers to re-use linguistic elements that occurred earlier in the context; structural priming (which Branigan & Pickering sometimes call *abstract priming*) is priming of linguistic structures, abstracted from the particular lexical items in those structures.

assume nondirectional and constraint-based generative capacities (i.e., specifying well-formed structures) that do not involve movement”.[8] HPSG is one of the frameworks they mention that fit this description.

3.2 Surface-oriented

The features and values in HPSG representations are motivated by straightforwardly observable linguistic phenomena. HPSG does not posit derivations of observable properties from abstract underlying structures. In this sense it is surface-oriented.

The evidence linguists use in formulating grammars consists of certain types of performance data, primarily judgments of acceptability and meaning. Accounts of the data necessarily involve some combination of grammatical and processing mechanisms. The closer the grammatical descriptions are to the observable phenomena, the less complex the processing component of the account needs to be.

For example, the grammatical theory of Kayne (1994), which posits a universal underlying order of specifier-head-complement, requires elaborate (and directional) transformational derivations to relate these underlying structures to the observable data in languages whose surface order is different (a majority of the language of the world). In the absence of experimental evidence that the production and comprehension of sentences with different constituent orders involve mental operations corresponding to the grammatical derivations Kayne posits, his theory of grammar seems to be incompatible with the competence hypothesis.

Experimental evidence supports this reasoning. As Branigan & Pickering (2017) conclude (p. 9), “[P]riming evidence supports the existence of abstract syntactic representations. It also suggests that these are shallow and monostratal in a way that corresponds at least roughly to the assumptions of … Pollard & Sag (1994) …. It does not support a second, underlying level of syntactic structure or the syntactic representation of empty categories associated with the movement of constituents in some transformational analyses.”

3.3 Informationally rich representations

The feature structure descriptions of HPSG include all types of linguistic information relevant to the well-formedness and interpretation of expressions. This

[8] Branigan & Pickering’s conclusions are controversial, as is evident from the commentaries accompanying their target article.

includes phonological, morphological, syntactic, semantic, and contextual information. They can also incorporate non-linguistic contextual information (e.g. social information), though this has not been extensively explored.

The local cooccurrence of these different types of information within a single representation facilitates modeling production and comprehension processes that make reference to more than one of them. The architecture of the grammar is thus well suited to the non-modularity and context-sensitivity of language processing. It is interesting in this regard to consider the conclusions of two papers by psycholinguists who surveyed experimental evidence and inferred what types of grammatical information was essential for processing.

The following series of quotes captures the essence of what MacDonald et al. (1994) wrote regarding lexical representation, based on a survey of a wide range of psycholinguistic studies:

- "[T]he lexical representation for a word includes a representation of the word's phonological form, orthographic form, semantics, grammatical features (including grammatical category), morphology (at least inflectional), argument structure, and X-bar structure." (p. 684)

- "[T]he connection structure of the lexicon encodes relationships among different types of lexical information."[9] (p. 684)

- "In addition to constraints that hold between various aspects of lexical representations, sentence and discourse contexts also constrain lexical representations during processing…" (p. 686)

With the possible exception of "X-bar structure", this sounds very much like a description of the types of information included in HPSG feature structure descriptions.

Over twenty years later, Branigan & Pickering (2017) came to the following conclusions about linguistic representations, based on priming studies:

- "The syntactic representations capture local relationships between a 'mother' and its constituent 'daughter(s)' (e.g., a VP comprising a verb and two NPs), independent of the larger context in which the phrase appears (e.g., that the VP occurs within a subordinate clause), or the internal structure of the

[9] A reviewer asked what feature of HPSG this maps into. The answer is straightforward: a word's phonological form, semantics, grammatical features, morphology, and argument structure are all represented together in one feature structure description, and the different pieces of the description may be linked through coindexing or tagging.

subphrases that constitute it (e.g., that the first NP comprises a determiner, adjective, and noun)." (p. 9)

- "[S]ome elements that are not phonologically represented may be syntactically represented." (p. 10)

- "Other priming evidence similarly indicates that some semantically specified elements are not specified syntactically." (p. 11)

- "[T]he semantic level of representation contains at least specifications of quantificational information, information structure, and thematic roles." (p. 11)

- "Evidence from priming supports a range of mappings between information encoded in the semantic representation and information encoded in the syntactic representation: between thematic roles and grammatical functions, between thematic roles and word order, between animacy and syntactic structure, and between event structures and syntactic structures." (p. 12)

The two lists are quite different. This is in part because the focus of the earlier paper was on lexical representations, whereas the later paper was on linguistic representations more generally. It may also be attributable to the fact that MacDonald et al. framed their paper around the issue of ambiguity resolution, while Branigan & Pickering's paper concentrated on what could be learned from structural priming studies. Despite these differences, it is striking that the conclusions of both papers about the mental representations employed in language processing are very much like those arrived at by work in HPSG.

3.4 Lexicalism

A great deal of the information used in licensing sentences in HPSG is stored in the lexical entries for words. A hierarchy of lexical types permits commonalities to be factored out to minimize what has to be stipulated in individual entries, but the information in the types gets into the representations of phrases and sentences through the words that instantiate those types. Hence, it is largely the information coming from the words that determines the well-formedness of larger expressions. Any lexical decomposition would have to be strongly motivated by the morphology.

Branigan & Pickering (2017) note that grammatical structures (what some might call *constructions*) such as V-NP-NP can prime the use of the same abstract structure, even in the absence of lexical overlap. But they also note that the priming is consistently significantly stronger when the two instances share the same verb, a fact known as *the lexical boost*. They write, "To explain abstract priming, lexicalist theories must assume that the syntactic representations […] are shared across lexical entries." The types in HPSG's lexicon provide just such representations. Branigan & Pickering go on to say that the lexical boost argues for "a representation that encodes a binding between constituent structure and the lemma … of the lexical entry for the head." In HPSG, this "binding" is simply the fact that the word providing the lexical boost (say, *give*) is an instantiation of a type specifying the structures it appears in (e.g. the ditransitive verb type).

Similarly, the fact, noted in Section 2.3 above, that a given structure may be more or less difficult to process depending on word choice is unsurprising in HPSG, so long as the processor has access to information about individual words and not just their types.

3.5 Underspecification

HPSG allows a class of linguistic structures that share some feature values to be characterized by means of feature structure descriptions that specify only the features whose values are shared. Such underspecification is very useful for a model of processing (particularly a model of the comprehender) because it allows partial descriptions of the utterance to be built up, based on the information that has been encountered. This property of the grammar makes it easy to incorporate into an incremental processing model.

4 Two phenomena of interest

4.1 Island constraints

Ever since Ross's seminal dissertation (1967) introduced the notion of "island constraints", linguists have sought explanations for their existence, often suggesting that they were motivated by processing considerations (notably Grosu 1972; Fodor 1983a; Deane 1991). The basic idea is that island constraints restrict the search space the parser needs to consider in looking for a gap to match a filler it has encountered, thereby facilitating processing. This then raises the question of whether island constraints need to be represented in grammar (language particular or universal), or can be attributed entirely to processing and/or other factors, such as pragmatics.

In principle, this question is orthogonal to the choice among theories of grammar. But in recent years, a controversy has arisen between some proponents of HPSG and certain transformational grammarians, with the former (e.g. Chaves 2012 and 2019, Chapter 16 of this volume; Hofmeister & Sag 2010; Hofmeister, Jaeger, Arnon, Sag & Snider 2013) arguing that certain island phenomena should be attributed entirely to extra-grammatical factors, and the latter (e.g. Phillips 2013 and Sprouse et al. 2012) arguing that island constraints are part of grammar.

I will not try to settle this dispute here. Rather, my point in this subsection is to note that a theory in which there is a close fit between the grammar and processing mechanisms allows for the possibility that some island phenomena should be attributed to grammatical constraints, whereas others should be explained in terms of processing. Indeed, if the basic idea that islands facilitate processing is correct, it is possible that some languages, but not others, have grammaticalized some islands, but not others. That is, in a theory in which the grammar is a tightly integrated component of a processing model, the question of whether a particular island phenomenon is due to a grammatical constraint is an empirical one whose answer might differ from language to language.

Early work on islands (e.g. Ross 1967 and Chomsky 1973) assumed that, in the absence of negative evidence, island constraints could not be learned and hence must be innate and therefore universal. But cross-linguistic variation in island constraints, even between closely related languages, has been noted since the early days of research on the topic (see e.g. Erteschik-Shir 1973 and Engdahl & Ejerhed 1982).

This situation is what one might expect if languages differ with respect to the extent to which the processing factors that motivate islandhood have been grammaticalized. In short, a theory with a tight fit between its grammatical machinery and its processing mechanisms allows for hybrid accounts of islands that are not available to theories without such a fit.

One example of such a hybrid is Chaves's (2012) account of Ross's Coordinate Structure Constraint. Following much earlier work, Chaves distinguishes between the "conjunct constraint", which prohibits a gap from serving as a conjunct in a coordinate structure (as in **What did you eat a sandwich and?*) and the "element constraint", which prohibits a gap from serving as an element of a larger conjunct (as in **What did you eat a sandwich and a slice of?*). The conjunct constraint, he argues, follows from the architecture of HPSG and is therefore built into the grammar. The element constraint, on the other hand, has exceptions and, he claims, should be attributed to extra-grammatical factors. See Chaves (2019), Chapter 16 of this volume for a more detailed discussion of islands.

4.2 Subject vs. object relative clauses

One of the most discussed phenomena in the literature on human sentence processing is the difference in processing complexity between relative clauses (RCs) in which the gap is the subject and those in which the gap is the object – or, as they are commonly called, "subject RCs" and "object RCs"; see, among many others, Wanner & Maratsos (1978), Gibson (1998), Traxler et al. (2002), and Gennari & MacDonald (2008). Relative clause processing complexity has been shown to be influenced by a number of factors other than the grammatical function of the gap, including the animacy and pronominality of the overt NP in the RC, as well as the frequency, animacy, and discourse properties of the head of the RC.[10] When these factors are controlled for, however, most psycholinguists accept that it has been established that subject RCs are generally easier to process than object RCs, at least in English.[11]

One approach to explaining this asymmetry has been based on the distance between the filler and the gap (see, among others, Wanner & Maratsos 1978; Gibson 1998; Hawkins 2004). In languages like English, with basic SVO clause order and RCs that follow the nouns they modify, the distance between the filler (the relativizer or head noun) and the gap is greater for an object gap than for a subject gap. If holding a filler in memory until the gap is encountered puts an extra burden on the processor, this could explain why object RCs are harder to process than subject RCs. This distance-based account makes an interesting prediction

[10] The stimuli in the experimental studies on this topic always have RCs with one overt NP, either in subject or object position and a gap corresponding to the other grammatical function. In most of the studies, that NP is non-pronominal and animate. See Reali & Christiansen (2007) and Roland et al. (2012) for evidence of the role of these factors in processing complexity.

[11] This processing difference corresponds to the top end of the "accessibility hierarchy" that Keenan & Comrie (1977) proposed as a linguistic universal. Based on a diverse sample of 50 languages, they proposed the hierarchy below, and hypothesized that any language allowing RC gaps at any point in the hierarchy would allow RC gaps at all points higher (to the left) on the hierarchy.

Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparison

Keenan & Comrie speculated that the generality of this hierarchy of relativizability lay in processing, specifically on the comprehension side. The extensive experimental evidence that has been adduced in support of this idea in the intervening decades has been concentrated on subject RCs vs. (direct) object RCs. The remainder of the hierarchy remains largely untested by psycholinguists.

for languages with different word orders. In languages like Japanese with SOV order and RCs that precede the nouns they modify, the distance relationships are reversed – that is, the gaps in object RCs are closer to their fillers than those in subject RCs. The same is true of Chinese, with basic SVO order and RCs that precede the nouns they modify. So the prediction of distance-based accounts of the subject/object RC processing asymmetry is that it should be reversed in these languages.

The experimental evidence on this prediction is somewhat equivocal. While Hsiao & Gibson (2003) found a processing preference for object RCs over subject RCs in Chinese, their findings were challenged by Lin & Bever (2006) and Vasishth et al. (2013), who claimed that Chinese has a processing preference for subject RCs. In Japanese, Miyamoto & Nakamura (2003) found that subject RCs were processed more easily than object RCs. The issue remains controversial, but, for the most part, the evidence has not supported the idea that the processing preference between subject RCs and object RCs varies across languages with different word orders.

The most comprehensive treatment of English RCs in HPSG is Sag (1997). Based entirely on distributional evidence, Sag's analysis treats (finite) subject RCs as fundamentally different from RCs whose gap does not function as the subject of the RC. The difference is that the SLASH feature, which encodes information about long-distance dependencies in HPSG, plays no role in the analysis of subject RCs. Non-subject RCs, on the other hand, involve a non-empty SLASH value in the RC.[12]

Sag deals with a wide variety of kinds of RCs. From the perspective of the processing literature, the two crucial kinds are exemplified by (5a) and (5b), from Gibson (1998).

(5) a. The reporter who attacked the Senator admitted the error.
    b. The reporter who the Senator attacked admitted the error.

A well-controlled experiment on the processing complexity of subject and object RCs must have stimuli that are matched in every respect except the role of the gap in the RC. Thus, the conclusion that object RCs are harder to process than subject RCs is based on a wide variety of studies using stimuli like (5). Sag's analysis of (5a) posits an empty SLASH value in the RC, whereas his analysis of (5b) posits a non-empty SLASH value.

[12] The idea that at least some subject gaps differ in this fundamental way from non-subject gaps goes back to Gazdar (1981).

There is considerable experimental evidence supporting the idea that unbounded dependencies – that is, what HPSG encodes with the SLASH feature – add to processing complexity; see, for example, Wanner & Maratsos (1978), King & Just (1991), Kluender & Kutas (1993), and Hawkins (1999). Combined with Sag's HPSG analysis of English RCs, this provides an explanation of the processing preference of subject RCs over object RCs. On such an account, the question of which other languages will exhibit the same preference boils down to the question of which other languages have the same difference in the grammar of subject and object RCs. At least for English, this is a particularly clear case in which the architecture of HPSG fits well with processing evidence.

5 Conclusion

This chapter opened with the observation that HPSG has not served as the theoretical framework for much psycholinguistic research. The observations in Sections 2 through 4 argue for rectifying that situation. The fit between the architecture of HPSG and what is known about human sentence processing suggests that HPSG could be used to make processing predictions that could be tested in the lab.

To take one example, the explanation of the processing asymmetry between subject and object RCs offered above is based on a grammatical difference in the HPSG analysis: all else being equal, expressions with non-empty SLASH values are harder to process than those with empty SLASH values. Psycholinguists could test this idea by looking for other cases of phenomena that look superficially very similar but whose HPSG analyses differ with respect to whether SLASH is empty. One such case occurs with pairs like Chomsky's famous minimal pair in (6).

(6) a. Chris is eager to please.
    b. Chris is easy to please.

Under the analysis of Pollard & Sag (1994), *to please* in (6b) has a non-empty SLASH value but an empty SLASH value in (6a). Processing (6a) should therefore be easier. This prediction could be tested experimentally, and modern methods such as eye-tracking could pinpoint the locus of any difference in processing complexity to determine whether it corresponds to the region where the grammatical analysis involves a difference in SLASH values.

The current disconnect between theoretical investigations of language structure and psycholinguistic studies is an unfortunate feature of our discipline. Because HPSG comports so well with what is known about processing, it could serve as the basis for a reconnection between these two areas of study.

Acknowledgements

A number of people made valuable suggestions on a preliminary outline and an earlier draft of this chapter, leading to improvements in content and presentation, as well as the inclusion of previously overlooked references. In particular, I received helpful comments from (in alphabetical order): Emily Bender, Bob Borsley, Rui Chaves, Danièle Godard, Jean-Pierre Koenig, and Stefan Müller. Grateful as I am for their advice, I take sole responsibility for any shortcomings in the chapter.

References

Altmann, Gerry & Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent refrence. *Cognition* 73(7). 247–264.

Altmann, Gerry & Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30(3). 191–238.

Arnold, Jennifer E., Carla Hudson Kam & Michael K. Tanenhaus. 2007. If you say thee uh- you're describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33. 914–930.

Bever, Thomas. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and the language development*, 279–362. New York: Wiley & Sons.

Branigan, Holly. 2007. Syntactic priming. *Language and Linguistics Compass* 1(1–2). 1–16.

Branigan, Holly & Martin Pickering. 2017. An experimental approach to linguistic representation. *Behavioral and Brain Sciences* 40. 1–61.

Bresnan, Joan, Anna Cueni, Tatjana Nikitina & Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: KNAW.

Bresnan, Joan, Shipra Dingare & Christopher D. Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 01 Conference*. Standford, CA: CSLI.

Chaves, Rui. 2019. Island phenomena and related matters. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 415–468. Berlin: Language Science Press. DOI:??

Chaves, Rui P. 2012. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30(2). 465–512.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart & Winston.

Crain, Stephen & Mark J. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language processing*, 320–358. Cambridge, UK: Cambridge University Press.

Deane, Paul D. 1991. Limits to attention: A cognitive theory. *Cognitive Linguistics* 2(1). 1–63.

Eberhard, Kathleen M., Michael J. Spivey-Knowlton, Julie C. Sedivy & Michael K. Tannenhaus. 1995. Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguist Research* 24. 409–436.

Engdahl, Elisabeth & Eva Ejerhed. 1982. *Readings on unbounded dependencies in Scandinavian languages*. Stockholm: Almqvist & Wiksell International.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT dissertation.

Flickinger, Dan, Carl Pollard & Tom Wasow. 2019. The evolution of HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 7–43. Berlin: Language Science Press. DOI:??

Fodor, Janet Dean. 1983a. Phrase structure parsing and the island constraints. *Linguisitcs and Philosophy* 6. 163–223.

Fodor, Jerry A. 1983b. *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The psychology of language: An introduction to psycholinguistics and Generative Grammar*. New York: McGraw-Hill Book Co.

Ford, Marilyn, Joan Bresnan & Ronald Kaplan. 1982. A competence-based theory of syntactic closure. In Joan Bresnan (ed.), *The mental representation of grammatical relations*, 727–796. Cambridge, MA: MIT Press.

Francis, Elaine. 2019. *Gradient acceptability and linguistic theory*. Ms. Purdue University, Under review.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.

Gennari, Silvia P. & Maryellen C. MacDonald. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 8. 161–187.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

Gollan, Tamar H., Timothy J. Slattery, Diane Goldenberg, Eva Van Assche, Wouter Duyck & Keith Rayner. 2011. Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General* 140(2). 186–209.

Grosu, A. 1972. *The strategic content of island constraints* (Working Papers in Linguistics 13). Ohio Sate University.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2). 244–285.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.

Hobbs, Jerry R. & Ralph Grishman. 1975. The automatic transformational analysis of English sentences: An implementation. *International Journal of Computer Mathematics* 5(1–4). 267–283.

Hofmeister, Philip, T. Florian Jaeger, Inbal Arnon, Ivan A. Sag & Neal Snider. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28(1–2). 48–87.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86(2).

Hsiao, Franny & Edward Gibson. 2003. Processing relative clauses in chinese. *Cognition* 90. 3–27.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

King, Jonathan & Marcel Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30. 580–602.

Kluender, Robert & Marta Kutas. 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5(2). 196–214.

Konieczny, Lars. 1996. *Human sentence processing: A semantics-oriented parsing approach*. Universität Freiburg Dissertation. IIG-Berichte 3/96.

Lin, Chien-Jer Charles & Thomas Bever. 2006. Subject preference in the processing of relative clauses in Chinese. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 254–260. Somerville, MA: Cascadilla Proceedings Project.

Linadarki, Evita. 2006. *Linguistic and statistical extensions of data oriented parsing*. University of Essex dissertation. Unpublished.

MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4(226). 1–16.

MacDonald, Maryellen C., Neal J. Pearlmutter & Mark. S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4). 676–703.

Marslen-Wilson, William & Lorraine Tyler. 1987. Against modularity. In Jay L. Garfield (ed.), *Modularity in knowledge representation and Natural Language Processing*, 37–62. Cambridge, MA: MIT Press.

Matsuki, Kazunaga, Tracy Chow, Mary Hare, Christoph Elman Jeffrey L. Scheepers & Ken McRae. 2011. Event-based plausibility immediately influences online language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(4). 913–934.

McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264(5588). 746–748.

McMurray, Bob, Meghan A. Clayards, Michael K. Tanenhaus & Richard Aslin. 2008. Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin* 15(6). 1064–1071.

Miyamoto, Edson T & Michiko Nakamura. 2003. *Subject/object asymmetries in the processing of relative clauses in Japanese*. Gina Garding & Mimu Tsujimura (eds.). Somerville, MA.

Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.

Momma, Shota & Colin Phillips. 2018. The relationship between parsing and generation. *Annual Review of Linguistics* 4. 233–254.

Phillips, Colin. 2013. On the nature of island constraints I: Language processing ad reductionist accounts. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and island effects*, 64–108. Cambridge: Cambridge University Press.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Reali, Florencia & Morten H. Christiansen. 2007. Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language* 57(1). 1–23.

Roland, Douglas, Gail Mauner, Carolyn O'Meara & Hongoak Yun. 2012. Discourse expectations and relative clause processing. *Joural of Memory and Language* 66(3). 470–508.

Ross, John R. 1967. *Constraints on variables in syntax*. MIT, Cambridge, Massachusetts. [Published in 1986 as *Infinite Syntax!* Norwood, NJ: Ablex Publishing] Ph.D. Dissertation.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Sag, Ivan A. & Thomas Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguistic Research* 44(1). 47–63.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.

Spevack, Samuel C., J. Benjamin Falandays & Michael J. Spivey. 2018. Interactivity of language. *Language and Linguistics Compass* 12(7). e12282.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012. A test of the relation between working memory capacity and syntactic island effects. *Language* 88(1). 82–123.

Stroop, John Ridley. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18(6). 643–662.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634. DOI:10.1126/science.7777863

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1996. Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In Toshio Inui & James L. McClelland (eds.), *Information integration in perception and communication* (Attention and Performance XVI), 457–478. Cambridge, MA: MIT Press.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton & Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634.

Tanenhaus, Michael K. & John C. Trueswell. 1995. Sentence comprehension. In Joanne L. Miller & Peter D. Eimas (eds.), *Handbook of cognition and perception*. San Diego, CA: Academic Press.

Traxler, Matthew J., Robin K. Morris & Rachel E. Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47(1). 69–90.

Traxler, Matthew J. & Kristen M. Tooley. 2007. Lexical mediation and context effects in sentence processing. *Brain Research* 1146. 59–74.

Trueswell, John C., Micahel K. Tanenhaus & Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(3). 528–553.

Vasishth, Shravan, Chen Zhong, Qiang Li & Gueilan Guo. 2013. Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS ONE* 8(10). e77006.

Wanner, Eric & Michael Maratsos. 1978. An ATN approach to comprehension. In Morris Halle, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, 119–161. Cambridge, MA: MIT Press.

Wasow, Thomas. 2015. Ambiguity avoidance is overrated. In Susanne Winkler (ed.), *Ambiguity: Language and communication*, 29–47. Berlin: De Gruyter.

Wasow, Thomas, Florian T. Jaeger & David Orr. 2011. Lexical variation in relativizer frequency. In Horst J. Simon & Heike Wiese (eds.), *Expecting the unexpected: Exceptions in grammar*, 175–195. Berlin: De Gruyter.

Chapter H

Computational linguistics and grammar engineering

Emily M. Bender

University of Washington

Guy Emerson

University of Cambridge

We discuss the relevance of HPSG for computational linguistics, and the relevance of computational linguistics for HPSG.

1 Introduction

From the inception of HPSG in the 1980s, there has been a close integration between theoretical and computational work (for an overview, see **chapters/chap-evolution** Chapter **??** of this volume). In this chapter, we discuss computational work in HPSG, starting with the infrastructure that supports it (both theoretical and practical) in Section 2. Next we describe several existing large-scale projects which build HPSG or HPSG-inspired grammars (see Section 3) and the deployment of such grammars in applications including both those within linguistic research and otherwise (see Section 4). Finally, we turn to linguistic insights gleaned from broad-coverage grammar development (see Section 5).

2 Infrastructure

2.1 Theoretical considerations

There are several properties of HPSG as a theory that make it well-suited to computational implementation. First, the theory is kept separate from the formalism:

Emily M. Bender & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 797–840. Berlin: Language Science Press. DOI:??

the formalism is expressive enough to encode a wide variety of possible theories. While some theoretical work does argue for or against the necessity of particular formal devices (e.g., the shuffle operator (Reape 1994)), much of it proceeds within shared assumptions about the formalism. This is in contrast to work in the context of the Minimalist Program (Chomsky 1995), where theoretical results are typically couched in terms of modifications to the formalism itself. From a computational point of view, the benefit of differentiating between theory and formalism is that the formalism is relatively stable. That enables the development and maintenance of software systems that target the formalism (Boguraev et al. 1988), such as software for parsing, generation, and grammar exploration (see Section 3 below for some examples).[1]

A second important property of HPSG that supports a strong connection between theoretical and computational work is an interest in both so-called "core" and so-called "peripheral" phenomena. Most implemented grammars are built with the goal of handling naturally occurring text.[2] This means that they will need to handle a wide variety of linguistic phenomena not always treated in theoretical syntactic work (Baldwin et al. 2005). A syntactic framework that excludes research on "peripheral" phenomena as uninteresting provides less support for implementational work than does one, like HPSG or Construction Grammar, that values such topics (for a comparison of HPSG and Construction Grammar, see Müller 2019c, Chapter P of this volume).

Finally, the type hierarchy characteristic of HPSG lends itself well to developing broad-coverage grammars which are maintainable over time (see Sygal & Wintner 2011). The use of the type hierarchy to manage complexity at scale comes out of the work of Flickinger (1987) and others at in the project where HPSG was originally developed. The core idea is that any given constraint is (ideally) expressed only once on types which serve as supertypes to all entities that bear that constraint.[3] Such constraints might represent broad generalizations that apply to many entities or relatively narrow, idiosyncratic properties that apply to

[1] There are implementations of Minimalism, notably Stabler 1997 and Herring 2016. (See also Torr et al. 2019 for a recent broad-coverage, treebank-trained parser in this framework.) However, implementing a theory requires fixing the formalism, and so these implementations are unlikely to be useful for testing theoretical ideas as the theory moves on.

[2] It is possible, but less common, to do implementation work strictly against test suites of sentences constructed specifically to focus on phenomena of interest.

[3] Originally this only applied to lexical entries in Flickinger's work. Now it also applies to phrase structure rules, lexical rules, and types below the level of the sign which are used in the definition of all of these. See **chapters/chap-evolution** Chapter ?? of this volume for further discussion.

only a few. By isolating any given constraint on one type (as opposed to repeating it in multiple places), we build grammars that are easier to update and adapt in light of new data that require refinements to constraints. Having a single locus for each constraint also makes the types a very useful target for documentation (Hashimoto et al. 2008) and grammar exploration (Letcher 2018).

2.2 Practical considerations

HPSG allows practical implementations because it uses a well-defined formalism. Furthermore, because HPSG is defined to be bi-directional, an implemented grammar can be used for both parsing and generation. In this section, we discuss how HPSG allows tractable algorithms, which enables linguists to empirically test hypotheses and which also enables HPSG grammars to be used in a range of applications, as we will see in Sections 4.1 and 4.2, respectively.

2.2.1 Computational complexity

One way to measure how easy or difficult it is to use a syntactic theory in practical computational applications is to consider the *computational complexity*[4] of parsing and generation algorithms (Gazdar & Pullum 1985). Computational complexity includes both how much memory and how much computational time a parsing algorithm needs to process a particular sentence.[5] Considering parsing time, longer sentences will take longer to process, but the more complex the algorithm is, the more quickly the amount of processing time increases. Parsing complexity can thus be measured by considering sentences containing n tokens, and then increasing n to see how the amount of time changes. This can be done based on the average amount of time for sentences in a corpus (average-case complexity), or based on the longest amount of time for all theoretically possible sentences (worst-case complexity).

At first sight, analyzing computational complexity would seem to paint HPSG in a bad light, because the formalism allows us to write grammars which can be arbitrarily complex; in technical terminology, the formalism is *Turing-complete*

[4] Computational complexity is related to the complexity hierarchy of language classes in formal language theory. More complex language classes tend to require parsing and generation algorithms with higher computational complexity, but this relationship is not exact. For example, the class of strictly local languages is a proper subset of the class of regular languages, but both classes can be parsed in linear time. Müller (2019b: ch. 17) discusses HPSG from the point of view of formal language theory.

[5] In this section, we only consider parsing algorithms, but a similar analysis can be done for generation (e.g., Carroll et al. 1999).

(Johnson 1988: Section 3.4). However, as discussed in the previous section, there is a clear distinction between theory and formalism. Although the HPSG formalism rules out the possibility of efficient algorithms that could cope with any possible feature-structure grammar, a particular theory (or a particular grammar) might well allow efficient algorithms.

Keeping processing complexity manageable is handled differently in other computationally-friendly frameworks, such as Combinatory Categorial Grammar (CCG),[6] or Tree Adjoining Grammar (TAG; Joshi 1987; Schabes et al. 1988). The formalisms of CCG and TAG inherently limit computational complexity: for both of them, as the sentence length n increases, worst-case parsing time is proportional to n^6 (Kasami et al. 1989). This is a deliberate feature of these formalisms, which aim to be just expressive enough to capture human language, and not any more expressive. Building this kind of constraint into the formalism itself highlights a different school of thought from HPSG. Indeed, Müller (2015: 64) explicitly argues in favor of developing linguistic analyses first, and improving processing efficiency second. As discussed above in Section 2.1, separating the formalism from the theory means that the formalism is stable, even as the theory develops.

It would be beyond the scope of this chapter to give a full review of parsing algorithms, but it is instructive to give an example. For grammars that have a context-free backbone (every analysis can be expressed as a phrase-structure tree plus constraints between mother and daughter nodes), it is possible to adapt the standard *chart-parsing* algorithm Kay (1973) for context-free grammars. The basic idea is to parse "bottom-up", starting by finding analyses for each token in the input, and then finding analyses for increasingly longer sequences of tokens (called *spans*, until the parser reaches the entire sentence.

For a context-free grammar, there is a finite number of nonterminal symbols, and each span is analyzed as a subset of the nonterminals. For a feature-structure grammar, each span must be analyzed as a set of feature structures,[7] which makes the algorithm more complicated. In principle, a grammar may allow an infinite number of possible feature structures, for example if it includes recursive unary

[6] For an introduction, see Steedman & Baldridge (2011). For a comparison with HPSG, see Kubota 2019, Chapter M of this volume.

[7] Much theoretical work in HPSG, including Pollard & Sag (1994), distinguishes between fully resolved feature structures and possibly underspecified feature structure descriptions. Much computational work, by contrast, operates entirely with partially specified feature structures, at both the level of grammar and the level of analyses licensed by the grammar. In keeping with this tradition, we use the term "feature structure" to refer to both fully specified and partially specified objects, and have no need for the term "feature structure description".

rules. However, if we can bound the number of possible feature structures as C, then the worst-case parsing time is proportional to $C^2 n^{\rho+1}$, where ρ is the maximum number of children in a phrase-structure rule (Carroll 1993: Section 3.2.3). This is less complex than for an arbitrary grammar (which means that this class of grammars is *not* Turing-complete), but C may nonetheless be very large.

But is the number of possible feature structures bounded in implemented HPSG grammars? For DELPH-IN grammars (see Section 3.2), the answer is yes. Assuming a system without relational constraints, the potential for unboundedness in the number of feature structures stems from the potential for recursion in feature paths: A listslist is a simple example,[8] and as another example, the elements on a COMPS list also include the feature COMPS.

However, in practice, such recursive paths do not need to be considered by the parsing algorithm. For example, selecting heads might place constraints on their complements' subjects (e.g., in raising/control constructions),raisingcontrol but no further than that (e.g., a complement's complement's subject). Similarly, while lists that are potentially unbounded in length are used in semantic representations, these are never involved in constraining grammaticality. The only lists that constrain grammaticality are valence lists, but in practical grammars these are never greater than length four or five.[9]

When parsing real corpora, it turns out that the average-case complexitycomplexity!average-case is much better than might be expected (Carroll 1994). On the one hand, grammatical constructions do not generally combine in the worst-case way, and on the other hand, when a grammar writer is confronted with multiple possible analyses for a particular construction, they may opt for the analysis that is more efficient for a particular parsing algorithm (Flickinger 2000). To measure the efficiency of grammars and parsing algorithms in practice, it can be helpful to use a test suite composed of a representative sample of sentences (Oepen & Flickinger 1998).

[8] More precisely, in the standard implementation of a list as a feature structure, the type *list* has two subtypes *null* and *non-empty-list*, and *non-empty-list* has the features FIRST and REST, where the value of REST is of type *list*. The value of REST can itself have the feature REST.

[9] In part, this is because DELPH-IN does not adopt proposals like the DEPS list of Bouma, Malouf & Sag (2001). Furthermore, in many DELPH-IN grammars, including the ERG, the SLASH list cannot have more than one element. If an unbounded SLASH list is required (such as to model cross-serial dependencies), the number of possible structures might still be bounded as a function of sentence length; this would allow us to bound worst-case parsing complexity, but it will be a higher bound.

2.2.2 Parse ranking

Various kinds of ambiguity are well-known in linguistics (such as modifier attachment and part-of-speech assignment), to the point that examples like (1) are stock in trade:

(1) a. I saw the kid with the telescope.
    b. Visiting relatives can be annoying.

A well-constructed grammar should be expected to return multiple parses for each ambiguous sentence.

However, people are naturally very good at resolving ambiguity, which means most ambiguity is not apparent, even to linguists. It is only with the development of large-scale grammars that the sheer scale of ambiguity has become clear. For example, (2) might seem unambiguous, but there is a second reading, where *my favorite* is the topicalized object of *speak*, which would mean that town criers generally speak the speaker's favorite thing (perhaps a language) clearly. There is also a third, even more implausible reading, where *my favorite town* is the topicalized object. Such implausible readings don't easily come to mind, and in fact, the 2018 version of the English Resource Grammar (ERG; Flickinger 2000; 2011) gives a total of 21 readings for this sentence. With increasingly long sentences, such ambiguities stack up very quickly. For (3), the first line of a newspaper article,[10] the ERG gives 35,094 readings.

(2) My favorite town criers speak clearly.

(3) A small piece of bone found in a cave in Siberia has been identified as the remnant of a child whose mother was a Neanderthal and father was a Denisovan, a mysterious human ancestor that lived in the region.

While exploring ambiguity can be interesting for a linguist, typical practical applications require just one parse per input sentence and specifically the parse that best reflects the intended meaning (or only the top few parses, in case the one put forward as "best" might be wrong). Thus, what is required is a *ranking* of the parses, so that the application can only use the most highly-ranked parse, or the top *N* parses.

[10] https://www.theguardian.com/science/2018/aug/22/offspring-of-neanderthal-and-denisovan-identified-for-first-time, accessed 16 August 2019

Parse ranking is not usually determined by the grammar itself, because of the difficulty of manually writing disambiguation rules.[11] Typically, a statistical system is used (Toutanova et al. 2002; 2005). First, a corpus is *treebanked*: for each sentence in the corpus, an annotator (often the grammar writer) chooses the best parse, out of all parses produced by the grammar. The set of all parses for a sentence is often referred to as the *parse forest*, and the selected best parse is often referred to as the *gold standard* or *gold parse*. Given the gold parses for the whole corpus, a statistical system is trained to predict the gold parse from a parse forest, based on many features[12] of the parse. From the example in (2), a number of different features all influence the preferred interpretation: the likelihood of a construction (such as topicalization), the likelihood of a valence frame (such as transitive *speak*), the likelihood of a collocation (such as *town crier*), the likelihood of a semantic relation (such as speaking a town), and so on.

Because of the large number of possible parses, it can be helpful to *prune* the search space: rather than ranking the full set of parses, ranking is restricted to a smaller set of parses. Carefully choosing how to restrict the parser's attention can drastically reduce processing time without hurting parsing accuracy, as long as the algorithm for selecting the subset includes the correct parse sufficiently frequently. One method, called *supertagging*,[13] exploits the fact that HPSG is a lexicalized theory: choosing the correct lexical entry brings in rich information that can be exploited to rule out many possible parses. Thus if the correct lexical entry can be chosen prior to parsing (e.g., on the basis of the prior and following words), the range of possible analyses the parser must consider is drastically reduced. Although there is a chance that the supertagger will predict the wrong lexical entry, using a supertagger can often improve parsing accuracy, by ruling out parses that the parse-ranking model might incorrectly rank too high. Supertagging was first applied to HPSG by Matsuzaki et al. (2007), building on previous work for TAG (Bangalore & Joshi 1999) and CCG (Clark & Curran 2004). To allow multi-word expressions (such as *by and large*), where the grammar assigns

[11] In fact, in earlier work, this task was undertaken by hand. One of the authors (Bender) had the job of maintaining rule weights in addition to developing the Jacy grammar (Siegel, Bender & Bond 2016) at YY Technologies in 2001–2002. No systematic methodology for determining appropriate weights was available and the system was both extremely brittle (sensitive to any changes in the grammar) and next to impossible to maintain.

[12] In the machine-learning sense of "feature", not the feature-structure sense.

[13] The term *supertagging*, due to Bangalore & Joshi (1999), refers to *part-of-speech tagging*, which predicts a part-of-speech for each input token, from a relatively small set of part-of-speech tags. Supertagging is "super", in that it predicts detailed lexical entries, rather than simple parts of speech.

a single lexical entry to multiple tokens, Dridan (2013) proposes an extension of supertagging, called *ubertagging*, which jointly predicts both a segmentation of the input and supertags for those segments. Dridan manages to increase parsing speed by a factor of four, while also improving parsing accuracy.

Finally, in order to train these statistical systems, we need to first annotate a treebank. When there are many parses for a sentence, it can be time-consuming to select the best parse. To efficiently use an annotator's time, it can be helpful to use *discriminants*, properties which hold for some parses but not for others (Carter 1997). For example, discriminants might include whether to analyze an ambiguous token as a noun or a verb, or where to attach a prepositional phrase. This approach to treebanking also means that annotations can be re-used when the grammar is updated (Oepen et al. 2004; Flickinger et al. 2017). For more on treebanking, see Section 4.1.4.

2.2.3 Semantic dependencies

In practical applications of HPSG grammars, the full phrase-structure trees and the full feature structures are often unwieldy, containing far more information than necessary for the task at hand. It is therefore often desirable to extract a concise semantic representation.

In computational linguistics, a popular approach to semantics is to represent the meaning of a sentence as a *dependency graph*, as this enables the use of graph-based algorithms.[14] Several types of dependency graph have been proposed based on Minimal Recursion Semantics (MRS; Copestake et al. 2005), with varying levels of simplification. Oepen & Lønning (2006) observe that if every predicate has a unique *intrinsic argument*, an MRS can be converted to a variable-free semantic representation, by replacing each reference to a variable with a reference to the corresponding predicate. They present Elementary Dependency Structures (EDS), semantic graphs which maintain predicate-argument structure but discard some scope information. (For many applications, scope information is less important than predicate-argument structure.) Copestake (2009) builds on this idea to create a more expressive graph-based representation called Dependency Minimal Recursion Semantics (DMRS), which is fully interconvertible with MRS.[15] This expressivity is achieved by adding annotations on the edges to

[14]In this section, we are concerned with *semantic* dependencies. For *syntactic* dependencies, see Hudson 2019, Chapter O of this volume. Some practical applications of HPSG use syntactic dependencies (including many applications of the Alpino grammar, discussed in Section 3.3.1).

[15]More precisely, for DMRS and MRS to be fully interconvertible, every predicate (except for

indicate scope information. Finally, DELPH-IN MRS Dependencies (DM; Ivanova et al. 2012) express predicate-argument structure purely in terms of the surface tokens, without introducing any abstract predicates.

For example, the English Resource Grammar (ERG) produces the MRS representation in (4) for the sentence *The cherry tree blossomed.* For simplicity, we have omitted some details, including features such as number and tense, individual constraints (ICONS), and the use of difference lists. By convention, DELPH-IN predicates beginning with an underscore correspond to a lexical item, and have a three-part format, consisting of a lemma, a part-of-speech tag, and (optionally) a sense. Predicates without an initial underscore are abstract predicates. The *qeq* constraints (equality modulo quantifiers) are scopal relationships, where quantifiers may possibly intervene (for details, see Copestake et al. 2005 or Koenig & Richter 2019, Chapter C of this volume).

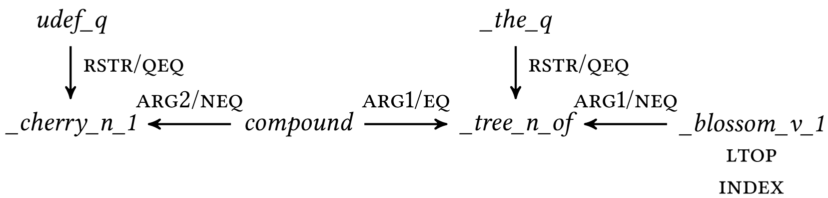
$$
(4)\quad \begin{bmatrix}
\textit{mrs} & \\
\text{HOOK} & \begin{bmatrix} \textit{hook} & \\ \text{LTOP} & \boxed{1} \\ \text{INDEX} & \boxed{2} \end{bmatrix} \\
\text{RELS} & \left\langle \begin{array}{l}
\begin{bmatrix} \textit{relation} & \\ \text{PRED} & \textit{\_the\_q} \\ \text{LBL} & \boxed{3} \\ \text{ARG0} & \boxed{4} \\ \text{RSTR} & \boxed{5} \\ \text{BODY} & \boxed{6} \end{bmatrix},
\begin{bmatrix} \textit{relation} & \\ \text{PRED} & \textit{compound} \\ \text{LBL} & \boxed{7} \\ \text{ARG0} & \boxed{8} \\ \text{ARG1} & \boxed{4} \\ \text{ARG2} & \boxed{9} \end{bmatrix},
\begin{bmatrix} \textit{relation} & \\ \text{PRED} & \textit{udef\_q} \\ \text{LBL} & \boxed{10} \\ \text{ARG0} & \boxed{9} \\ \text{RSTR} & \boxed{11} \\ \text{BODY} & \boxed{12} \end{bmatrix}, \\
\begin{bmatrix} \textit{relation} & \\ \text{PRED} & \textit{\_cherry\_n\_1} \\ \text{LBL} & \boxed{13} \\ \text{ARG0} & \boxed{9} \end{bmatrix},
\begin{bmatrix} \textit{relation} & \\ \text{PRED} & \textit{\_tree\_n\_of} \\ \text{LBL} & \boxed{7} \\ \text{ARG0} & \boxed{4} \end{bmatrix},
\begin{bmatrix} \textit{relation} & \\ \text{PRED} & \textit{\_blossom\_v\_1} \\ \text{LBL} & \boxed{1} \\ \text{ARG0} & \boxed{2} \\ \text{ARG1} & \boxed{4} \end{bmatrix}
\end{array} \right\rangle \\
\text{HCONS} & \left\langle \begin{bmatrix} \textit{qeq} & \\ \text{HARG} & \boxed{5} \\ \text{LARG} & \boxed{7} \end{bmatrix},
\begin{bmatrix} \textit{qeq} & \\ \text{HARG} & \boxed{11} \\ \text{LARG} & \boxed{13} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

For readability, it can be easier to express an MRS in a more abstract mathematical form, as shown in (5). This is equivalent to the feature structure in (4).

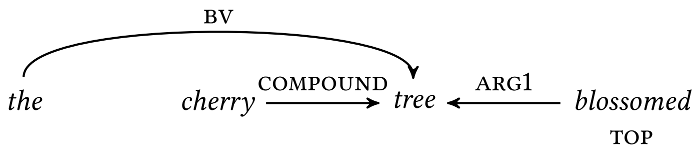
quantifiers) must have an intrinsic argument, and every variable must be the intrinsic argument of exactly one predicate.

$$
(5)\quad
\begin{array}{l}
\textsc{index}: e_1 \\
\qquad l_1: \textit{\_the\_q}\,(x_1, h_1, h_2)\,,\, h_1 \ \textsc{qeq}\ l_4 \\
\qquad l_2: \textit{udef\_q}\,(x_2, h_3, h_4)\,,\, h_3 \ \textsc{qeq}\ l_3 \\
\qquad l_3: \textit{\_cherry\_n\_1}(x_2) \\
\qquad l_4: \textit{\_tree\_n\_of}(x_1)\,,\, \textit{compound}\,(e_2, x_1, x_2) \\
\textsc{ltop}, l_5: \textit{\_blossom\_v\_1}(e_1, x_1)
\end{array}
$$

The corresponding Dependency Minimal Recursion Semantics (DMRS) representation is shown in (6). This captures all of the information in the MRS in (5). Predicates are represented as nodes, while semantic roles and scopal constraints are represented as directed edges, called *dependencies* or *links*. Each dependency has two labels. The first is an argument label, such as ARG1, ARG2, or RSTR (the restriction of a quantifier). The second is a scopal constraint, such as QEQ[16], EQ (the linked nodes share a label in the MRS, which is generally true for modifiers), or NEQ (the linked nodes don't share a label).

(6) 

Finally, the corresponding DELPH-IN MRS Dependencies (DM) representation is shown in (7). This is a simplified version of MRS, where all nodes are tokens in the sentence. Some abstract predicates are dropped (such as *udef_q*), while others are converted to dependencies (such as *compound*). Some scopal information is dropped (such as EQ vs NEQ). The label BV stands for the "bound variable" of a quantifier, equivalent to the RSTR/QEQ of DMRS.

(7) 

The existence of such dependency graph formalisms, as well as software packages to manipulate such graphs (e.g., Ivanova et al. 2012, Copestake et al. 2016, Hershcovich et al. 2019, or PyDelphin[17]), has made it easier to use HPSG grammars in a number of practical tasks, as we will discuss in Section 4.2.

[16] An alternative notation is to write /H instead of /QEQ.

[17] https://github.com/delph-in/pydelphin/, accessed 16 August 2019

3 Development of HPSG resources

In this section we describe various projects that have developed computational resources on the basis of or inspired by HPSG. As we'll discuss in Section 4 below, such resources can be used both in linguistic hypothesis testing as well as in various practical applications. The intended purpose of the resources influences the form that they take. The CoreGram Project (Section 3.1) and Babel (Section 3.3.3) primarily target linguistic hypothesis testing, the Alpino and Enju parsers (Section 3.3.1 and 3.3.2) primarily target practical applications, and the DELPH-IN Consortium (Section 3.2) attempts to balance these two goals.

3.1 CoreGram

The CoreGram[18] Project aims to produce large-scale HPSG grammars, which share a common "core" grammar (Müller 2015). At the time of writing, large grammars have been produced for German (Müller 2007), Danish (Müller & Ørsnes 2013), Persian (Müller & Ghayoomi 2010), Maltese (Müller 2009), and Mandarin (Müller & Lipenkova 2013). Smaller grammars are also available for English, Yiddish, Spanish, French, and Hindi.

All grammars are implemented in the TRALE system (Meurers et al. 2002; Penn 2004), which accommodates a wide range of technical devices proposed in the literature, including: phonologically empty elements, relational constraints, implications with complex antecedents, and cyclic feature structures. It also accomodates macros and an expressive morphological component. Melnik (2007) observes that, compared to other platforms like the LKB (see Section 3.2 below), this allows grammar engineers to directly implement a wider range of theoretical proposals.

An important part of CoreGram is the sharing of grammatical constraints across grammars. Some general constraints hold for all grammars, while others hold for a subset of the grammars, and some only hold for a single grammar. Müller (2015) describes this as a "bottom-up approach with cheating" (p. 43) – the aim is to analyze each language on its own terms (hence "bottom-up"), but to re-use analyses from existing grammars if possible (hence "with cheating"). The use of a core set of constraints is motivated not just for practical reasons, but also for theoretical ones. By developing multiple grammars in parallel, analyses can be improved by cross-linguistic comparison. The constraints encoded in the core grammar can be seen as an hypothesis about the structure of human language.

[18] https://hpsg.hu-berlin.de/Projects/CoreGram.html

as we will discuss in Section 4.1.1.

CoreGram grammar development aims to incrementally increase coverage of each language. To measure progress, grammars are evaluated against test suites, collections of sentences each annotated with a grammaticality judgment (Oepen et al. 1997; Müller 2004b). This allows a grammarian to check for unexpected side effects when modifying a grammar, to avoid situations when implementing an analysis of one phenomenon would break the analysis of another phenomenon. This is particularly important when modifying a constraint that is used by several grammars. To help achieve these aims, grammar development is supported by a range of software tools, including the test suite tool [incr tsdb()] (Oepen 2001; see also Section 3.2), and the graphical debugging tool Kahina (Dellert et al. 2010; 2013).

3.2 The DELPH-IN Consortium

The DELPH-IN[19] Consortium was established in 2001 to facilitate the development of large-scale, linguistically motivated HPSG grammars for multiple languages in tandem with the software required for developing them and deploying them in practical applications. At the time that DELPH-IN was founded, the English Resource Grammar (ERG; Flickinger 2000; 2011) had been under development already for 8 years and the Verbmobil project (Wahlster 2000) had also spurred the development of grammars for German (GG; Müller & Kasper 2000; Crysmann 2003) and Japanese (Jacy; Siegel, Bender & Bond 2016). Project DeepThought (Callmeier, Eisele, Schäfer & Siegel 2004) was exploring methodologies for combining deep and shallow processing in practical applications across multiple languages. This inspired the development of the LinGO Grammar Matrix (Bender, Flickinger & Oepen 2002), which began as a core grammar, consisting of constraints hypothesized to be cross-linguistically useful, abstracted out of the ERG with reference to Jacy and GG. The goal of the Grammar Matrix is to serve as a starting point for the development of new grammars making it easy to reuse what has been learned in the development of existing grammars. In the years since, it has been extended to include "libraries" of analyses of cross-linguistically variable phenomena (e.g., Drellishak 2009; Bender et al. 2010).

DELPH-IN provides infrastructure (version control repositories, mailing lists, annual meetings) and an emphasis on open-source distribution of resources. Both of these support the collaboration of a global network of researchers working on

[19]This stands for DEep Linguistic Processing in HPSG INitiative; see http://www.delph-in.net, accessed 16 August 2019

interoperable components. These include repositories of linguistic knowledge, that is, both grammars and meta-grammars (including the Matrix and CLIMB, Fokkens 2014); processing engines that apply that knowledge for parsing and generation (discussed further below); software for supporting the development of grammar documentation (e.g., Hashimoto et al. 2008), software for creating treebanks (Oepen et al. 2004; Packard 2015; see also Section 4.1.4 below), and parse ranking models trained on them (Toutanova et al. 2005; see also Section 2.2.2 above), and software for robust processing, i. e. using the knowledge encoded in the grammars to return analyses for sentences even if the grammar deems them ungrammatical (Zhang & Krieger 2011; Buys & Blunsom 2017; Chen et al. 2018).

A key accomplishment of the DELPH-IN Consortium is the standardization of a formalism for the declaration of grammars (Copestake 2002a), a formalism for the semantic representations (Copestake et al. 2005), and file formats for the storage and interchange of grammar outputs (e.g., the forest that results from parsing a sentence, as well as the results of treebanking (Oepen 2001; Oepen et al. 2004)). These standards facilitate the development of multiple different parsing and generation engines which can all process the same grammars (including, so far, the LKB (Copestake 2002b), PET (Callmeier 2000), ACE[20], and Agree (Slayden 2012)), of multiple software systems for processing bulk grammar output ([incr tsdb()] (Oepen 2001), art[21], and PyDelphin[22]) and of multilingual downstream systems which can be adapted to additional languages by plugging in different grammars. These tools and standards have in turn helped support a thriving community of users who furthermore accumulate and share information about best practices. Melnik (2007: 234) credits this community and the information it shares as a key factor that makes the grammar engineering with DELPH-IN ecosystem more accessible to HPSG linguists, compared to other platforms like TRALE (see Section 3.1 above).

The DELPH-IN community maintains research interests in both linguistics and practical applications. The focus on linguistics means that DELPH-IN grammarians strive to create grammars which capture linguistic generalizations and model grammaticality. This, in turn, leads to grammars with lower ambiguity than one finds with treebank-trained grammars and, importantly, grammars which produce well-formed strings in generation. The focus on practical applications leads to several kinds of additional research goals. Practical applications require robust processing, which in turn requires methods for handling unknown words (e.g.,

[20] http://sweaglesw.org/linguistics/ace/, accessed 16 August 2019

[21] https://sweaglesw.org/linguistics/libtsdb/art.html, accessed 16 August 2019

[22] https://github.com/delph-in/pydelphin/, accessed 16 August 2019

Adolphs et al. 2008), methods for managing extra-grammatical mark-up in text such as in Wikipedia pages (e.g., Flickinger, Oepen & Ytrestøl 2010) and strategies for processing inputs that are ungrammatical, at least according to the grammar (e.g., Zhang & Krieger 2011, see also Section 4.2.3). Processing large quantities of text motivates performance innovations, such as supertagging or ubertagging (e.g., Matsuzaki et al. 2007; Dridan 2013, see also Section 2.2.2) to speed up processing times. Naturally occurring text can include very long sentences which can run up against processing limits. Supertagging helps some here, too, but other strategies include , or the task of breaking a long sentence into smaller ones without loss of meaning (Muszyńska 2016). Working with real-world text (rather than curated testsuites designed for linguistic research only) requires the integration of external components such as morphological analyzers (e.g., Marimon 2013) and named entity recognizers (e.g., Waldron et al. 2006; Schäfer et al. 2008). As described in Section 2.2.2, working with real-world applications requires parse ranking (e.g., Toutanova et al. 2005), and similarly ranking of generator outputs (known as *realization ranking*; e.g., Velldal 2009). Finally, research on embedding broad-coverage grammars in practical applications inspires work towards making sure that the semantic representations can serve as a suitable interface for external components (e.g., Flickinger et al. 2005). These efforts are also valuable from a strictly linguistic point of view, i. e. one not concerned with practical applications. First, the broader the coverage of a grammar, the more linguistic phenomena it can be used to explore. Second, external constraints on the form of semantic representations provide useful guide points in the development of semantic analyses.

3.3 Other HPSG and HPSG-inspired broad-coverage grammars

3.3.1 Alpino

Alpino[23] is a broad-coverage grammars of Dutch (Bouma, van Noord & Malouf 2001; van Noord & Malouf 2005; van Noord 2006). The main motivation is practical: to provide coverage and accuracy comparable to state-of-the-art parsers for English. Nonetheless, it also includes theoretically interesting analyses, such as for cross-serial dependencies (Bouma & van Noord 1998). In addition to using hand-written rules, lexical information (such as subcategorisation frames) has also been extracted from two existing lexicons, Celex (Baayen et al. 1995) and Parole (Kruyt & Dutilh 1997).

[23] http://www.let.rug.nl/vannoord/alp/Alpino/, accessed 16 August 2019

Alpino produces syntactic dependency graphs, following the annotation format of the Spoken Dutch Corpus (Oostdijk 2000). These dependencies are constructed directly in the feature-structure formalism, exploiting the fact that a feature structure can be formalised as a directed acyclic graph. Each lexical entry encodes a partial dependency graph, and these graphs are composed through phrase structure rules to give a dependency graph for a whole sentence.

Although these dependencies differ from the semantic dependencies discussed in Section 2.2.3, a common motivation is to make the representations easier to use in practical applications. To harmonize with other computational work on dependency parsing, Bouma & van Noord (2017) have also produced a mapping from this format to Universal Dependencies (UD; Nivre et al. 2016), as discussed in Section 4.1.4 below. Alpino uses a statistical model trained on a dependency treebank, and in fact the same statistical model can be used in both parsing and generation (de Kok et al. 2011).

3.3.2 Enju

Enju[24] (Miyao et al. 2005) is a broad-coverage grammar of English, semi-automatically acquired from the Penn Treebank (Marcus et al. 1993). This approach aims to reduce the cost of writing a grammar by leveraging existing resources. The basic idea is that, by viewing Penn Treebank trees as partial specifications of HPSG analyses, it is possible to infer lexical entries.

Miyao et al. converted the relatively flat trees in the Penn Treebank to binary-branching trees, and percolated head information through the trees. They also had to convert analyses for certain constructions, including subject-control verbs, auxiliary verbs, coordination, and extracted arguments. Each converted tree can then be combined with a small set of hand-written HPSG schemata, to induce a lexical entry for each word in the sentence.

Development of Enju has focused on performance in practical applications, and the grammar is supported by an efficient parser (Tsuruoka et al. 2004; Matsuzaki et al. 2007), using a probabilistic model for feature structures (Miyao & Tsujii 2008). Enju has been used in a variety of NLP tasks, as will be discussed in Section 4.2.2.

3.3.3 Babel

Babel is a broad-coverage grammar of German (Müller 1996; 1999). One interesting feature of this grammar is that it makes extensive use of discontinuous con-

[24] http://www.nactem.ac.uk/enju/, accessed 29 August 2019

stituents (Müller 2004a). Although this makes the worst-case parsing complexity much worse, parsing speed doesn't seem to suffer in practice. This mirrors the findings of Carroll (1994), discussed in Section 2.2.1 above.

4 Deployment of HPSG resources

There are several different ways in which computational resources based on HPSG are used. In Section 4.1, we first consider applications furthering linguistic research, including both language documentation and linguistic hypothesis testing. Then, in Section 4.2, we consider applications outside of linguistics.

4.1 Language documentation and linguistic hypothesis testing

As described by Müller (1999), Bender (2008) and Bender et al. (2011), grammar engineering, that is the building of grammars in software, is an essential technique for testing linguistic hypotheses at scale. By "at scale", we mean both against large quantities of data and as integrated models of language that handle multiple phenomena at once. In this section, we overview how this is done in the CoreGram and Grammar Matrix projects for cross-linguistic hypothesis testing, and in the AGGREGATION project in the context of language documentation.[25]

4.1.1 CoreGram

As described in Section 3.1, the CoreGram project develops grammars for a diverse set of languages, and shares constraints across grammars, in a bottom-up fashion, so that more similar languages share more constraints. Still, there are constraints shared across all of the grammars in the project which can be seen as an hypothesis about properties shared by all languages. Whenever the CoreGram project expands to cover a new language, it can be seen as a test of this hypothesis.

For example, the most general constraint set allows a language to have V2 word order (as exemplified by Germanic languages), but rules out verb-penultimate word order, as discussed by Müller (2015) (see also Müller 2019a, Chapter 10 of this volume). It also includes constraints for argument structure and linking (see

[25]Grammar engineering isn't specific to HPSG and in fact has a history going back to at least the early 1960s (Kay 1963; Zwicky et al. 1965; Petrick 1965; Friedman et al. 1971) and modern work in Lexical Functional Grammar (Butt et al. 1999), Combinatory Categorial Grammar (Baldridge et al. 2007), Grammatical Framework (Ranta 2009), and others. For reflections on grammar engineering for linguistic hypothesis testing in LFG, see Butt et al. 1999 and King 2016.

Wechsler, Koenig & Davis 2019, Chapter 9 of this volume), as well as for information structure (see Kuthy 2019, Chapter D of this volume).

4.1.2 Grammar Matrix

As noted in Section 3.2, the LinGO Grammar Matrix (Bender et al. 2002; 2010) was initially developed in the context of Project DeepThought with the goal of speeding up the development of DELPH-IN-style grammars for additional languages. It consists of a shared core grammar and a series of "libraries" of analyses for cross-linguistically variable phenomena. Both of these constitute linguistic hypotheses: the constraints in the core grammar are hypothesized to be cross-linguistically useful. However, in the course of developing grammars based on the Matrix for specific languages, it is not uncommon to find reasons to refine the core grammar. The libraries, in turn, are intended to cover the attested range of variation for the phenomena they model. Languages that are not covered by the analyses in the libraries provide evidence that the libraries need to be extended or refined.

Grammar Matrix grammar development is less tightly coordinated than that of CoreGram (see Section 3.1): in the typical use case, grammar developers start from the Grammar Matrix, but with their own independent copy of the Matrix core grammar. This impedes somewhat the ability of the Matrix to adapt to the needs of various languages (unless grammar developers report back to the Matrix developers). On the other hand, the Matrix libraries represent an additional kind of linguistic hypothesis testing: each library on its own represents one linguistic phenomenon, but the libraries must be interoperable with each other. This is the cross-linguistic analogue of how monolingual implemented grammars allow linguists to ensure that analyses of different phenomena are interoperable (Müller 1999: p.439–440; Bender 2008): the Grammar Matrix customization system allows its developers to test cross-linguistic libraries of analyses for interactions with other phenomena (Bender et al. 2011; Bender 2016). Without computational support — i. e. a computer keeping track of the constraints that make up each analysis, compiling them into specific grammars, and testing those grammars against testsuites — this problem space would be too complex for exploration.

4.1.3 AGGREGATION

In many ways, the most urgent need for computational support for linguistic hypothesis testing is the description of endangered languages. Implemented grammars can be used to process transcribed but unglossed text in order to find relevant examples more quickly, both of phenomena that have already been analyzed and of phenomena that are as yet not well-understood.[26] Furthermore, treebanks constructed from implemented grammars can be tremendously valuable additions to language documentation (see Section 4.1.4 below). However, the process of building an implemented grammar is time-consuming, even with the start provided by a multilingual grammar engineering project like CoreGram, ParGram (Butt et al. 2002; King et al. 2005), the GF Resource Grammar Library Ranta (2009), or the Grammar Matrix.

This is the motivation for the AGGREGATION[27] project, which starts from two observations: (1) descriptive linguists produce extremely rich annotations on data in the form of interlinear glossed text (IGT); and (2) the Grammar Matrix's libraries are accessed through a customization system which elicits a grammar specification in the form of a series of choices describing either high-level typological properties or specific constraints on lexical classes and lexical rules. The goal of AGGREGATION is to automatically produce such grammar specifications on the basis of information encoded in IGT, to be used by the Grammar Matrix customization system to produce language-particular grammars. AGGREGATION uses different approaches for different linguistic subsystems. For example, it learns morphotactics by observing morpheme order in the training data, and the grouping of affixes together into position classes based on measures of overlap of stems they attach to (Wax 2014; Zamaraeva et al. 2017). For many kinds of syntactic information, it leverages syntactic structure projected from the translation line (English, easily parsed with current tools) through the gloss line (which facilitates aligning the language and translation lines) to the language line (Xia & Lewis 2007; Georgi 2016). Using this projected information, the AGGREGATION system can detect case frames for verbs, word order patterns, etc. (Bender et al. 2013; Zamaraeva et al. 2019).[28]

4.1.4 Treebanks and sembanks

A particularly valuable type of resource that can be derived from HPSG grammars are treebanks and sembanks. A *treebank* is a collection of text where each sentence is associated with a syntactic representation. A *sembank* has semantic representations (in some cases in addition to the syntactic ones). Treebanks and

[26] This methodology of using an implemented grammar as a sieve to sift the interesting examples out of corpora is demonstrated for English by Baldwin et al. (2005).

[27] http://depts.washington.edu/uwcl/aggregation/, accessed 16 August 2019

[28] The TypeGram project (Hellan & Beermann 2014) is in a similar spirit. TypeGram provides methods of creating HPSG grammars by encoding specifications of valence and inflection in particularly rich IGT and then creating grammars based on those specifications.

sembanks can be used for linguistic research, as the analyses allow for more detailed structure-based searches for phenomena of interest (Rohde 2005; Ghodke & Bird 2010; Kouylekov & Oepen 2014).[29] In the context of language documentation and description, searchable treebanks can also be a valuable addition, helping readers connect prose descriptions of linguistic phenomena to multiple examples in the corpus (Bender et al. 2012). In natural language processing, treebanks and sembanks are critical source material for training stochastic and neural parsers (see Section 4.2.3).

Traditional treebanks are created by doing a certain amount of preprocessing on data, including possibly chunking or CFG parsing, and then hand-correcting the result (Marcus et al. 1993; Banarescu et al. 2013). While this approach is a means to encode human insight about linguistic structure for later automatic processing, it is both inefficient and potentially error-prone. The Alpino project (van der Beek et al. 2002; see also Section 3.3.1 above) addresses this by first parsing the text with a broad-coverage HPSG-inspired grammar of Dutch and then having annotators select among the parses. The selection process is facilitated by allowing the annotators to mark potential lexical entries for words in the sentence at hand as correct, possibly correct, or wrong and to pre-mark some constituent boundaries. These constraints reduce the search space for the parser and consequently also the range of analyses the annotator has to consider before choosing one. A facility for adding one-off lexical entries to handle e.g., misspellings helps increase grammar coverage. Disambiguation is handled with the aid of *discriminants* i. e. properties true of some but not all trees in the parse forest (Carter 1997). Finally, the annotators may further edit analyses deemed insufficient. Though the underlying grammar is based on HPSG, the treebank stores dependency graphs instead. The Alpino parser was similarly used to construct the Lassy treebanks of written Dutch (van Noord et al. 2013). In more recent work, these dependency representations have been mapped to the Universal Dependencies (UD) annotation standards (Nivre et al. 2016) to produce a UD treebank for Dutch (Bouma & van Noord 2017).

The Redwoods project (Oepen et al. 2004) also produces grammar-driven treebanks, in this case for English and without any post-editing of analyses.[30] As with Alpino, this is done by first parsing the corpus with the grammar and calcuating the discriminants for each parse forest. Finally, the treebanking software

[29] The WeSearch interface of Kouylekov & Oepen (2014) can be accessed at http://wesearch.delph-in.net/deepbank/search.jsp (accessed 16 August 2019).

[30] There are also Redwoods-style treebanks for other languages, including the Hinoki Treebank of Japanese (Bond et al. 2004) and the Tibidabo Treebank of Spanish (Marimon 2015).

stores not only the final full HPSG analysis that was selected, but also the decisions the annotator made about each discriminant. Thus when the grammar is updated to include for example a refinement to the semantic representations, the corpus can be reparsed and the decisions replayed, leaving only a small amount of further annotation work to be done to handle any additional ambiguity introduced. The activity of treebanking in turn provides useful insight into grammatical analyses, including sources of spurious ambiguity and phenomena that are not yet properly handled and thus informs and spurs further grammar development. A downside to strictly grammar-based treebanking is that only items for which the grammar finds a reasonable parse can be included in the treebank. For many applications, this is not a drawback, so long as there are sufficient and sufficiently varied sentences that do receive analyses.

Finally, there are also automatically annotated treebanks. These are not as reliable as manually annotated treebanks, but they can be considerably larger. WikiWoods[31] covers 55m sentences of English (900m tokens). It was produced by Flickinger, Oepen & Ytrestøl (2010) and Solberg (2012) from the July 2008 dump of the full English Wikipedia, using the ERG and PET, with parse ranking trained on the manually treebanked subcorpus WeScience (Ytrestøl et al. 2009). As with the Redwoods treebanks, WikiWoods is updated with each release of the ERG.

4.2 Downstream applications

In this section, we discuss the use of HPSG grammars for practical tasks. There is a large number of applications, and we focus on several important applications here. In Section 4.2.1, we cover educational applications where a grammar is used directly. In Section 4.2.2, we cover applications where a grammar is used to provide features to help solve tasks in Natural Language Processing (NLP). Finally, in Section 4.2.3, we cover applications where a grammar is used to provide data for machine learning systems.[32]

4.2.1 Education

Precise syntactic analyses can be useful in language teaching, in order to automatically identify errors and give feedback to the student. In order to model

[31] http://moin.delph-in.net/WikiWoods, accessed 16 August 2019

[32] The DELPH-IN community maintains an updated list of applications of DELPH-IN software and resources at http://moin.delph-in.net/DelphinApplications (accessed 16 August 2019).

common mistakes, a grammar can be extended with so-called *mal-rules*. A mal-rule is like a normal rule, in that it licenses a construction, and can be treated the same during parsing – however, given a parse, the presence of mal-rules indicates that the student needs to be given feedback (Bender et al. 2004; Flickinger & Yu 2013; Morgado da Costa et al. 2016). A large scale system implementing this kind of computer-aided teaching has been developed by the Education Program for Gifted Youth at Stanford University, using the ERG (Suppes et al. 2014). This system has reached tens of thousands of elementary and middle school children, and has been found to improve the school results of underachieving children.

Another way to use a precision grammar is to automatically produce teaching materials. Given a semantic representation, a grammar can generate one or more sentences. Flickinger (2017) uses the ERG to produce practice exercises for a student learning first-order logic. For each exercise, the student is presented with an English sentence and is supposed to write down the corresponding first-order logical form. By using a grammar, the system can produce syntactically varied questions and automatically evaluate the student's answer.

4.2.2 NLP tasks

Much NLP work focuses on specific *tasks*, where a system is presented with some input, and required to produce an output, with a clearly-defined metric to determine how well the system performs. HPSG grammars have been used in a range of such tasks, where the syntactic and semantic analyses provide useful features.

*Information retrieval*information retrieval is the task of finding relevant documents for a given query. For example, Schäfer et al. (2011) present a tool for searching the ACL Anthology, using the ERG. *Information extraction*information extraction is the task of identifying useful facts in a collection of documents. For example, Reiplinger et al. (2012) aim to identify definitions of technical concepts from English text, in order to automatically construct a glossary. They find that using the ERG reduces noise in the candidate definitions. Miyao et al. (2008) aim to identify protein-protein interactions in the English biomedical literature, using Enju.

For these tasks, some linguistic phenomena are particularly important, such as negation and hedging (including adverbs like *possibly*, modals like *may*, and verbs of speculation like *suggest*). When it comes to identifying facts asserted in a document, a clause that has been negated or hedged should be treated with caution. MacKinlay et al. (2012) consider the biomedical domain, evaluating on the BioNLP 2009 Shared Task (Kim et al. 2009), where they outperform previous approaches for negation, but not for speculation. Velldal et al. (2012) consider negation and speculation in biomedical text, evaluating on the CoNLL 2010 Shared Task (Farkas et al. 2010), where they outperform previous approaches. Packard et al. (2014) propose a general-purpose method for finding the scope of negation in an MRS, evaluating on the *SEM 2012 Shared Task (Morante & Blanco 2012). They find that converting the output of the ERG with a relatively simple set of rules achieves high performance on this English dataset, and combining this approach with a purely statistical system outperforms previous approaches. Zamaraeva et al. (2018) use the ERG for detection and then use that information to refine the (machine-learning) features in a system that classifies English pathology reports and improve system performance. A common finding from these studies is that a system using the output of the ERG tends to have high precision (items identified by the system tend to be correct) but low recall (items are often overlooked by the system). One reason for low recall is that the grammar does not cover all sentences in natural text. As we will see in Section 4.2.3, recent work on robust parsing may help to close this coverage gap.

Negation resolution is also included in Oepen et al.'s (2017) Shared Task on Extrinsic Parser Evaluation. As mentioned in Section 2.2.3, dependency graphs can provide a useful tool in NLP tasks, and this shared task aims to evaluate the use of dependency graphs (both semantic and syntactic), for three downstream applications: biomedical information extraction, negation resolution, and fine-grained opinion analysis. Some participating teams use DM dependencies[33] (Schuster et al. 2017; Chen et al. 2017). The results of this shared task suggest that, compared to other dependency representations, DM is particularly useful for negation resolution.

Another task where dependency graphs have been used is *summarization*. Most existing work on this task focuses on so-called *extractive summarization*: given an input document, a system forms a summary by extracting short sections of the input. This is in contrast to *abstractive summarization*, where a system generates new text based on the input document. Extractive summarization is limited, but widely used because it is easier to implement. Fang et al. (2016) show how a wide-coverage grammar like the ERG makes it possible to implement an abstractive summarizer with state-of-the-art performance. After parsing the input document into logical propositions, the summarizer prunes the set of propositions using a cognitively inspired model. A summary is then generated based on the pruned set of propositions. Because no text is directly extracted from the input document, it is possible to generate a more concise summary.

Finally, no discussion of NLP tasks would be complete without including *ma-*

[33]DM stands for DELPH-IN MRS dependencies; see (7).

chine translation. A traditional grammar-based approach uses three grammars: a grammar for the source language, a grammar for the target language, and a *transfer grammar*, which converts semantic representations for the source language to semantic representations for the target language (Oepen et al. 2007; Bond et al. 2011). Translation proceeds in three steps: parse the source sentence, transfer the semantic representation, and generate a target sentence. The transfer grammar is needed both to find appropriate lexical items, and also to convert semantic representations when languages differ in how an idea might be expressed. The difficulty in writing a transfer grammar that is robust enough to deal with arbitrary input text means that statistical systems might be preferred. Horvat (2017) explores the use of statistical techniques, skipping out the transfer stage: a target-language sentence is generated directly from a semantic representation for the source language. Goodman (2018) explores the use of statistical techniques within the paradigm of parsing, transferring, and generating.

4.2.3 Data for machine learning

In Section 4.2.2, we described how HPSG grammars can be directly incorporated into NLP systems. Another use of HPSG grammars in NLP is to generate data on which a statistical system can be trained.

For example, one limitation of using an HPSG grammar in an NLP system is that the grammar is unlikely to cover all sentences in the data (Flickinger et al. 2012). One way to overcome this coverage gap is to train a statistical system to produce the same output as the grammar. The idea is that the trained system will be able to generalise to sentences that the grammar does not cover. Oepen et al. (2014), Oepen et al. (2015), and Oepen et al. (2019) present shared tasks on semantic dependency parsing, including both DM dependencies and Enju predicate-argument structures. As of 2015, the best-performing systems in these shared tasks could already produce dependency graphs almost as accurately as grammar-based parsers (for sentences where the grammar has coverage). Similarly, Buys & Blunsom (2017) develop a parser for EDS and DMRSDependency Minimal Recursion Semantics (DMRS) which performs almost as well as a grammar-based parser, but has full coverage, and can run 70 times faster.

In fact, in more recent work, the difference in performance has been effectively closed. Chen et al. (2018) consider parsing to EDS and DMRS graphs, and actually achieve slightly higher accuracy with their system, compared to a grammar-based parser. Unlike the previous statistical approaches, Chen et al. do not just train on the desired dependency graphs, but also use information in the phrase-structure tree. They suggest that using this information allows their system to learn compositional rules mirroring composition in the grammar, and thereby allows their system to generalise better.

Another application of HPSG-derived dependency graphs is for *distributional semantics*. Here, the aim is to learn the meanings of words from a corpus, exploiting the fact that the context of a word tells us something about its meaning. This is known as the *distributional hypothesis*, an idea with roots in American structuralism (Harris 1954) and British lexicology (Firth 1951; 1957). Most work on distributional semantics learns a *vector space model*, where the meaning of each word is represented as a point in a high-dimensional vector space (for an overview, see Erk 2012 and Clark 2015). However, Emerson (2018) argues that vector space models cannot capture various aspects of meaning, including logical structure, as well as phenomena like polysemy. Instead, Emerson presents a distributional model which can learn truth-conditional semantics, using a parsed corpus like WikiWoods (see Section 4.1.4). This approach relies on the semantic analyses given by a grammar, as well as the infrastructure to parse a large amount of text.

Finally, there are also applications using grammars not to parse, but to generate. Kuhnle & Copestake (2018) consider the task of *visual question answering*, where a system is presented with an image and a question about the image, and must answer the question. This task requires language understanding, reference resolution, and grounded reasoning, in a way that is relatively well-defined. However, for many existing datasets, there are biases in the questions which mean that high performance can be achieved without true language understanding. For this reason, there is increasing interest in artificial datasets, which are controlled to make sure that high performance requires true understanding. Kuhnle & Copestake present ShapeWorld, a configurable system for generating artificial data. The system generates an abstract representation of a scene (coloured shapes in different configurations), and then generates an image and a caption based on this representation. The use of a broad-coverage grammar is crucial in allowing the system to be configurable and scale across a variety of syntactic constructions.

5 Linguistic insights

In Section 4.1 above, we described multiple ways in which computational methods can be used in the service of linguistic research, especially in testing linguistic hypotheses. Here, we highlight a few ways in which grammar engineering work in HPSG has turned up linguistic insights that had not previously been discovered through non-computational means.[34]

5.1 Ambiguity

As discussed in Section 2.2.2, the scale of ambiguity has become clear now that broad-coverage precision grammars are available. By taking both coverage and precision seriously, it is possible to investigate ambiguity on a large scale, quantifying the sources of ambiguity and the information needed to resolve it. For example, Toutanova et al. (2002; 2005) found that in the Redwoods treebank (3rd Growth), roughly half of the ambiguity was lexical, and half syntactic. They also showed how combining sources of information (such as both semantic and syntactic information) is important for resolving ambiguity, and argue that using multiple kinds of information in this way is consistent with probabilistic approaches in psycholinguistics.

5.2 Long-tail phenomena

One of the strengths of HPSG as a theoretical framework is that it allows for the analysis of both "core" and "peripheral" phenomena within a single, integrated model. Indeed, by implementing large-scale grammars across a range of languages, it becomes possible to investigate the extent to which a particular phenomenon should be considered "core", "peripheral", or something in between (Müller 2014).

In fact, when working with actual data and large-scale grammars, it quickly becomes apparent just how long the long-tail of "peripheral" phenomena is. Furthermore, the sustained development of broad-coverage linguistic resources makes it possible to bring into view more and more low-frequency phenomena (or low-frequency variations on relatively high-frequency phenomena). A case in point is the range of raising and control valence frames found in the ERG (Flickinger 2000; 2011). As of the 2018 release, the ERG includes over 60 types for raising and control predicates, including verbs, adjectives, and nouns, many of which are not otherwise discussed in the syntactic literature. These include such low-frequency types as the one for *incumbent*, which requires an expletive *it* subject, an obligatory *(up)on* PP complement, and an infinitival VP complement, and which establishes a control relation between the object of *on* and the VP's missing subject:[35]

(8) It is incumbent on you to speak plainly.

[34] For similar reflections from the point of view of LFG, see King (2016).

[35] Our thanks to Dan Flickinger for this example.

5.3 Analysis-order effects

Grammar engineering means making analyses specific and then being able to build on them. This has both benefits and drawbacks: on the one hand, it means that additional grammar engineering work can build directly on the results of previous work. It also means that any additional grammar engineering work is constrained by the work it is building on. Fokkens (2014) observes this phenomenon and notes that it introduces artifacts: the form an implemented grammar takes is partially the result of the order in which the grammar engineer considered phenomena to implement. This is probably also true for non-computational work, as theoretical ideas developed with particular phenomena (and indeed languages) in mind influence the questions with which researchers approach additional phenomena. Fokkens proposes that the methodology of meta-grammar engineering can be used to address this problem: using her CLIMB methodology, rather than deciding between analyses of a given phenomenon without input from later-studied phenomena, the grammar engineer can maintain multiple competing analyses through time and break free, at least partially, of the effects of the timeline of grammar development. The central idea is that the grammar writer develops a meta-grammar, like the Grammar Matrix customization system (see Section 4.1.2), but for a single language. This customization system maintains alternate analyses of particular phenomena which are invoked via grammar specifications so the different versions of the grammar can be compiled and tested.

6 Summary

In this chapter, we have attempted to illuminate the landscape of computational work in HPSG. We have discussed how HPSG as a theory supports computational work, described large-scale computational projects that use HPSG, highlighted some applications of implemented grammars in HPSG, and explored ways in which computational work can inform linguistic research. This field is very active and our overview necessarily incomplete. Nonetheless, it is our hope that the pointers and overview provided in this chapter will serve to help interested readers connect with on-going research in computational linguistics using HPSG.

Acknowledgements

We'd like to thank Stephan Oepen for helpful comments on an early draft of this chapter and Stefan Müller for detailed comments as volume editor.

References

Adolphs, Peter, Stephan Oepen, Ulrich Callmeier, Berthold Crysmann, Dan Flickinger & Bernd Kiefer. 2008. Some fine points of hybrid natural language parsing. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis & Daniel Tapias (eds.), *Proceedings of the sixth international conference on language resources and evaluation (LREC'08)*. Marrakech, Morocco: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2008/.

Baayen, R Harald, Richard Piepenbrock & Leon Gulikers. 1995. *The CELEX lexical database*. Distributed by the Linguistic Data Consortium, University of Pennsylvania.

Baldridge, Jason, Sudipta Chatterjee, Alexis Palmer & Ben Wing. 2007. DotCCG and VisCCG: Wiki and programming paradigms for improved grammar engineering with OpenCCG. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar Engineering across Frameworks 2007* (Studies in Computational Linguistics ONLINE), 5–25. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/, accessed 2018-2-25.

Baldwin, Timothy, John Beavers, Emily M. Bender, Dan Flickinger, Ara Kim & Stephan Oepen. 2005. Beauty and the beast: What running a broad-coverage precision grammar over the BNC taught us about the grammar — and the corpus. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives* (Studies in Generative Grammar 85), 49–69. Mouton de Gruyter.

Banarescu, Laura, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer & Nathan Schneider. 2013. Abstract Meaning Representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, 178–186. Sofia, Bulgaria: Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W13-2322.

Bangalore, Srinivas & Aravind K. Joshi. 1999. Supertagging: an approach to almost parsing. *Computational Linguistics* 25(2). 237–265. http://aclweb.org/anthology/J99-2004.

Bender, Emily M. 2008. Grammar engineering for linguistic hypothesis testing. In Nicholas Gaylord, Alexis Palmer & Elias Ponvert (eds.), *Proceedings of the Texas Linguistics Society X Conference: Computational linguistics for less-studied languages*, 16–36. Stanford CA: CSLI Publications ONLINE.

Bender, Emily M. 2016. Linguistic typology in natural language processing. *Linguistic Typology* 20(3). 645–660.

Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson & Safiyyah Saleem. 2010. Grammar customization. *Research on Language & Computation* 8(1). 23–72. http://dx.doi.org/10.1007/s11168-010-9070-1. 10.1007/s11168-010-9070-1.

Bender, Emily M., Dan Flickinger & Stephan Oepen. 2011. Grammar engineering and linguistic hypothesis testing: Computational support for complexity in syntactic analysis. In E.M. Bender & Arnold J.E. (eds.), *Language from a cognitive perspective: Grammar, usage and processing*, 5–29. Stanford, CA: CSLI Publications.

Bender, Emily M., Dan Flickinger, Stephan Oepen, Annemarie Walsh & Timothy Baldwin. 2004. Arboretum. Using a precision grammar for grammar checking in CALL. In *Proceedings of the InSTIL Symposium on NLP and Speech Technologies in Advanced Language Learning Systems*. Venice, Italy.

Bender, Emily M., Daniel P. Flickinger & Stephan Oepen. 2002. The Grammar Matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In John Carroll, Nelleke Oostdijk & Richard Sutcliffe (eds.), *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, 8–14. Taipei, Taiwan.

Bender, Emily M., Sumukh Ghodke, Timothy Baldwin & Rebecca Dridan. 2012. From database to treebank: enhancing hypertext grammars with grammar engineering and treebank search. In Sebastian Nordhoff & Karl-Ludwig G. Poggeman (eds.), *Electronic grammaticography*, 179–206. Honolulu: University of Hawaii Press.

Bender, Emily M, Michael Wayne Goodman, Joshua Crowgey & Fei Xia. 2013. Towards creating precision grammars from interlinear glossed text: inferring large-scale typological properties. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 74–83. Sofia, Bulgaria: Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W13-2710.

Boguraev, Bran, John Carroll, Ted Briscoe & Claire Grover. 1988. Software support for practical grammar development. In *Proceedings of the 12th International Conference on Computational Linguistics (COLING)*, 54–58. International Committee on Computational Linguistics (ICCL). http://aclweb.org/anthology/C88-1012.

Bond, Francis, Sanae Fujita, Chikara Hashimoto, Kaname Kasahara, Shigeko Nariyama, Eric Nichols, Akira Ohtani, Takaaki Tanaka & Shigeaki Amano. 2004. The Hinoki treebank: a treebank for text understanding. In *Proceedings of the 1st International Joint Conference on Natural Language Processing (IJCNLP)*, 158–167. Springer.

Bond, Francis, Stephan Oepen, Eric Nichols, Dan Flickinger, Erik Velldal & Petter Haugereid. 2011. Deep open-source machine translation. *Machine Translation* 25(2). 87–105.

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press.

Bouma, Gosse & Gertjan van Noord. 2017. Increasing return on annotation investment: the automatic construction of a Universal Dependency Treebank for Dutch. In *Proceedings of the NoDaLiDa 2017 workshop on Universal Dependencies (UDW 2017)*, 19–26. Association for Computational Linguistics (ACL). https://aclweb.org/anthology/W17-0403.

Bouma, Gosse, Gertjan van Noord & Robert Malouf. 2001. Alpino: wide-coverage computational analysis of Dutch. In Walter Daelemans, Khalil Sima'an, Jorn Veenstra & Jakub Zavrel (eds.), *Computational linguistics in the Netherlands 2000: Selected papers from the Eleventh CLIN Meeting* (Language and Computers 37). Amsterdam/New York, NY: Rodopi.

Butt, Miriam, Helge Dyvik, Tracy Holloway King, Hiroshi Masuichi & Christian Rohrer. 2002. The Parallel Grammar Project. In *Proceedings of COLING-2002 Workshop on Grammar Engineering and Evaluation*, 1–7.

Butt, Miriam, Tracy Holloway King, María-Eugenia Niño & Frédérique Segond. 1999. *A grammar writer's cookbook* (CSLI Lecture Notes 95). Stanford, CA: CSLI Publications.

Buys, Jan & Phil Blunsom. 2017. Robust incremental neural semantic graph parsing. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, long papers*, 1215–1226. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/P17-1112.

Callmeier, Ulrich. 2000. PET: A platform for experimentation with efficient HPSG processing techniques. *Natural Language Engineering* 6(1). 99–108. Special Issue on Efficient Processing with HPSG: Methods, Systems, Evaluation.

Callmeier, Ulrich, Andreas Eisele, Ulrich Schäfer & Melanie Siegel. 2004. The DeepThought core architecture framework. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC)*. Lisbon, Portugal: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2004/pdf/603.pdf.

Carroll, John. 1994. Relating complexity to practical performance in parsing with wide-coverage unification grammars. In James Pustejovsky (ed.), *32th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 287–294. Las Cruses: Association for Computational Linguistics.

Carroll, John Andrew. 1993. *Practical unification-based parsing of natural language*. University of Cambridge dissertation. https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-314.ps.gz.

Carroll, John, Ann Copestake, Dan Flickinger & Victor Poznański. 1999. An efficient chart generator for (semi-)lexicalist grammars. In *Proceedings of the 7th European Workshop on Natural Language Generation*, 86–95.

Carter, David. 1997. The TreeBanker. a tool for supervised training of parsed corpora. In Dominique Estival, Alberto Lavelli, Klaus Netter & Fabio Pianesi (eds.), *Proceedings of the 1997 ACL workshop on computational environments for grammar development and linguistic engineering (ENVGRAM)*, 9–15. Madrid, Spain: Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W97-1502.

Chen, Yufei, Junjie Cao, Weiwei Sun & Xiaojun Wan. 2017. Peking at EPE 2017: a comparison of tree approximation, transition-based and maximum subgraph models for semantic dependency analysis. In *Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation, at the 4th International Conference on Dependency Linguistics and the 15th International Conference on Parsing Technologies*, 60–64. Nordic Language Processing Laboratory. http://svn.nlpl.eu/epe/2017/public/proceedings.pdf.

Chen, Yufei, Weiwei Sun & Xiaojun Wan. 2018. Accurate SHRG-based semantic parsing. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, long papers*, 408–418. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/P18-1038.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Clark, Stephen. 2015. Vector space models of lexical meaning. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 2nd, chap. 16, 493–522. Wiley. http://www.cl.cam.ac.uk/~sc609/pubs/sem_handbook.pdf.

Clark, Stephen & James R. Curran. 2004. The importance of supertagging for wide-coverage CCG parsing. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING)*. International Committee on Computational Linguistics (ICCL). http://aclweb.org/anthology/C04-1041.

Copestake, Ann. 2002a. Definitions of typed feature structures. In Stephan Oepen, Dan Flickinger, Jun-ichi Tsujii & Hans Uszkoreit (eds.), *Collaborative language engineering* (CSLI Lecture Notes 118), 227–230. Stanford, CA: CSLI Publications.

Copestake, Ann. 2002b. *Implementing typed feature structure grammars* (CSLI Lecture Notes 110). Stanford, CA: CSLI Publications.

Copestake, Ann. 2009. Slacker semantics: why superficiality, dependency and avoidance of commitment can be the right way to go. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 1–9. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/E09-1001.

Copestake, Ann, Guy Emerson, Michael Wayne Goodman, Matic Horvat, Alexander Kuhnle & Ewa Muszyńska. 2016. Resources for building applications with Dependency Minimal Recursion Semantics. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*, 1240–1247. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2016/pdf/634_Paper.pdf.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Crysmann, Berthold. 2003. On the efficient implementation of German verb placement in HPSG. In *Proceedings of RANLP 2003*, 112–116. Borovets, Bulgaria.

de Kok, Daniël, Barbara Plank & Gertjan van Noord. 2011. Reversible stochastic attribute-value grammars. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, short papers*, 194–199. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/P11-2034.

Dellert, Johannes, Kilian Evang & Frank Richter. 2010. *Kahina, a debugging framework for logic programs and TRALE*. Presentation at the HPSG 2010 Conference.

Dellert, Johannes, Kilian Evang & Frank Richter. 2013. Kahina: A hybrid trace-based and chart-based debugging system for grammar engineering. In Denys Duchier & Yannick Parmentier (eds.), *Proceedings of the workshop on high-level methodologies for grammar engineering (HMGE 2013), Düsseldorf*, 75–86.

Drellishak, Scott. 2009. *Widespread but not universal: Improving the typological coverage of the Grammar Matrix*. University of Washington Doctoral dissertation.

Dridan, Rebecca. 2013. Ubertagging: joint segmentation and supertagging for English. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1201–1212. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/D13-1120.

Emerson, Guy. 2018. *Functional distributional semantics: learning linguistically informed representations from a precisely annotated corpus*. University of Cambridge dissertation. https://www.repository.cam.ac.uk/bitstream/handle/1810/284882/thesis.pdf.

Erk, Katrin. 2012. Vector space models of word meaning and phrase meaning: a survey. *Language and Linguistics Compass* 6(10). 635–653. https://onlinelibrary.wiley.com/doi/epdf/10.1002/lnco.362.

Fang, Yimai, Haoyue Zhu, Ewa Muszyńska, Alexander Kuhnle & Simone Teufel. 2016. A proposition-based abstractive summariser. In *Proceedings of the 26th International Conference on Computational Linguistics (COLING)*, 567–578. International Committee on Computational Linguistics (ICCL). http://aclweb.org/anthology/C16-1055.

Farkas, Richárd, Veronika Vincze, György Móra, János Csirik & György Szarvas. 2010. The CoNLL-2010 shared task: learning to detect hedges and their scope in natural language text. In *Proceedings of the 14th Conference on Computational Natural Language Learning (CoNNL), shared task*, 1–12. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W10-3001.

Firth, John Rupert. 1951. Modes of meaning. *Essays and Studies* 4. 118–149. Reprinted in: Firth, 1957, *Papers in Linguistics*, 190–215, Oxford University Press.

Firth, John Rupert. 1957. A synopsis of linguistic theory 1930–1955. In John Rupert Firth (ed.), *Studies in linguistic analysis* (Special volume of the Philological Society), chap. 1, 1–32. Blackwell.

Flickinger, Dan, Jan Tore Lønning, Helge Dyvik, Stephan Oepen & Francis Bond. 2005. SEM-I rational MT: enriching deep grammars with a semantic interface for scalable machine translation. In *Proceedings of Machine Translation Summit X*, 165–172. Asia-Pacific Association for Machine Translation. http://www.mt-archive.info/MTS-2005-Flickinger.pdf.

Flickinger, Dan, Stephan Oepen & Emily M. Bender. 2017. Sustainable Development and Refinement of Complex Linguistic Annotations at Scale. In Nancy

Ide & James Pustejovsky (eds.), *Handbook of Linguistic Annotation*, 353–377. Dordrecht: Springer. DOI:10.1007/978-94-024-0881-2_14

Flickinger, Dan, Stephan Oepen & Gisle Ytrestøl. 2010. WikiWoods: syntacto-semantic annotation for English Wikipedia. In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC)*, 1665–1671. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2010/pdf/432_Paper.pdf.

Flickinger, Dan, Stephan Oepen & Gisle Ytrestøl. 2010. Wikiwoods: syntacto-semantic annotation for english wikipedia. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, 1665–1671. Valletta, Malta: European Language Resources Association (ELRA).

Flickinger, Dan & Jiye Yu. 2013. Toward more precision in correction of grammatical errors. In *Proceedings of the 17th Conference on Computational Natural Language Learning (CoNLL): shared task*, 68–73. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W13-3609.

Flickinger, Dan, Yi Zhang & Valia Kordoni. 2012. DeepBank: a dynamically annotated treebank of the Wall Street Journal. In *Proceedings of the 11th International Workshop on Treebanks and Linguistic Theories (TLT)*, 85–96. University of Lisbon. http://tlt11.clul.ul.pt/ProceedingsTLT11.tgz.

Flickinger, Daniel. 2017. Generating English paraphrases from logic. In Martijn Wieling, Gosse Bouma & Gertjan van Noord (eds.), *From semantics to dialectometry: festschrift in honour of john nerbonne*, 99–108. Springer.

Flickinger, Daniel P. 1987. *Lexical rules in the hierarchical lexicon*. Stanford University dissertation.

Flickinger, Daniel P. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6(1). 15–28. Special Issue on Efficient Processing with HPSG: Methods, Systems, Evaluation.

Flickinger, Daniel P. 2011. Accuracy vs. robustness in grammar engineering. In Emily M. Bender & Jennifer E. Arnold (eds.), *Language from a cognitive perspective: Grammar, usage, and processing*, 31–50. Stanford, CA: CSLI Publications.

Fokkens, Antske Sibelle. 2014. *Enhancing empirical research for linguistically motivated precision grammars*. Department of Computational Linguistics, Universität des Saarlandes dissertation.

Friedman, Joyce, Thomas H. Bredt, Robert W. Doran, Bary W. Pollack & Theodore S. Martner. 1971. *A computer model of Transformational Grammar* (Mathematical Linguistics and Automatic Language Processing 9). New York: Elsevier.

Gazdar, Gerald & Geoffrey K. Pullum. 1985. Computationally relevant properties of natural languages and their grammars. *New Generation Computing* (3). 237–306. Reprinted in Savitch, Bach, Marsh, and Safran-Naveh (eds.), *The Formal Complexity of Natural Language.*

Georgi, Ryan. 2016. *From Aari to Zulu: massively multilingual creation of language tools using interlinear glossed text.* University of Washington dissertation.

Ghodke, Sumukh & Steven Bird. 2010. Fast query for large treebanks. In *Proceedings of the 8th Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), long papers*, 267–275. Los Angeles, California: Association for Computational Linguistics (ACL). http://aclweb.org/anthology/N10-1034.

Goodman, Michael Wayne. 2018. *Semantic operations for transfer-based machine translation* dissertation. https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/42432/Goodman_washington_0250E_18405.pdf.

Harris, Zellig Sabbetai. 1954. Distributional structure. *Word* 10. 146–162. Reprinted in: Harris, 1970, *Papers in Structural and Transformational Linguistics*, 775–794, Reidel; Harris, 1981, *Papers on Syntax*, 3-22, Reidel.

Hashimoto, Chikara, Francis Bond, Takaaki Tanaka & Melanie Siegel. 2008. Semi-automatic documentation of an implemented linguistic grammar augmented with a treebank. *Language Resources and Evaluation* 42(2). 117–126.

Hellan, Lars & Dorothee Beermann. 2014. Inducing grammars from IGT. In Zygmunt Vetulani & Joseph Mariani (eds.), *Human language technology challenges for computer science and linguistics*, 538–547. Cham: Springer.

Herring, Joshua. 2016. *Grammar construction in the Minimalist Program.* Indiana University dissertation.

Hershcovich, Daniel, Marco Kuhlmann, Stephan Oepen & Tim O'Gorman. 2019. *Mtool. The Swiss Army Knife of meaning representation.* https://github.com/cfmrp/mtool. Accessed on October 6, 2019.

Horvat, Matic. 2017. *Hierarchical statistical semantic translation and realization.* University of Cambridge dissertation. http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-913.pdf.

Hudson, Dick. 2019. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1073–1076. Berlin: Language Science Press. DOI:??

Ivanova, Angelina, Stephan Oepen, Lilja Øvrelid & Dan Flickinger. 2012. Who did what to whom?: a contrastive study of syntacto-semantic dependencies. In *Proceedings of the 6th Linguistic Annotation Workshop*, 2–11. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W12-3602.

Johnson, Mark. 1988. *Attribute-value logic and the theory of grammar* (CSLI Lecture Notes 16). Stanford, CA: CSLI Publications.

Joshi, Aravind K. 1987. Introduction to Tree Adjoining Grammar. In Alexis Manaster-Ramer (ed.), *The mathematics of language*, 87–114. Amsterdam: John Benjamins Publishing Co.

Kasami, Tadao, Hiroyuki Seki & Mamoru Fujii. 1989. Generalized context-free grammars and multiple context-free grammars. *Systems and Computers in Japan* 20(7). 43–52.

Kay, Martin. 1963. Rules of interpretation. an approach to the problem of computation in the semantics of natural language. In Cicely M. Popplewell (ed.), *Proceedings of IFIP Congress 62*, 318–21. Amsterdam, The Netherlands: North-Holland Publishing Company.

Kay, Martin. 1973. The MIND system. In R. Rustin (ed.), *Courant Computer Science Symposium 8: natural language processing*. Algorithmics Press.

Kim, Jin-Dong, Tomoko Ohta, Sampo Pyysalo, Yoshinobu Kano & Jun'ichi Tsujii. 2009. Overview of BioNLP'09 shared task on event extraction. In *Proceedings of the workshop on current trends in biomedical natural language processing (BioNLP): shared task*, 1–9. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W09-1401.

King, Tracy Holloway. 2016. Theoretical linguistics and grammar engineering as mutually constraining disciplines. In *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 339–359. CSLI Publications.

King, Tracy Holloway, Martin Forst, Jonas Kuhn & Miriam Butt. 2005. The feature space in parallel grammar writing. *Research on Language and Computation, Special Issue on Shared Representations in Multilingual Grammar Engineering* 3(2). 139–163.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Kouylekov, Milen & Stephan Oepen. 2014. RDF triple stores and a custom SPARQL front-end for indexing and searching (very) large semantic networks. In *Proceedings of the 25th International Conference on Computational Linguistics (COLING), system demonstrations*, 90–94. Dublin, Ireland: Dublin City University & Association for Computational Linguistics. http://aclweb.org/anthology/C14-2020.

Kruyt, Johanna G & MWF Dutilh. 1997. A 38 million words Dutch text corpus and its users. *Lexikos* 7. 229–244.

Kubota, Yusuke. 2019. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1013–1069. Berlin: Language Science Press. DOI:??

Kuhnle, Alexander & Ann Copestake. 2018. Deep learning evaluation using deep linguistic processing. In *Proceedings of the workshop on generalization in the age of deep learning*, 17–23. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W18-1003.

Kuthy, Kordula De. 2019. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 709–743. Berlin: Language Science Press. DOI:??

Letcher, Ned. 2018. *Discovering syntactic phenomena with and within precision grammars*. University of Melbourne dissertation.

MacKinlay, Andrew, David Martinez & Timothy Baldwin. 2012. Detecting modification of biomedical events using a deep parsing approach. *BMC Medical Informatics and Decision Making* 12(Supplement 1). S4. http://www.biomedcentral.com/1472-6947/12/S1/S4/. Proceedings of the ACM 5th International Workshop on Data and Text Mining in Biomedical Informatics (DTMBio 2011).

Marcus, Mitchell P., Beatrice Santorini & Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics* 19. 313–330.

Marimon, Montserrat. 2013. The Spanish DELPH-IN grammar. *Language Resources and Evaluation* 47(2). 371–397. DOI:10.1007/s10579-012-9199-7

Marimon, Montserrat. 2015. Tibidabo: a syntactically and semantically annotated corpus of Spanish. *Corpora* 10(3). 259–276.

Matsuzaki, Takuya, Yusuke Miyao & Jun'ichi Tsujii. 2007. Efficient HPSG parsing with supertagging and CFG-filtering. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, 1671–1676. Association for the Advancement of Artificial Intelligence (AAAI). https://www.aaai.org/Papers/IJCAI/2007/IJCAI07-270.pdf.

Melnik, Nurit. 2007. From "hand-written" to computationally implemented HPSG theories. *Research on Language and Computation* 5(2). 199–236.

Meurers, Walt Detmar, Gerald Penn & Frank Richter. 2002. A web-based instructional platform for constraint-based grammar formalisms and parsing. In Dragomir Radev & Chris Brew (eds.), *Effective tools and methodologies for teaching NLP and CL*, 18–25. Association for Computational Linguistics. Proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, PA.

Miyao, Yusuke, Takashi Ninomiya & Jun'ichi Tsujii. 2005. Corpus-oriented grammar development for acquiring a Head-Driven Phrase Structure Grammar from the Penn Treebank. In Keh-Yih Su, Oi Yee Kwong, Jn'ichi Tsujii & Jong-Hyeok Lee (eds.), *Natural language processing IJCNLP 2004* (Lecture Notes in Artificial Intelligence 3248), 684–693. Berlin: Springer Verlag.

Miyao, Yusuke, Rune Sætre, Kenji Sagae, Takuya Matsuzaki & Jun'ichi Tsujii. 2008. Task-oriented evaluation of syntactic parsers and their representations. *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics*. 46–54. http://aclweb.org/anthology/P08-1006.

Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.

Morante, Roser & Eduardo Blanco. 2012. *SEM 2012 shared task: resolving the scope and focus of negation. In *Proceedings of *SEM 2012: the 1st Joint Conference on Lexical and Computational Semantics*, 265–274. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/S12-1035.

Morgado da Costa, Luis, Francis Bond & Xiaoling He. 2016. Syntactic well-formedness diagnosis and error-based coaching in computer assisted language learning using machine translation. In *Proceedings of the 3rd Workshop on Natural Language Processing Techniques for Educational Applications (NLPTEA2016)*, 107–116. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W16-4914.

Müller, Stefan. 1996. The Babel-System: An HPSG fragment for German, a parser, and a dialogue component. In *Proceedings of the Fourth International Conference on the Practical Application of Prolog*, 263–277. London.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2004a. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation* 2(2). 209–257.

Müller, Stefan. 2004b. *Example sentences and making them useful for theoretical and computational linguistics*. Paper presented at the DGfS Jahrestagung: AG Empirische Fundierung der Modellbildung in der Syntax.

Müller, Stefan. 2007. *Head-Driven Phrase Structure Grammar: Eine Einführung*. 1st edn. (Stauffenburg Einführungen 17). Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2009. A Head-Driven Phrase Structure Grammar for Maltese. In Bernard Comrie, Ray Fabri, Beth Hume, Manwel Mifsud, Thomas Stolz & Martine Vanhove (eds.), *Introducing Maltese linguistics: Papers from the 1st International Conference on Maltese Linguistics (Bremen/Germany, 18–20 October, 2007)* (Studies in Language Companion Series 113), 83–112. Amsterdam: John Benjamins Publishing Co.

Müller, Stefan. 2014. Kernigkeit: Anmerkungen zur Kern-Peripherie-Unterscheidung. In Antonio Machicao y Priemer, Andreas Nolda & Athina Sioupi (eds.), *Zwischen Kern und Peripherie* (studia grammatica 76), 25–39. Berlin: de Gruyter.

Müller, Stefan. 2015. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.91

Müller, Stefan. 2019a. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

Müller, Stefan. 2019b. *Grammatical theory: From Transformational Grammar to constraint-based approaches.* 3rd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.5281/zenodo.3364215

Müller, Stefan. 2019c. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

Müller, Stefan & Masood Ghayoomi. 2010. PerGram: A TRALE implementation of an HPSG fragment of Persian. In *Proceedings of 2010 IEEE International Multiconference on Computer Science and Information Technology – Computational Linguistics Applications (CLA'10). Wisła, Poland, 18–20 October 2010*, vol. 5, 461–467. Polnish Information Processing Society.

Müller, Stefan & Walter Kasper. 2000. HPSG analysis of German. In Wolfgang Wahlster (ed.), *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence), 238–253. Berlin: Springer Verlag.

Müller, Stefan & Janna Lipenkova. 2013. ChinGram: A TRALE implementation of an HPSG fragment of Mandarin Chinese. In Huei-ling Lai & Kawai Chui (eds.), *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, 240–249. Taipei, Taiwan: Department of English, National Chengchi University.

Müller, Stefan & Bjarne Ørsnes. 2013. *Danish in Head-Driven Phrase Structure Grammar* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. In preparation.

Muszyńska, Ewa. 2016. Graph- and surface-level sentence chunking. In *Proceedings of the ACL 2016 Student Research Workshop*, 93–99. Berlin, Germany: Association for Computational Linguistics (ACL). http://anthology.aclweb.org/P16-3014.

Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty & Daniel Zeman. 2016. Universal Dependencies v1: a multilingual treebank collection. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*. Portorož, Slovenia: European Language Resources Association (ELRA).

Oepen, Stephan. 2001. *[incr tsdb()] — competence and performance laboratory. user manual*. Technical Report. Saarbrücken, Germany: COLI.

Oepen, Stephan, Omri Abend, Jan Hajič, Daniel Hershcovich, Marco Kuhlmann, Tim O'Gorman, Nianwen Xue, Jayeol Chun, Milan Straka & Zdeňka Urešová. 2019. MRP 2019: Cross-framework Meaning Representation Parsing. In *Proceedings of the shared task on cross-framework meaning representation parsing at the 2019 Conference on Natural Language Learning*, 1–27. Hong Kong, China.

Oepen, Stephan & Daniel P. Flickinger. 1998. Towards systematic grammar profiling: Test suite technology ten years after. *Journal of Computer Speech and Language* 12(4). 411–436. http://www.delph-in.net/itsdb/publications/profiling.ps.gz, accessed 2018-2-25. (Special Issue on Evaluation).

Oepen, Stephan, Daniel P. Flickinger, Kristina Toutanova & Christopher D. Manning. 2004. LinGO Redwoods: A rich and dynamic treebank for HPSG. *Research on Language and Computation* 2(4). 575–596.

Oepen, Stephan, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Silvie Cinková, Dan Flickinger, Jan Hajic & Zdenka Uresova. 2015. SemEval 2015 task 18: broad-coverage semantic dependency parsing. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 915–926. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/S15-2153.

Oepen, Stephan, Marco Kuhlmann, Yusuke Miyao, Daniel Zeman, Dan Flickinger, Jan Hajic, Angelina Ivanova & Yi Zhang. 2014. SemEval 2014 task 8: broad-coverage semantic dependency parsing. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 63–72. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/S14-2008.

Oepen, Stephan & Jan Tore Lønning. 2006. Discriminant-based MRS banking. In *Proceedings of the 5th International Conference on Language Resources and Eval-*

uation (LREC), 1250–1255. European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2006/pdf/364_pdf.pdf.

Oepen, Stephan, Klaus Netter & Judith Klein. 1997. TSNLP — Test suites for natural language processing. In John Nerbonne (ed.), *Linguistic databases*, 13–36. Stanford, CA: CSLI Publications.

Oepen, Stephan, Lilja Øvrelid, Jari Björne, Richard Johansson, Emanuele Lapponi, Filip Ginter & Erik Velldal. 2017. The 2017 Shared Task on Extrinsic Parser Evaluation: towards a reusable community infrastructure. In *Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation, at the 4th International Conference on Dependency Linguistics and the 15th International Conference on Parsing Technologies*, 1–16. Nordic Language Processing Laboratory. http://svn.nlpl.eu/epe/2017/public/proceedings.pdf.

Oepen, Stephan, Erik Velldal, Jan Tore Lønning, Paul Meurer, Victoria Rosén & Dan Flickinger. 2007. Towards hybrid quality-oriented machine translation: On linguistics and probabilities in MT. In *Proceedings of 11th conference on theoretical and methodological issues in machine translation*, 144–153. Skövde, Sweden.

Oostdijk, Nelleke. 2000. The Spoken Dutch Corpus. overview and first evaluation. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC)*. European Language Resources Association (ELRA). http://lrec-conf.org/proceedings/lrec2000/pdf/110.pdf.

Packard, Woodley. 2015. *Full forest treebanking*. University of Washington MA thesis.

Packard, Woodley, Emily M Bender, Jonathon Read, Stephan Oepen & Rebecca Dridan. 2014. Simple negation scope resolution through deep parsing: a semantic solution to a semantic problem. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, long papers*, 69–78. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/P14-1007.

Penn, Gerald. 2004. Balancing clarity and efficiency in typed feature logic through delaying. In Donia Scott (ed.), *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), main volume*, 239–246. Barcelona, Spain.

Petrick, Stanley Roy. 1965. *A recognition procedure for Transformational Grammars*. Massachusetts Institute of Technology. Dept. of Modern Languages dissertation. http://hdl.handle.net/1721.1/13013, accessed 2018-2-25.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Ranta, Aarne. 2009. The GF resource grammar library. *Linguistic Issues in Language Technology (LiLT)* 2. 1–62. http://journals.linguisticsociety.org/elanguage/lilt/article/download/214/214-501-1-PB.pdf.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Reiplinger, Melanie, Ulrich Schäfer & Magdalena Wolska. 2012. Extracting glossary sentences from scholarly articles: a comparative evaluation of pattern bootstrapping and deep analysis. In *Proceedings of the ACL-2012 special workshop on rediscovering 50 years of discoveries*, 55–65. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W12-3206.

Rohde, Douglas LT. 2005. *Tgrep2 user manual, version 1.15*. http://www.cs.cmu.edu/afs/cs.cmu.edu/project/cmt-55/OldFiles/lti/Courses/722/Spring-08/Penn-tbank/Tgrep2/tgrep2_manual.pdf.

Schabes, Yves, Anne Abeillé & Aravind K. Joshi. 1988. *Parsing strategies with 'lexicalized' grammars: Application to Tree Adjoining Grammars*. Technical Report MS-CIS-88-65. University of Pennsylvania Department of Computer & Information Science.

Schäfer, Ulrich, Bernd Kiefer, Christian Spurk, Jörg Steffen & Rui Wang. 2011. The ACL anthology searchbench. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, systems demonstrations*, 7–13. Association for Computational Linguistics (ACL). http://aclweb.org/anthology/P11-4002.

Schäfer, Ulrich, Hans Uszkoreit, Christian Federmann, Torsten Marek & Yajing Zhang. 2008. Extracting and querying relations in scientific papers. In Andreas R. Dengel, Karsten Berns, Thomas M. Breuel, Frank Bomarius & Thomas R. Roth-Berghofer (eds.), *KI 2008: advances in artificial intelligence*, 127–134. Berlin, Heidelberg: Springer.

Schuster, Sebastian, Éric Villemonte de La Clergerie, Marie Candito, Benoît Sagot, Christopher Manning & Djamé Seddah. 2017. Paris and Stanford at EPE 2017: downstream evaluation of graph-based dependency representations. In *Proceedings of the 2017 Shared Task on Extrinsic Parser Evaluation, at the 4th International Conference on Dependency Linguistics and the 15th International Conference on Parsing Technologies*, 47–59. Nordic Language Processing Laboratory. http://svn.nlpl.eu/epe/2017/public/proceedings.pdf.

Siegel, Melanie, Emily M. Bender & Francis Bond. 2016. *Jacy: An implemented grammar of Japanese* (CSLI Studies in Computational Linguistics). Stanford, CA: CSLI Publications.

Slayden, Glenn C. 2012. *Array TFS storage for unification grammars*. University of Washington MA thesis.

Solberg, Lars Jørgen. 2012. *A corpus builder for Wikipedia*. University of Oslo MA thesis. https://www.duo.uio.no/bitstream/handle/10852/34914/thesis.pdf.

Stabler, Edward. 1997. Derivational minimalism. In C. Retoré (ed.), *Logical aspects of computational linguistics* (Lecture Notes in Computer Science 1328), 68–95. Berlin: Springer Verlag.

Steedman, Mark & Jason Baldridge. 2011. Combinatory Categorial Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 181–224. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Suppes, Patrick, Tie Liang, Elizabeth E Macken & Daniel P Flickinger. 2014. Positive technological and negative pre-test-score effects in a four-year assessment of low socioeconomic status K-8 student learning in computer-based math and language arts courses. *Computers & Education* 71. 23–32.

Sygal, Yael & Shuly Wintner. 2011. Towards modular development of typed unification grammars. *Computational Linguistics* 37(1). 29–74.

Torr, John, Milos Stanojevic, Mark Steedman & Shay B. Cohen. 2019. Wide-coverage neural A* parsing for minimalist grammars. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 2486–2505. Florence, Italy: Association for Computational Linguistics (ACL). https://aclweb.org/anthology/P19-1238.

Toutanova, Kristina, Christopher D. Manning, Dan Flickinger & Stephan Oepen. 2005. Stochastic HPSG parse disambiguation using the Redwoods corpus. *Research on Language & Computation* 3(1). 83–105.

Toutanova, Kristina, Christopher D. Manning, Stuart M. Shieber, Dan Flickinger & Stephan Oepen. 2002. Parse disambiguation for a rich HPSG grammar. In *Proceedings of the 1st Workshop on Treebanks and Linguistic Theories (TLT)*. Sozopol, Bulgaria: BulTreeBank Group. http://bultreebank.org/wp-content/uploads/2017/05/paper17.pdf.

Tsuruoka, Yoshimasa, Yusuke Miyao & Jun'ichi Tsujii. 2004. Towards efficient probabilistic HPSG parsing: integrating semantic and syntactic preference to guide the parsing. In *Proceedings of the IJCNLP workshop beyond shallow analyses: formalisms and statistical modeling for deep analyses*. Association for Computational Linguistics (ACL).

van der Beek, Leonoor, Gosse Bouma, Rob Malouf & Gertjan van Noord. 2002. The Alpino Dependency Treebank. In *Computational Linguistics in the Netherlands 2001: selected papers from the twelfth CLIN meeting* (Language and Computers 45), 8–22. Rodopi.

van Noord, Gertjan. 2006. At last parsing is now operational. In *Actes de la 13ème conf'erence sur le Traitement Automatique des Langues Naturelles (TALN)*, 20–42. l'Association pour Traitement Automatique des LAngues (ATALA). http://talnarchives.atala.org/TALN/TALN-2006/taln-2006-invite-002.pdf.

van Noord, Gertjan, Gosse Bouma, Frank van Eynde, Daniël de Kok, Jelmer van der Linde, Ineke Schuurman, Erik Tjong Kim Sang & Vincent Vandeghinste. 2013. Large scale syntactic annotation of written Dutch: Lassy. In Peter Spyns & Jan Odijk (eds.), *Essential speech and language technology for Dutch: results by the STEVIN programme*, 147–164. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-30910-6_9. DOI:10.1007/978-3-642-30910-6_9

van Noord, Gertjan & Robert Malouf. 2005. Wide coverage parsing with stochastic attribute value grammars. Unpublished draft. An earlier version was presented at the IJCNLP workshop *B*eyond Shallow Analyses: Formalisms and statistical modeling for deep analyses.

Velldal, Erik. 2009. *Empirical realization ranking*. University of Oslo, Department of Informatics dissertation.

Velldal, Erik, Lilja Øvrelid, Jonathon Read & Stephan Oepen. 2012. Speculation and negation: rules, rankers, and the role of syntax. *Computational Linguistics* 38(2). 369–410. https://www.mitpressjournals.org/doi/full/10.1162/COLI_a_00126.

Wahlster, Wolfgang (ed.). 2000. *Verbmobil: Foundations of speech-to-speech translation* (Artificial Intelligence). Berlin: Springer Verlag.

Waldron, Benjamin, Ann Copestake, Ulrich Schäfer & Bernd Kiefer. 2006. Preprocessing and tokenisation standards in DELPH-IN tools. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*. Genoa, Italy: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2006/pdf/214_pdf.pdf.

Wax, David. 2014. *Automated grammar engineering for verbal morphology*. University of Washington MA thesis.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Xia, Fei & William D. Lewis. 2007. Multilingual structural projection across interlinear text. In *Proceedings of the 6th Conference of the North American Chapter of the Associaton for Computational Linguistics (NAACL), long papers*, 452–459. Rochester, New York: Association for Computational Linguistics (ACL).

Ytrestøl, Gisle, Dan Flickinger & Stephan Oepen. 2009. Extracting and annotating Wikipedia sub-domains: towards a new eScience community resource. In *Proceedings of the 7th International Workshop on Treebanks and Linguistic Theories (TLT)*, 185–197. Utrecht University. https://dspace.library.uu.nl/handle/1874/296811.

Zamaraeva, Olga, Kristen Howell & Emily M. Bender. 2019. Handling cross-cutting properties in automatic inference of lexical classes: a case study of Chintang. In *Proceedings of the 3rd Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL)*. University of Colorado Boulder. https://scholar.colorado.edu/cgi/viewcontent.cgi?article=1009&context=scil-cmel.

Zamaraeva, Olga, Kristen Howell & Adam Rhine. 2018. Improving feature extraction for pathology reports with precise negation scope detection. In *Proceedings of the 27th International Conference on Computational Linguistics (COLING)*, 3564–3575. International Committee on Computational Linguistics (ICCL). http://aclweb.org/anthology/C18-1302.

Zamaraeva, Olga, František Kratochvíl, Emily M Bender, Fei Xia & Kristen Howell. 2017. Computational support for finding word classes: a case study of Abui. In *Proceedings of the 2nd Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL)*, 130–140. Association for Computational Linguistics (ACL). https://aclweb.org/anthology/W17-0118.

Zhang, Yi & Hans-Ulrich Krieger. 2011. Large-scale corpus-driven PCFG approximation of an HPSG. In *Proceedings of the 12th International Conference on Parsing Technologies*, 198–208. Dublin, Ireland: Association for Computational Linguistics (ACL). http://aclweb.org/anthology/W11-2923.

Zwicky, Arnold M., Joyce Friedman, Barbara C. Hall & Donald E. Walker. 1965. The MITRE syntactic analysis procedure for Transformational Grammars. In *Proceedings – FALL Joint Computer Conference*, 317–326. DOI:10.1109/AFIPS.1965.108

Chapter I

Grammar in dialogue

Andy Lücking
Université de Paris, Goethe-Universität Frankfurt

Jonathan Ginzburg
Université de Paris

Robin Cooper
Göteborgs Universitet

> "It takes two to make a truth."
> Austin (1950: 124, footnote 1)

This chapter portrays some phenomena, technical developments and discussions that are pertinent to analysing natural language use in face-to-face interaction from the perspective of HPSG and closely related frameworks. The use of the CONTEXT attribute in order to cover basic pragmatic meaning aspects is sketched. With regard to the notion of common ground, it is argued how to complement CONTEXT by a dynamic update semantics. Furthermore, this chapter discusses challenges posed by dialogue data such as clarification requests to constrained-based, model-theoretic grammars. Responses to these challenges in terms of a type-theoretical underpinning (TTR, a Type Theory with Records) of both the semantic theory and the grammar formalism are reviewed. Finally, the dialogue theory *KoS* that emerged in this way from work in HPSG is sketched.

1 Introduction

The archaeologists Ann Wesley and Ray Jones are working in an excavation hole, and Ray Jones is looking at the excavation map. Suddenly, Ray discovers a feature

Andy Lücking, Jonathan Ginzburg & Robin Cooper. 2019. Grammar in dialogue. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 841–884. Berlin: Language Science Press. DOI:??

that catches his attention. He turns to his colleague Ann and initiates the following exchange (the example is slightly modified from Goodwin 2003: 222; underlined text is used to indicate overlap, italic comments in double round brackets are used to describe non-verbal actions, numbers in brackets quantify the duration of pauses):

(1)	1.	RAY:	Doctor Wesley?
	2.		(0.7) ((*Ann turns and walks towards Ray*))
	3.	ANN:	EHHH HEHH ((*Cough*))
	4.		Yes Mister Jones.
	5.	RAY:	I was gonna see:
	6.	ANN:	°Eh heh huh huh
	7.		°eh heh huh huh
	8.	RAY:	Uh::m,
	9.	ANN:	Ha huh HHHuh
	10.	RAY:	((*Points with trowel to an item on the map*)) I think I finally found **this** feature ((*looks away from map towards a location in the surrounding*))
	11.		(0.8) Cause I: hit the **nail**
	12.	((*Ann looks at map, Ray looks at Ann, Ann looks at Ray*))	

Contrast the archaeological dialogue from (1) with a third person perspective text on a related topic. In a recent archaeology paper, the excavation of gallery grave Falköping stad 5 is described, among others (Blank et al. 2018: 4):

> During excavation the grave was divided in different sections and layers and the finds were documented in these units. The bone material lacking stratographic and spatial information derives from the top layer […]. Both the antechamber and the chamber contained artefacts as well as human and animal skeletal remains, although most of the material was found in the chamber.

The differences between the archaeological dialogue and the paper are obvious and concern roughly the levels of *medium* (spoken vs. written), *situatedness* (degree of context dependence), *processing speed* (online vs. offline) and *standardisation* (compliance with standard language norms) (Klein 1985). Attributing differences between dialogue and text simply to the medium (i.e. spoken vs. written) is tempting but insufficient. The corresponding characterising features seem to form a continuum, as discussed under the terms *conceptual orality* and *conceptual literacy* in the (mainly German-speaking) literature for some time (Koch &

Oesterreicher 1985). For example, much chat communication, although realised by written inscriptions, exhibits many traits of (conceptually) spoken communication, as investigated, for instance, by means of chat corpora (Beißwenger et al. 2012). Face-to-face dialogue stands out due to a high degree of context dependence manifested in shared attention (Tomasello 1998; see also turns 2 and 12 between Ann and Ray), non-verbal actions such as hand and arm gestures (Kendon 2004; McNeill 2000; turn 10; cf. Lücking 2019, Chapter K of this volume for a brief overview of non-verbal communication means), disfluencies (Ginzburg et al. 2014; turns 5 to 8), non-sentential utterances (Fernández & Ginzburg 2002; Fernández et al. 2007; turns 1, 4, and 5), laughter (Ginzburg et al. 2015; turn 9), shared knowledge of interlocutors (Clark et al. 1983; turns 10–12), turn-taking (Sacks et al. 1974; Heldner & Edlund 2010; Levinson & Torreira 2015; e.g. question-answering in turns 1 and 4) and indirect reference (turn 10, where Ray points to an item on the map but refers to an archaeological artefact in the excavation hole). Note that such instances of deferred reference (Nunberg 1993) in situated communication actually differ from bridging anaphora (Clark 1975) in written texts, although they seem to be closely related at first glance. Bridging is a kind of indirect reference, too, where a definite noun phrase refers back to an antecedent entity which is not given in a strict sense, like *the goalkeeper* in *I watched the football match yesterday. The goalkeeper did an amazing save in overtime*. However, bridging NPs does not give rise to an index or demonstratum, which is the "deferring base" in case of indirect deixis (cf. Lücking 2018).

Since these phenomena are usually abstracted away from the linguistic knowledge encoded by a grammar, linguistics is said to exhibit a "written language bias" (Linell 2005). In fact, many of the phenomena exemplified above provide serious challenges to current linguistic theory, as has been argued by Ginzburg (2012), Ginzburg & Poesio (2016) and Kempson et al. (2016). So the question is: how serious is this bias? Is there a single language system with two modes, written and spoken (but obeying the qualifications we made above with respect to conceptual orality and literacy)? Or do written and spoken communication even realise different language systems? Responses can be given from different standpoints. When the competence/performance distinction was proposed (Chomsky 1965), one could claim that linguistic knowledge is more purely realised by the high degree of standardisation manifested in written text, while speech is more likely to be affected by features attributed to performance (e.g. processing issues such as short term memory limitations or impaired production/perception). Once one attaches more importance to dialogical phenomena, one can also claim that there is a single, basic language system underlying written and spoken communication which bifurcates only in some cases, with interactivity and deixis being salient examples (such a position is delineated but not embraced by Klein 1985; in fact, Klein remains neutral on this issue). Some even claim that "grammar is a system that characterizes talk in interaction" (Ginzburg & Poesio 2016: 1).[1] This position is strengthened by the primacy of spoken language in both ontogenetic and language acquisition areas (on acquisition see Ginzburg 2019, Chapter F of this volume).

Advances in dialogue semantics are compatible with the latter two positions, but their ramifications are inconsistent with the traditional competence/performance distinction (Ginzburg & Poesio 2016; Kempson et al. 2016). Beyond investigating phenomena which are especially related to people engaging in face-to-face interaction, dialogue semantics contributes to the theoretical (re)consideration of the linguistic competence that grammars encode. Some of the challenges posed by dialogue for the notion of linguistic knowledge – exemplified by non-sentential utterances such as clarification questions and reprise fragments (Fernández & Ginzburg 2002; Fernández et al. 2007) – are also main actors in arguing *against* doing semantics within a unification-based framework (like Pollard & Sag 1987) and have implications for doing semantics in constraint-based frameworks (like Pollard & Sag 1994; see Section 3.1 below). In light of this, the relevant arguments are briefly reviewed below. As a consequence, we show how dialogue phenomena can be captured with a framework that leaves "classical" HPSG (i.e. HPSG as documented throughout this handbook). To this end, TTR (a Type Theory with Records) is introduced in Section 3.3. TTR is a strong competitor to other formalisms since it provides an account of semantics that covers dialogue phenomena from the outset. TTR also allows for "emulating" an HPSG kind of grammar, giving rise to a unified home for sign-based SYNSEM interfaces bridging to dialogue gameboards (covered in Section 4). To begin with, however, we give a brief historical review of pragmatics within HPSG.

[1] The sign structure used in HPSG is partly motivated by the bilateral notion of sign of de Saussure. In this respect it is interesting to note that also de Saussure advocated the primacy of spoken language:

> Sprache und Schrift sind zwei verschiedene Systeme von Zeichen; das letztere besteht nur zu dem Zweck, um das erstere darzustellen. Nicht die Verknüpfung von geschriebenem und gesprochenem Wort ist Gegenstand der Sprachwissenschaft, sondern nur das letztere, das gesprochene Wort allein ist ihr Objekt. (de Saussure 2001: 28)
> (*Language and writing are two different systems of signs; the latter exists only for the purpose of representing the former. It is not the combination of the written and the spoken word that is the subject of linguistics, but only the latter, the spoken word alone, is its object.*)

In this respect, de Saussure acts as an early exponent *against* any written language bias.

2 From CONTEXT to update semantics for dialogue

HPSG's interface to pragmatics is the CONTEXT attribute. The CONTEXT attribute accommodates contextual constraints that have to be fulfilled in order for an expression to be used appropriately or felicitously (Austin 1962), to use a term from speech act theory (Pollard & Sag 1994: 27). The CONTEXT attribute has been used and extended to model the content of indexical and pronominal expressions (see Section 2.1), information packaging (Section 2.2) and shared background assumptions concerning standard meanings (Section 2.3). A further step from such pragmatic phenomena to dialogue semantics is achieved by making signs encode their dialogue context, leading to an architectural revision in terms of *update semantics* (see Section 2.4).

2.1 C-INDS and BACKGROUND

The CONTEXT attribute introduces two sub-attributes, CONTEXTUAL-INDICES (C-INDS) and BACKGROUND. The C-INDS attribute values provide pointers to circumstantial features of the utterance situation such as speaker, addressee and time and location of speaking. Within the BACKGROUND attribute, assumptions such as presuppositions or conventional implicatures are expressed in terms of *psoas*, *parameterised state of affairs* (see Section 3.2 for some alternative semantic representation formats). For instance, it is part of the background information of the pronoun *she* of the "natural gender language" English that its referent is female (this does not hold for "grammatical gender languages" like French or German). In the HPSG format of Pollard & Sag (1994: 20), this constraint is expressed as in (2), where "HEAD *noun*[CASE *nom*]" abbreviates a head structure of type *noun* which bears a case attribute with value *nom* (nominative):

(2)
$$\begin{bmatrix} \textit{word} \\ \text{PHON} & \langle \textit{she} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} & \begin{bmatrix} \textit{local} \\ \text{CATEGORY} & \begin{bmatrix}\textit{cat}\\ \text{HEAD} & \textit{noun}[\text{CASE } \textit{nom}] \\ \text{SUBCAT} & \langle\ \rangle\end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix}\textit{ppro} \\ \text{INDEX} & \boxed{1}\begin{bmatrix}\textit{ref}\\ \text{PER} & \textit{3rd}\\ \text{NUM} & \textit{sing} \\ \text{GEN} & \textit{fem}\end{bmatrix} \\ \text{RESTR} & \{\}\end{bmatrix} \\ \text{CONTEXT} & \begin{bmatrix}\textit{context}\\ \text{BACKGROUND} & \left\{\begin{bmatrix}\textit{psoa}\\ \text{RELN} & \textit{female}\\ \text{INST} & \boxed{1}\end{bmatrix}\right\}\end{bmatrix}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

The CONTENT value is of type *ppro* (*personal-pronoun*), which is related to the NP type (+p, −a) from *Government and Binding* theory (Chomsky 1981) and interacts with HPSG's binding theory (see Branco 2019, Chapter 21 of this volume; see also Wechsler 2019, Chapter 6 of this volume). The CONTENT/CONTEXT description in (2) claims that whatever the referent of the pronoun is, it has to be female.

The contextual indices that figure as values for the C-INDS attribute provide semantic values for indexical expressions. For instance, the referential meaning of the singular first person pronoun *I* is obtained by identifying the semantic index with the contextual index "speaker".[2] This use of CONTEXT is illustrated in (3), which is part of the lexical entry of *I*.

[2]There are also indirect uses of *I*, where identification with the circumstantial speaker role would lead to wrong results. An example is the following:

Here it is the truck, not the speaker, or rather the author of the note, that is for rent. Such examples of the German cognate of "I", namely *Ich*, are collected and discussed in Kratzer (1978).

(3)
$$\begin{bmatrix} \textit{word} & \\ \textsc{phon} & \langle\, i \,\rangle \\ \textsc{synsem} & \begin{bmatrix} \textit{synsem} & \\ \textsc{local} & \begin{bmatrix} \textit{local} & \\ \textsc{content} & \begin{bmatrix} \textit{ppro} & \\ \textsc{index} & \boxed{1} \begin{bmatrix} \textit{ref} & \\ \textsc{ref} & \textit{1st} \\ \textsc{num} & \textit{sing} \end{bmatrix} \\ \textsc{restr} & \{\} \end{bmatrix} \\ \textsc{context} & \begin{bmatrix} \textit{context} & \\ \textsc{c-inds} & \begin{bmatrix} \textit{psoa} & \\ \textsc{spkr} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Inasmuch as the contextual anchors (see Barwise & Perry 1983: 72–73 or Devlin 1991: 52–63 on anchors in Situation Semantics) indicated by a boxed notation from (3) provide a semantic value for the speaker in a directly referential manner (see Marcus 1961 and Kripke 1980 on the notion of direct reference with regard to proper names), they also provide semantic values for the addressee (figuring in the content of *you*) as well as the time (*now*) and the place (*here*) of speaking.[3] Hence, the CONTEXT attribute accounts for the standard indexical expressions and provides a present tense marker needed for a semantics of tenses along the lines of *Discourse Representation Theory* (Kamp & Reyle 1993; see Partee 1973 on the preeminent role of an indexical time point). We will not discuss this issue further here (see Van Eynde 1998; 2000, Bonami 2002 and Costa & Branco 2012 for HPSG work on tense and aspect), but move on to briefly recapture other phenomena usually ascribed to pragmatics (see also Kathol et al. 2011: Section 5.2).

2.2 Information structure

Focus, expressed by sentence accent in English, can be used for information packaging that may lead to truth-conditional differences even when the surface structures (i.e. strings; see Section 1 on a brief juxtaposition of spoken and written language) are the same (Halliday 1967). An example is given in (4), taken from Krifka (2008: 246), where capitalisation indicates main accent and subscript "F" labels the focused constituent (see also Wasow 2019, Chapter G of this volume on incremental processing also with respect to aspects of information structure):

[3]Of these, in fact, only the speaker is straightforwardly given by the context; all others can potentially involve complex inference.

(4) a. John only showed Mary [the PICTures]$_F$.
 b. John only showed [MARY]$_F$ the pictures.

An analysis of examples like (4) draws on an interplay of phonology, semantics, pragmatics and constituency and hence emphasises in particular the advantages of the *fractal* architecture of HPSG (Johnson & Lappin 1999). HPSG has the fractal property since information about phonetic, syntactic and semantic aspects is present in every sign, from words to phrases and clauses (Pollard 1997: 5) – see also Kubota (2019), Chapter M of this volume, Borsley & Müller (2019), Chapter L of this volume, Müller (2019), Chapter P of this volume, Arnold (2019), Chapter N of this volume and Hudson (2019), Chapter O of this volume for a comparison of HPSG to other grammar theories; a benchmark source is Müller (2016).

At the core of information structure is a distinction between *given* and *new* information. Accordingly, information structure is often explicated in terms of dynamic semantics (ranging from *File Change Semantics* by Heim 2002 and *Discourse Representation Theory* by Kamp & Reyle 1993 to information state update semantics proper by Traum & Larsson 2003) – see for instance Krifka (2008) or Vallduví (2016) for a discussion and distinction of various notions bound up with information structure such as *focus*, *topic*, *ground* and *comment* seen from the perspective of dialogue content and dialogue management. The most influential approach to information structure within HPSG is that of Engdahl & Vallduví (1996). Here a distinction between *focus*, that is, new information, and *ground*, the given information, is made (Engdahl & Vallduví 1996: 3). The *ground* is further bifurcated into LINK and TAIL, which connect to the preceding discourse in different ways (basically, the link corresponds to a discourse referent or file, and the tail corresponds to a predication which is already subsumed by the interlocutors' information states). The information packaging of the content values of a sentence is driven by phonetic information in terms of A-accent and B-accent (Jackendoff 1972: Chapter 6), where "A-stressed" constituents are coindexed with FOCUS elements and "B-stressed" are coindexed with LINK elements – see also Kuthy (2019), Chapter D of this volume. The CONTEXT extension for information structure on this account is given in (5):

(5)
$$\begin{bmatrix} \textit{context} & \\ \text{C-INDS} & [\,] \\ \text{BACKGR} & \{\} \\ \text{INFO-STR} & \begin{bmatrix} \textit{info-struc} & \\ \text{FOCUS} & \textit{set(content)} \\ \text{GROUND} & \begin{bmatrix} \textit{ground} & \\ \text{LINK} & \textit{set(content)} \\ \text{TAIL} & \textit{set(content)} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Part of the analysis of the sample sentences from (4) is that in (4a), the CONTENT value of the indirect object NP *the pictures* is the focused constituent, while it is the CONTENT value of the direct object NP *Mary* in (4b). The FOCUS-LINK-TAIL approach works via structure sharing: the values of FOCUS, LINK and TAIL get instantiated by whatever means the language under consideration uses in order to tie up information packages (whether syntactic, phonological or something else besides). If prosodic information is utilised for signalling information structure, a grammar has to account for the fact that prosodic constituency is not isomorphic to syntactic constituency, that is, prosodic structures cannot be built up in parallel to syntactic trees. Within HPSG, the approach to *prosodic constituency* of Klein (2000) employs *metrical trees* independent from syntactic trees, but grammatic composition remains syntax-driven. The latter assumption is given up in the work of Haji-Abdolhosseini (2003). Starting from Klein's work, an architecture is developed that generalises over prosody-syntax mismatches: on this account, syntax, phonology and information structure are parallel features of a common list of domain objects (usually the inflected word forms). Information structure realised by prosodic stress is also part of the speech-gesture interfaces within multimodal extensions of HPSG (cf. Lücking 2019, Chapter K of this volume).

2.3 Mutual beliefs

A strictly pragmatic view on meaning and reference is presented by Green (1996). Green provides a CONTEXT extension for the view that restrictions on the index actually are background assumptions concerning standard uses of referential expressions. One of the underlying observations is that people can, for example, use the word *dog* to refer to, say, toy dogs or even, given appropriate context information, to a remote control (we will come back to this example shortly). The fact that the word *dog* can be used without further ado successfully to refer to instances of the subspecies *Canis lupus familiaris*[4] is due to shared assumptions

[4] Green (1996: Example (73)) actually restricts the standard use of *dog* to the family *Canis* (regiven in our example (6)), which seems to be too permissive. The *Canis* family also include

about the standard meaning of *dog*. Green represents this account in terms of mutual beliefs between EXPERIENCER and STANDARD as part of the background condition of the CONTEXT of referential NPs. Drawing on work by Cohen & Levesque (1990), *mutually-believe* is a recursive relation such that the experiencer believes a proposition, believes that the standard believes the proposition too, believes that the standard believes that the experiencer believes the proposition, and so on. When a proposition is mutually believed within a speech community, it is *normally believed.* The semantic part of the lexical structure of *dog* is given in (6). The analysis of proper names is pursued in a similar manner, amounting to the requirement that for a successful use of a proper name, the interlocutors have to know that the intended referent of this name actually bears the name in question.

(6)
$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & \langle\ \textit{dog}\ \rangle \\ \text{CONTENT} & \begin{bmatrix} \textit{ref} & \\ \text{INDEX} & \boxed{1} \end{bmatrix} \\ \text{CONTEXT} & \begin{bmatrix} \textit{context} & \\ \text{C-INDS} & \begin{bmatrix} \textit{psoa} & \\ \text{SPKR} & \boxed{2} \\ \text{ADDR} & \boxed{3} \end{bmatrix} \\ \text{BACKGROUND} & \left\{ \begin{bmatrix} \textit{mutually-believe} & \\ \text{EXPERIENCER} & \boxed{2} \\ \text{STANDARD} & \boxed{3} \\ \text{SOA} & \begin{bmatrix} \textit{normally-believe} & \\ \text{EXPERIENCER} & \textit{English speakers} \\ \text{SOA} & \begin{bmatrix} \textit{canis} & \\ \text{INST} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$

Adding beliefs to CONTEXT provides the representational means to integrate (at least some kinds of) presuppositions, illocutionary force and deferred reference (Nunberg 1978) into grammar. However, a fuller model of speech acts and meaning transfers is still needed (Kathol et al. 2011: 94).

Taking a closer look at the argument underlying adding mutual beliefs to CONTEXT, one notices a striking similarity of shared assumptions about standard uses with *community membership* as a source for common ground (but see Footnote 4 for a hint on a possible refinement). However, community membership is just one of three sources of information on which the common ground between two

foxes, coyotes and wolves, which are, outside of biological contexts, usually not described as being dogs. This indicates that the EXPERIENCER group should be further restricted and allowed to vary over different language communities and genres.

interlocutors (scaling up to multilogue is obvious) can be based, according to Clark & Marshall (1981) and Clark et al. (1983):

> The first is *perceptual evidence*, what the two have jointly experienced or are jointly experiencing at the moment. The second is *linguistic evidence*, what the two have jointly heard said or are now jointly hearing as participants in the same conversation. The third is *community membership*. They take as common ground everything they believe is universally, or almost universally, known, believed, or supposed in the many communities and subcommunities to which they mutually believe they both belong. (Clark et al. 1983: 247)

Reconsidering the "*dog*-used-to-refer-to-remote-control" example mentioned above: in order for this kind of reference to happen, one can imagine a preparatory sequence like the following:

(7) Can you please give me the … what's the name? … the … ah, let's call it "dog" … can you please give me the dog?

In this monologue, the speaker establishes a name for the remote control. After this name-giving, the situationally re-coined term can be used referentially (see Lücking et al. 2006 on situated conventions). Obviously, the felicity of reference is due to *linguistic evidence* provided and agreed upon in dialogical exchange. Dialogue contexts (Lee-Goldman 2011) and the dynamics of common ground is a dimension which is absent in the static CONTEXT representations surveyed above. This is where dynamic update semantics enters the stage.

2.4 Towards an update semantics for dialogue

Starting from Stalnakerian contexts (Stalnaker 1978; see also Lewis 1979), that is, contexts which consist of mutually known propositions (also corresponding roughly to the mutual belief structures employed by Green 1996, cf. Section 2.3), Ginzburg argues in a series of works that this context actually has a more elaborate structure (Ginzburg 1994; 1996; 1997). One motivation for this refinement is found in data like (8), an example given by Ginzburg (1994: 2) from the London-Lund corpus (Svartvik 1990).

(8)

1.	A:	I've been at university.
2.	B:	Which university?
3.	A:	Cambridge.
4.	B:	Cambridge, um.

5.   what did you read?
6. A: History and English.
7. B: History and English.

There is nothing remarkable about this dialogical exchange; it is a mundane piece of natural language interaction. However, given standard semantic assumptions and a *given-new* information structuring as sketched in Section 2.2, (8) poses two problems. The first problem is that one and the same word, namely *Cambridge*, plays a different role in different contexts as exemplified by turns 2 to 3 on the one hand and turns 3 to 4 on the other hand. The reason is that the first case instantiates a question-answering pair, where *Cambridge* provides the requested referent. The second case is an instance of *accept*: speaker B not only signals that she heard what A said (what is called *acknowledge*), but also that she updates her information state with a new piece of information (namely that A studied in Cambridge).

The second problem is that neither of B's turns 4 and 7 is redundant, although neither of them contribute new information (or *foci*) in the information-structural sense of Section 2.2: the turns just consist of a replication of A's answer. The reason for non-redundancy obviously is that in both cases the repetition manifests an *accept* move in the sense just explained.

In order to make grammatical sense out of such dialogue data – eventually in terms of linguistic competence – contextual background rooted in language is insufficient, as discussed. The additional context structure required to differentiate the desired interpretation of (8) from redundant and co-text-insensitive ones is informally summarised by Ginzburg (1994: 4) in the following way:

- FACTS: a set of commonly agreed upon facts;
- QUD ("question under discussion"): a partially ordered set that specifies the currently discussable questions. If *q* is topmost in QUD, it is permissible to provide any information specific to *q*.
- LATEST-MOVE: the content of *latest move* made: it is permissible to make whatever moves are available as reactions to the latest move.

Intuitively, turn 2 from the question-answer pair in turns 2 and 3 from (8) directly introduces a *question under discussion* – a semi-formal analysis is postponed to Section 4, which introduces the required background notions of *dialogue gameboards* and *conversational rules* which regiment dialogue gameboard updating. Given that in this case the *latest move* is a question, turn 3 is interpreted as an answer relating to the most recent question under discussion. This answer, however, is not simply added to the dialogue partners' common knowledge, that is, the *facts*. Rather, the receiver of the answer first has to *accept* the response offered to him – this is the dialogue reading of "It takes two to make a truth". After acceptance, the answer can be *grounded* (see Clark 1996: Chapter 4 for a discussion of common ground), that is, *facts* is *updated* with the proposition bound up with the given answer, the resolved question under discussion is removed from the QUD list (*downdating*) – in a nutshell, this basic mechanism is also the motor of the dialogue progressing. This mechanism entails an additional qualification compared to a static mutual belief context: dialogue update does not abstract over the individual dialogue partners. A dialogue move does not present the same content to each of the dialogue partners, nor does the occurrence of a move lead automatically to an update of the common ground (or mutual beliefs). Dialogue semantics accounts for this fact by distinguishing *public* from *private* information. Public information consists of observable linguistic behaviour and its conventional interpretations, collected under the notion of *dialogue gameboard* (DGB). The DGB can be traced back to the *commitment-stores* of Hamblin (1970) that keep track of the commitments made at each turn by each speaker.

Private information is private since it corresponds to interlocutors' mental states (MS). The final ingredient is that the (fourfold) dynamics between the interlocutors' dialogue game boards and mental states unfolds in time, turn by turn. In sum, a minimal participant-sensitive model of dialogue contributions is a tuple of DGB and MS series of the form $\langle \text{DGB} \times \text{MS} \rangle^{+}$ for each dialogue agent. Here the tuple represents a temporarily ordered sequence of objects of a given type (i.e. DGB and MS in case of dialogue agents' information state models) which is witnessed by a *string* of respective events which at least of length 1, as required by the "Kleene +" (see Cooper & Ginzburg 2015: Section 2.7 on a type-theoretical variant of the string theory of events of Fernando 2011).

Guided by a few dialogue-specific semantic phenomena, we moved from various extensions to CONTEXT to minimal participant models and updating/downdating dynamics. In Sections 3 and 4, further progress which mainly consists of inverting the theory's strategic orientation is reviewed: instead of extending HPSG in order to cover pragmatics and dialogue semantics, it is argued that there are reasons to start with an interactive semantic framework and then embed an HPSG variant therein.

In order to move on, a remaining issue has to be resolved: what happens if an addressee for some reason refuses to accept a contribution of the previous speaker? In this case, the addressee (now taking the speaker role) poses a *clarifi-*

cation request. Clarification potential plays an important methodological role in the dialogue semantic business, as exemplified in Section 3.1 below.

3 Type-theoretical pragmatics and dialogue semantics

A minimal primer for the rich type theory TTR is given in Section 3.3. But why should (dialogue) semantics make use of a type theory at all? In what follows, two sources of motivation are presented, the one drawing on semantic data gained from the clarification potential of reprise fragments (Section 3.1), the other resulting from HPSG's struggle with connecting to semantic theories (Section 3.2).

3.1 Subsentential meanings: unification and constraint-satisfaction vs. reprise content

In (9), B poses a clarification request in terms of a reprise fragment concerning the verb used by A (Ginzburg 2012: 115):

(9)
1. A: Did Bo finagle a raise?
2. B: Finagle?

The reprise fragment has at least two interpretations: it can query the phonetic component of the verb ("did I hear correctly that you said 'finagle'?"), or it can query the meaning of the verb ("what does 'finagle' mean?"). Both queried aspects are available as part of the PHON-SYNSEM structure of signs, emphasizing the significance of HPSG's fractal design (cf. the remark on fractality in Section 2.2). However, when B uses the reprise fragment to clarify the content of the expression reprised, then B queries *only* the meaning of the reprised fragment (Purver & Ginzburg 2004; Ginzburg & Purver 2012) – in our example (9), this is *finagle*. This can be seen when answers are given that target the head verb or the verb phrase (head verb plus direct object argument *a raise*):

(10) Finagle?
   a. Yeah, like wangle.
   b. Yeah, he wangled a wage increase.

From the continuations in (10) only the first one provides an answer to B's clarification question in (9). The second continuation can also answer a clarification request, but this clarification request is *finagle a raise?* That is, "[a] nominal fragment reprise question queries exactly the standard semantic content of the fragment being reprised", which is the strong version of the *Reprise Content Hypothesis* put forth by Purver & Ginzburg (2004: 288).[5] In case of the example given in (9), the content of the head verb is queried, and not the meaning of the verb phrase (verb plus direct object) or the sentence (verb plus direct object and subject), since they correspond to constructions that are larger than the reprised fragment. In other words, a reprise fragment allows us to access the meaning of any expression regardless of its syntactic degree of embedding. However, this is not what follows from unification-based semantics. Due to structure sharing, certain slots of a head are *identified* with semantic contributions of modifier or argument constructions (see Wechsler, Koenig & Davis 2019, Chapter 9 of this volume). In the case of *finagle a raise*, this means that once the content of the VP is composed, the patient role (or whatever semantic composition means are employed – see Koenig & Richter 2019, Chapter C of this volume for an overview) of the verb *finagle* is instantiated by the semantic index contributed by *a raise*. At this stage one cannot recover the V content from the VP content – unification appears to be too strong a mechanism to provide contents at all levels as required by reprise fragments.

However, as Richter (2000: Chapter 2) argues, unification is only required in order to provide a formal foundation for the *language-as-partial-information* paradigm of Pollard & Sag (1987) and its spin-offs. The *language-as-collection-of-total-objects* paradigm underlying Pollard & Sag (1994) and its derivatives is not in need of employing unification. Rather, grammars following this paradigm are model-theoretic, constraint-based grammars, resting on *Relational Speciate Re-entrant Language* (RSRL) as formal foundation (Richter 2000 via precursors like Penn 1999). The formalism RSRL in its most recent implementation (Richter 2004) has the advantage that the models it describes can be interpreted in different ways.[6] On the one hand, it is compatible with the idea that grammars accumulate constraints that describe classes of (well-formed) linguistic objects, which in turn classify models of linguistic tokens (King 1999). On the other hand, it is compatible with the view that grammars describe linguistic types, where types are construed as equivalence classes of utterance tokens (Pollard 1999). On these accounts, a related argument applies nonetheless: once the constraints are accumulated that describe total objects with the PHON string *finagle a raise*, the superset of total objects corresponding to just *finagle* is not available any more. The implications of clarification data for any kind of grammar, in particular for

[5]The weak version (Purver & Ginzburg 2004: 287) only claims that a nominal fragment reprise question queries a part of the standard semantic content of the fragment being reprised.
[6]Richter (2019) *p.c.*; see also Richter (2019), Chapter 3 of this volume.

semantics, seem to be that some mechanism is needed that keeps track of the semantic contribution of each constituent of complex linguistic objects such as the verb *finagle* within the verb phrase *finagle a raise*. We do not know of any such attempts within constraint-based grammars and of the possible formal intricacies that may be involved, however. In the following, therefore, the HPSG$_{\text{TTR}}$/KoS framework that provides trackable constituents by means of labelled representations and a dialogue gameboard architecture is introduced. We should emphasize to the reader that at this point we leave the formal background of standard HPSG as documented in this book. We want to point this out since the subsequently-used representations look deceptively similar to attribute-value matrices (the risk of confusion is known from the essentially identical representations employed within unification- and constraint-based HPSG variants). We see this as a consequence of the dynamics of theories when their empirical domain is extended; at best, it adds to the formal and conceptual controversies and developments that take place in HPSG anyway, as briefly sketched in the beginning of this paragraph. However, HPSG$_{\text{TTR}}$ aims at adopting most of HPSG's desirable features such as its fractal architecture, its sign-based set-up and its linking facility between different layers of grammatical description. To begin with, we want to further motivate the point of departure in terms of HPSG's semantic objects.

3.2 Semantic objects: data structures vs. types

Aiming at a declarative characterisation of natural languages, the model theoretic set-up of HPSG has to define models for its domain of linguistic objects (Levine & Meurers 2006: Section 3; see also Richter 2019, Chapter 3 of this volume). In particular with regard to the values of the CONTENT and CONTEXT attribute, the crucial question is "how types in the [feature] logic should correspond to the semantic types being represented" (Penn 2000: 70). In order to provide an answer to this crucial question, one has to clarify what a semantic type is. This question, however, is perhaps even more far-reaching and intricate than the initial one and following it further would lead us to undertake a considerable diversion and probably even turn away from the actual point of the initial question (but for a recent related discussion on the status of propositions see King et al. 2014). A pragmatic interpretation of the crucial question probably is this: how do the types in the feature logic correspond to the semantic types employed in semantic theories? There is a justification for this restatement from the actual semantic practice in HPSG (cf. Koenig & Richter 2019, Chapter C of this volume).

For the purpose of the present discussion, a semantic theory can be conceived as consisting of two components, *semantic representations* and an extensional *domain* or *universe* within which the semantic representations are interpreted (Zimmermann 2011; Kempson 2011). That is, another reformulation of the question is how the HPSG model theory is related to a semantic model theory. Further concreteness can be obtained by realising that both kinds of theories aim to talk about the same extensional domain. Given this, the question becomes: how do HPSG's semantic representations correspond to the semantic representation of the semantic theory of choice? A closely related point is made by Penn (2000: 63): "A model-theoretic denotation could be constructed so that nodes, for example, are interpreted in a very heterogeneous universe of entities in the world, functions on those entities, abstract properties that they may have such as number and gender and whatever else is necessary – the model theories that currently exist for typed feature structures permit that […]". Formulating things in this way has a further advantage: the question is independent from other and diverging basic model theoretic assumptions made in various versions of HPSG, namely whether the linguistic objects to model are types (Pollard & Sag 1994) or tokens (Pollard & Sag 1987) and whether they are total objects (Pollard & Sag 1994) or partial information (Carpenter 1992). However, such a semantic model-theoretic denotation of nodes is not available in many of the most influential versions of HPSG: the semantic structures of the HPSG version developed by Pollard & Sag (1994) rests on a situation-theoretic framework. However, the (parameterised) states of affairs used as semantic representations lack a direct model-theoretic interpretation; they have to be translated into situation-theoretic formulæ first (such a translation from typed feature structures to situation theory is developed by Ginzburg & Sag 2000). That is, the semantic structures do not encode semantic entities; rather they are data structures that represent descriptions which in turn correspond to semantic objects. This is also the conclusion drawn by Penn. The quotation given above continues: "[…] but at that point feature structures are not being used as a formal device to represent knowledge but as a formal device to represent data structures that encode formal devices to represent knowledge" (Penn 2000: 63; see also the discussion given by Ginzburg 2012: Section 5.2.2).

There are two options in order to unite typed feature structures and semantic representations. The first is to use logical forms instead of (P)SOAs and by this means connect directly to truth-conditional semantics. This option makes use of what Penn (see above) calls a *heterogeneous universe*, since syntactic attributes receive a different extensional interpretation than semantic attributes (now consisting of first or second order logic formulæ). The second option is to resort to a homogeneous universe and take PHON-SYNSEM structures as objects in the world, as is done in type-theoretical frameworks – signs nonetheless stand out from ordinary objects due to their CONT part, which makes them representational entities in the first place.

The first option, using logical forms instead of situation-semantic (P)SOAS, was initiated by Nerbonne (1992). The most fully worked out semantics for HPSG from this strand has been developed by Richter and Sailer, by providing a mechanism to use the higher-order Ty2 language for semantic descriptions (Richter & Sailer 1999). This approach has been worked out in terms of *Lexical Resource Semantics* (LRS) where logical forms are constructed in parallel with attribute-value matrices (Richter & Sailer 2004).

At this point we should insert a word on HPSG's most popular underspecification mechanism, namely *(Robust) Minimal Recursion Semantics* (Copestake, Flickinger, Pollard & Sag 2005; Copestake 2007). (R)MRS formulæ may have unfilled argument slots so that they can be assembled in various ways. However, resolving such underspecified representations is not part of the grammar formalism, so (R)MRS representations do not provide an autonomous semantic component for HPSG. Therefore, they do not address the representation problem under discussion as LRS does.

The second option, using the type-theoretical framework TTR, has been developed by Cooper (2008; 2014; 2019) and Ginzburg (2012). TTR, though looking similar to feature structures, directly provides semantic entities, namely types (Ginzburg 2012: Sec. 5.2.2). TTR also has a model-theoretic foundation (Cooper 2019), so it complies with the representation-domain format we drew upon above.

Turning back to the issue discussed in Section 3.1, there is a difference between the two semantic options. Relevant observations are reported by Purver & Ginzburg (2004) concerning the clarification potential of noun phrases. They discuss data like the following (bold face added):

(11) a. TERRY: Richard hit the ball on the car.
NICK: **What ball**? *[⇝ What ball do you mean by 'the ball'?]*
TERRY: James [last name]'s football.

(BNC file KR2, sentences 862, 865–866)

b. RICHARD: No I'll commute every day
ANON 6: **Every day?** *[⇝ Is it **every** day you'll commute?]*
*[⇝ Is it every **day** you'll commute?]*
*[⇝ Which days do you mean by **every day**?]*
RICHARD: as if, er Saturday and Sunday
ANON 6: And all holidays?
RICHARD: Yeah [pause]

As testified in (11), the accepted answers which are given to the clarification requests are in terms of an *individual* with regard to *the ball* (11a) and in terms of *sets* with regard to *every day* in (11b). The expressions put to a clarification request (*the ball* and *every day*, respectively) are analysed as *generalised quantifiers* in semantics (Montague 1974). A generalised quantifier, however, denotes a *set of sets*, which is at odds with its clarification potential in dialogue. Accordingly, in a series of works, a theory of quantified noun phrases (QNPs) has been developed that refrains from type raising and that analyses QNPs in terms of the intuitively expected and clarificationally required denotations of types *individual* and *sets of individuals*, respectively (Purver & Ginzburg 2004; Ginzburg & Purver 2012; Ginzburg 2012; Cooper 2013; Lücking & Ginzburg 2018; Cooper 2019). Since this dialogue-friendly improvement has been given in terms of the second, type-theoretical option and is lacking in the first, logical form-based option (which usually involves generalised quantifier analyses), there is an empirical advantage for the former over the latter, at least from a pragmatic, dialogue semantics viewpoint.

There are further distinguishing features, however. Types are intensional entities, so they directly provide belief objects, as touched upon in Section 2.3, which are needed for intensional readings as figuring in attitude reports such as in the assertion that *Flat Earthers believe that the earth is flat* (see also Cooper 2005a and Cooper 2019 on attitude reports in TTR).

Furthermore, TTR is not susceptible to the *slingshot argument* (Barwise & Perry 1983: 24–26): explicating propositional content on a Fregean account (Frege 1892) – that is, denoting the true or the false – in terms of sets of possible worlds is too coarse-grained, since two sentences which are both true (or false) but have nonetheless different meanings cannot be distinguished. In this regard, TTR provides a *structured theory of meaning*, where types are not traded for their extensions. Accordingly, a brief introduction to TTR is given in Section 3.3 and the architecture of the dialogue theory *KoS* incorporating a type-theoretic HPSG variant is sketched in Section 4.

3.3 A brief primer on TTR

TTR, which builds on ideas in the intuitionistic Type Theory of Martin-Löf (1984) and its application to natural language semantics (see Ranta 2015), provides semantic objects at both the token and the type level and structures to organise these objects, namely records and record types (see Cooper 2005b, Cooper 2005a, Cooper 2012, Cooper 2017, and Cooper & Ginzburg 2015 for expositions). Records consist of fields of pairs of labels and objects, and record types consist of fields of pairs of labels and types, which both can be nested (Cooper 2019). Take for instance the schematic record in (12):

(12) $$\begin{bmatrix} l_0 & = & \begin{bmatrix} l_1 & = & o_1 \\ l_2 & = & o_2 \end{bmatrix} \\ l_3 & = & o_3 \end{bmatrix}$$

Here, o_1, o_2 and o_3 are (real-world) objects, which are labelled by l_1, l_2 and l_3, respectively (o_1 and o_2 are additionally part of a sub-record labelled l_0). Records can be *witnesses* for record types. For instance, the record from (12) is a witness for the record type in (13) only in the case that the objects from the record are of the type required by the record type (i.e. $o_1 : T_1$, $o_2 : T_2$, $o_3 : T_3$), where objects and types are paired by same labelling.

(13) $$\begin{bmatrix} l_0 & : & \begin{bmatrix} l_1 & : & T_1 \\ l_2 & : & T_2 \end{bmatrix} \\ l_3 & : & T_3 \end{bmatrix}$$

The colon notation indicates a basic notion in TTR: a *judgement*. A judgement of the form $a : T$ means that object a is of type T, or, put differently, that a is a witness for T. Judgements are used to capture basic classifications like *Marc Chagall is an individual* (*mc* : *Ind*), as well as propositional descriptions of situations like *The cat is on the mat* for the situation depicted in Figure 1, where Fritz the cat sits on mat m33. The record type for the example sentence (ignoring the semantic contribution of the definite article for the sake of exposition[7]) will be (14):

(14) $$\begin{bmatrix} \text{x} & : & \mathit{Ind} \\ \text{c1} & : & \mathit{cat(x)} \\ \text{y} & : & \mathit{Ind} \\ \text{c2} & : & \mathit{mat(y)} \\ \text{c3} & : & \mathit{on(x,y)} \end{bmatrix}$$

Note that the types labelled "c1", "c2", and "c3" in (14) are *dependent types*, since the veridicality of judgements involving these types depends on the objects that are assigned to the basic types labelled "x" and "y". A *witness* for the record type in (14) will be a *record* that provides suitable objects for each field of the record type (and possibly more). Obviously, the situation depicted in Figure 1 (adapted from Lücking 2018: 270) is a witness for the type in (14). The participants of the depicted situation can be thought of as situations themselves which show Fritz to be a cat, m33 to be a mat and Fritz to be on m33. The scene in the figure then

[7] This record type corresponds to *a cat is on a mat.*

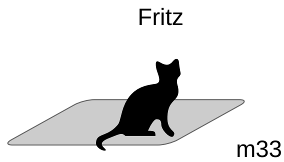


Figure 1: Fritz the cat sits on a mat.

corresponds to the following record, which is of the type expressed by the record type from (14):

(15) $\begin{bmatrix} \textsc{x} & = & \textit{Fritz} \\ \textsc{c1} & = & \textit{cat situation} \\ \textsc{y} & = & \textit{m33} \\ \textsc{c2} & = & \textit{mat situation} \\ \textsc{c3} & = & \textit{relation situation} \end{bmatrix}$

Using type constructors, various types can be build out of basic and complex (dependent) types, such as set types and list types. In order to provide two (slightly simplified) examples of type constructors that will be useful later on, we just mention *function types* and *singleton types* here.

(16) Function type

a. If T_1 and T_2 are types, then $(T_1 \rightarrow T_2)$ is a type, namely the type of functions that map T_1 to T_2.

b. If a function f is of type $(T_1 \rightarrow T_2)$ then f's domain is $\{a \mid a : T_1\}$ and its range is included in $\{a \mid a : T_2\}$.

The characterisation in (16) is that of a standard extensional notion of function. Given that TTR is an intensional semantic theory – that is, two types are different even if their extension is the same – other notions of function types could be developed.

(17) Singleton type

a. If T is a type and $a : T'$ (i.e. object a is of type T'), then T_a is a type.

b. $b : T_a$ (i.e. object b is of type T_a) iff $b : T$ and $b = a$.

That is, a singleton type is singleton since it is the type of specific object.

Since types are semantic objects in their own right (types are not defined by or reduced to their extensions), not only an object o of type T can be the value of a label, but also type T itself. One way of expressing this is in terms of *manifest* *fields*. A type-manifest field is notated in the following way: $l = T : T'$, specifying that l is the type T. Analogously, object-manifest fields can be expressed by restricting the value of a label to a certain object.

For more comprehensive and formal elaborations of TTR, see the references given at the beginning of this section, in particular Cooper (2019).

4 Putting things together: HPSG$_{\text{TTR}}$ and dialogue game boards

Signs as construed within HPSG can be reconstructed as record types of a specific kind (Cooper 2008). For instance, (18) shows the record type (the judgement colon indicates that we now talk about TTR objects) for a general sign according to Pollard & Sag (1994) (where *PhonType*, *CategoryType* and *SemType* denote obvious types – see the Appendix for a minimal HPSG fragment defined in terms of TTR).

(18) $$\begin{bmatrix} \textsc{phon} & : & \textit{list(PhonType)} \\ \textsc{synsem} & : & \begin{bmatrix} \textsc{local} & : & \begin{bmatrix} \textsc{cat} & : & \textit{CategoryType} \\ \textsc{content} & : & \textit{SemType} \\ \textsc{context} & : & \textit{SemType} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Signs are extended by an interface to circumstantial features of the utterance situation in terms of the DGB-PARAMS attribute, which corresponds to the C-INDS from Section 2.1. The attribute's name abbreviates *dialogue gameboard parameters* , since its values have to be instantiated (that is, witnessed) in the process of grounding. Thus, if the content of an NP α is part of DGB-PARAMS, then α gets a referential interpretation. However, NPs need not be used referentially; there are what Donnellan (1966) calls *attributive uses* as in *The thief (whoever he is) stole my credit card*. To this end, there is a "coercion" operation from DGB-PARAMS to Q-PARAMS (*quantificational parameters*) involving an abstraction from individuals to α's descriptive condition (Purver & Ginzburg 2004; see the Appendix for the respective operation).

These HPSG$_{\text{TTR}}$ signs figure as constituents within an architecture known as *dialogue gameboard*, giving rise to a grammar-dialogue interface within the dialogue theory *KoS* (Ginzburg 1994; 1996; 2003; 2012). A Dialogue Game Board (DGB) is an information-state based sheet for describing communicative interactions. The DGB from KoS tracks the interlocutors (*spkr* and *addr* fields), a record of the dialog history (*Moves*), dialogue moves that are in the process of grounding (*Pending*), the question(s) currently under discussion (*QUD*), the assumptions shared among the interlocutors (*Facts*) and the dialogue participant's view of the visual situation and attended entities (*VisualSit*). The TTR representation of a DGB following Ginzburg (2012) is given in (19), where *LocProp* is the type of a *locutionary proposition* (see (21) below) and *poset* abbreviates "partially ordered set".

(19)
$$\begin{bmatrix} \textsc{spkr} & : & \mathit{Ind} \\ \textsc{addr} & : & \mathit{Ind} \\ \textsc{utt-time} & : & \mathit{Time} \\ \textsc{c-utt} & : & \mathit{addressing(spkr, addr, utt\text{-}time)} \\ \textsc{facts} & : & \mathit{set(Prop)} \\ \textsc{visualsit} & : & \mathit{RecType} \\ \textsc{pending} & : & \mathit{list(LocProp)} \\ \textsc{moves} & : & \mathit{list(LocProp)} \\ \textsc{qud} & : & \mathit{poset(Question)} \end{bmatrix}$$

TTR, like many HPSG variants (e.g. Pollard & Sag 1987 and Pollard & Sag 1994), employs a situation semantic domain (Cooper 2019). This involves propositions being modelled in terms of types of situations, not in terms of sets of possible worlds. Since TTR is a type theory, it offers at least two explications of proposition. On the one hand, propositions can be identified with types (Cooper 2005a). On the other hand, propositions can be developed in an explicit Austinian (Austin 1950) way, where a proposition is individuated in terms of a situation and situation type (Ginzburg 2011: 845) – this is the truth-making (and Austin's original) interpretation of "It takes two to make a truth", since on Austin's conception a situation type can only be truth-evaluated against the situation it is about. We follow the latter option here. The type of propositions and the relation to a situation semantics conception of "true" (Barwise & Perry 1983) is given in (20):

(20) a. $\mathit{Prop} =_{def} \begin{bmatrix} \textsc{sit} & : & \mathit{Record} \\ \textsc{sit-type} & : & \mathit{RecType} \end{bmatrix}$

b. A proposition $p = \begin{bmatrix} \textsc{sit} & = & s \\ \textsc{sit-type} & = & T \end{bmatrix}$ is true iff $s : T$.

A special kind of proposition, namely *locutionary propositions* (*LocProp*) (Ginzburg 2012: 172), can be defined as follows:

(21) $\mathit{LocProp} =_{def} \begin{bmatrix} \textsc{sign} & : & \mathit{Record} \\ \textsc{sign-type} & : & \mathit{RecType} \end{bmatrix}$

Locutionary propositions are sign objects utilized to explicate clarification potential (see Section 3.1) and grounding.

Given the dialogue-awareness of signs just sketched, a content for interjections such as "EHHH HEHH" which constitutes turn 3 from the exchange between Ann and Ray in (1) at the beginning of this chapter can be given. Intuitively, Ann signals with these sounds that she heard Ray's question, which in turn is neither grounded nor clarified at this point of dialogue but is waiting for a response, what is called *pending*. This intuition can be made precise by means of the following lexical entry (which is closely related to the meaning of *mmh* given by Ginzburg 2012: 163):

(22)
$$\left[\begin{array}{l} \textsc{phon} : \langle\ \textsc{ehh hehh}\ \rangle \\ \textsc{cat} : \left[\textsc{head}{=}\textit{interjection}\ :\ \textit{syncat}\right] \\ \textsc{dgb-params} : \left[\begin{array}{ll} \textsc{spkr} & : \textit{Ind} \\ \textsc{addr} & : \textit{Ind} \\ \textsc{pending} & : \textit{LocProp} \\ \textsc{c2} & : \textit{address(spkr,addr,pending)} \end{array}\right] \\ \textsc{cont}{=}\textit{Understand}(\textsc{spkr,addr,dgb-params.pending}) : \textit{IllocProp} \end{array}\right]$$

Knowing how to use feedback signals such as the one in (22) can be claimed to be part of linguistic competence. It is difficult to imagine how to model this aspects of linguistic knowledge if not by means of *grammar in dialogue*.

Dialogue gameboard structures as defined in (19) as well as lexical entries for interjections such as (22) are still *static*. The mechanism that is responsible for the dynamics of dialogue and regiments the interactive evolution of DGBs is *conversational rules*. A conversational rule is a mapping between an input and an output information state, where the input DGB is constrained by a type labelled *preconditions* (PRE) and the output DGB is subject to EFFECTS. That is, a conversational rule can be notated in the following form, where *DGBType* is the type of dialogue gameboards defined in (19).

(23)
$$\left[\begin{array}{ll} \textsc{pre} & : \textit{DGBType} \\ \textsc{effects} & : \textit{DGBType} \end{array}\right]$$

Several basic conversational rules are defined in Ginzburg (2012: Chapter 4) and some of them, namely those needed to analyse example (8) discussed above, are re-given below (with "Fact update/QUD-downdate" being simplified, however). *IllocProp* abbreviates "Illocutionary Proposition", *IllocRel* "Illocutionary Relation", *poset* "Partially Ordered Set", *AbSemObj* "Abstract Semantic Object" and *QSPEC* "Question-under-Discussion-Specific". With regard to the partially ordered QUD set, we use "$\langle u, X\rangle$" to denote the upper bound u for subset X. For details, we have to refer the reader to Ginzburg (2012); we believe the following list to convey at least a solid impression of how dialogue dynamics works in KoS, however.

- Free Speech:

$$\begin{bmatrix} \textsc{pre} & : & \begin{bmatrix}\textsc{qud}=\langle\rangle & : & \textit{poset(Question)}\end{bmatrix} \\ \textsc{effects} & : & \textit{TurnUnderspec} \wedge_{\textit{merge}} \begin{bmatrix} \textsc{r} : \textrm{AbSemObj} \\ \textrm{R} : \textrm{IllocRel} \\ \textrm{LatestMove}=\textrm{R}(\textsc{spkr},\textsc{addr},\textsc{r}) : \textrm{IllocProp} \end{bmatrix} \end{bmatrix}$$

- QSPEC:

$$\begin{bmatrix} \textsc{pre} & : & \begin{bmatrix}\textsc{qud}=\left\langle \textsc{q},\textrm{Q}\right\rangle & : & \textit{poset(Question)}\end{bmatrix} \\ \textsc{effects} & : & \textit{TurnUnderspec} \wedge_{\textit{merge}} \begin{bmatrix} \textsc{r} : \textrm{AbSemObj} \\ \textrm{R} : \textrm{IllocRel} \\ \textrm{LatestMove}=\textrm{R}(\textsc{spkr},\textsc{addr},\textsc{r}) : \textrm{IllocProp} \\ \textsc{c1} : \textrm{Qspecific}(\textsc{r},\textsc{q}) \end{bmatrix} \end{bmatrix}$$

- Ask QUD-incrementation:

$$\begin{bmatrix} \textsc{pre} & : & \begin{bmatrix} \textsc{q} & : & \textit{Question} \\ \textrm{LatestMove}=\textrm{Ask}(\textsc{spkr},\textsc{addr},\textsc{q}) & : & \textit{IllocProp} \end{bmatrix} \\ \textsc{effects} & : & \begin{bmatrix}\textsc{qud}=\left\langle \textsc{q},\textsc{pre.qud}\right\rangle & : & \textit{poset(Question)}\end{bmatrix} \end{bmatrix}$$

- Assert QUD-incrementation:

$$\begin{bmatrix} \textsc{pre} & : & \begin{bmatrix} \textsc{p} & : & \textit{Proposition} \\ \textrm{LatestMove}=\textrm{Assert}(\textsc{spkr},\textsc{addr},\textsc{p}) & : & \textit{IllocProp} \end{bmatrix} \\ \textsc{effects} & : & \begin{bmatrix}\textsc{qud}=\left\langle \textsc{p}?,\textsc{pre.qud}\right\rangle & : & \textit{poset(Question)}\end{bmatrix} \end{bmatrix}$$

- Accept:

$$\begin{bmatrix} \textsc{pre} & : & \begin{bmatrix} \textsc{spkr} & : & \textit{Ind} \\ \textsc{addr} & : & \textit{Ind} \\ \textsc{p} & : & \textit{Prop} \\ \textrm{LatestMove}=\textrm{Assert}(\textsc{spkr},\textsc{addr},\textsc{p}) & : & \textit{IllocProp} \\ \textsc{qud}=\left\langle \textsc{p}?,\textsc{subqud}\right\rangle & : & \textit{poset(Question)} \end{bmatrix} \\ \textsc{effects} & : & \begin{bmatrix} \textsc{spkr}=\textsc{pre.addr} & : & \textit{Ind} \\ \textsc{addr}=\textsc{pre.sprk} & : & \textit{Ind} \\ \textrm{LatestMove}=\textrm{Accept}(\textsc{spkr},\textsc{addr},\textsc{p}) & : & \textit{IllocProp} \end{bmatrix} \end{bmatrix}$$

- Fact update/QUD-downdate (simplified):

– $$\left[\begin{array}{ll}\textsc{pre} & : \left[\begin{array}{lll}\textsc{q} & : & \textit{Question}\\ \textsc{p} & : & \textit{Prop}\\ \textsc{LatestMove}{=}\textsc{Accept}(\textsc{spkr},\textsc{addr},\textsc{p}) & : & \textit{IllocProp}\\ \textsc{qud}{=}\langle\textsc{q},\textsc{subqud}\rangle & : & \textit{poset(Question)}\\ \textsc{qbg} & : & \textit{Qspecific(p,q)}\end{array}\right]\\ \textsc{effects} & : \left[\begin{array}{lll}\textsc{facts}{=}\textsc{pre.facts} \cup \{\textsc{p}\} & : & \textit{Set(Prop)}\\ \textsc{qud}{=}\textsc{pre.qud} \setminus \{\textsc{q}\} & & \end{array}\right]\end{array}\right]$$

– $$\left[\begin{array}{ll}\textsc{pre} & : \left[\begin{array}{lll}\textsc{p} & : & \textit{Prop}\\ \textsc{LatestMove}{=}\textsc{Accept}(\textsc{spkr},\textsc{addr},\textsc{p}) & : & \textit{IllocProp}\\ \textsc{qud}{=}\langle\textsc{p}?,\textsc{subqud}\rangle & : & \textit{poset(Question)}\end{array}\right]\\ \textsc{effects} & : \left[\begin{array}{lll}\textsc{facts}{=}\textsc{pre.facts} \cup \{\textsc{p}\} & : & \textit{Set(Prop)}\\ \textsc{qud}{=}\textsc{pre.qud} \setminus \{\textsc{p}?\} & & \end{array}\right]\end{array}\right]$$

Having dialogue game boards and conversational rules at one's disposal, we can apply KoS' analytical tools to the dialogue example from (8) above. We make the following simplifying assumptions: if the *n*th move is an assertion, we refer to the asserted proposition in terms of "p(*n*)". The corresponding question *whether p(n)* is notated "p?(*n*)". If the *n*th move is a question, we refer to the question in terms of "q(*n*)". Additionally, we assume that subsentential utterances project to Austinian propositions by resolving elliptical expressions in context in terms of their missing semantic constituents which are available as the contents of the maximal elements in QUD (that is, they are addressable via the path QUD.FIRST.CONT; cf. Ginzburg 2012).

Turn	**DGB dynamics**	**Utterance / Conversational rule(s)**
init.	$\left[\begin{array}{lll}\textsc{participants} & = & \{\text{A,B}\}\\ \textsc{Moves} & = & \langle\ \rangle\\ \textsc{qud} & = & \langle\ \rangle\\ \textsc{facts} & = & \textit{cg0}\end{array}\right]$	—
1.	$\left[\begin{array}{lll}\textsc{spkr} & = & \textit{A}\\ \textsc{addr} & = & \textit{B}\\ \textsc{Moves} & = & \langle\textsc{Assert}(\text{A,B},\textsc{p}(1))\rangle\\ \textsc{qud} & = & \langle\textsc{p}?(1)\rangle\\ \textsc{facts} & = & \textit{cg0}\end{array}\right]$	"I've been at university." Free Speech + Assert QUD-incrementation

2.	$\begin{bmatrix} \textsc{spkr} & = B \\ \textsc{addr} & = A \\ \textsc{Moves} & = \left\langle \textsc{Ask}(B,A,q(2)),\ \textsc{Assert}(A,B,p(1)) \right\rangle \\ \textsc{qud} & = \left\langle q(2) \right\rangle \\ \textsc{facts} & = \mathit{cg0} \cup \left\{ p(1) \right\} \end{bmatrix}$	"Which university?" Accept + Ask QUD-incrementation
3.	$\begin{bmatrix} \textsc{spkr} & = A \\ \textsc{addr} & = B \\ \textsc{Moves} & = \left\langle \begin{array}{l} \textsc{Assert}(A,B,p(3)), \\ \textsc{Ask}(B,A,q(2)),\ \textsc{Assert}(A,B,p(1)) \end{array} \right\rangle \\ \textsc{qbg} & = \mathit{About(p(3),q(2))} \\ \textsc{qud} & = \left\langle p?(3),\ q(2) \right\rangle \\ \textsc{facts} & = \mathit{cg0} \cup \left\{ p(1) \right\} \end{bmatrix}$	"Cambridge." QSPEC (via *About* relation) + Assert QUD-incrementation
4.	$\begin{bmatrix} \textsc{spkr} & = B \\ \textsc{addr} & = A \\ \textsc{Moves} & = \left\langle \begin{array}{l} \textsc{Accept}(B,A,p(3)),\ \textsc{Assert}(A,B,p(3)), \\ \textsc{Ask}(B,A,q(2)),\ \textsc{Assert}(A,B,p(1)) \end{array} \right\rangle \\ \textsc{qud} & = \left\langle \quad \right\rangle \\ \textsc{facts} & = \mathit{cg0} \cup \left\{ p(3),\ p(1) \right\} \end{bmatrix}$	"Cambridge, um." Accept + Fact update/QUD-downdate
5.	$\begin{bmatrix} \textsc{spkr} & = B \\ \textsc{addr} & = A \\ \textsc{Moves} & = \left\langle \begin{array}{l} \textsc{Ask}(B,A,q(5)), \\ \textsc{Accept}(B,A,p(3)),\ \textsc{Assert}(A,B,p(3)), \\ \textsc{Ask}(B,A,q(2)),\ \textsc{Assert}(A,B,p(1)) \end{array} \right\rangle \\ \textsc{qud} & = \left\langle q(5) \right\rangle \\ \textsc{facts} & = \mathit{cg0} \cup \left\{ p(3),\ p(1) \right\} \end{bmatrix}$	"what did you read?" Free Speech + Ask QUD-incrementation
6.	$\begin{bmatrix} \textsc{spkr} & = A \\ \textsc{addr} & = B \\ \textsc{Moves} & = \left\langle \begin{array}{l} \textsc{Assert}(A,B,p(6)),\ \textsc{Ask}(B,A,q(5)), \\ \textsc{Accept}(B,A,p(3)),\ \textsc{Assert}(A,B,p(3)), \\ \textsc{Ask}(B,A,q(2)),\ \textsc{Assert}(A,B,p(1)) \end{array} \right\rangle \\ \textsc{qbg} & = \mathit{About(p(6),q(5))} \\ \textsc{qud} & = \left\langle p?(6),\ q(5) \right\rangle \\ \textsc{facts} & = \mathit{cg0} \cup \left\{ p(3),\ p(1) \right\} \end{bmatrix}$	"History and English." QSPEC (via *About* relation) + Assert QUD-incrementation

7.	$\begin{bmatrix} \textsc{spkr} & = & B \\ \textsc{addr} & = & A \\ \textsc{Moves} & = & \left\langle \begin{array}{l} \textsc{Accept}(\text{B,A},\textsc{p}(6)), \\ \textsc{Assert}(\text{A,B},\textsc{p}(6)),\ \textsc{Ask}(\text{B,A},\textsc{q}(5)), \\ \textsc{Accept}(\text{B,A},\textsc{p}(3)),\ \textsc{Assert}(\text{A,B},\textsc{p}(3)), \\ \textsc{Ask}(\text{B,A},\textsc{q}(2)),\ \textsc{Assert}(\text{A,B},\textsc{p}(1)) \end{array} \right\rangle \\ \textsc{qud} & = & \langle\ \rangle \\ \textsc{facts} & = & \mathit{cg0} \cup \left\{\textsc{p}(6),\ \textsc{p}(3),\ \textsc{p}(1)\right\} \end{bmatrix}$	“History and English.” Accept + Fact update/QUD-downdate

Note that the dialogical exchange leads to an increase of the common ground of the interlocutors A and B: after chatting, the common ground contains the propositions *that A has been at university* (p(1)), *that A has been at Cambridge University* (p(3)) and *that A read History and English* (p(6)).

On these grounds, a lexical entry for “hello” can be spelled out. “Hello” realises a greeting move (which is its content) and must be used discourse-initially (the MOVES list and the QUD set have to be empty):

(24)
$$\begin{bmatrix} \textsc{phon} : \langle\ \textsc{hello}\ \rangle \\ \textsc{cat} : \begin{bmatrix} \textsc{head}{=}\mathit{INTERJECTION} & : & \mathit{syncat} \end{bmatrix} \\ \textsc{dgb-params} : \begin{bmatrix} \textsc{spkr} & : & \mathit{Ind} \\ \textsc{addr} & : & \mathit{Ind} \\ \textsc{moves}{=}\langle\ \ \rangle & : & \mathit{list(IllocProp)} \\ \textsc{qud}{=}\{\} & : & \mathit{poset(Question)} \end{bmatrix} \\ \textsc{cont}{=}\textsc{Greet}(\textsc{spkr},\textsc{addr}) : \mathit{\textsc{IllocProp}} \end{bmatrix}$$

Discourse-dynamically, “hello” puts a greeting move onto the MOVES list of the dialogue gameboard, thereby initiates an interaction and invites for a countergreeting (the requirement for countergreeting is exactly that a greeting move is the element of the otherwise empty list of dialogue moves) – giving rise to an *adjacency pair* as part of the local management system for dialogues investigated in conversational analysis (Schegloff & Sacks 1973).

The discourse particle “yes” can be used to answer a polar yes/no question. In this use, “yes” has a propositional content p that asserts the propositional content of the polar question p?, which has to be the maximal element in QUD (Ginzburg 2012: Chapter 2, 231 *et seq.*). That is, “yes” affirmatively resolves a given polar question. Polar questions, in turn, are 0-ary propositional abstracts (Ginzburg 2012: 231), that is, the polar question p? corresponding to a proposition p is a function mapping an empty record to p: $\lambda r : [].p$. Thus, applying p? to an empty record [] returns p, which is exactly what “yes” does. The affirmative particle (used to answer a yes/no question) is a propositional lexeme which applies a polar question which is maximal in QUD to an empty record (cf. Ginzburg 2012: 232):

(25)
$$\begin{bmatrix} \text{PHON} : \langle\ \text{YES}\ \rangle \\ \text{CAT} : \begin{bmatrix} \text{HEAD=}\textit{PARTCL} \ :\ \textit{syncat} \end{bmatrix} \\ \text{DGB-PARAMS} : \begin{bmatrix} \text{QUD=}\begin{bmatrix} \text{MAX} & : & \textit{PolQuestion} \\ \text{REST} & : & \textit{set(Question)} \end{bmatrix} \ :\ \textit{poset(Question)} \end{bmatrix} \\ \text{CONT=DGB-PARAMS.QUD.MAX([ ]): } \textit{Prop} \end{bmatrix}$$

Due to its involvedness in DGB-PARAMS.QUD, "yes" directly interacts with accept and downdating, as described above. For more on this, see Ginzburg (2012).

5 Outlook

Given a basic framework for formulating and analysing content in dialogue context, there are various directions to explore, including the following ones.

- One of the main challenges of dialogue semantics is the integration of *non-verbal communication* means, like gaze, gestures, body posture, timing and non-language vocal sounds (e.g. laughter; Ginzburg et al. 2015; Tian et al. 2016). Since non-verbal communication means are informative, not only does a (dialogue) semantic representation have to be developed, but also the rules of their interaction with speech have to be formulated.

- Strictly speaking, dialogue is the interaction between *two* interlocutors. How can one scale up to *multilogue*, where the number of participants is at least three (Ginzburg & Fernández 2005)? Given the increased number of participants, problems that emerge include *grounding by proxy*, where a representative represents the dialogue gameboard of a group (Eshghi & Healey 2016) and of course *turn taking*.

- People do not process natural language input sentence-wise. Rather, processing begins with the initial sound and proceeds word for word or even on smaller units like affixes and phonemes – that is, processing is incremental (e.g. Sedivy et al. 1999; see also Wasow 2019, Chapter G of this volume). This is a key ingredient in the efficient (relatively gap-free and interruption-less) managing of turn taking. One direction of dialogue theories therefore is to bring psycholinguistics and formal semantics closer together by devising incremental grammar and dialogue gameboard models (Hough et al. 2015; Demberg et al. 2013; Poesio & Rieser 2011).

Finally, we want to mention two other dialogue-theoretic frameworks that have been worked out to a substantial degree, namely PTT (Traum 1994; Poesio 1995; Poesio & Traum 1997; Poesio & Rieser 2010), and *Segmented Discourse Representation Theory* (SDRT) (Asher 1993; Asher & Lascarides 2003; 2013; Hunter & Asher 2015). The phenomena and outlook directions discussed in this chapter apply to all theories of dialogue semantics, of course.

Appendix: An HPSG$_{\text{TTR}}$ fragment

The appendix provides a fragment of HPSG$_{\text{TTR}}$. The grammar framework used is oriented at a *Head-driven Phrase Structure Grammar* variant (Sag et al. 2003), namely its TTR implementation (Cooper 2008). We use HPSG because its architecture satisfies the property of *incremental correspondence* (Johnson & Lappin 1999) – utterance representations encode phonological, syntactic, semantic and contextual information *fractally*. This is crucial *inter alia* for any treatment of clarification interaction (cf. Section 3.1). We use HPSG$_{\text{TTR}}$ because the type-theoretical version allows us to directly incorporate semantic objects (cf. Section 3.2).

TTR has a counterpart to unification, namely the *merge* construction.

(26) a. If R_1 and R_2 are record types, then $R_1 \wedge_{merge} R_2$ is a record type and is called the *merge* of R_1 and R_2.

b. Since merge types are complicated to define (but see Cooper 2012), we follow the strategy of Cooper (2017) and illustrate the working of merges by means of some examples:

(i) $\begin{bmatrix} \text{A} & : & T \\ \text{B} & : & R \end{bmatrix} \wedge_{\textit{MERGE}} \begin{bmatrix} \text{C} & : & S \end{bmatrix} = \begin{bmatrix} \text{A} & : & T \\ \text{B} & : & R \\ \text{C} & : & S \end{bmatrix}$

(ii) $\begin{bmatrix} \text{A} & : & T \end{bmatrix} \wedge_{\textit{MERGE}} \begin{bmatrix} \text{A} & : & R \end{bmatrix} = \begin{bmatrix} \text{A} & : & T \wedge_{merge} R \end{bmatrix}$

Structure sharing is indicated by a “tag type” notation. Tag types are defined in terms of manifest fields.[8] The notational convention is exemplified in (27) by means of head-specifier agreement, where the tag type from (27a) abbreviates the structure in (27b):

[8] *NB:* technically, tag types apply singleton types to record types, instead of to objects, thereby making use of a revision of the notion of singleton types introduced by Cooper (2013: 4, footnote 3).

(27) a. $\left[\text{CAT} : \begin{bmatrix} \text{HEAD} : \left[\text{AGR}_{\boxed{1}} : \mathit{Agr}\right] \\ \text{SPR} : \left\langle\left[\text{CAT} : \left[\text{HEAD} : \left[\text{AGR}=\boxed{1} : \mathit{Agr}\right]\right]\right]\right\rangle \end{bmatrix}\right]$

b. $\left[\text{CAT} : \begin{bmatrix} \text{HEAD} : \left[\text{AGR} : \mathit{Agr}\right] \\ \text{SPR} : \left\langle\left[\text{CAT} : \left[\text{HEAD} : \left[\text{AGR}=/\text{CAT.HEAD.AGR} : \mathit{Agr}\right]\right]\right]\right\rangle \end{bmatrix}\right]$

The tag type notation alludes to the box notation common in HPSG work.

Agr is defined as usual:

(28) $\mathit{Agr} := \begin{bmatrix} \text{NUM} : \mathit{Num} \\ \text{PERS} : \mathit{Per} \\ \text{GEN} : \mathit{Gen} \end{bmatrix}$

A basic *sign* is a pairing of phonetic, syntactic and semantic information and follows the geometry in (29):

(29) $\mathit{sign} := \begin{bmatrix} \text{PHON} : \mathit{Phoneme} \\ \text{CAT} : \mathit{SynCat} \\ \text{DGB-PARAMS} : \mathit{RecType} \\ \text{CONT} : \mathit{SemObj} \end{bmatrix}$

Signs employ DGB-PARAMS, which host referential meanings that are witnessed among interlocutors. Quantificational abstraction is achieved by coercing parts of DGB-PARAMS to Q-PARAMS:

(30) If DGB-PARAMS : R_2 and for two record types R_0 and R_1 lacking any mutual dependencies[9] $R_2 = R_0 \wedge_{merge} R_1$, then R_0 can be moved to Q-PARAMS, resulting in the following structure:

$\begin{bmatrix} \text{DGB-PARAMS} : R_1 \\ \text{CONT} = \left[\text{Q-PARAMS} : R_0\right] \end{bmatrix}$

A word is a sign with constituent type (CXTYPE) *word*. Using the merge operation, the word extension on signs can represented compactly as in (31a), which expands to the structure given in (31b):

(31) a. $\mathit{word} := \mathit{sign} \wedge_{merge} \left[\text{CXTYPE} : \mathit{word}\right] : \mathit{RecType}$

b. $\begin{bmatrix} \text{CXTYPE} : \mathit{word} \\ \text{PHON} : \mathit{Phoneme} \\ \text{CAT} : \mathit{SynCat} \\ \text{DGB-PARAMS} : \mathit{RecType} \\ \text{CONT} : \mathit{SemObj} \end{bmatrix}$

[9]None of the labels occurring in R_0 occur in R_1 and vice versa.

Words – that is, cxtype *word* – are usually the result of lexical rules, whose input are lexemes. Lexemes differ from words in their constituent type:

(32) $\textit{lexeme} \coloneqq \textit{sign} \wedge_{\textit{merge}} \left[\textsc{cxtype} : \textit{lexeme}\right] : \textit{RecType}$

A phrasal sign can be seen as a word with daughters:

(33) a. $\textit{phrase} \coloneqq \textit{sign} \wedge_{\textit{merge}} \left[\begin{array}{ll} \textsc{cxtype} & : \textit{phrase} \\ \textsc{dtrs} & : \left[\textsc{nhd-dtrs} : \textit{List(Sign)}\right] \end{array}\right] : \textit{RecType}$

b. $\left[\begin{array}{ll} \textsc{cxtype} & : \textit{phrase} \\ \textsc{phon} & : \textit{List(Phoneme)} \\ \textsc{cat} & : \textit{SynCat} \\ \textsc{dgb-params} & : \textit{RecType} \\ \textsc{cont} & : \textit{SemObj} \\ \textsc{dtrs} & : \left[\textsc{nhd-dtrs} : \textit{List(Sign)}\right] \end{array}\right]$

A headed phrase is a phrase with a prominent daughter, i.e. the head daughter:

(34) a. $\textit{hd-phrase} \coloneqq \textit{phrase} \wedge_{\textit{merge}} \left[\textsc{dtrs} : \left[\textsc{hd-dtr} : \textit{Sign}\right]\right] : \textit{RecType}$

b. $\left[\begin{array}{ll} \textsc{cxtype} & : \textit{phrase} \\ \textsc{phon} & : \textit{List(Phoneme)} \\ \textsc{cat} & : \textit{SynCat} \\ \textsc{dgb-params} & : \textit{RecType} \\ \textsc{cont} & : \textit{SemObj} \\ \textsc{dtrs} & : \left[\begin{array}{ll} \textsc{hd-dtr} & : \textit{Sign} \\ \textsc{nhd-dtrs} & : \textit{List(Sign)} \end{array}\right] \end{array}\right]$

The head daughter is special since it (as a default, at least) determines the syntactic properties of the mother construction. This aspect of headedness is captured in terms of the *head-feature principle* (HFP), which can be implemented by means of tag types as follows:

(35) $\text{HFP} \coloneqq \left[\begin{array}{ll} \textsc{cxtype} & : \textit{phrase} \\ \textsc{cat} & : \left[\textsc{head}_{\boxed{2}} : \textit{PoS}\right] \\ \textsc{dtrs} & : \left[\textsc{hd-dtr} : \left[\textsc{cat} : \left[\textsc{head}{=}\boxed{2} : \textit{PoS}\right]\right]\right] \end{array}\right]$

The fact that the daughters' locutions combine to the mother's utterance is captured in terms of a "phon principle" (we use a slash notation in order to indicate paths starting at the outermost level of a feature structure):

(36) $\text{PHON} \coloneqq \left[\begin{array}{ll} \textsc{cxtype} & : \textit{phrase} \\ \textsc{phon} & : \textit{List(/dtrs.hd.dtr/phon, /dtrs.nhd.dtrs/pos1.phon,} \\ & \quad \textit{..., /dtrs.nhd.dtrs/posn.phon)} \end{array}\right]$

Since semantic composition rests on predication rather than unification, there is no analog to the semantic compositionality principle of Sag et al. (2003) in our account. There is, however, something akin to semantic inheritance: we need to keep track of the contextual and quantificational paramaters contributed by the daughters of a phrase. This is achieved in terms of a *dgb-params principle* (*DGBPP*) in (37) which unifies the daughters' DGB-PARAMS into the mother's DGB-PARAMS (see Ginzburg 2012: 126 *et seq.* for a similar principle):

(37) DGBPP :=

$$\begin{bmatrix} \text{CXTYPE} & : & \textit{phrase} \\ \text{DGB-PARAMS} & : & \begin{bmatrix} \text{/DTRS.HD-DTR.DGB-PARAMS} \wedge_{\textit{MERGE}} \text{/DTRS.NHD-DTRS.POS1.DGB-PARAMS} \wedge_{\textit{MERGE}} \\ \ldots\wedge_{\textit{MERGE}} \text{/DTRS.NHD-DTRS.POS}n\text{.DGB-PARAMS} \end{bmatrix} \\ \text{DTRS} & : & \begin{bmatrix} \text{HD-DTR} & : & \begin{bmatrix}\text{Q-PARAMS} : \textit{RecType}\end{bmatrix} \\ \text{NHD-DTRS} & : & \left\langle \text{POS1} : \begin{bmatrix}\text{Q-PARAMS} : \textit{RecType}\end{bmatrix}, \ldots, \textit{posn} : \begin{bmatrix}\text{Q-PARAMS} : \textit{RecType}\end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

A headed phrase is well-formed when it is a headed phrase and it obeys the head feature principle, the phon principle and the dgb-params principle, which is expressed by extending *hd-phrase* by the following constraint:

(38) $\textit{hd-phrase} := \textit{hd-phrase} \wedge_{\textit{merge}} \text{HFP} \wedge_{\textit{merge}} \text{PHON} \wedge_{\textit{merge}} \text{DGBPP}$

Using this set-up, lexical entries, lexical rules and syntactic constructions can be formulated straightforwardly.

References

Arnold, Doug. 2019. HPSG and Lexical Functional Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1071–1072. Berlin: Language Science Press. DOI:??

Asher, Nicholas. 1993. *Reference to abstract objects in discourse* (Studies in Linguistics and Philosophy 50). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of conversation*. Cambridge: Cambridge University Press.

Asher, Nicholas & Alex Lascarides. 2013. Strategic conversation. *Semantics and Pragmatics* 6(2). 1–62. DOI:10.3765/sp.6.2

Austin, John L. 1950. Truth. In *Proceedings of the Aristotelian Society. Supplementary*, vol. xxiv, 111–128. Reprinted in John L. Austin: *Philosophical Papers*. 2. ed. Oxford: Clarendon Press, 1970.

Austin, John L. 1962. *How to do things with words*. 2nd edn. Cambridge, MA: Harvard University Press.

Barwise, Jon & John Perry. 1983. *Situations and attitudes* (The David Hume Series of Philosophy and Cognitive Science Reissues). Stanford: CSLI Publications.

Beißwenger, Michael, Maria Ermakova, Alexander Geyken, Lothar Lemnitzer & Angelika Storrer. 2012. DeRiK: A German reference corpus of computer-mediated communication. In *Proceedings of Digital Humanities 2012*. http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/derik-a-german-reference-corpus-of-computer-mediated-communication/.

Blank, Malou, Anna Tornberg & Corina Knipper. 2018. New perspectives on the late neolithic of south-western Sweden. An interdisciplinary investigation of the gallery grave Falköping Stad 5. *Open Archaeology* 4(1). 1–35.

Bonami, Olivier. 2002. A syntax-semantics interface for tense and aspect in french. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *The proceedings of the 8th international conference on head-driven phrase structure grammar*, 31–50. Stanford: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2/.

Borsley, Robert D. & Stefan Müller. 2019. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 937–1012. Berlin: Language Science Press. DOI:??

Branco, António. 2019. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 613–665. Berlin: Language Science Press. DOI:??

Carpenter, Bob. 1992. *The logic of typed feature structures*. Cambridge, MA: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Clark, Herbert H. 1975. Bridging. In *Proceedings of the 1975 workshop on theoretical issues in natural language processing* (TINLAP '75), 169–174. Cambridge, Massachusetts: Association for Computational Linguistics. DOI:10.3115/980190.980237

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press.

Clark, Herbert H. & Catherine R. Marshall. 1981. Definite reference and mutual knowledge. In Aravind K. Joshi, Bonnie L. Webber & Ivan A. Sag (eds.), *Elements of discourse understanding*, 10–63. Cambridge: Cambridge University Press.

Clark, Herbert H., Robert Schreuder & Samuel Buttrick. 1983. Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior* 22(2). 245–258.

Cohen, Philip R. & Hector J. Levesque. 1990. Rational interaction as the basis for comunication. In Philip R. Cohen, Jerry R. Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 221–256. Cambridge, MA: MIT Press.

Cooper, Robin. 2005a. Austinian truth, attitudes and type theory. *Research on Language and Computation* 3(2-3). 333–362. DOI:10.1007/s11168-006-0002-z

Cooper, Robin. 2005b. Records and record types in semantic theory. *Journal of Logic and Computation* 15(2). 99–112.

Cooper, Robin. 2008. Type theory with records and unification-based grammar. In Fritz Hamm & Stephan Kepser (eds.), *Logics for linguistic structures* (Trends in Linguistics: Studies and Monographs 201), 9–33. Berlin & New York: Mouton de Gruyter.

Cooper, Robin. 2012. Type theory and semantics in flux. In Ruth Kempson, Tim Fernando & Nicholas Asher (eds.), *Philosophy of linguistics* (Handbook of Philosophy of Science 6), 271–323. Oxford & Amsterdam: Elsevier.

Cooper, Robin. 2013. Clarification and generalized quantifiers. *Dialogue & Discourse* 4(1). 1–25. http://elanguage.net/journals/dad/article/view/1554, accessed 2015-1-13. DOI:10.5087/dad.2013.101

Cooper, Robin. 2014. Phrase structure rules as dialogue update rules. In *Proccedings of DialWatt – The 18th workshop on the semantics and pragmatics of dialogue* (SemDial 2014), 26–34. Edinburgh.

Cooper, Robin. 2017. Adapting Type Theory with Records for natural language semantics. In Stergios Chatzikyriakidis & Zhaohui Luo (eds.), *Modern perspectives in Type-Theoretical Semantics* (Studies in Linguistics and Philosophy 98), 71–94. Cham: Springer International Publishing AG. DOI:10.1007/978-3-319-50422-3_4

Cooper, Robin. 2019. *From perception to linguistic communication: Using a theory of types with records (TTR) to model linguistic action and content.* https://github.com/robincooper/ttl. MS Gothenburg University.

Cooper, Robin & Jonathan Ginzburg. 2015. Type theory with records for natural language semantics. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 2nd edn., chap. 12, 375–407. Oxford, UK: Wiley-Blackwell.

Copestake, Ann. 2007. Semantic composition with (Robust) Minimal Recursion Semantics. In *Proceedings of the workshop on deep linguistic processing* (DeepLP'07), 73–80. Prague, Czech Republic.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(4). 281–332.

Costa, Francisco & António Branco. 2012. Backshift and tense decomposition. In Stefan Müller (ed.), *Proceedings of the 19th International Conference on Head-Driven Phrase Structure Grammar, Chungnam National University Daejeon*, 86–106. CSLI Publications. http://cslipublications.stanford.edu/HPSG/2012/costa-branco.pdf.

Demberg, Vera, Frank Keller & Alexander Koller. 2013. Incremental, predictive parsing with psycholinguistically motivated Tree-Adjoining Grammar. *Computational Linguistics* 39(4). 1025–1066.

Devlin, Keith. 1991. *Logic and information*. Cambridge: Cambridge University Press.

Donnellan, Keith S. 1966. Reference and definite descriptions. *The Philosophical Review* 75(3). 281–304.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Elisabet Engdahl & Enric Vallduví (eds.), *Edinburgh working papers in cognitive science* (Studies in HPSG 12), 1–31. Edinburgh: University of Edinburgh.

Eshghi, Arash & Patrick G. T. Healey. 2016. Collective contexts in conversation: grounding by proxy. *Cognitive Science* 40(2). 299–324. DOI:10.1111/cogs.12225

Fernández, Raquel & Jonathan Ginzburg. 2002. Non-sentential utterances: A corpus study. *Traîtement Automatique de Languages* 43(2). 13–42.

Fernández, Raquel, Jonathan Ginzburg & Shalom Lappin. 2007. Classifying non-sentential utterances in dialogue: a machine learning approach. *Computational Linguistics* 33(3). 397–427. DOI:10.1162/coli.2007.33.3.397

Fernando, Tim. 2011. Constructing situations and time. *Journal of Philosophical Logic* 40(3). 371–396. DOI:10.1007/s10992-010-9155-1

Frege, Gottlob. 1892. Über Sinn und Bedeutung. In Mark Textor (ed.), *Funktion – Begriff – Bedeutung* (Sammlung Philosophie 4). Göttingen: Vandenhoeck & Ruprecht.

Ginzburg, Jonathan. 1994. An update semantics for dialogue. In Harry Bunt, Reinhard Muskens & Gerrit Rentier (eds.), *Proceedings of the first international workshop on computational semantics*. Tilburg, The Netherlands: Katholieke Universiteit Brabant.

Ginzburg, Jonathan. 1996. Dynamics and the semantics of dialogue. In Jerry Seligman & Dag Westerstøahl (eds.), *Logic, language and computation*, vol. 1 (CSLI Lecture Notes 58), 221–237. Stanford, CA: CSLI Publications.

Ginzburg, Jonathan. 1997. *On some semantic consequences of turn taking*. Paul Dekker, Martin Stokhof & Yde Venema (eds.). ILLC, University of Amsterdam.

Ginzburg, Jonathan. 2003. Disentangling public from private meaning. In Jand van Kuppevelt & Ronnie W. Smith (eds.), *Current and new directions in discourse & dialogue*, 183–211. Dordrecht: Kluwer Academic Publishers.

Ginzburg, Jonathan. 2011. Situation semantics and the ontology of natural language. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 1 (Handbücher zur Sprach- und Kommunikationswissenschaft), 830–851. De Gruyter Mouton. DOI:10.1515/9783110226614.830

Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford, UK: Oxford University Press.

Ginzburg, Jonathan. 2019. Acquisition. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 773–774. Berlin: Language Science Press. DOI:??

Ginzburg, Jonathan, Ellen Breitholtz, Robin Cooper, Julian Hough & Ye Tian. 2015. Understanding laughter. In Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 137–146. Amsterdam, Netherlands. http://semanticsarchive.net/Archive/mVkOTk2N/AC2015-proceedings.pdf, accessed 2018-9-10.

Ginzburg, Jonathan & Raquel Fernández. 2005. Scaling up from dialogue to multilogue: Some principles and benchmarks. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics* (ACL '05), 231–238. Ann Arbor, Michigan: Association for Computational Linguistics. DOI:http://dx.doi.org/10.3115/12198

Ginzburg, Jonathan, Raquel Fernández & David Schlangen. 2014. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics* 7(9). 1–64. DOI:10.3765/sp.7.9

Ginzburg, Jonathan & Massimo Poesio. 2016. Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology* 7. 1938. DOI:10.3389/fpsyg.2016.01938

Ginzburg, Jonathan & Matthew Purver. 2012. Quantification, the Reprise Content Hypothesis, and Type Theory. In Lars Borin & Staffan Larsson (eds.), *From quantification to conversation. Festschrift for Robin Cooper on the occasion of his 65th birthday*. London: College Publications.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford: CSLI Publications.

Goodwin, Charles. 2003. Pointing as situated practice. In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 2, 217–241. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Green, Georgia M. 1996. The structure of CONTEXT: The representation of pragmatic restrictions in HPSG. In *Proceedings of the 5th annual meeting of the For-

mal Linguistics Society of the Midwest (Studies in the Linguistic Sciences 24), 215–232.

Haji-Abdolhosseini, Mohammad. 2003. A constraint-based approach to information structure and prosody correspondence. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 143–162. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-2-25.

Halliday, Michael A. K. 1967. Notes on transitivity and theme in English. Part ii. *Journal of Linguistics* 3(2). 199–244.

Hamblin, Charles Leonard. 1970. *Fallacies*. London: Methuen & Co Ltd.

Heim, Irene. 2002. File change semantics. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings* (Linguistics: The Essential Readings 1), 223–248. Oxford, UK: Wiley-Blackwell.

Heldner, Mattias & Jens Edlund. 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics* 38(4). 555–568.

Hough, Julian, Casey Kennington, David Schlangen & Jonathan Ginzburg. 2015. Incremental semantics for dialogue processing: Requirements, and a comparison of two approaches. In *Proceedings of the 11th International Conference on Computational Semantics*, 206–216. London, UK: Association for Computational Linguistics.

Hudson, Dick. 2019. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1073–1076. Berlin: Language Science Press. DOI:??

Hunter, Julie & Nicholas Asher. 2015. Integrating non-linguistic events into discourse structure. In *Proceedings of the 11th international conference on computational semantics* (IWCS 2015), 184–194. London, UK.

Jackendoff, Ray. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Johnson, David E. & Shalom Lappin. 1999. *Local constraints vs. economy* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.

Kathol, Andreas, Adam Przepiórkowski & Jesse Tseng. 2011. Advanced topics in Head-Driven Phrase Structure Grammar. In *Non-transformational syntax: Formal and explicit models of grammar*, 54–111. Oxford, UK: Wiley-Blackwell. DOI:10.1002/9781444395037.ch2

Kempson, Ruth. 2011. Formal semantics and representationalism. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international* *handbook of natural language meaning*, vol. 1 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 33), 216–241. Berlin: De Gruyter Mouton. DOI:10.1515/9783110226614.946

Kempson, Ruth, Ronnie Cann, Eleni Gregoromichelaki & Stergios Chatzikyriakidis. 2016. Language as mechanisms for interaction. *Theoretical Linguistics* 42(3-4). 203–276. DOI:10.1515/tl-2016-0011

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge, MA: Cambridge University Press.

King, Jeffrey C., Scott Soames & Jeff Speaks. 2014. *New thinking about propositions*. Oxford, UK: Oxford University Press.

King, Paul J. 1999. Towards truth in Head-driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in head-driven phrase structure grammar*, vol. 2 (Arbeitspapiere des SFB 340 132), 301–352. Eberhard-Karls Universität Tübingen.

Klein, Ewan. 2000. Prosodic constituency in HPSG. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism), chap. 10, 169–200. Stanford, CA: CSLI Publications.

Klein, Wolfgang. 1985. Gesprochene Sprache – geschriebene Sprache. *Zeitschrift für Literaturwissenschaft und Linguistik* 59. 9–35.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In Olaf Deutschmann, Hans Flasche, Bernhard König, Margot Kruse, Walter Pabst & Wolf-Dieter Stempel (eds.), *Romanistisches jahrbuch*, 15–43. Berlin & New York: Walter de Gruyter.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Kratzer, Angelika. 1978. *Semantik der Rede: Kontexttheorie, Modalwörter, Konditionalsätze* (Monographien: Linguistik und Kommunikationswissenschaft 38). Königstein/Ts.: Scriptor.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3–4). 243–276. DOI:10.1556/ALing.55.2008.3-4.2

Kripke, Saul A. 1980. *Naming and necessity*. Oxford: Blackwell.

Kubota, Yusuke. 2019. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1013–1069. Berlin: Language Science Press. DOI:??

Kuthy, Kordula De. 2019. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 709–743. Berlin: Language Science Press. DOI:??

Lee-Goldman, Russell. 2011. *Context in constructions*. Berkeley: University of California dissertation. https://escholarship.org/uc/item/02m6b3hx, accessed 2018-11-24.

Levine, Robert D. & Walt Detmar Meurers. 2006. Head-Driven Phrase Structure Grammar: Linguistic approach, formal foundations, and computational realization. In Keith Brown (ed.), *The encyclopedia of language and linguistics*, 2nd edn., 237–252. Oxford: Elsevier Science Publisher B.V. (North-Holland).

Levinson, Stephen C. & Francisco Torreira. 2015. Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology* 6(731). 10–26.

Lewis, David. 1979. Scorekeeping in a language game. In Rainer Bäuerle, Urs Egli & Arnim von Stechow (eds.), *Semantics from different points of view* (Springer Series in Language and Communication 6), 172–187. Berlin, Heidelberg & New York: Springer.

Linell, Per. 2005. *The written language bias in linguistics: Its nature, origins and transformations* (Routledge Advances in Communication and Linguistic Theory). New York: Routledge.

Lücking, Andy. 2018. Witness-loaded and witness-free demonstratives. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives. Syntax, semantics and pragmatics* (Linguistische Arbeiten 568). Berlin: De Gruyter.

Lücking, Andy. 2019. Gesture. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 887–934. Berlin: Language Science Press. DOI:??

Lücking, Andy & Jonathan Ginzburg. 2018. 'most people but not Bill': Integrating sets, individuals and negation into a cognitively plausible account of noun phrase interpretation. In *Proceedings of Cognitive Structures: Linguistic, Philosophical and Psychological Perspectives* (CoSt'18).

Lücking, Andy, Hannes Rieser & Marc Staudacher. 2006. SDRT and multi-modal situated communication. In David Schlangen & Raquel Fernández (eds.), *Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue* (Brandial'06), 72–79. Potsdam: Universitätsverlag Potsdam.

Marcus, R. Barcan. 1961. Modalities and intensional languages. *Synthese* 13(4). 303–322.

Martin-Löf, Per. 1984. *Intuitionistic type theory* (Studies in Proof Theory). Napoli: Bibliopolis.

McNeill, David (ed.). 2000. *Language and gesture*. Cambridge, UK: Cambridge University Press.

Montague, Richard. 1974. The proper treatment of quantification in ordinary English. In Richmond Thomason (ed.), *Formal philosophy*. New Haven: Yale UP.

Müller, Stefan. 2016. *Grammatical theory. From transformational grammar to constraint-based approaches* (Textbooks in Language Science 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2019. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

Nerbonne, John. 1992. *Constraint-based semantics*. Tech. rep. 92-18. Saarbrücken: DFKI. DOI:10.22028/D291-24846

Nunberg, Geoffrey. 1978. *The pragmatics of reference*. Indiana University Bloomington, Indiana University Linguistics Club dissertation.

Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16(1). 1–43. DOI:10.1007/BF00984721

Partee, Barbara Hall. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70(18). 601–609.

Penn, Gerald. 1999. An RSRL formalization of Serbo-Croatian second position clitic placement. In Valia Kordoni (ed.), *Tübingen studies in head-driven phrase structure grammar*, 177–197. Universität Tübingen.

Penn, Gerald. 2000. *The algebraic structure of attributed type signatures*. Pittsburgh, PA: School of Computer Science, Language Technologies Institute, Carnegie Mellon University dissertation.

Poesio, Massimo. 1995. A model of conversation processing based on micro conversational events. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 698–703. Pittsburgh: The Cognitive Science Society.

Poesio, Massimo & Hannes Rieser. 2010. Completions, coordination, and alignment in dialogue. *Dialogue & Discourse* 1(1). 1–89. DOI:10.5087/dad.2010.001

Poesio, Massimo & Hannes Rieser. 2011. An incremental model of anaphora and reference resolution based on resource situations. *Dialogue & Discourse* 2(1). 235–277. DOI:10.5087/dad.2011.110

Poesio, Massimo & David Traum. 1997. Conversational actions and discourse situations. *Computational Intelligence* 13(3). 309–347.

Pollard, Carl. 1997. Lectures on the foundations of HPSG. Unpublished manuscript: Ohio State University.

Pollard, Carl J. 1999. Strong generative capacity in HPSG. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 281–297. Stanford, CA: CSLI Publications.

Pollard, Carl & Ivan A. Sag. 1987. *Information-based syntax and semantics*. Stanford, CA: CSLI Publications.

Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications.

Purver, Matthew & Jonathan Ginzburg. 2004. Clarifying noun phrase semantics. *Journal of Semantics* 21(3). 283–339. DOI:10.1093/jos/21.3.283

Ranta, Aarne. 2015. Constructive type theory. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 2nd edn., chap. 11, 343–374. Oxford, UK: Wiley-Blackwell.

Richter, Frank. 2000. *A mathematical formalism for linguistic theories with an application in head-driven phrase structure grammar*. Eberhard-Karls-Universität Tübingen dissertation.

Richter, Frank. 2004. *Foundations of lexcial ressource semantics*. Eberhard-Karls-Universität Tübingen. Habilitation.

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 45–46. Berlin: Language Science Press. DOI:??

Richter, Frank & Manfred Sailer. 1999. LF conditions on expressions of Ty2: an HPSG analysis of negative concord in Polish. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in HPSG*, 247–282. Stanford, CA: CSLI Publications.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of Lexical Resource Semantics. In Arne Beckmann & Norbert Preining (eds.), *ESSLLI 2003 – course material I*, vol. 5 (Collegium Logicum), 87–143. Vienna: Kurt Gödel Society Wien.

Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735.

Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. Stanford, CA: CSLI Publications.

de Saussure, Ferdinand. 2001. *Grundfragen der Allgemeinen Sprachwissenschaft*. Charles Bally & Albert Sechehaye (eds.). 3rd edn. (de Gruyter Studienbuch). Berlin: de Gruyter. Translation of *Course de linguistique générale*. Lausanne & Paris, 1916.

Schegloff, Emanuel A. & Harvey Sacks. 1973. Opening up closings. *Semiotica* 8(4). 289–327. DOI:https://doi.org/10.1515/semi.1973.8.4.289

Sedivy, Julie C., Michael K. Tanenhaus, Craig G. Chambers & Gregory N. Carlson. 1999. Achieving incremental semantic interpretation through contextual representation. *Cognition* 71(2). 109–147. DOI:http://dx.doi.org/10.1016/S0010-0277(99)00025-6

Stalnaker, Robert. 1978. Assertion. *Syntax and Semantics* 9. 315–332.

Svartvik, Jan (ed.). 1990. *The London Corpus of spoken English: Description and research* (Lund Studies in English 82). Lund: Lund University Press.

Tian, Ye, Chiara Mazzoconi & Jonathan Ginzburg. 2016. When do we laugh? In *Proceedings of the 17th Annual SIGdial Meeting on Discourse and Dialogue* (SIGDIAL 2016). Los Angeles, USA.

Tomasello, Michael. 1998. Reference: Intending that others jointly attend. *Pragmatics & Cognition* 6(1/2). 229–243. Special Issue: The Concept of Reference in the Cognitive Sciences.

Traum, David. 1994. *A computational theory of grounding in natural language conversation*. Computer Science Dept., University of Rochester dissertation.

Traum, David & Staffan Larsson. 2003. The information state approach to dialogue management. In Jan Kuppevelt, Ronnie W. Smith & Nancy Ide (eds.), *Current and new directions in discourse and dialogue* (Text, Speech and Language Technology 22), 325–353. Amsterdam: Springer.

Vallduví, Enric. 2016. Information structure. In Maria Aloni & Paul Dekker (eds.), *The Cambridge handbook of formal semantics*, 728–755. Cambridge: Cambridge University Press.

Van Eynde, Frank. 1998. Tense, aspect and negation. In Frank Van Eynde & Paul Schmidt (eds.) (Studies in Machine Translation and Natural Language Processing 10), 209–280. Office for Official Publications of the European Communities.

Van Eynde, Frank. 2000. A constraint-based semantics for tenses and temporal auxiliaries. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism), 231–249. Stanford, CA: CSLI Publications.

Wasow, Tom. 2019. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 775–796. Berlin: Language Science Press. DOI:??

Wechsler, Stephen. 2019. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 131–154. Berlin: Language Science Press. DOI:??

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-

Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Zimmermann, Thomas Ede. 2011. Model-theoretic semantics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 1 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 33), 762–802. Berlin: De Gruyter Mouton. DOI:10.1515/9783110226614.94

Chapter J

Sign languages

Markus Steinbach

Georg-August-Universität Göttingen

Anke Holler

Georg-August-Universität Göttingen

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla

Markus Steinbach & Anke Holler. 2019. Sign languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 885–886. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

$$\begin{bmatrix} \textit{some-type} & \\ \textsc{feat-a} & \boxed{10}\begin{bmatrix} \textit{type-a} & \\ \textsc{feat-aa} & \textit{type-aa} \\ \textsc{feat-ab} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head} & \textit{type-aba} \\ \textsc{feat-abc} & \textit{type-abc} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix} \\ \textsc{feat-b} & \boxed{10}\textit{type-b} \end{bmatrix}$$

$$\begin{bmatrix} \textit{some-type} & \\ \textsc{feat-a} & \boxed{10}\begin{bmatrix} \textit{type-a} & \\ \textsc{feat-aa} & \textit{type-aa} \\ \textsc{feat-ab} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{head} & \textit{type-aba} \\ \textsc{feat-abc} & \textit{type-abc} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix} \\ \textsc{feat-b} & \boxed{10}\textit{type-b} \end{bmatrix}$$

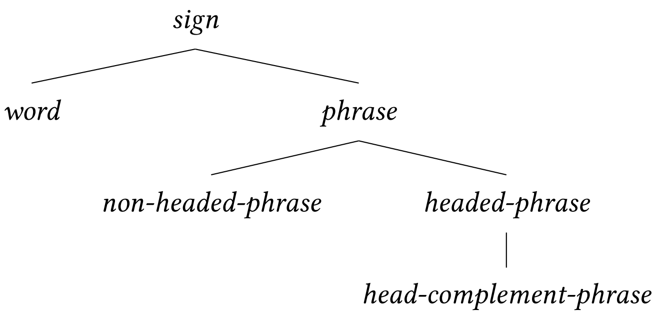


Figure 1: Type hierarchy for *sign*

Abbreviations

Acknowledgements

Chapter K

Gesture

Andy Lücking

Université de Paris, Goethe-Universität Frankfurt

The received view in (psycho)linguistics, dialogue theory and gesture studies is that co-verbal gestures, i.e. hand and arm movement, are part of the utterance and contribute to its content (Kendon 1980; McNeill 1992). The relationships between gesture and speech obey regularities that need to be defined in terms of not just the relative timing of gesture to speech, but also the linguistic form of that speech: for instance, prosody and syntactic constituency and headedness (Loehr 2007; Ebert et al. 2011; Alahverdzhieva et al. 2017). Consequently, speech-gesture integration is captured in grammar by means of a gesture-grammar interface. This chapter provides basic snapshots from gesture research, reviews constraints on speech-gesture integration and summarises their implementations into HPSG frameworks. Pointers to future developments conclude the exposition. Since there are already a couple of overviews on gesture such as Özyürek (2012), Wagner et al. (2014) and Abner et al. (2015), this chapter aims at distinguishing itself by providing a guided tour of research that focuses on using (mostly) standard methods for semantic composition in constraint-based grammars like HPSG to model gesture meanings.

1 Why gestures?

People talk with their whole body. A verbal utterance is couched in an intonation pattern that, via prosody, articulation speed or stress, function as *paralinguistic* signals (e.g. Birdwhistell 1970). The temporal dimension of paralinguistics

Andy Lücking. 2019. Gesture. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 887–934. Berlin: Language Science Press. DOI:??

gives rise to *chronemic* codes (Poyatos 1975; Bruneau 1980). *Facial expressions* are commonly used to signal emotional states (Ekman & Friesen 1978), even without speech (Argyle 1975), and are correlated to different illocutions of the speech acts performed by a speaker (Domaneschi et al. 2017). Interlocutors use *gaze* as a mechanism to achieve joint attention (Argyle & Cook 1976) or provide social signals (Kendon 1967). Distance and relative direction of speakers and addressees are organised according to culture-specific radii into social spaces (*proxemics*, Hall 1968). Within the inner radius of private space, tactile codes of *tacesics* (Kauffman 1971) are at work. Since the verbal and nonverbal communication means of face to face interaction may occur simultaneously, *synchrony* (i.e. the mutual overlap or relative timing of verbal vs. non-verbal communicative actions) is a feature of the multimodal utterance itself; it contributes, for instance, to identifying the word(s) that are affiliated to a gesture (Wiltshire 2007). A special chronemic case is signalling at the right moment – or, for that matter, missing the right moment (an aspect of communication dubbed *kairemics* by Lücking & Pfeiffer 2012: 600). Besides the manifold areas of language use, the conventionalised, symbolic nature of language secures language's primacy in communication, however (de Ruiter 2004). For thorough introductions into semiotics and multimodal communication see Nöth (1990), Posner et al. (1997–2004) or Müller et al. (2013–2014).

The most conspicuous non-verbal communication means of everyday interaction are hand and arm movements, known as *gestures* (in a more narrow sense which is also pursued from here on). In seminal works, McNeill (1985; 1992) and Kendon (1980; 2004) argue that co-verbal gestures, i.e. hand and arm movements, can be likened to words in the sense that they are part of a speaker's utterance and contribute to discourse. Accordingly, integrated speech-gesture production models have been devised (Kita & Özyürek 2003; de Ruiter 2000; Krauss et al. 2000) that treat utterance production as a multimodal process (see Section 4.4 for a brief discussion). Given gestures' imagistic and often spontaneous character, it is appealing to think of them as "postcards from the mind" (de Ruiter 2007: 21). Clearly, given this entrenchment in speaking, the fact that one can communicate meaning with non-verbal signals has repercussions to areas hitherto taken to be purely linguistic (in the sense of being related to the verbal domain). This section highlights some phenomena particularly important for grammar, including, for instance, *mixed syntax* (Slama-Cazacu 1976), or *pro-speech gesture*:

(1) He is a bit [*circular movement of index finger in front of temple*].

In (1), a gesture replaces a position that is usually filled by a syntactic constituent. The gesture is emblematically related to the property of *being mad* so that the mixed utterance from (1) is equivalent to the proposition that the referent of *he* is a bit mad.

Figure 1: *Die Skulptur die hat 'n* [*BETONsockel*] ('The sculpture has a [CONCRETE base]') [V5, 0:39]

The gesture shown in Figure 1 depicts the shape of a concrete base, which the speaker introduces into discourse as an attribute of a sculpture:[1]

(2) Die Skulptur die hat 'n [BETONsockel].
The sculpture it has a [CONCRETE base].
'The sculpture has a concrete base.'

The following representational conventions obtain: square brackets roughly indicate the portion of speech which overlaps temporally with the gesture (or more precisely, with the gesture stroke; see Figure 5 below) and upper case is used to mark main stress or accent. So both timing and intonation give clues that the gesture is related to the noun *Betonsockel* 'concrete base'. From the gesture, but not from speech, we get that the concrete base of the sculpture has the shape of a flat cylinder – thus, the gesture acts as a nominal modifier. There is a further complication, however: the gesture is incomplete with regard to its interpretation – it just depicts about half of cylinder. Thus, gesture interpretation may involve processes known from gestalt theory (see Lücking 2016 on a *good continuation* constraint relevant to (2)/Figure 1).

The speaker of the datum in Figure 2 uses just a demonstrative adverb in order to describe the shape of a building he is talking about:

[1]The examples in Figures 1, 2, 3, 4, 9 and 11 are drawn from the (German) *Speech and Gesture Alignment* corpus (SaGA, Lücking et al. 2010) and are quoted according to the number of the dialogue they appear in and their starting time in the respective video file (e.g. "V9, 5:16" means that the datum can be found in the video file of dialogue V9 at minute 5:16). Examples/Figures 4 and 9 have been produced especially for this volume; all others have also been used in Lücking (2013) and/or Lücking (2016).

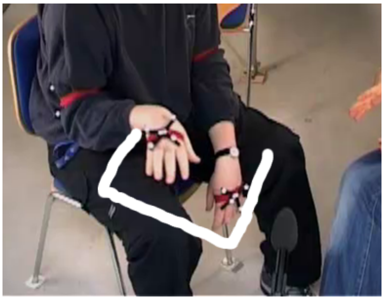


Figure 2: *Dann ist das Haus halt SO* [] (‘The house is like THIS []’) [V11, 2:32]

(3) Dann ist das Haus halt SO [].
Then is the house just like THIS [].
‘The house is like this [].’

The demonstrative shifts the addressee’s attention to the gesture, which accomplishes the full shape description, namely a cornered U-shape. In contrast to the example in Figure 1, the utterance associated with Figure 2 is not even interpretable without the gesture.

A lack of interpretability is shared by exophorically used demonstratives, which are *incomplete* without a demonstration act like a pointing gesture (Kaplan 1989: 490). For instance, Claudius would experience difficulties in understanding how serious Polonius is about his (Polonius’) conjecture about the reason of Hamlet’s (alleged) madness, if Polonius had not produced pointing gestures (Shakespeare, *Hamlet, Prince of Denmark* Act II, Scene 2; the third occurrence of *this* is anaphoric and refers back to Polonius’ conjecture):

(4) POLONIUS (*points to his head and shoulders*): Take this from this if this be otherwise.

In order for Claudius to interpret Polonius’ multimodal utterance properly, he has to associate correctly the two pointing gestures with the first two occurrences of *this* (cf. the problems discussed by Kupffer 2014). Polonius facilitates such an interpretation by means of a temporal coupling of pointing gestures and their associated demonstratives – a relationship that is called *affiliation*. The role of synchrony in multimodal utterances is further illustrated by the following example, (5), and Figure 3 (taken from Lücking 2013: 189):

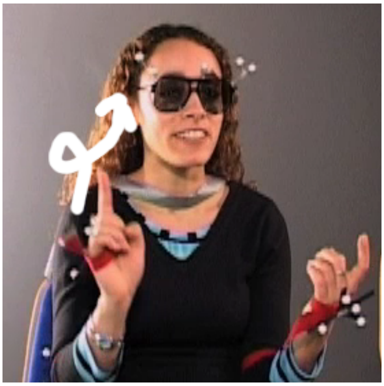


Figure 3: *Ich g[laube das sollen TREP]pen sein* (‘I think those should be STAIRcases’) [V10, 3:19]

(5) Ich g[laube das sollen TREP]pen sein.
I think those should STAIRcases be
‘I think those should be staircases.’

The first syllable of the German noun *Treppen* (*staircases*) carries main stress, indicated by capitalization. The square brackets indicate the temporal overlap between speech and gesture stroke, which is shown in Figure 3. The gesture attributes a property to the noun it attaches to: from the multimodal utterance, the observer retrieves the information that the speaker talks about spiral staircases. This interpretation assumes that the common noun is the affiliate of the gesture. Obviously, mere temporal synchrony is too weak to be an indicator of affiliation. In fact, there are speech-gesture affiliations without temporal overlap between gesture and verbal affiliate at all (e.g. Lücking et al. 2004). Therefore, temporal overlap or vicinity is just one indicator of affiliation. A second one is intonation: a gesture is usually related to a stressed element in speech (McClave 1994; Nobe 2000; Loehr 2004; 2007). As a result, multimodal communication gives rise to a complex “peak pattern” (Tuite 1993; Loehr 2004; Jannedy & Mendoza-Denton 2005).

The interpretation of a gesture changes with different affiliations. Suppose the gesture from Figure 3 is produced in company to stressed *glaube* (*think*) instead of *staircases*:

(6) Ich G[LAUbe das sollen Trep]pen sein.
I THINK those should staircases be
‘I think those should be staircases.’

Now the spiral movement is interpreted as a metaphorical depiction of a psychological process. Thus, the interpretation of a gesture depends on the integration point (affiliation), which in turn is marked by temporal vicinity, prosody and syntactic constituency of the candidate affiliate (Alahverdzhieva et al. 2017).

The crucial observations in any case are that gestures contribute to propositional content and take part in pragmatic processes. Interestingly, gestures share the latter aspect with laughter, which also has propositional content (Ginzburg et al. 2015), for instance, when referring to real world events. Thus, a multimodal utterance may express a richer content than speech alone, as in (5), or a content equivalent to speech, as in (6); it can even express less than speech or contradict speech:[2]

> The nonverbal act can repeat, augment, illustrate, accent, or contradict the words; it can anticipate, coincide with, substitute for or follow the verbal behavior; and it can be unrelated to the verbal behavior. (Ekman & Friesen 1969: 53)

Contradictions or speech-gesture mismatches can occur when saying "right" but pointing left (as can be observed in everyday life but also been found in SaGA, e.g. in dialogue V24, at 4:50). A more complex case is given in (7) and Figure 4, where the speaker talks about a "rectangular arch" (which is of course a *contradictio in adiecto* in itself), but produces a roundish movement with the extended index finger of her right hand (the object she talks about is an archway). Note that the gesture just overlaps with "rectangular": its temporal extension in (7) is again indicated by means of square brackets within the original German utterances. The main stress is on the first syllable of the adjective and the noun receives secondary stress. The dots ("..") mark a short pause, so the gesture starts before "rechteckiger".

(7)	so'n	so'ne Art	[.. RECHTecki]ger	BOgen
	such an	such kind of	.. RECTangular	ARrch

‘kind of rectangular arch’

An obvious interpretation of this mismatch is that "rectangular" is a slip of the tongue; interestingly, we found no "slip of the hand" in our data so far (which may be a hint to a possibly imagistic origin of gestures, as assumed in some production models; cf. Section 4.4).

[2] In case of contradiction or speech-gesture mismatch, the resulting multimodal utterance is perceived as ill-formed and induces N400 effects (Wu & Coulson 2005; Kelly et al. 2004).

Figure 4: “so’n so’ne Art [.. RECHTecki]ger BOgen” (*kind of RECTangular ARch*) [V4, 1:47].

Moving from sentence to dialogue, *interactive gestures* are bound up with turn management, among other things (Bavelas et al. 1992; 1995). For instance, pointing gestures can be used to indicate the next speaker (Rieser & Poesio 2009). Interestingly, speaker-indicating pointings are typically not produced with an outstretched index finger, but with an open hand (an example is given in Figure 14 in Section 3.6). Thus, irrespective of the question whether grammar is inherently multimodal, dialogue theory has to deal at least with certain non-verbal interaction means in any case (see also Lücking, Ginzburg & Cooper (2019), Chapter I of this volume).

While there is ample evidence that at least some gestures contribute to the content of the utterance they co-occur with, does this also mean that they are part of the content *intended to be communicated*? A prominent counter-example is gesturing on the telephone (see Bavelas et al. 2008 for an overview of a number of respective studies). Since such gestures are not observable for the addressee, they cannot reasonably be taken to be a constituent of the content intended for communication. Rather, “telephone gestures” seem to be speaker-oriented, presumably facilitating word retrieval. The fact that it is difficult to suppress gesturing even in absence of an addressee speaks in favour of a multimodal nature if not of language, then at least of speaking and surely interacting. Furthermore, the lion’s share of everyday gestures seems to consist of rather sloppy movements that do not contribute to the content of the utterance in any interesting sense, though they might signal other information like speaker states. In this sense they are contingent, as opposed to being an obligatory semantic component (Lücking 2013). Gestures (or some other demonstration act) can become obligatory when they are produced within the scope of a demonstrative expression (recall (3)/Figure 2). A concurrent use with demonstratives is also one of the hallmarks collected by Cooperrider (2017) in order to distinguish *foreground* from *background* gestures (the other hallmarks are absence of speech, co-organization with speaker gaze and speaker effort). This distinction reflects two traditions within gesture studies: according to one tradition most prominently bound up with the work of McNeill (1992), gesture is a *byproduct* of speaking and therefore opens a "window into the speaker's mind". The other tradition, represented early on by Goodwin (2003) and Clark (1996), conceives gestures as a *product* of speaking, that is, as interaction means designed with a communicative intention. Since a gesture cannot be both a byproduct and a product at the same time, as noted by Cooperrider (2017), a bifurcation that is rooted in the cause and the production process of the gesture has to be acknowledged (e.g. gesturing on the phone is only puzzling from the product view, but not from the byproduct one). We will encounter this distinction again when briefly reviewing speech-gesture production models in Section 4.4. Gestures of both species are covered in the following.

2 Kinds of gestures

Pointing at an object seems to be a different kind of gesture than mimicking drinking by moving a bent hand (i.e. virtually holding something) towards the mouth while slightly rotating the back of hand upwards. And both seem to be different from actions like scratching or nose-picking. On such grounds, gestures are usually assigned to one or more classes of a taxonomy of gesture classes. Gestures that fulfil a physiological need (such as scratching, nose-picking, foot-shaking or pen-fiddling) have been called *adaptors* (Ekman & Friesen 1969) and are not dealt with further here (but see Żywiczyński et al. 2017 for evidence that adaptors may be associated with turn transition points in dialogue). Gestures that have an intrinsic relation to speech and what is communicated have been called *regulators* and *illustrators* (Ekman & Friesen 1969) and cover a variety of gesture classes. These gesture classes are characterized by the function performed by a gesture and the meaning relation the gesture bears to its content. A classic taxonomy consists of the following inventory (McNeill 1992):

- iconic (or representational) gestures. Spontaneous hand and arm movements that are commonly said to be based on some kind of resemblance relation.[3] Iconic gestures employ a mode of representation such as *draw-*

[3]But see footnote 8 in Section 3.5 for pointers to critical discussions of resemblance as a sign-bearing relation.

ing, *modelling*, *shaping* or *placing* (Streeck 2008; Müller 1998).

- deictic gestures (pointing). Typically hand and arm movements that perform a demonstration act. In which way pointing is standardly accomplished is subject to culture-specific conventions (Wilkins 2003). In principle, any extended body part, artefact or locomotor momentum will serve the demonstrative purpose. Accordingly, there are deictic systems that involve lip-pointing (Enfield 2001) and nose-pointing (Cooperrider & Núñez 2012). Furthermore, under certain circumstances, pointing with the eyes (gaze-pointing) is also possible (Hadjikhani et al. 2008). Note further that the various deictic means can be interrelated. For instance, manual pointing can be differentiated by cues of head and gaze (Butterworth & Itakura 2000). Furthermore, pointing with the hand can be accomplished by various hand shapes: Kendon & Versante (2003) distinguish *index finger pointing*, (with a *palm down* and a *palm vertical* variant) *thumb pointing*, and *open hand pointing* (again with various palm orientations). Kendon & Versante (2003: 109) claim that "the form of pointing adopted provides information about how the speaker wishes the object being indicated to be regarded". For instance, pointing with the thumb is usually used when the precise location of the intended referent is not important (Kendon & Versante 2003: 121–125), while the typical use of index finger palm down pointing is to single out an object (Kendon & Versante 2003: 115). Open hand pointing has a built-in metonymic function since the object pointed at is introduced as an example for issues related to the current discourse topic (what in semantic parlance can be conceived as the *question under discussion*; see, e.g. Ginzburg 2012). For instance, with 'open hand palm vertical', one indicates the *type* of the object pointed at instead of the object itself (Kendon & Versante 2003: 126).

- beats (rhythmic gestures, baton). Hand and arm movements that are coupled to the intonational or rhythmic contour of the accompanying speech. Beats lack representational content but are usually used for an emphasising effect. "The typical beat is a simple flick of the hand or fingers up and down, or back and forth" (McNeill 1992: 15). Hence, a beat is a gestural means to accomplish what is usually expressed by vocal stress, rhythm or speed in speech.

- emblem (lexicalized gestures). In contrast to the other classes, emblems are special in that they follow a fully conventionalized form-meaning relation. A common example in Western countries is the thumbs-up gesture, signalling "approval or encouragement" (Merriam Webster online dictionary[4]). Emblems may also be more local and collected within a dictionary like the dictionary of everyday gestures in Bulgaria (Kolarova 2011).

Reconsidering gestures that have been classified as beats, among other gestures, Bavelas et al. (1992) observed that many of the stroke movements accomplish functions beyond rhythmic structuring or emphasis. Rather, they appear to contribute to dialogue management and have been called *interactive gestures*. Therefore, these gestures should be added to the taxonomy:

- interactive gestures. Hand and arm movements that accomplish the function "of helping the interlocutors coordinate their dialogue" (Bavelas et al. 1995: 394). Interactive gestures include pointing gestures that serve turn allocation ("go ahead, it's your turn") and gestures that are bound up with speaker attitudes or the relationship between speaker and addressee. Examples can be found in 'open palm/palm upwards' gestures used to indicate the information status of a proposition ("as you know") or the mimicking of quotation marks in order to signal a report of direct speech (although this also has a clear iconic aspect).

The gesture classes should not be considered as mutually exclusive categories, but rather as dimensions according to which gestures can be defined, allowing for multi-dimensional cross-classifications (McNeill 2005; Gerwing & Bavelas 2004). For instance, it is possible to superimpose pointing gestures with iconic traits. This has been found in the study on pointing gestures described in Kranstedt et al. (2006a), where two participants at a time were involved in an identification game: one participant pointed at one of several parts of a toy airplane scattered over a table, the other participant had to identify the pointed object. When pointing at a disk (a wheel of the toy airplane), some participants used index palm down pointing, but additionally turned around their index finger in a circle – that is, the pointing gesture not only locates the disk (deictic dimension) but also depicted its shape (iconic dimension). See Özyürek (2012) for an overview of various gesture classification schemes.

In addition to classifying gestures according to the above-given functional groups, a further distinction is usually made with regard to the ontological place of their referent: representational and deictic gestures can relate to concrete or to

[4] https://www.merriam-webster.com/dictionary/thumbs-up, accessed 20th August 2018. The fact that emblems can be lexicalized in dictionaries emphasizes their special, conventional status among gestures.

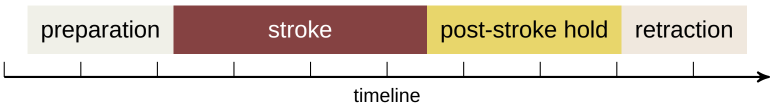


Figure 5: Gesture phases

abstract objects or scenes. For instance, an iconic drawing gesture can metaphorically display the notion "genre" via a conduit metaphor (McNeill 1992: 14):

(8) It [was a Sylves]ter and Tweety cartoon.
both hands rise up with open palm handshape, palms facing; brackets indicate segments concurrent with the gesture stroke (see Figure 5).

The gesture in (8) virtually holds an object, thus depicting the abstract concept of the genre of being a Sylvester and Tweety cartoon as a bounded container. Accordingly, gestures can be cross-classified into *concrete* and *abstract* or *metaphorical* ones (see the volume of Cienki & Müller 2008 on gesture and metaphor).

On the most basic, kinematic level, the movement of a prototypical gesture follows an "anatomic triple": gestures have to be partitioned into at least a preparation, a stroke, and a retraction phase (Kendon 1972). The gesture phases are shown in the diagram in Figure 5. The stroke is the movement part that carries the gesture's meaning. It can be "frozen", leading to a post-stroke hold. If a stroke has to wait for its affiliated expression(s), a pre-stroke hold can also arise (Kita et al. 1998). The preparation and retraction phases bring hand and arms into and out of the stroke, respectively. Unless stated otherwise, when talking about gestures in what follows (and in hindsight concerning the examples given in Section 1), the stroke phase, which is the "gesture proper" or the "semantically interpretable" phase, is referred to.

Perhaps it should be noted that the spontaneous, usually co-verbal hand and arm movements considered in this chapter are different from the signed signs of sign languages (see Steinbach & Holler 2019, Chapter J of this volume) and pantomime (neither spontaneous nor co-verbal).[5]

3 Gestures in HPSG

Integrating a gesture's contribution into speech was initiated in computer science (Bolt 1980). Coincidentally, these early works used typed feature structure

[5] In languages like German, the difference between free gesticulation and sign language signs is also reflected terminologically: the former are called *Gesten*, the latter *Gebärden*.

descriptions akin to the descriptive format used in HPSG grammars. Though linguistically limited, the crucial invention has been a *multimodal chart parser*, that is, an extension of chart parsing that allows the processing of input in two modalities (namely speech and gesture). Such approaches are reviewed in Section 3.2. Afterwards, a more elaborate gesture representation format is introduced that makes it possible to encode the observable form of a gesture in terms of kinematically derived attribute-value structures (Section 3.3). Following the basic semiotic distinction between deictic (or indicating or pointing) gestures and iconic (or representational or imagistic) gestures, the analysis of each class of gestures is exemplified in Sections 3.4 and 3.5, respectively. To begin with, however, some basic phenomena that should be covered by a multimodal grammar are briefly summarized in Section 3.1.

3.1 Basic empirical phenomena of grammatical gesture integration

With regard to grammar-gesture integration, three main phenomena have to be dealt with:

- What is the meaning of a gesture? On which grounds should semantic representations or truth conditions be assigned to hand and arm movements?
- What is the affiliate of a gesture, that is, its verbal attachment site?
- What is the result of multimodal integration, that is, the outcome of composing verbal and non-verbal meanings?

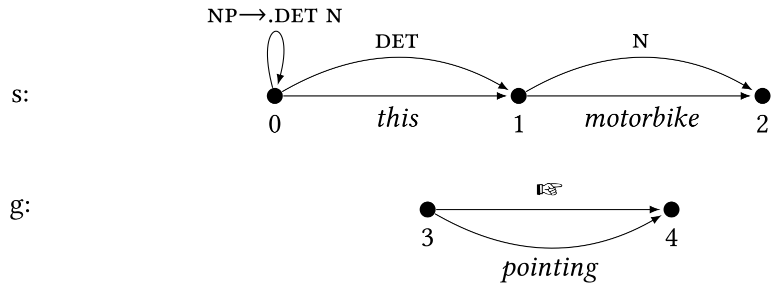
Given the linguistic significance of gestures as sketched in the preceding sections, formal grammar- and semantic-oriented accounts of speech-gesture integration have recently been developed that try to deal with (at least one of) the three basic phenomena, though with different priorities, including Alahverdzhieva (2013), Alahverdzhieva & Lascarides (2010), Ebert (2014), Giorgolo (2010), Giorgolo & Asudeh (2011), Lücking (2013; 2016), Rieser (2008; 2011; 2015), Rieser & Poesio (2009) and Schlenker (2018). It should be noted that the first basic question does not have to be considered a question for grammar, but can be delegated to a foundational theory of gesture meaning. Here gestures turn out to be like words again, where "semantic theory" can refer to explaining meaning (foundational) or specifying meaning (descriptive) (Lewis 1970: 19). In any case, the HPSG-related approaches are briefly reviewed below.

3.2 Precursors

Using typed feature structure descriptions to represent the form and meaning of gestures goes back to computer science approaches to human-computer interaction. For instance, the *QuickSet* system (Cohen et al. 1997) allows users to operate on a map and move objects or lay out barbed wires (the project was funded by a grant from the US army) by giving verbal commands and manually indicating coordinates. The system processes voice and pen (gesture) input by assigning signals from both media representations in the form of attribute-value matrices (AVMs) (Johnston 1998; Johnston et al. 1997). For instance, *QuickSet* will move a vehicle to a certain location on the map when asked to *Move this*[☞] *motorbike to here*[☞], where '☞' represents an occurrence of touch gesture (i.e. pen input).

Since a conventional constrained-based grammar for speech-only input rests on a "unimodal" parser, Johnston (1998) and Johnston et al. (1997) developed a *multimodal chart parser*, which is still a topic of computational linguistics (Alahverdzhieva et al. 2012) (see also Bender & Emerson 2019, Chapter H of this volume). A multimodal chart parser consists of two or more layers and allows for layer-crossing charts. The multimodal NP *this*[☞] *motorbike*, for instance, is processed in terms of a multimodal chart parser covering a speech (s) and a gesture (g) layer:

(9)



A multimodal chart or *multichart* is defined in terms of sets of identifiers from both layers. Possible multicharts from (9) include the following ones:

(10) multichart 1: {[s,0,1], [g,3,4]}
multichart 2: {[s,1,2], [g,3,4]}
…

The basic rule for integrating spatial gestures with speech commands is the *basic integration scheme* (Johnston 1998; Johnston et al. 1997), reproduced in (11):

(11)
$$\begin{bmatrix}
\text{LHS}: & \begin{bmatrix} \text{CAT}: & \textit{command} \\ \text{MODALITY}: & \boxed{2} \\ \text{CONTENT}: & \boxed{1} \\ \text{TIME}: & \boxed{3} \end{bmatrix} \\
\text{RHS}: & \begin{bmatrix} \text{DTR1}: & \begin{bmatrix} \text{CAT}: & \textit{located_command} \\ \text{MODALITY}: & \boxed{6} \\ \text{CONTENT}: & \boxed{1}\left[\textit{location}\ \boxed{5}\right] \\ \text{TIME}: & \boxed{7} \end{bmatrix} \\ \text{DTR2}: & \begin{bmatrix} \text{CAT}: & \textit{spatial_gesture} \\ \text{CONTENT}: & \boxed{5} \\ \text{MODALITY}: & \boxed{9} \\ \text{TIME}: & \boxed{10} \end{bmatrix} \end{bmatrix} \\
\text{CONSTRAINTS}: & \left\{ \begin{array}{l} \text{OVERLAP}(\boxed{7},\boxed{10}) \vee \text{FOLLOW}(\boxed{7},\boxed{10},4\text{S}) \\ \text{TOTAL-TIME}(\boxed{7},\boxed{10},\boxed{3}) \\ \text{ASSIGN-MODALITY}(\boxed{6},\boxed{9},\boxed{2}) \end{array} \right\}
\end{bmatrix}$$

The AVM in (11) implements a mother-daughter structure along the lines of a context-free grammar rule, where a left-hand side (LHS) expands to a right-hand side (RHS). The right-hand side consists of two constituents (daughters DTR1 and DTR2), a verbal expression (*located_command*) and a gesture. The semantic integration between both modalities is achieved in terms of structure sharing, see tag $\boxed{5}$: the spatial gesture provides the location coordinate for the verbal command.

The bimodal integration is constrained by a set of restrictions, mainly regulating the temporal relationship between speech and gesture (see tags $\boxed{7}$ and $\boxed{10}$ in the CONSTRAINTS set): the gesture may overlap with its affiliated word in time, or follow it in at most four seconds (see the 4S under CONSTRAINTS). An integration scheme highly akin to that displayed in (11) also underlies current grammar-oriented approaches to deictic and iconic gestures (see Sections 3.4 and 3.5 below).

3.3 Representing gestures with AVMs

Representing the formal features of gestures in terms of attribute-value matrices has been initiated in robotics (Kopp et al. 2004). A representation format that captures the "phonological", physical-kinematic properties of a gesture is designed according to the moveable junctions of arms and hands. For instance, the representation of the gesture in Figure 3 according to the format used in Lücking et al. (2010) is given in (12):

(12)
$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{right hand}} \\
\textsc{handshape} & \left[\begin{array}{ll} \textsc{shape} & \textit{G} \\ \textsc{path} & \textit{0} \\ \textsc{dir} & \textit{0} \end{array}\right] \\
\textsc{palm} & \left[\begin{array}{ll} \textsc{orient} & \textit{PAB>PAB/PUP>PAB} \\ \textsc{path} & \textit{0} \\ \textsc{dir} & \textit{0} \end{array}\right] \\
\textsc{boh} & \left[\begin{array}{ll} \textsc{orient} & \textit{BUP>BTB/BUP>BUP} \\ \textsc{path} & \textit{arc>arc>arc} \\ \textsc{dir} & \textit{MR>MF>ML} \end{array}\right] \\
\textsc{wrist} & \left[\begin{array}{ll} \textsc{position} & \textit{P-R} \\ \textsc{path} & \textit{line} \\ \textsc{dir} & \textit{MU} \\ \textsc{dist} & \textit{D-EK} \\ \textsc{extent} & \textit{small} \end{array}\right] \\
\textsc{sync} & \left[\begin{array}{ll} \textsc{config} & \textit{BHA} \\ \textsc{rel.mov} & \textit{LHH} \end{array}\right]
\end{array}\right]$$

The formal description of a gestural movement is given in terms of the handshape, the orientations of the palm and the back of the hand (BOH), the movement trajectory (if any) of the wrist and the relation between both hands (synchronicity, SYNC). The handshape is drawn from the fingerspelling alphabet of American Sign Language, as illustrated in Figure 6. The orientations of palm and back of hand are specified with reference to the speaker's body (e.g. *PAB* encodes "palm away from body" and *BUP* encodes "back of hand upwards"). Movement features for the whole hand are specified with respect to the wrist: the starting position is given and the performed trajectory is encoded in terms of the described path and the direction and extent of the movement. Position and extent are given with reference to the *gesture space*, that is, the structured area within the speaker's immediate reach (McNeill 1992: 86–89) – see the left-hand side of Figure 7. Originally, McNeill considered the gesture space as "a shallow disk in front of the speaker, the bottom half flattened when the speaker is seated" (McNeill 1992: 86). However, also acknowledging the distance of the hand from the speaker's body (feature DIST) turns the shallow disk into a three-dimensional space, giving rise to the three-dimensional model displayed on the right-hand side of Figure 7. The gesture space regions known as *center-center*, *center* and *periphery*, possibly changed by location modifiers (*upper right*, *right*, *lower right*, *upper left*, *left*, *lower left*), are now modelled as nested cuboids. Thus, gesture space is structured according to all three body axes: the sagittal, the longitudinal and the transverse axes. Annotations straightforwardly transfer to the three-dimensional gesture

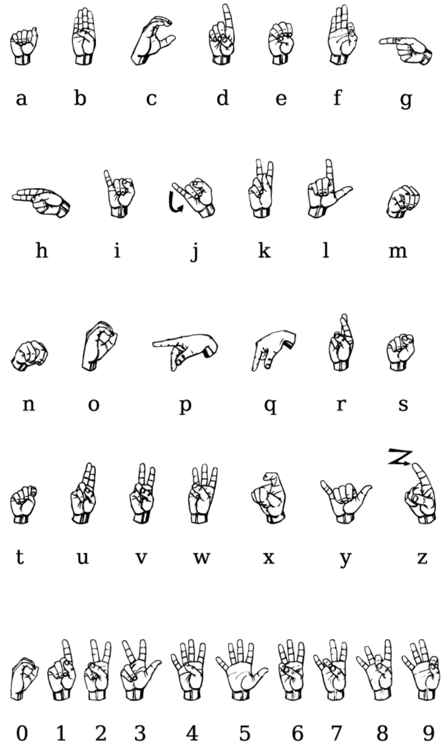


Figure 6: American Sign Language fingerspelling alphabet (image released to Public Domain by user Ds13 in the English Wikipedia at 18th December 2004, https://commons.wikimedia.org/wiki/File:Asl_alphabet_gallaudet.png)

space model. Such a three-dimensional gesture space model is assumed throughout this chapter. Complex movement trajectories through the vector space can describe a rectangular or a roundish path (or mixtures of both). Both kinds of movements are distinguished in terms of *line* or *arc* values of feature PATH. An example illustrating the difference is given in Figure 8. A brief review of gesture annotation can be found in Section 4.1.

3.4 Pointing Gestures

Pointing gestures are *the* prototypical referring device: they probably pave a way to reference in both evolutionary and language acquisition perspectives (Bruner

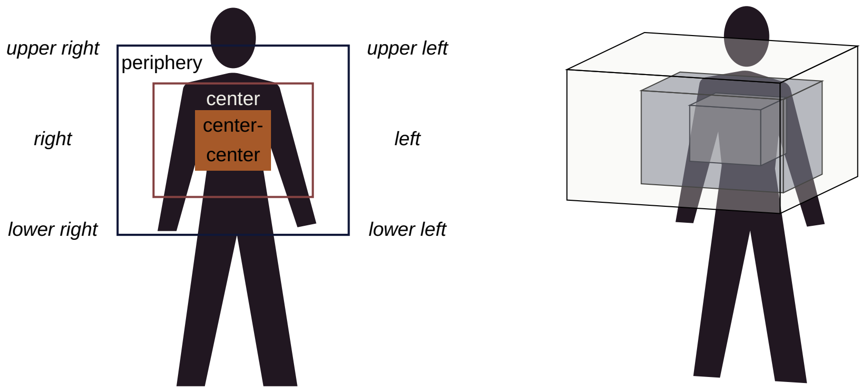


Figure 7: Gesture Space (left hand side is simplified from McNeill 1992: 89). Although originally conceived as a structured "shallow disk" McNeill (1992: 86), adding distance information gives rise to a three-dimensional gesture space model as illustrated on the right-hand side.

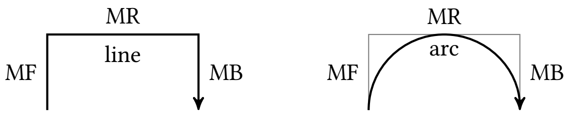


Figure 8: The same sequence of direction labels can give rise to an open rectangle or a semicircle, depending on the type of concatenation (Lücking 2016: 385).

1998; Masataka 2003; Matthews et al. 2012); they are predominant inhabitants of the "deictic level" of language, interleaving the symbolic (and the iconic) levels (Levinson 2008, see also Bühler 1934); they underlie reference in *Naming Games* in computer simulation approaches (Steels 1995) (for a semantic assessment of naming and categorisation games, see Lücking & Mehler 2012).

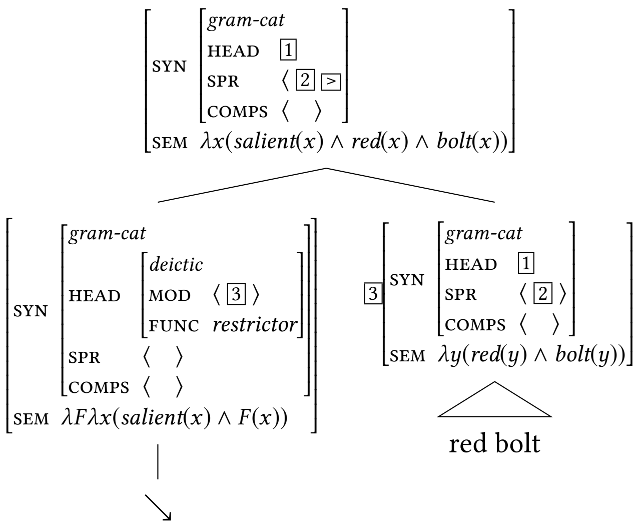
With regard to deictic gestures, Fricke (2012: Sec. 5.4) argues that deictic words within noun phrases – her prime example is German *so* 'like this' – provide a *structural*, that is, *language-systematic* integration point between the vocal plane of conventionalized words and the non-vocal plane of body movement. Therefore, with this conception, not only utterance production but *grammar* is inherently multimodal.

The referential import of the pointing gesture has been studied experimentally in some detail (Bangerter & Oppenheimer 2006; Kranstedt et al. 2006b,a; van der Sluis & Krahmer 2007). As a result, it turns out that pointings do not rely on a direct "laser" or "beam" mechanism (McGinn 1981). Rather, they serve a (more or less rough) locating function (Clark 1996) that can be modelled in terms of a *pointing cone* (Kranstedt et al. 2006b; Lücking et al. 2015). These works provide an answer to the first basic question (cf. Section 3.1): pointing gestures have a "spatial meaning" which focuses or highlights a region in relation to the direction of the pointing device. Such a spatial semantic model has been introduced in Rieser (2004) under the name of *region pointing*, where the gesture adds a locational constraint to the restrictor of a noun phrase. In a related way, two different functions of a pointing gesture have been distinguished by Kühnlein et al. (2002), namely singling out an object (13a) or making an object salient (13b).

(13) a. $\lambda F \lambda x(x = c \wedge F(x))$
  b. $\lambda F \lambda x(\mathit{salient}(x) \wedge F(x))$

The approach is expressed in lambda calculus and couched in an HPSG framework. The derivation of the instruction *Take the red bolt* plus a pointing gesture is exemplified in (14). A pointing gesture is represented by means of "↘" and takes a syntactic position within the linearized inputs according to the start of the stroke phase. For instance, the pointing gesture in (14a) occurred after *the* has been articulated but before *red* is finished. The derivation of the multimodal N′ constituent is shown in (14b).

(14) a. Take [the ↘[N′ [N′ red bolt]]]
  b.



The spatial model is also adopted in Lascarides & Stone (2009), where the region denoted by pointing is represented by a vector $\vec{p}$. This region is an argument

Figure 9: *Und man*[*chmal ist da auch ein EISverkäufer*] ('and some[times there's an ICE cream guy]'), [V5, 7:20]

to function v, however, which maps the projected cone region to $v(\vec{p})$, the space-time talked about, which may be different from the gesture space (many more puzzles of local deixis are collected by Klein 1978 and Fricke 2007).

Let us illustrate some aspects of pointing gesture integration by means of the real world example in (15) and Figure 9, taken from dialogue V5 of the SaGA corpus.

(15) Und man[chmal ist da auch ein EISverkäufer].
And some[times] is there also an ICE cream guy.
'And sometimes there's an ice cream guy'

The context in which the gesture appears is the following: the speaker describes a route which goes around a pond. He models the pond with his left hand, a post-stroke hold (cf. Figure 5) held over several turns. After having drawn the route around the pond with his right hand, the pointing gesture in Figure 9 is produced. The pointing indicates the location of an ice cream vendor in relation to the pond modelled in gesture space. Such instances of indirect or proxy pointing have been interpreted as *dual points* by Goodwin (2003); in standard semantics they are analysed in terms of *deferred reference*, where one thing is indicated but another but related thing is referred to (Quine 1950; Nunberg 1993). The "duality" or "deference" involved in the datum consists of a mapping from the location indicated in gesture space onto a spatial area of the described real world situation. Such mappings are accounted for by the function v that shifts the pointing cone area from gesture space $\vec{p}$ to some other space $v(\vec{p})$ (Lascarides & Stone 2009). So, the deictic gesture locates the ice cream vendor. Since it is held during nearly the whole utterance, its affiliate expression "Eisverkäufer" (*ice cream* *guy*) is picked out due to carrying primary accent (indicated by capitalization).[6] Within HPSG, such constraints can be formulated within an interface to metrical trees from the phonological model of Klein (2000) or phonetic information packing from Engdahl & Vallduví (1996) – see also Kuthy (2019), Chapter D of this volume. The well-developed basic integration scheme of Alahverdzhieva et al. (2017: 445) rests on a strict speech and gesture overlap and is called the *Situated Prosodic Word Constraint*, which allows the combination of a speech daughter (S-DTR) and a gesture daughter (G-DTR) – see Figure 10. The Situated Prosodic Word Constraint applies to both deictic and iconic gestures. Under certain conditions, including when a deictic gesture is direct (i.e. $\vec{p} = v(\vec{p})$), however, the temporal and prosodic constraints can be relaxed for pointings.

$$
\begin{bmatrix}
\textit{word} & \\
\text{OVERLAP} & \langle @7, @8 \rangle \\
\text{TIME} & @7 \cup @8 \\
\text{PHON} & @3 \\
\text{SYNSEM} & \begin{bmatrix} \text{CAT} & @5 \\ \text{CONT} & \begin{bmatrix} \text{RELS} & \boxed{C_{rel}} \oplus \boxed{S_{rel}} \oplus \boxed{G_{rel}} \\ \text{HCONS} & \boxed{S_{hc}} \oplus \boxed{G_{hc}} \end{bmatrix} \end{bmatrix} \\
\text{S-DTR} & \begin{bmatrix} \textit{word} & \\ \text{TIME} & @7 \\ \text{PHON} & @3\textit{nuclear\_or\_pre-nuclear} \\ \text{SYNSEM} & \begin{bmatrix} \text{CAT} & @5 \\ \text{CONT} & \begin{bmatrix} \text{RELS} & \boxed{S_{rel}} \\ \text{HCONS} & \boxed{S_{hc}}\textit{hcons} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\text{G-DTR} & \begin{bmatrix} \textit{depicting\_or\_deictic} & \\ \text{TIME} & @8 \\ \text{SYNSEM} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{G-FEATURE} & \textit{value} \\ \ldots & \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{RELS} & \boxed{G_{rel}} \\ \text{HCONS} & \boxed{G_{hc}} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\
\text{C-CONT.RELS} & \boxed{C_{rel}}
\end{bmatrix}
$$

Figure 10: *Situated Prosodic Word Constraint* (Alahverdzhieva et al. 2017: 445)

[6] Semantically, other integration points are possible, too, most notably with "da" (*there*). However, the intonation-based integration point patterns well with observations of the affiliation behaviour of iconic gestures, as indicated with respect to examples (5) and (6) in Section 1. Concerning deictic gestures, a constraint that favours affiliation to deictic words over affiliation to stressed words (if they differ at all) seems conceivable nonetheless.

In order to deal with gestures that are affiliated with expressions that are larger than single words, Alahverdzhieva et al. (2017) also develop a phrase or sentence level integration scheme, where the stressed element has to be a semantic head (in the study of Mehler & Lücking 2012, 18.8% of the gestures had a phrasal affiliate). In this account, the affiliation problem (the second desideratum identified in Section 3.1) has a well-motivated solution on both the word and the phrasal levels, at least for temporally overlapping speech-gesture occurrences (modulo the conditioned relaxations for pointings). Semantic integration of gesture location and verbal meaning (the third basic question from Section 3.1) is brought about using the underspecification mechanism of *Robust Minimal Recursion Semantics* (RMRS), a refinement of *Minimal Recursion Semantics* (MRS) (Copestake et al. 2005), where basically scope as well as arity of elementary expressions is underspecified (Copestake 2007) – see the RELS and HCONS features in Figure 10. For some background on (R)MRS see the above given references, or see Koenig & Richter (2019), Chapter C of this volume.

A dialogue-oriented focus on pointing is taken in Lücking (2018): here, pointing gestures play a role in formulating processing instructions that guide the addressee in where to look for the referent of demonstrative noun phrases.

3.5 Iconic Gestures

There is nearly no semantic work on the grounds according to which the meanings assigned to iconic gestures should be assigned to them in the first place (this is the first basic question from Section 3.1). Semantic modelling usually focuses on the interplay of (in this sense presumed) gesture content with speech content, that is, on the third of the basic questions from Section 3.1. Schlenker (2018: 296) is explicit in this respect: "It should be emphasized that we will not seek to explain how a gesture […] comes to have the content that it does, but just ask how this content interacts with the logical structure of a sentence".[7] Two exceptions, however, can be found in the approaches of Rieser (2010) and Lücking (2013; 2016). Rieser (2010) tries to extract a "depiction typology" out of a speech-and-gesture corpus where formal gesture features are correlated with topological clusters consisting of geometrical constructs. Thus, he tries to address the first basic question from Section 3.1 in terms of an empirically extracted gesture typology. These geometrical objects are used in order to provide a possibly underspecified semantic representation for iconic gestures, which is then integrated into word meaning via lambda calculus (Hahn & Rieser 2010; Rieser 2011). The work of Lücking (2013;

[7] The omission indicated by "[…]" just contains a reference to an example in the quoted paper.

2016) is inspired by Goodman's notion of *exemplification* (Goodman 1976), that is, iconic gestures are connected to semantic predicates in terms of a reversed denotation relation: the meaning of an iconic gesture is given in terms of the set of predicates which have the gesture event within their denotation. In order to make this approach work, common perceptual features for predicates are extracted from their denotation and represented as part of a lexical extension of their lexemes, serving as an interface between hand and arm movements and word meanings. This conception in turn is motivated by psychophysic theories of the perception of biological events (Johansson 1973), draws on philosophical similarity conceptions beyond isomorphic mappings (Peacocke 1987),[8] and, using a somewhat related approach, has been proven to work in robotics by means of imagistic description trees (Sowa 2006). These perceptual features serve as the integration locus for iconic gestures, using standard unification techniques. The integration scheme for achieving this is the following one (Lücking 2013: 249) (omitting the time constraint used in the basic integration scheme in 11):

(16)
$$\begin{bmatrix} \textit{sg-ensemble} & \\ \text{PHON} & \boxed{12} \\ \text{CAT} & \boxed{2} \\ \text{CONT} & \boxed{3}\begin{bmatrix}\text{RESTR} & \left\langle \ldots, \boxed{5}\begin{bmatrix}\textit{pred} & \\ \text{CVM} & \boxed{1}\end{bmatrix}, \ldots \right\rangle\end{bmatrix} \\ \text{S-DTR} & \begin{bmatrix}\textit{verbal-sign} & \\ \text{PHON} & \boxed{12}\begin{bmatrix}\text{ACCENT} & \boxed{6}\end{bmatrix} \\ \text{CAT} & \boxed{2} \\ \text{CONT} & \boxed{3}\end{bmatrix} \\ \text{G-DTR} & \begin{bmatrix} \textit{gesture-vec} & \\ \text{AFF} & \left\langle\begin{bmatrix}\text{PHON} & \begin{bmatrix}\text{ACCENT} & \boxed{6}\,\textit{marked}\end{bmatrix}\end{bmatrix}\right\rangle \\ \text{TRAJ} & \boxed{1} \\ \text{CONT} & \begin{bmatrix}\text{MODE} & \textit{exemplification} \\ \text{EX-PRED} & \boxed{5}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Comparable to a modifier, a gesture attaches to an affiliate via feature AFF, which in turn is required to carry intonational accent, expressed in terms of information packaging developed by Engdahl & Vallduví (1996) (cf. Kuthy 2019, Chapter D of this volume). The semantic contribution of a gesture is contributed via the new semantic mode *exemplification*, that is, a gesture displays a predication from the RESTR list of its affiliate. The exemplification interface is established

[8] That mere resemblance, usually associated with iconic signs, is too empty a notion to provide the basis for a signifying relation has been emphasised on various occasions (Burks 1949; Bierman 1962; Eco 1976; Goodman 1976; Sonesson 1998).

using the format of vector semantics developed by Zwarts & Winter (2000) and Zwarts (2003) in order to capture the semantic contribution of locative prepositions, motion verbs and shape adjectives, among other things. This involves two steps: on the one hand, the representation of a gesture (cf. Section 3.3) is mapped onto a vectorial representation; on the other hand, the content of place and form predicates is enriched by abstract psychophysic information in the sense of Johansson (1973) (see above), also spelled out in terms of vector representations. Both steps are illustrated by means of the simple example shown in Figure 11, where the speaker produces a semicircle in both speech and gesture.

Figure 11: “und [oben haben die so’n HALBkreis]” (*and on the top they have such a SEMIcircle*), [V20, 6:36].

The kinematic gesture representation of the movement carried out (CARRIER) by the wrist – *move up*, *move left*, *move down*, which are concatenated (“⊕”) by movement steps in a bent (“$\oplus_{\smile}$”, as opposed to rectangular “$\oplus_{\llcorner}$”) way (cf. also Figure 8) – is translated via vectorising function **V** into a vector trajectory (TRAJ(ECTORY)) from the three-dimensional vector space, cf. Figure 7:[9]

(17)
$$\begin{bmatrix} \textit{gesture-vec} & \\ \textsc{traj} & \left[\mathbf{V}\boxed{1}\mathbf{=UP} \oplus_{\smile} \mathbf{-RT} \oplus_{\smile} \mathbf{-UP}\right] \\ \textsc{carrier} & \begin{bmatrix} \textit{gesture} & \\ \textsc{morph} & \left[\textsc{wrist.mov}\ \boxed{1}\textit{mu} \oplus_{\smile} \textit{ml} \oplus_{\smile} \textit{md}\right] \end{bmatrix} \end{bmatrix}$$

The lexical entry for *semicircle* is endowed with a *conceptual vector meaning* attribute CVM. Within CVM it is specified (or underspecified) what kind of vector (VEC) is at stake (axis vector, shape vector, place vector), and how it looks, that

[9]Vectors within gesture space can be conceived of as equivalence classes over concrete movement annotation predicates.

is, which PATH it describes. A semicircle can be defined as an axis vector whose path is a 180° trajectory. Accordingly, 180° is the root of a type hierarchy which hosts all vector sequences within gesture space that describe a half circle. This information is added in terms of a form predicate to the restriction list of *semicircle*, as shown in the speech daughter's (S-DTR) content (CONT) value in (18). Licensed by the speech-gesture integration scheme in (16), the half-circular gesture trajectory from (17) and its affiliate expression *semicircle* can enter into an ensemble construction, as shown in (18):

(18)
$$\begin{bmatrix}
\multicolumn{2}{l}{\textit{sg-ensemble}} \\
\text{PHON} & \left\langle \boxed{1}\begin{bmatrix}\text{SEG} & \textit{semicircle}\end{bmatrix}\right\rangle \\
\text{CAT} & \boxed{2}\begin{bmatrix}\text{HEAD} & \textit{noun}\end{bmatrix} \\
\text{CONT} & \boxed{3}\begin{bmatrix}\text{INDEX} & i \\ \text{RESTR} & \left\langle \begin{bmatrix}\multicolumn{2}{l}{\textit{geom-obj-pred}} \\ \text{RELN} & \textit{semicircle} \\ \text{INST} & i\end{bmatrix}, \boxed{5}\begin{bmatrix}\multicolumn{2}{l}{\textit{form-pred}} \\ \text{RELN} & \textit{round2} \\ \text{ARG} & i \\ \text{CVM} & \begin{bmatrix}\text{VEC} & \textit{axis-path}(i,\boldsymbol{v}) \\ \text{PATH} & \boxed{7}\end{bmatrix}\end{bmatrix} \right\rangle\end{bmatrix} \\
\text{S-DTR} & \begin{bmatrix}\multicolumn{2}{l}{\textit{verbal-sign}} \\ \text{PHON} & \boxed{1}\begin{bmatrix}\text{ACCENT} & \boxed{6}\end{bmatrix} \\ \text{CAT} & \boxed{2} \\ \text{CONT} & \boxed{3}\end{bmatrix} \\
\text{G-DTR} & \begin{bmatrix}\multicolumn{2}{l}{\textit{gesture-vec}} \\ \text{AFF} & \left\langle\begin{bmatrix}\text{PHON} & \begin{bmatrix}\text{ACCENT} & \boxed{6}\textit{marked}\end{bmatrix}\end{bmatrix}\right\rangle \\ \text{TRAJ} & \boxed{7}\boldsymbol{UP} \oplus_{\smile} -\boldsymbol{RT} \oplus_{\smile} -\boldsymbol{UP} \\ \text{CONT} & \begin{bmatrix}\text{MODE} & \textit{exemplification} \\ \text{EX-PRED} & \boxed{5}\end{bmatrix}\end{bmatrix}
\end{bmatrix}$$

By extending lexical entries with frame information from frame semantics (Fillmore 1982), also the exemplification of non-overtly-expressed predicates becomes feasible (Lücking 2013: Sec. 9.2.1); a datum showing this case has already been given with the *spiral staircases* in (5)/Figure 3. A highly improved version of the "vectorisation" of gestures with a translation protocol has been spelled out in Lücking (2016), but within the semantic framework of a *Type Theory with Records* (Cooper 2019; Cooper & Ginzburg 2015; cf. also Lücking, Ginzburg & Cooper 2019, Chapter I of this volume).

The richer formal, functional and representational features of iconic gestures as compared to deictic gestures (cf. Section 3.4) is accounted for in Alahverdzhieva et al. (2017) by assigning a formal predicate to each "phonological" feature of a gesture representation (cf. Section 3.3). These formal gesture predicates are highly underspecified, using *Robust Minimal Recursion Semantics* (RMRS) (Copestake 2007). That is, they can be assigned various predications (which are assumed to be constrained by iconicity with differing arity in the gesture resolution process).

Let us illustrate this by means of Example 1 from Alahverdzhieva et al. (2017), which is due to Loehr (2004) and re-given in (19), adapted to the representational conventions followed in this chapter.

(19) [So he mixes MUD]
The speaker performs a circular movement with the right hand over the upwards, open palm of the left hand

Using a variant of a kinematic representation format for gestures (cf. Section 3.3), the right hand from example 19 is notated as follows (Alahverdzhieva et al. 2017: 440):

(20)
$$\begin{bmatrix} \textit{depict-literal} & \\ \textsc{hand-shape} & \textit{bent} \\ \textsc{palm-orient} & \textit{towards-down} \\ \textsc{finger-orient} & \textit{towards-down} \\ \textsc{hand-location} & \textit{lower-periphery} \\ \textsc{hand-movement} & \textit{circular} \end{bmatrix}$$

Each feature value pair from the gesture's representation in (20) is mapped onto an RMRS-based underspecified representation (Alahverdzhieva et al. 2017: 442):

(21)
$$\begin{array}{l} l_0 : a_0 : [\mathcal{G}](h) \\ l_1 : a_1 : \textit{hand\_shape\_bent}(i_1) \\ l_2 : a_2 : \textit{palm\_orient\_towards\_down}(i_2) \\ l_3 : a_3 : \textit{finger\_orient\_towards\_down}(i_3) \\ l_4 : a_4 : \textit{hand\_location\_lower\_periphery}(i_4) \\ l_5 : a_5 : \textit{hand\_movement\_circular}(i_5) \\ h =_q l_n \textit{ where } 1 \leq n \leq 5 \end{array}$$

Note that all predicates mapped from the gesture in (21) fall within the scope of the scopal operator $[\mathcal{G}]$; this prevents an individual introduced by a depicting gesture from being an antecedent of a pronoun in speech.

Regimented by the *Situated Prosodic Word Constraint* from Figure 10, the underspecified semantic description of the gesture in (21) and its affiliated noun

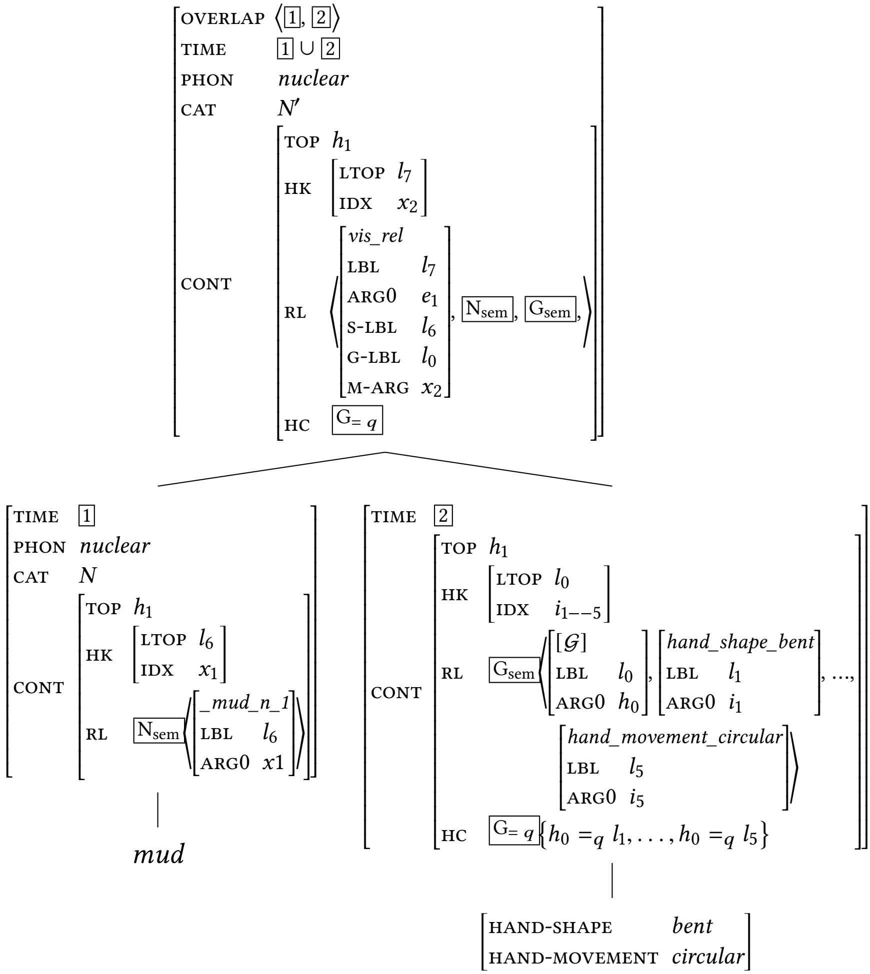


Figure 12: Derivation tree for depicting gesture and its affiliate noun *mud* (Alahverdzhieva et al. 2017: 447)

mud can enter into the multimodal construction given in Figure 12 (where the gesture features are partly omitted for the sake of brevity).

The underspecified RMRS predicates derived from gesture annotations are interpreted according to a type hierarchy rooted in those underspecified RMRS predicates. For example, the circular hand movement of the "mud gesture" can give rise to two slightly different interpretations: on the one hand, the circular hand movement can depict – in the context of the example – that mud is being mixed from an observer viewpoint (McNeill 1992). This reading is achieved by following the left branch of (22), where the gesture contributes a conjunction of predications that express that a substance rotates. When integrated with speech, the substance resolves to the mud and the rotating event to the mixing. On the other hand, the gesture can depict seen from the character viewpoint (McNeill 1992), which corresponds to the predication from the right branch of (22). Here the rotating event is brought about by agent j' which is required to be coreferential with *he*, the subject of the utterance.

(22)

$$\mathit{hand\_movement\_circular}(\mathrm{i})$$

$$\mathit{substance}(x') \wedge \mathit{rotate}(e', x') \qquad \mathrm{rotate}(e', j', x')$$

In addition to addressing (solving) the three basic questions identified in Section 3.1 – roughly, foundation of gesture meaning, regimenting affiliation, and characterisation of semantic integration – another issue has received attention recently, namely the projection behaviour of gestures when interacting with logical operators (Ebert 2014; Schlenker 2018). For instance, the unembedded gesture in (23) triggers the inference that the event being described actually happened in the manner in which it was gesticulated (Schlenker 2018: 303):

(23) John [*slapping gesture*] punished his son.
⇒ John punished his son by slapping him.

That is, (23) more or less corresponds to what semantic speech-gesture integration approaches, as briefly reviewed above, would derive as the content of the multimodal utterance.

Embedding the slapping gesture under the *none*-quantifier triggers, according to Schlenker (2018: 303), the following inference:

(24) None of these 10 guys [*slapping gesture*] punished his son.
⇒ for each of these 10 guys, if he had punished his son, this would have involved some slapping.

The universal inference patterns with presupposition. Unlike presupposition, however, Schlenker (2018: 303) claims that the inference is conditionalized on the at-issue contribution of (24), expressed by the *if*-clause. He then develops a notion of "cosupposition", which rests on an expression's local context that entails the content of its affiliated gesture. So far, there is no connection from such projections to HPSG, however.

Beyond being involved in pragmatic processes like inferring, gestures also take part in "micro-evolutionary" developments. Iconic gestures in particular are involved in a short-term dynamic phenomenon: on repeated co-occurrence, iconic gestures and affiliated speech can fuse into a *multimodal ensemble* (Kendon 2004; Lücking et al. 2008; Mehler & Lücking 2012). The characteristic feature of such an ensemble is that their gestural part, their verbal part, or even both parts can be simplified without changing the meaning of the ensemble. Ensembles, thus, are the result of a process of sign formation as studied, for instance, in experimental semiotics (Galantucci & Garrod 2011). Such grammaticalisation processes eventually might lead to conventional signs. However, most conventional, emblematic everyday gestures seem to be the result of circumventing a taboo: something you should not name is gesticulated (Posner 2002).

3.6 Other gestures

As noted in the taxonomy reviewed in Section 2, there are gestures that, unlike the deictic and iconic ones discussed in the previous sections, do not contribute to propositional content, but serve functions bound up with dialogue management. Such gestures have been called *interactive gestures* (Bavelas et al. 1992). Two examples are given in Figures 13 and 14, which have been discussed by Bavelas et al. (1995).

The "delivery gesture" in Figure 13 is used to underline an argument, or to refer to the fact that the current issue is known to the interlocutors. In the latter function, the gesture is also termed *shared information gesture*.

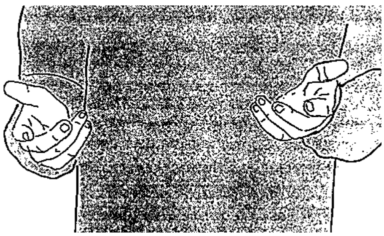


Figure 13: "Here's my point."

The ‘open hand’ pointing gesture in Figure 14 acts as a turn-taking device: it can function as a turn-assigning gesture (underlined by the caption of Figure 14), or, when used to point at the current speaker, it can also indicate that the gesturer wants to take the turn and address the current turn holder.

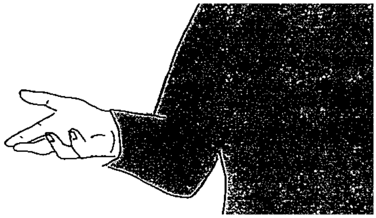


Figure 14: “You go ahead.”

So far there is no account of interactive gestures in HPSG. Given their entrenchment in dialogue processes, their natural home seems to be in a dialogue theory, anyway (see Lücking, Ginzburg & Cooper 2019, Chapter I of this volume). Accordingly, what is presumably the only formal approach to some of these gestures has been spelled out within the dialogical framework PTT in Rieser & Poesio (2009).

4 Gesture and …

Besides being of a genuine linguistic, theoretical interest, gesture studies are a common topic in various areas of investigation, some of which are briefly pointed at below.

4.1 … tools, annotation, corpora

Since gestures are signs in the visual modality, they have to be videotaped. Gesture annotation is carried out on the recorded video films. The main tools that allow for video annotation are, in alphabetical order, Anvil (https://www.anvil-software.org/, Kipp 2014), ELAN (https://tla.mpi.nl/tools/tla-tools/elan/, Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, Sloetjes & Wittenburg 2008) and EXMARaLDA (https://exmaralda.org/, Schmidt 2012).

Annotation should follow an annotation standard which is specified in an annotation scheme. Various annotation schemes for gestures and speech-gesture integration have been proposed, partly differing in annotation foci, including the following ones: annotation schemes that focus on form description and gestures classification in terms of a taxonomy like the one introduced in Section 2 have been developed by R. Breckenridge Church, published in the appendix of McNeill (1992); CoGEST (Gibbon et al. 2003); FORM (Martell et al. 2002) and the SaGA annotation (Lücking et al. 2013). The form of gestures and their timing with speech is the object of the coding scheme of Kipp et al. (2007). An interaction-oriented scheme has been proposed by Allwood et al. (2007), which is formulated on the level of turns and dialogue management. A detailed annotation scheme for the form and function of gestures has been developed in terms of "annotation decision trees" within the NEUROGES system (Lausberg & Sloetjes 2009).

Annotated videos of real life interactions give rise to multimodal corpora. Among those that include data on gestures are the following ones.

The multimodal SmartKom Corpus (Schiel et al. 2003), which grew out of the SmartKom project (Wahlster 2006), comprises recording sessions of various Wizard-of-Oz experiments (that is, human-computer interaction where the human participant is made to believe that the system she or he interacts with is autonomous while in fact it is, at least partly, operated by another human). Recordings are extended basically by a transliteration and labelling of natural speech, labelling of gestures and annotation of user states (in the corpus' first release). The first public release, SKP 1.0, contains 90 recording sessions of 45 users. The multimodal SmartKom corpus as well as further SmartKom resources are hosted at the *Bavarian Archive for Speech Signals* (https://www.bas.uni-muenchen.de/Bas/).

The AMI Meeting Corpus (Carletta et al. 2006) consists of 100 hours of meeting recordings. The meetings were recorded in English but include mostly non-native speakers. The AMI Meeting Corpus provides orthographic transcriptions, but also has a couple of further annotations, including dialogue acts, named entities, head gesture, hand gesture, gaze direction, movement and emotional states.

The SaGA ("Speech and Gesture Alignment") corpus consists of 24 German route direction dialogues obtained after a bus ride through a virtual town (Lücking et al. 2010). Audio and video data from the direction-giver were recorded. The SaGA corpus consists of 280 minutes of video material containing 4,961 iconic/deictic gestures, approximately 1,000 discourse gestures and 39,435 word tokens (Lücking et al. 2013). Gesture annotation has been carried out in great detail, following a kinematic, form-based approach (cf. the above-given remark on annotation schemes). Part of the SaGA corpus is available from the *Bavarian Archive for Speech Signals* (https://www.bas.uni-muenchen.de/Bas).

The DUEL ("Disfluency, exclamations and laughter in dialogue") corpus (Hough et al. 2016) comprises 24 hours of natural, face-to-face dialogue in German, French and Mandarin Chinese. It includes audio, video and body tracking data and is transcribed and annotated for disfluency, laughter and exclamations.

The FIGURE (derived from "Frankfurt Image GestURE") corpus (Lücking et al. 2016) is built on recordings of 50 participants with various mother tongues (though mostly German) spontaneously producing gestures in response to five or six terms from a total of 27 stimulus terms, which have been compiled mainly from image schemata (Lakoff 1987). The gestures have been kinematically annotated by means of a variant of the SaGA annotation scheme. The FIGURE annotation is available from the Text Technology Lab Frankfurt (https://www.texttechnologylab.org/applications/corpora).

4.2 … robots and virtual agents

In the context of Human-Computer Interaction (HCI) or Human-Robot Interaction (HRI), gesture plays an important role (in fact, the formal modelling of deictic and iconic gestures has been initiated in these fields, cf. Section 3.2). One reason for this prominence of gesture in technical areas is that people who interact with a robot evaluate it more positively when the robot displays non-verbal behaviours such as hand and arm gestures along with speech (see e.g. Salem et al. 2012). Within HCI/HRI, two kinds of distinctions have to be made. The first is a distinction between "robot" in the sense of virtual avatars and "robot" in the (probably more common) sense of physical devices (only the latter will be henceforth called a "robot"). The second distinction discerns gesture generation from gesture recognition. Given this simple systematization, altogether four divisions of gesture and virtual avatars/robots arise (references are just exemplary and preferably from earlier HCI/HRI times): (i) gesture generation by robots (e.g. Le et al. 2011); (ii) gesture recognition by robots (e.g. Triesch & von der Malsburg 1998); (iii) gesture generation by virtual avatars (e.g. Cassell et al. 2000); and (iv) gesture recognition in VR/AR (e.g. Weissmann & Salomon 1999). For a more detailed overview see Lücking & Pfeiffer (2012). Enabling humans to act and interact in virtual rooms (e.g. Pfeiffer et al. 2018) can be seen as recent extension of gesture use in HCI/HRI.

In order to plan and design the speech/gesture output of a virtual avatar or a robot, a multimodal representation format is required. To this end, the *Multimodal Utterances Representation Markup Language* for conversational robots (MURML) has been developed (Kranstedt et al. 2002). A similar purpose is served by the *Extensible MultiModal Annotation* (EMMA; Johnston 2009).

4.3 … learning

Following a "gesture as a window to the mind" view, gestures must be a prime object of educational theory and practice, and they are indeed, as demonstrated by research of Cook & Goldin-Meadow (2006) and colleagues. Effectiveness of gestures has been studied in math lessons (Goldin-Meadow et al. 2001), in the acquisition of counting competence (Alibali & DiRusso 1999) and in bilingual education (Breckinridge Church et al. 2004), among other areas. The fairly unanimous result is that gestures can indeed reflect students' conceptualisations and provide insights into cognitive processes involved in learning. Therefore, they can be used as a teaching device as well as an indicator of learning progress and understanding.

4.4 … aphasia

Current models of utterance production are speech-gesture production models, assuming a (more or less) integrated generation of multimodal utterances. Based on such models, one expects an effect on gesture performance when speech production is impaired, as is the case with aphasic speakers. Aphasia is an acquired speech disorder, which can be caused by a stroke, ischaemia, haemorrhage, craniocerebral trauma and further brain-damaging diseases. Different speech-gesture production models make slightly different predictions for speakers suffering from aphasia and can be evaluated accordingly (de Ruiter & de Beer 2013). Indeed, observing the gesture behaviour of aphasic speakers is one aspect of gesture and aphasia (Jakob et al. 2011; Kong et al. 2017; Sekine & Rose 2013). With the exception of the growth point theory, speech-gesture production models are based on Levelt's (1989) model.

The *Growth Point model* (McNeill & Duncan 2000) assumes that the "seed" of an utterance is an inherently multimodal idea unit that comprises imagistic as well as symbolic proto-representations which unfold into gesture and speech respectively in the process of articulation (see also Röpke 2011 on the growth point's entrenchment in contexts and frames).

The *Sketch model* (de Ruiter 2000) reflects explicitly different kinds of gestures (see Section 2). Its name is due to the sketch component, an abstract spatio-temporal representation alongside Levelt's preverbal message. Independently from each other, the sketch is sent to a gesture planner, while the preverbal message is processed by the formulator.

According to the *Lexical Access model* of Krauss et al. (2000), iconic gestures are related to words and are used in order to facilitate speaker-internal word retrieval rather than communicating pictorial information.

The *Interface model* (Kita & Özyürek 2003) assumes that the processes for speech and gesture generation negotiate with each other and therefore can influence each other during the production phase.

Other aspects include the use of gesture in speech therapy. Very much in line with the lexical access model, gestures have been used in order to facilitate word retrieval in what can be called *multimodal therapy* (Rose 2006). Following a different strategy, gestures are also used in order to enhance the communicative range of patients: they learn to employ gestures instead of words in order to communicate at least some of their needs and thoughts more fluently (Cubelli et al. 1991; Caute et al. 2013).

However, just counting on gestures in therapy does not automatically lead to success (Auer & Bauer 2011). The type and severity of aphasia, the individual traits of the aphasic speaker and the kinds of gestures impaired or still at disposal, among other factors, seem to constitute a complex network for which currently no generally applicable clinical pathway can be given.

5 Outlook

What are (still) challenging issues with respect to grammar-gesture integration, in particular from a semantic point of view? Candidates include:

- gestalt phenomena: the trajectories described by a gesture are often incomplete and have to be completed by drawing on gestalt principles or everyday knowledge (Lücking 2016).
- negligible features: not all formal features of a gesture are meaning-carrying features in the context of utterance. For instance, in a dynamic gesture the handshape often (though not always) does not provide any semantic information (cf. also examples (17) and (21)/(22)). How can we distinguish between significant and negligible gesture features?
- "semantic endurance": due to holds, gestures can show their meaning contributions for some period of time and keeps available for semantic attachment. This may call for a more sophisticated algebraic treatment of speech-gesture integration than offered by typed feature structures (Rieser 2015).

Finally, the empirical domain of "gesture" has to be extended to other nonverbal signals, in particular propositional ones such as laughter (Ginzburg et al.

2015), facial expressions or gaze (see Section 1 for a brief list of non-verbal signals), in isolation as well as in mutual combination. Thus, there is still some way to go in order to achieve a fuller understanding of natural language interaction and thereby natural languages.

References

Abner, Natasha, Kensy Cooperrider & Susan Goldin-Meadow. 2015. Gesture for linguists: A handy primer. *Language and Linguistics Compass* 9(11). 437–451. DOI:10.1111/lnc3.12168

Alahverdzhieva, Katya. 2013. *Alignment of speech and co-speech gesture in a constraint-based grammar*. School of Informatics, University of Edinburgh dissertation.

Alahverdzhieva, Katya, Dan Flickinger & Alex Lascarides. 2012. Multimodal grammar implementation. In Eric Fosler-Lussier, Ellen Riloff & Srinivas Bangalore (eds.), *Proceedings of the 2012 conference of the North American Chapter of the Association for Computational Linguistics: human language technologies* (NAACL-HLT 2012), 582–586. Montreal, Canada.

Alahverdzhieva, Katya & Alex Lascarides. 2010. Analysing language and coverbal gesture in constraint-based grammars. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 5–25. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Alahverdzhieva, Katya, Alex Lascarides & Dan Flickinger. 2017. Aligning speech and co-speech gesture in a constraint-based grammar. *Journal of Language Modelling* 5(3). 421–464.

Alibali, Martha W. & Alyssa A. DiRusso. 1999. The function of gesture in learning to count: More than keeping track. *Cognitive Development* 14(1). 37–56. DOI:10.1016/S0885-2014(99)80017-3

Allwood, Jens, Loredana Cerrato, Kristiina Jokinen, Costanza Navarretta & Patrizia Paggio. 2007. The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language Resources and Evaluation* 41. 273–287. DOI:10.1007/s10579-007-9061-5

Argyle, Michael. 1975. *Bodily communication*. New York, NY: Methuen & Co.

Argyle, Michael & Mark Cook. 1976. *Gaze and mutual gaze*. Cambridge, UK: Cambridge University Press.

Auer, Peter & Angelika Bauer. 2011. Multimodality in aphasic conversation: Why gestures sometimes do not help. *Journal of Interactional Research in Communication Disorders* 2(2). 215–243. DOI:10.1558/jircd.v2i2.215

Bangerter, Adrian & Daniel M. Oppenheimer. 2006. Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture* 6(1). 85–102.

Bavelas, Janet B., Nicole Chovil, Linda Coates & Lori Roe. 1995. Gestures specialized for dialogue. *Personality and Social Psychology Bulletin* 21(4). 394–405.

Bavelas, Janet B., Nicole Chovil, Douglas A. Lawrie & Allan Wade. 1992. Interactive gestures. *Discourse Processes* 15(4). 469–489. DOI:10.1080/01638539209544823

Bavelas, Janet B., Jennifer Gerwing, Chantelle Sutton & Danielle Prevost. 2008. Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language* 58(2). 495–520. DOI:10.1016/j.jml.2007.02.004

Bender, Emily M. & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 797–840. Berlin: Language Science Press. DOI:??

Bierman, Arthur K. 1962. That there are no iconic signs. *Philosophy and Phenomenological Research* 23(2). 243–249.

Birdwhistell, Ray L. 1970. *Kinesics and context: Essays on body motion communication*. Philadelphia: University of Pennsylvania Press.

Bolt, Richard A. 1980. ‘‘put-that-there”: voice and gesture at the graphics interface. *SIGGRAPH Comput. Graph.* 14. 262–270. DOI:http://doi.acm.org/10.1145/965105.807503

Breckinridge Church, Ruth, Saba Ayman-Nolley & Shahrzad Mahootian. 2004. The role of gesture in bilingual education: Does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism* 7(4). 303–319. DOI:10.1080/13670050408667815

Bruneau, Thomas J. 1980. Chronemics and the verbal-nonverbal interface. In Mary Ritchie Key (ed.), *The relationship of verbal and nonverbal communication*, 101–118. The Hague, The Netherlands: Mouton.

Bruner, Jerome S. 1998. Routes to reference. *Pragmatics & Cognition* 6(1/2). 209–227. Special Issue: The Concept of Reference in the Cognitive Sciences.

Bühler, Karl. 1934. *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Stuttgart: Gustav Fischer Verlag. Re-edition Stuttgart: UTB, Lucius & Lucius, 1999.

Burks, Arthur W. 1949. Icon, index, and symbol. *Philosophy and Phenomenological Research* 9(4). 673–689.

Butterworth, George & Shoji Itakura. 2000. How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology* 18(1). 25–50.

Carletta, Jean, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma & Pierre Wellner. 2006. The AMI Meeting Corpus: a pre-announcement. In Steve Renals & Samy Bengio (eds.), *Machine learning for multimodal interaction: second international workshop, mlmi 2005, edinburgh, uk, july 11-13, 2005, revised selected papers* (Lecture Notes in Computer Science 3869), 28–39. Berlin & Heidelberg: Springer. DOI:10.1007/11677482_3

Cassell, Justin, Matthew Stone & Hao Yan. 2000. Coordination and context-dependence in the generation of embodied conversation. In *Proceedings of the first international conference on natural language generation*, 171–178.

Caute, Anna, Tim Pring, Naomi Cocks, Madeline Cruice, Wendy Best & Jane Marshall. 2013. Enhancing communication through gesture and naming therapy. *Journal of Speech, Language, and Hearing Research* 56(1). 337–351. DOI:10.1044/1092-4388(2012/11-0232)

Cienki, Alan J. & Cornelia Müller (eds.). 2008. *Metaphor and gesture* (Gesture studies 3). Amsterdam: John Benjamins.

Clark, Herbert H. 1996. *Using language.* Cambridge: Cambridge University Press.

Cohen, Philip R., Michael Johnston, David McGee, Sharon Oviatt, Jay Pittman, Ira Smith, Liang Chen & Josh Clow. 1997. QuickSet: Multimodal interaction for distributed applications. In *Proceedings of the fifth ACM international conference on multimedia* (MULTIMEDIA '97), 31–40. Seattle, Washington, USA. DOI:10.1145/266180.266328

Cook, Susan Wagner & Susan Goldin-Meadow. 2006. The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development* 7(2). 211–232. DOI:10.1207/s15327647jcd0702_4

Cooper, Robin. 2019. *From perception to linguistic communication: Using a theory of types with records (TTR) to model linguistic action and content.* https://github.com/robincooper/ttl. MS Gothenburg University.

Cooper, Robin & Jonathan Ginzburg. 2015. Type theory with records for natural language semantics. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 2nd edn., chap. 12, 375–407. Oxford, UK: Wiley-Blackwell.

Cooperrider, Kensy. 2017. Foreground gesture, background gesture. *Gesture* 16(2). 176–202. DOI:10.1075/gest.16.2.02coo

Cooperrider, Kensy & Rafael Núñez. 2012. Nose-pointing: Notes on a facial gesture of Papua New Guinea. *Gesture* 12(2). 103–129. DOI:doi:10.1075/gest.12.2.01coo

Copestake, Ann. 2007. Semantic composition with (Robust) Minimal Recursion Semantics. In *Proceedings of the workshop on deep linguistic processing* (DeepLP'07), 73–80. Prague, Czech Republic.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(4). 281–332.

Cubelli, Roberto, Piera Trentini & Carmelo G. Montagna. 1991. Re-education of gestural communication in a case of chronic global aphasia and limb apraxia. *Cognitive Neuropsychology* 8(5). 369–380. DOI:10.1080/02643299108253378

Domaneschi, Filippo, Marcello Passarelli & Carlo Chiorri. 2017. Facial expressions and speech acts: Experimental evidences on the role of the upper face as an illocutionary force indicating device in language comprehension. *Cognitive Processing* 18(3). 285–306. DOI:10.1007/s10339-017-0809-6

Ebert, Cornelia. 2014. *The non-at-issue contributions of gestures.* Workshop on Demonstration and Demonstratives, April 11–12 2014, Stuttgart.

Ebert, Cornelia, Stefan Evert & Katharina Wilmes. 2011. Focus marking via gestures. In Ingo Reich, Eva Horch & Dennis Pauly (eds.), *Proceedings of sinn & bedeutung 15*, 193–208. Saarbrücken, Germany: Saarland University Press.

Eco, Umberto. 1976. *A theory of semiotics.* Bloomington: Indiana University Press.

Ekman, Paul & Wallace V. Friesen. 1969. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica* 1(1). 49–98.

Ekman, Paul & Wallace V. Friesen. 1978. *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Enfield, Nick J. 2001. Lip-pointing: A discussion of form and function with reference to data from Laos. *Gesture* 1(2). 185–212.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Elisabet Engdahl & Enric Vallduví (eds.), *Edinburgh working papers in cognitive science* (Studies in HPSG 12), 1–31. Edinburgh: University of Edinburgh.

Fillmore, Charles. 1982. Frame Semantics. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 111–137. Seoul: Hanshin Publishing Co.

Fricke, Ellen. 2007. *Origo, Geste und Raum* (Linguistik – Impulse & Tendenzen 24). Berlin: De Gruyter.

Fricke, Ellen. 2012. *Grammatik multimodal. Wie Wörter und Gesten zusammenwirken* (Linguistik – Impulse und Tendenzen 40). Berlin: De Gruyter.

Galantucci, Bruno & Simon Garrod. 2011. Experimental semiotics: A review. *Frontiers in Human Neuroscience* 5(11). DOI:10.3389/fnhum.2011.00011

Gerwing, Jennifer & Janet B. Bavelas. 2004. Linguistic influences on gesture's form. *Gesture* 4(2). 157–195.

Gibbon, Dafydd, Ulrike Gut, Benjamin Hell, Karin Looks, Alexandra Thies & Thorsten Trippel. 2003. A computational model of arm gestures in conversation. In ISCA Archive (ed.), *Proceedings of the 8th European conference on speech communication and technology* (EUROSPEECH 2003), 813–816. Geneva, Switzerland. http://www.isca-speech.org/archive/eurospeech_2003, accessed 2018-9-10.

Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation.* Oxford, UK: Oxford University Press.

Ginzburg, Jonathan, Ellen Breitholtz, Robin Cooper, Julian Hough & Ye Tian. 2015. Understanding laughter. In Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 137–146. Amsterdam, Netherlands. http://semanticsarchive.net/Archive/mVkOTk2N/AC2015-proceedings.pdf, accessed 2018-9-10.

Giorgolo, Gianluca. 2010. A formal semantics for iconic spatial gestures. In Maria Aloni, Harald Bastiaanse, Tikitu de Jager & Katrin Schulz (eds.), *Logic, language and meaning* (Lecture Notes in Computer Science 6042), 305–314. Berlin: Springer.

Giorgolo, Gianluca & Ash Asudeh. 2011. Multimodal communication in LFG: gestures and the correspondence architecture. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2011 conference*, 257–277. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/16/, accessed 2018-9-30.

Goldin-Meadow, Susan, Howard Nusbaum, Spencer D. Kelly & Susan Wagner. 2001. Explaining math: Gesturing lightens the load. *Psychological Science* 12(6). 516–522.

Goodman, Nelson. 1976. *Languages of art. An approach to a theory of symbols.* 2nd edn. Idianapolis: Hackett Publishing Company, Inc.

Goodwin, Charles. 2003. Pointing as situated practice. In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 2, 217–241. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Hadjikhani, Nouchine, Rick Hoge, Josh Snyder & Beatrice de Gelder. 2008. Pointing with the eyes: The role of gaze in communicating danger. *Brain and Cognition* 68(1). 1–8. DOI:http://dx.doi.org/10.1016/j.bandc.2008.01.008

Hahn, Florian & Hannes Rieser. 2010. Explaining speech gesture alignment in MM dialogue using gesture typology. In Paweł Łupowski & Matthew Purver (eds.), *Proceedings of the 14th workshop on the semantics and pragmatics of dialogue: Aspects of semantics and pragmatics of dialogue* (SemDial 2010), 99–111. Poznan, Poland.

Hall, Edward T. 1968. Proxemics. *Current Anthropology* 9(2-3). 83–108.

Hough, Julian, Ye Tian, Laura de Ruiter, Simon Betz, Spyros Kousidis, David Schlangen & Jonathan Ginzburg. 2016. DUEL: a multi-lingual multimodal dialogue corpus for disfluency, exclamations and laughter. In *Proceedings of the 10th international conference on language resources and evaluation* (LREC 2016), 1784–1788. Portorož, Slovenia.

Jakob, Hanna, Daniela Bartmann, Georg Goldenberg, Wolfram Ziegler & Katharina Hogrefe. 2011. Zusammenhang von Spontansprachproduktion und Gesten bei Patienten mit Aphasie. *Aphasie und verwandte Gebiete* 30(3). 20–38.

Jannedy, Stefanie & Norma Mendoza-Denton. 2005. Structuring information through gesture and intonation. *Interdisciplinary Studies on Information Structure (ISIS)* (3). 199–244. http://opus.kobv.de/ubp/volltexte/2006/877/, accessed 2018-9-30. URN: urn:nbn:de:kobv:517-opus-8774.

Johansson, Gunnar. 1973. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics* 14(2). 201–211. DOI:10.3758/BF03212378

Johnston, Michael. 1998. Unification-based multimodal parsing. In *Proceedings of the 36th annual meeting of the Association for Computational Linguistics and 17th international conference on computational linguistics – Volume 1* (ACL '98), 624–630. Montreal, Quebec, Canada: Association for Computational Linguistics. DOI:10.3115/980845.980949

Johnston, Michael. 2009. Building multimodal applications with EMMA. In *Proceedings of the 2009 international conference on multimodal interfaces* (ICMI-MLMI '09), 47–54. Cambridge, Massachusetts, USA. DOI:10.1145/1647314.1647325

Johnston, Michael, Philip R. Cohen, David McGee, Sharon L. Oviatt, James A. Pittman & Ira Smith. 1997. Unification-based multimodal integration. In *Proceedings of the 35th annual meeting of the Association for Computational Linguistics and eighth conference of the european chapter of the association for computational linguistics*, 281–288. Madrid, Spain: Association for Computational Linguistics. DOI:10.3115/976909.979653

Kaplan, David. 1989. Demonstratives. In Joseph Almog, John Perry & Howard Wettstein (eds.). In collab. with Ingrid Deiwiks & Edward N. Zalta, *Themes from Kaplan*, 481–563. New York & Oxford: Oxford University Press.

Kauffman, Lynn E. 1971. Tacesics, the study of touch: A model for proxemic analysis. *Semiotica* 4(2). 149–161. DOI:10.1515/semi.1971.4.2.149

Kelly, Spencer D., Corinne Kravitz & Michael Hopkins. 2004. Neural correlates of bimodal speech and gesture comprehension. *Brain and Language* 89(1). 253–260. DOI:10.1016/S0093-934X(03)00335-3

Kendon, Adam. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica* 26(1). 22–63. DOI:10.1016/0001-6918(67)90005-4

Kendon, Adam. 1972. Some relationships between body motion and speech. An analysis of an example. In Aron Wolfe Siegman & Benjamin Pope (eds.), *Studies in dyadic communication*, chap. 9, 177–210. Elmsford, NY: Pergamon Press.

Kendon, Adam. 1980. Gesticulation and speech: Two aspects of the process of utterance. In Mary Ritchie Key (ed.), *The relationship of verbal and nonverbal communication* (Contributions to the Sociology of Language 25), 207–227. The Hague: Mouton.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge, MA: Cambridge University Press.

Kendon, Adam & Laura Versante. 2003. Pointing by hand in "neapolitan". In Sotaro Kita (ed.), *Pointing: where language, culture, and cognition meet*, chap. 6, 109–137. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Kipp, Michael. 2014. ANVIL: A universal video research tool. In Jacques Durand, Ulrike Gut & Gjert Kristofferson (eds.), *Handbook of corpus phonology*, chap. 21, 420–436. Oxford, UK: Oxford University Press.

Kipp, Michael, Michael Neff & Irene Albrecht. 2007. An annotation scheme for conversational gestures: how to economically capture timing and form. *Journal on Language Resources and Evaluation - Special Issue on Multimodal Corpora* 41(3-4). 325–339.

Kita, Sotaro & Aslı Özyürek. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language* 48(1). 16–32. DOI:10.1016/S0749-596X(02)00505-3

Kita, Sotaro, Ingeborg van Gijn & Harry van der Hulst. 1998. Movement phases in signs and co-speech gestures, and their transcription by human coders. In Ipke Wachsmuth & Martin Fröhlich (eds.), *Gesture and sign language in human-computer interaction*, 23–35. Berlin & Heidelberg: Springer.

Klein, Ewan. 2000. Prosodic constituency in HPSG. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism), chap. 10, 169–200. Stanford, CA: CSLI Publications.

Klein, Wolfgang. 1978. Wo ist hier? Präliminarien zu einer Untersuchung der lokalen Deixis. *Linguistische Berichte* 58. 18–40.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Kolarova, Zornitza. 2011. *Lexikon der bulgarischen Alltagsgesten.* Technische Universität Berlin dissertation.

Kong, Anthony Pak-Hin, Sam-Po Law & Gigi Wan-Chi Chak. 2017. A comparison of coverbal gesture use in oral discourse among speakers with fluent and nonfluent aphasia. *Journal of Speech, Language, and Hearing Research* 60(7). 2031–2046. DOI:10.1044/2017_JSLHR-L-16-0093

Kopp, Stefan, Paul Tepper & Justine Cassell. 2004. Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proceedings of the 6th international conference on multimodal interfaces* (ICMI '04), 97–104. State College, PA, USA: ACM. DOI:10.1145/1027933.1027952

Kranstedt, Alfred, Stefan Kopp & Ipke Wachsmuth. 2002. MURML: a multimodal utterance representation markup language for conversational agents. In *Proceedings of the aamas02 workshop on Embodied Conversational Agents – let's specify and evaluate them.* Bologna, Italy.

Kranstedt, Alfred, Andy Lücking, Thies Pfeiffer, Hannes Rieser & Ipke Wachsmuth. 2006a. Deictic object reference in task-oriented dialogue. In Gert Rickheit & Ipke Wachsmuth (eds.), *Situated communication*, 155–207. Berlin: Mouton de Gruyter.

Kranstedt, Alfred, Andy Lücking, Thies Pfeiffer, Hannes Rieser & Ipke Wachsmuth. 2006b. Deixis: How to determine demonstrated objects using a pointing cone. In Sylvie Gibet, Nicolas Courty & Jean-Francois Kamp (eds.), *Gesture in human-computer interaction and simulation: 6th International Gesture Workshop, gw 2005, Berder Island, France, May 18-20, 2005, revised selected papers* (Lecture Notes in Computer Science 3881), 300–311. Berlin: Springer.

Krauss, Robert M., Yihsiu Chen & Rebecca F. Gottesmann. 2000. Lexical gestures and lexical access: A process model. In David McNeill (ed.), *Language and gesture*, chap. 13, 261–283. Cambridge, UK: Cambridge University Press.

Kühnlein, Peter, Manja Nimke & Jens Stegmann. 2002. *Towards an HPSG-based formalism for the integration of speech and co-verbal pointing.* Talk presented at The first congress of the ISGS, *Gesture – The Living Medium*, University of Texas at Austin.

Kupffer, Manfred. 2014. Does context change? In Daniel Gutzmann, C'ecile Meier & Jan Köpping (eds.), *Approaches to meaning. Composition, value, and interpretation* (Current Research in the Semantics/Pragmatics Interface 1), 25–44. Leiden, NL: Brill.

Kuthy, Kordula De. 2019. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 709–743. Berlin: Language Science Press. DOI:??

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.

Lascarides, Alex & Matthew Stone. 2009. A formal semantic analysis of gesture. *Journal of Semantics* 26(4). 393–449. DOI:10.1093/jos/ffp004

Lausberg, Hedda & Han Sloetjes. 2009. Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods* 41(3). 841–849.

Le, Quoc Anh, Souheil Hanoune & Catherine Pelachaud. 2011. Design and implementation of an expressive gesture model for a humanoid robot. In *Proceedings of the 11th international conference on humanoid robots* (IEEE-RAS 2011), 134–140. DOI:10.1109/Humanoids.2011.6100857

Levelt, Willem J. M. 1989. *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.

Levinson, Stephen C. 2008. Deixis. In *The handbook of pragmatics*, chap. 5, 97–121. Blackwell.

Lewis, David. 1970. General semantics. *Synthese*. Semantics of Natural Language II 22(1/2). 18–67.

Loehr, Daniel. 2004. *Gesture and intonation*. Washington, D.C.: Georgetown University dissertation.

Loehr, Daniel. 2007. Apects of rhythm in gesture in speech. *Gesture* 7(2). 179–214.

Lücking, Andy. 2013. *Ikonische Gesten. Grundzüge einer linguistischen Theorie*. Berlin: De Gruyter. Zugl. Diss. Univ. Bielefeld (2011).

Lücking, Andy. 2016. Modeling co-verbal gesture perception in type theory with records. In Maria Ganzha, Leszek Maciaszek & Marcin Paprzycki (eds.), *Proceedings of the 2016 federated conference on computer science and information systems* (Annals of Computer Science and Information Systems 8), 383–392. IEEE. DOI:10.15439/2016F83

Lücking, Andy. 2018. Witness-loaded and witness-free demonstratives. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical* *demonstratives. Syntax, semantics and pragmatics* (Linguistische Arbeiten 568). Berlin: De Gruyter.

Lücking, Andy, Kirsten Bergman, Florian Hahn, Stefan Kopp & Hannes Rieser. 2013. Data-based analysis of speech and gesture: the Bielefeld Speech and Gesture Alignment Corpus (SaGA) and its applications. *Journal on Multimodal User Interfaces* 7(1-2). 5–18. DOI:10.1007/s12193-012-0106-8

Lücking, Andy, Kirsten Bergmann, Florian Hahn, Stefan Kopp & Hannes Rieser. 2010. The Bielefeld speech and gesture alignment corpus (SaGA). In *Multimodal corpora: Advances in capturing, coding and analyzing multimodality* (LREC 2010), 92–98. Malta. DOI:10.13140/2.1.4216.1922

Lücking, Andy, Jonathan Ginzburg & Robin Cooper. 2019. Grammar in dialogue. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 841–884. Berlin: Language Science Press. DOI:??

Lücking, Andy & Alexander Mehler. 2012. What's the scope of the Naming Game? Constraints on semantic categorization. In Thomas C. Scott-Phillips, Mónica Tamariz, Erica A. Cartmill & James R. Hurford (eds.), *Proceedings of the 9th international conference on the evolution of language* (Evolang IX), 196–203. Kyoto, Japan.

Lücking, Andy, Alexander Mehler & Peter Menke. 2008. Taking fingerprints of speech-and-gesture ensembles: Approaching empirical evidence of intrapersonal alignmnent in multimodal communication. In *Proceedings of the 12th workshop on the semantics and pragmatics of dialogue* (LonDial'08), 157–164. King's College London.

Lücking, Andy, Alexander Mehler, Désirée Walther, Marcel Mauri & Dennis Kurfürst. 2016. Finding recurrent features of image schema gestures: the FIGURE corpus. In *Proceedings of the 10th international conference on language resources and evaluation* (LREC 2016), 1426–1431. Portorož (Slovenia).

Lücking, Andy & Thies Pfeiffer. 2012. Framing multimodal technical communication. In Alexander Mehler & Laurent Romary (eds.). In collab. with Dafydd Gibbon, *Handbook of technical communication* (Handbooks of Applied Linguistics 8), chap. 18, 591–644. Berlin: De Gruyter Mouton.

Lücking, Andy, Thies Pfeiffer & Hannes Rieser. 2015. Pointing and reference reconsidered. *Journal of Pragmatics* 77. 56–79. DOI:10.1016/j.pragma.2014.12.013

Lücking, Andy, Hannes Rieser & Jens Stegmann. 2004. Statistical support for the study of structures in multi-modal dialogue: Inter-rater agreement and synchronization. In Jonathan Ginzburg & Enric Vallduví (eds.), *Proceedings of* *the eighth workshop on the semantics and pragmatics of dialogue* (Catalog '04), 56–63. Barcelona.

Martell, Craig, Chris Osborn, Jesse Friedman & Paul Howard. 2002. FORM: a kinematic annotation scheme and tool for gesture annotation. In *Proceedings of multimodal resources and multimodal systems evaluation*, 15–22. Las Palmas, Spain.

Masataka, Nobuo. 2003. From index-finger extension to index-finger pointing: Ontogenesis of pointing in preverbal infants. In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 4, 69–84. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Matthews, Danielle, Tanya Behne, Elena Lieven & Michael Tomasello. 2012. Origins of the human pointing gesture: A training study. *Developmental Science* 15(6). 817–829. DOI:10.1111/j.1467-7687.2012.01181.x

McClave, Evelyn. 1994. Gestural beats: The rhythm hypothesis. *Journal of Psycholinguistic Research* 23(1). 45–66.

McGinn, Colin. 1981. The mechanism of reference. *Synthese* 49(2). 157–186.

McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review* 92(3). 350–371.

McNeill, David. 1992. *Hand and mind – What gestures reveal about thought*. Chicago: Chicago University Press.

McNeill, David. 2005. *Gesture and thought*. Chicago: University of Chicago Press.

McNeill, David & Susan D. Duncan. 2000. Growth points in thinking-for-speaking. In David McNeill (ed.), *Language and gesture*, chap. 7, 141–161. Cambridge, MA: Cambridge University Press.

Mehler, Alexander & Andy Lücking. 2012. Pathways of alignment between gesture and speech: Assessing information transmission in multimodal ensembles. In Gianluca Giorgolo & Katya Alahverdzhieva (eds.), *Proceedings of the international workshop on formal and computational approaches to multimodal communication under the auspices of ESSLLI 2012, Opole, Poland, 6–10 August*.

Müller, Cornelia. 1998. *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich* (Körper – Kultur – Kommunikation 1). Berlin: Berlin Verlag. Zugl. Diss. Frei Universität Berlin (1996).

Müller, Cornelia, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Sedinha Tessendorf (eds.). 2013–2014. *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction*. 2 vols. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 38/1 & 2). Berlin: De Gruyter Mouton.

Nobe, Shuichi. 2000. Where do *most* spontaneous representational gestures actually occur with respect to speech? In David McNeill (ed.), *Language and gesture*, 186–198. Cambridge, MA: Cambridge University Press.

Nöth, Winfried. 1990. *Handbook of semiotics*. Bloomington, Indianapolis: Indiana University Press.

Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16(1). 1–43. DOI:10.1007/BF00984721

Özyürek, Aslı. 2012. Gesture. In Roland Pfau, Markus Steinbach & Bencie Woll (eds.), *Sign language: An international handbook*, vol. 37 (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK)), chap. 27, 626–646. Berlin: De Gruyter Mouton.

Peacocke, Christopher. 1987. Depiction. *The Philosophical Review* 96(3). 383–410.

Pfeiffer, Thies, Carolin Hainke, Leonard Meyer, Maik Fruhner & Moritz Niebling. 2018. Proceedings der pre-conference-workshops der 16. e-learning fachtagung informatik co-located with 16th e-learning conference of the german computer society. In Daniel Schiffner (ed.) (DeLFI 2018 2250). Frankfurt, Germany: CEUR Workshop Proceedings.

Posner, Roland. 2002. Alltagsgesten als Ergebnis von Ritualisierung. In Jan C. Joerden (ed.), *Ritualisierte Tabuverletzung, Lachkultur und das Karnevaleske*, vol. 6 (Studien zur Ethik in Ostmitteleuropa), 395–421. Frankfurt am Main: Peter Lang.

Posner, Roland, Klaus Robering & Thomas A. Sebeok (eds.). 1997–2004. *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur / Semiotics. A Handbook on the Sign-Theoretic Foundations of Nature and Culture*. 4 vols. (Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 13/1–4). Berlin: De Gruyter.

Poyatos, Fernando. 1975. Cross-cultural study of paralinguistic "alternants" in face-to-face interaction. In Adam Kendon, Harris M. Richards & Mary Ritchie Key (eds.), *Organization of behavior in face-to-face interaction*, 285–314. The Hague: Mouton.

Quine, Willard Van Orman. 1950. Identity, ostension, and hypostasis. *The Journal of Philosophy* 47(22). 621–633.

Rieser, Hannes. 2004. Pointing in dialogue. In *Proceedings of the eighth workshop on the semantics and pragmatics of dialogue* (Catalog '04), 93–100. Barcelona.

Rieser, Hannes. 2008. Aligned iconic gesture in different strata of mm route-description. In *LonDial 2008: The 12th workshop on the semantics and pragmatics of dialogue* (SemDial), 167–174. King's College London.

Rieser, Hannes. 2010. On factoring out a gesture typology from the Bielefeld Speech-Gesture-Alignment Corpus. In Stefan Kopp & Ipke Wachsmuth (eds.), *Proceedings of GW 2009: Gesture in embodied communication and human-computer interaction*, 47–60. Berlin & Heidelberg: Springer.

Rieser, Hannes. 2011. How to disagree on a church window's shape using gesture. In Klaus Hölker & Carla Marello (eds.), *Dimensionen der Analyse von Texten und Diskursen. Festschrift für János Sándor Petöfi*, 231–246. Berlin: LIT Verlag.

Rieser, Hannes. 2015. When hands talk to mouth. Gesture and speech as autonomous communicating processes. In Christine Howes & Staffan Larsson (eds.), *Proceedings of the 19th workshop on the semantics and pragmatics of dialogue* (SemDial 2015: goDIAL), 122–130. Gothenburg, Sweden.

Rieser, Hannes & Massimo Poesio. 2009. Interactive gestures in dialogue: A PTT model. In *Proceedings of the 10th annual meeting of the special interest group in discourse and dialogue* (SIGDIAL 2009), 87–96. Queen Mary University of London.

Röpke, Insa. 2011. Watching the growth point grow. In *Proceedings of the second conference on gesture and speech in interaction* (GESPIN 2011). http://coral2.spectrum.uni-bielefeld.de/gespin2011/final/Roepke.pdf.

Rose, Miranda L. 2006. The utility of arm and hand gestures in the treatment of aphasia. *International Journal of Speech-Language Pathology* 8(2). 92–109. DOI:10.1080/14417040600657948

de Ruiter, Jan P. & Carola de Beer. 2013. A critical evaluation of models of gesture and speech production for understanding gesture in aphasia. *Aphasiology* 27(9). 1015–1030. DOI:10.1080/02687038.2013.797067

de Ruiter, Jan Peter. 2000. The production of gesture and speech. In David McNeill (ed.), *Language and gesture*, chap. 14, 284–311. Cambridge, UK: Cambridge University Press.

de Ruiter, Jan Peter. 2004. On the primacy of language in multimodal communication. In Jean-Claude Martin, Elisabeth Den, Peter Kühnlein, Lou Boves, Patrizia Paggio & Roberta Catizone (eds.), *Proceedings of the fourth international conference on language resources and evaluation and the workshop on multimodal corpora: Models of human behaviour for the specification and evaluation of multimodal input and output interfaces* (LREC 2004), 38–41. Lisbon, Portugal.

de Ruiter, Jan Peter. 2007. Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture* 7(1). 21–38.

Salem, Maha, Stefan Kopp, Ipke Wachsmuth, Katharina Rohlfing & Frank Joublin. 2012. Generation and evaluation of communicative robot gesture. *International Journal of Social Robotics* 4(2). 201–217. DOI:10.1007/s12369-011-0124-9

Schiel, Florian, Silke Steininger & Uli Türk. 2003. *The SmartKom multimodal corpus at BAS*. Tech. rep. 34. München: Ludwig-Maximilians-Universität München.

Schlenker, Philippe. 2018. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. DOI:10.1007/s10988-017-9225-8

Schmidt, Thomas. 2012. EXMARaLDA and the FOLK tools. In *Proceedings of lrec*. ELRA. http://www.lrec-conf.org/proceedings/lrec2012/pdf/529_Paper.pdf.

Sekine, Kazuki & Miranda L. Rose. 2013. The relationship of aphasia type and gesture production in people with aphasia. *American Journal of Speech-Language Pathology* 22(4). 662–672. DOI:10.1044/1058-0360(2013/12-0030)

Slama-Cazacu, Tatiana. 1976. Nonverbal components in message sequence: "Mixed Syntax". In William C. McCormick & Stephan A. Wurm (eds.), *Language and man. Anthropological issues* (World Anthropology), 217–227. The Hague & Paris: Mouton.

Sloetjes, Han & Peter Wittenburg. 2008. Annotation by category – ELAN and ISO DCR. In *Proceedings of the 6th international conference on language resources and evaluation* (LREC 2008).

van der Sluis, Ielka & Emiel Krahmer. 2007. Generating multimodal references. *Discourse Processes* 44(3). 145–174. Special Issue: Dialogue Modelling: Computational and Empirical Approaches. DOI:10.1080/01638530701600755

Sonesson, Göran. 1998. That there are many kinds of iconic signs. *Visio* 1(1). 33–54.

Sowa, Timo. 2006. *Understanding coverbal iconic gestures in shape descriptions*. Berlin: Akademische Verlagsgesellschaft. Zugl. Diss. Univ. Bielefeld.

Steels, Luc. 1995. A self-organizing spatial vocabulary. *Artificial Life* 2(3). 319–332.

Steinbach, Markus & Anke Holler. 2019. Sign languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 885–886. Berlin: Language Science Press. DOI:??

Streeck, Jürgen. 2008. Depicting by gesture. *Gesture* 8(3). 285–301. DOI:10.1075/gest.8.3.02str

Triesch, Jochen & Christoph von der Malsburg. 1998. Robotic gesture recognition. In Ipke Wachsmuth & Martin Fröhlich (eds.), *Gesture and sign language in human-computer interaction*, 233–244. Berlin & Heidelberg: Springer.

Tuite, Kevin. 1993. The production of gesture. *Semiotica* 93(1/2). 83–105.

Wagner, Petra, Zofia Malisz & Stefan Kopp. 2014. Gesture and speech in interaction: An overview. *Speech Communication* 57. 209–232. DOI:10.1016/j.specom.2013.09.008

Wahlster, Wolfgang (ed.). 2006. *SmartKom: foundations of multimodal dialogue systems*. Berlin & Heidelberg: Springer.

Weissmann, John & Ralf Salomon. 1999. Gesture recognition for virtual reality applications using data gloves and neural networks. In *Proceedings of the international joint conference on neural networks* (IJCNN'99. 3), 2043–2046. DOI:10.1109/IJCNN.1999.832699

Wilkins, David. 2003. Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, chap. 8, 171–215. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Wiltshire, Anne. 2007. Synchrony as the underlying structure of gesture: The relationship between speech sound and body movement at the micro level. In Robyn Loughnane, Cara Penry Williams & Jana Verhoeven (eds.), *In between wor(l)ds: Transformation and translation*, vol. 6 (Postgraduate Research Papers on Language and Literature), 235–253. Melbourne: School of Languages & Linguistics, The University of Melbourne.

Wu, Ying Choon & Seanna Coulson. 2005. Meaningful gestures: electrophysiological indices of iconic gesture comprehension. *Psychophysiology* 42(6). 654–667.

Zwarts, Joost. 2003. Vectors across spatial domains: from place to size, orientation, shape, and parts. In *Representing direction in language and space* (Explorations in Language and Space 1), chap. 3, 39–68. Oxford, NY: Oxford University Press.

Zwarts, Joost & Yoad Winter. 2000. Vector space semantics: a model-theoretic analysis of locative prepositions. *Journal of Logic, Language, and Information* 9(2). 169–211.

Żywiczyński, Przemysław, Sławomir Wacewicz & Sylwester Orzechowski. 2017. Adaptors and the turn-taking mechanism. The distribution of adaptors relative to turn borders in dyadic conversation. *Interaction Studies* 18(2). 276–298. DOI:10.1075/is.18.2.07zyw

Part V

The broader picture

Chapter L

HPSG and Minimalism

Robert D. Borsley

University of Essex and Bangor University

Stefan Müller

Humboldt-Universität zu Berlin

This chapter compares work done in Head-Driven Phrase Structure Grammar with work done under the heading *Minimalist Program*. We discuss differences in the respective approaches and the outlook of theories. We have a look at the procedural/constraint-based views on grammar and discuss the differences in complexity of the structures that are assumed. We also address psycholinguistic issues like processing and language acquisition.

1 Introduction

The Minimalist framework, which was first outlined by Chomsky in the early 1990s (Chomsky 1993; 1995b), still seems to be the dominant approach to syntax. It is important, therefore, to consider how HPSG compares with this framework. In a sense both frameworks are descendants of the transformation-generative approach to syntax, which Chomsky introduced in the 1950s. HPSG is a result of the questioning of transformational analyses that emerged in the late 1970s. This led to Lexical Functional Grammar (Bresnan & Kaplan 1982) and Generalized Phrase Structure Grammar (Gazdar et al. 1985), and then in the mid 1980s to HPSG (Pollard & Sag 1987, see Flickinger, Pollard & Wasow (2019), Chapter 2 of this volume for more on the origins of HPSG). Minimalism in contrast remains committed to transformational, i.e. movement, analyses. It is simpler in some respects than the earlier Government and Binding framework (Chomsky 1981), but as we will see below, it involves a variety of complexities.

Robert D. Borsley & Stefan Müller. 2019. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 937–1012. Berlin: Language Science Press. DOI:??

The relation between the two frameworks is clouded by the discourse that surrounds Minimalism. At one time "virtual conceptual necessity" was said to be its guiding principle. A little later, it was said to be concerned with the "perfection of language", with "how closely human language approaches an optimal solution to design conditions that the system must meet to be usable at all" (Chomsky 2002: 58). Much of this discourse seems designed to suggest that Minimalism is quite different from other approaches and should not be assessed in the same way. In the words of Postal (2003: 19), it looks like "an attempt to provide certain views with a sort of privileged status, with the goal of placing them at least rhetorically beyond the demands of serious argument or evidence". However, the two frameworks have enough in common to allow meaningful comparisons.

Both frameworks seek to provide an account of what is and is not possible both in specific languages and in language in general. Moreover, both are concerned not just with local relations such as that between a head and its complement or complements but also with non-local relations such as those in the following:

(1) a. The student knows the answer.
    b. It seems to be raining.
    c. Which student do you think knows the answer?

In (1a), *the student* is subject of *knows* and is responsible for the fact that *knows* is a third person singular form, but *the student* and *knows* are not sisters if *knows* and *the answer* form a VP. In (1b) the subject is *it* because the complement of *be* is *raining* and *raining* requires an expletive subject, but *it* and *raining* are obviously not sisters. Finally, in (1c), *which student* is understood as the subject of *knows* and is responsible for the fact that it is third person singular, but again the two elements are structurally quite far apart. Both frameworks provide analyses for these and other central syntactic phenomena, and it is quite reasonable to compare them and ask which is the more satisfactory.[1]

Although HPSG and Minimalism have enough in common to permit comparisons, there are obviously many differences. Some are more important than others, and some relate to the basic approach and outlook, while others concern the nature of grammatical systems and syntactic structures. In this chapter we will explore the full range of differences.

The chapter is organized as follows: in Section 2, we look at differences of approach between the two frameworks. Then in Section 3, we consider the quite different views of grammar that the two frameworks espouse, and in Section 4,

[1] As noted below, comparison is complicated somewhat by the fact that Minimalists typically provide only sketches of analyses in which various details are left quite vague.

we look at the very different syntactic structures which result. Finally, in Section 5, we consider how the two frameworks relate to psycholinguistic issues, especially processing and language acquisition.

2 Differences of approach and outlook

This section deals with some higher level differences between the two frameworks. We start with sociological differences in Section 2.1. We go on with the degree of formalization and the range of data that is covered (Section 2.2). Section 2.3 discusses the quality of empirical work. Finally, Section 2.4 deals with arguments for invisible entities and innate knowledge.

2.1 Sociological differences

A difference between Mainstream Generative Grammar[2] and HPSG that should not be underestimated is the fact that MGG has one very important key figure, the by far most-cited linguist (and scientist in general) Noam Chomsky. HPSG is different. This is how Carl Pollard put it:

> There is no Chomsky-like figure who is always assumed to be basically on the right track no matter what s/he proposes. HPSG research is normal science: the testing of hypotheses that appear plausible given accepted assumptions. The goal is not to fill in the details of a vague theory which is assumed to be basically right, but to successively replace empirical hypothesis with ones that make better predictions. (Pollard 1997: Section 1.2.1)

While Ivan Sag is perceived as the leading figure by outsiders, the fact that Ivan Sag's recent switch to Sign-Based Construction Grammar (Sag 2012) is not unanimously adopted by the HPSG community (see Müller (2019d), Chapter P of this volume and Müller (2018a: Section 10.6.2) for discussion), that empty elements are assumed by various researchers (Bender 2000; Müller 2014; Borsley 2009) even though Ivan Sag argued against them from 1994 (Sag & Fodor 1994) and that defaults are not used by all researchers even though they play a prominent role in Ivan Sag's work (e.g., in Ginzburg & Sag 2000) shows that Pollard's claim is correct. Of course Ivan Sag is the most influential HPSG grammarian but this influence is based on properly argued empirically interesting research.

[2]We follow Culicover & Jackendoff (2005: 3) in using the term *Mainstream Generative Grammar* (MGG) to refer to work in Minimalism and the earlier Government & Binding framework.

2.2 Formalization and exhaustivity

As many of the chapters in this volume emphasize, HPSG is a framework which places considerable emphasis on detailed formal analyses of the kind that one might expect within Generative Grammar.[3] Thus, it is not uncommon to find lengthy appendices setting out formal analyses. See, for example, Sag's (1997) paper on English relative clauses, Van Eynde's (2015) book on predicative constructions and especially Ginzburg & Sag (2000), which has a 50 page appendix. One consequence of this is that HPSG has had considerable influence in computational linguistics. Sometimes theoretical work comes paired with computer implementations, which show that the analyses are consistent and complete (e.g., all publications coming out of the CoreGram project (Müller 2015b) and the HPSG textbook for German that comes with implementations corresponding to the individual chapters of the book (Müller 2007b)). For more on the relation of HPSG and computational linguistics see Bender & Emerson (2019), Chapter H of this volume.

In Minimalism things are very different. Detailed formal analyses are virtually non-existent. There appear to be no appendices like those in Sag (1997) and Ginzburg & Sag (2000). In fact the importance of formalization has long been downplayed in Chomskyan work (e.g., by Chomsky in an interview with Huybregts & Riemsdijk (1982: 73) and in discussions between Pullum (1989) and Chomsky (1990: 146)), and this view seems fairly standard within Minimalism (see also the discussion in Müller (2016a: Section 3.6.2)). Chomsky & Lasnik (1995: 28) attempt to justify the absence of detailed analyses when they suggest that providing a rule system from which some set of phenomena can be derived is not "a real result" since "it is often possible to devise one that will more or less work". Instead, they say, "the task is now to show how the phenomena […] can be deduced from the invariant principles of UG with parameters set in one of the permissible ways". Postal (2004: 5) comments that what we see here is the "notion that descriptive success is not really that hard and so not of much importance". He points out that if this were true, one would expect successful descriptions to be abundant

[3] We follow Ginzburg & Sag (2000: 2) in counting HPSG among Generative Grammar in the sense it was defined by Chomsky (1965: 4), namely as a framework that provides an explicit characterization of the theories developed within it. When we refer to work in Government & Binding or Minimalism, we follow Culicover & Jackendoff (2005: 3) in using the term *Mainstream Generative Grammar*. It should be kept in mind that there is another meaning associated with the term *generative*. A generative grammar in the latter sense generates a set (Chomsky 1957: 13. HPSG is not generative in this sense but rather model-theoretic. See Pullum & Scholz (2001) for differences between generative-enumerative and model theoretic approaches. See also Richter (2019), Chapter 3 of this volume and Wasow (2019), Chapter G of this volume.

within transformational frameworks. He argues that actual transformational descriptions are quite poor, and justifies this assessment with detailed discussions of Chomskyan work on strong crossover phenomena and passives in Chapters 7 and 8 of his book.

There has also been a strong tendency within Minimalism to focus on just a subset of the facts in whatever domain is being investigated. As Culicover & Jackendoff (2005: 535) note, "much of the fine detail of traditional constructions has ceased to garner attention". This tendency has sometimes been buttressed by a distinction between core grammar, which is supposedly a fairly straightforward reflection of the language faculty, and a periphery of marked constructions, which are of no great importance and which can reasonably be ignored. However, as Culicover (1999) and others have argued, there is no evidence for a clear cut distinction between core and periphery. It follows that a satisfactory approach to grammar needs to account both for such core phenomena as *wh*-interrogatives, relative clauses, and passives but also with more peripheral phenomena such as the following:

(2) a. It's amazing the people you see here.
  b. The more I read, the more I understand.
  c. Chris lied his way into the meeting.

These exemplify the nominal extraposition construction (Michaelis & Lambrecht 1996), the comparative correlative construction (Culicover & Jackendoff 1999; Borsley 2011), and the *X's Way* construction (Salkoff 1988; Sag 2012). As has been emphasized in other chapters, the HPSG system of types and constraints is able to accommodate broad linguistic generalizations and highly idiosyncratic facts and everything in between.

The general absence in Minimalism of detailed formal analyses is quite important. It means that Minimalists may not be fully aware of the complexity of the structures they are committed to and this allows them to sidestep the question whether it is really justified. It also allows them to avoid the question of whether the very simple conception of grammar that they favour is really satisfactory. Finally, it may be that they are unaware of how many phenomena remain unaccounted for. These are all important matters.

The general absence of detailed formal analyses has also led to Minimalism having little impact on computational linguistics. There has been some work that has sought to implement Minimalist ideas (Stabler 2001; Fong & Ginsburg 2012; Fong 2014), but Minimalism has not had anything like the productive relation with computational work that HPSG has enjoyed. Existing Minimalist implementations are rather toy grammars analyzing very simple sentences, some are not faithful to the theories they are claimed to be implementing,[4] and some do not even parse natural language but require pre-segmented, pre-formatted input. For example, Stabler's test sentences have the form as in (3).

(3) a. the king will -s eat
    b. the king have -s eat -en
    c. the king be -s eat -ing
    d. the king -s will -s have been eat -ing the pie

See Müller (2019c: Section 4.7.2) for discussion.

As far as large-scale coverage is concerned, the more recent work by Torr, Stanojevic, Steedman & Cohen (2019) is an exception to what was said above. They state that their parser is the first one to take up the Sproat & Lappin Challenge to the Minimalist community (2005). The work of the authors is impressive and they really implemented a wide-coverage statistically trained parser based on Transformational Grammar but what they did is different from standard Minimalism since they assume "around 45" versions of Move and Merge (p. 2488) in comparison to the two versions usually assumed in Minimalism (Move and Merge or Internal and External Merge).[5] Torr & Stabler (2016) explain some of the schemata that are assumed: there are versions of Merge that combine a head with a complement and versions that combine a head with a specifier (see Müller (2013) for a comparison of Minimalist Grammars with HPSG. Müller notes that the respective schemata correspond to the Specifier-Head Schema and the Head-Complement Schema, respectively). Torr & Stabler (2016: 4) assume four schemata for adjunction (HPSG has one such schema and use underspecification with respect to order, see Müller (2019a), Chapter 10 of this volume). They assume a special rule for rightward movement (p. 5) corresponding to **Keller95a**'s (**Keller95a**) and Müller's (1999b) Head-Extra Schema for extraposition. In addition the authors assume two schemata for head movement. HPSG assumes a

[4] Fong's grammars are simple Definite Clause Grammars, that is, context-free phrase structure grammars, and hence nowhere near an implementation of Minimalism, contrary to claims by Berwick, Pietroski, Yankama & Chomsky (2011: 1221). Lin's parsers PrinciPar and MiniPar (1993; 2003) are based on GB and Miniamlism but according to Lin (1993: 116) and Torr et al. (2019: 2487), they are not transformational but use a SLASH passing mechanism like the one developed in GPSG (Gazdar 1981) and standardly used in HPSG (see Borsley & Crysmann 2019, Chapter 14 of this volume).

[5] Torr explained in p.c. 2019 that these 45 rules can be folded into two Merge functions and two Move functions. But in the end this is just a clever way of hiding complexity. It is like Chomsky **ref** revising the theory with Move and Merge into one with just one operation Merge but assuming two subcases of Internal and External Merge.

lexical rule or a unary branching schema applying to words or coordinations of words (Müller 2017 and Müller (2019a), Chapter 10 of this volume). Across the Board extraction is taken care of by four special schemata. See Abeillé & Chaves (2019), Chapter 17 of this volume for the treatment of coordination in HPSG. The treatment of Across the Board Extraction (Ross 1967b) is non-standard Minimalism. For the analysis of examples like (4) in which one filler corresponds to two gaps in two conjuncts, the authors build on Kobele (2008) who uses a SLASH passing mechanism going back to Sag (1983) and Gazdar (1981). While Koeble assumes the SLASH passing mechanism of GPSG, Torr & Stabler (2016) suggest an analyis of (4) with two instances of *who* in object positions, which are later unified into one when the second conjunct is merged into the main structure.

(4) Who$_i$ did Jack say Mary likes _$_i$ and Pete hates $_i$?

An interesting property of the analysis is that *who Pete hates* forms a discontinuous constituent: *who* is combined with *hates* despite its sentence-initial position. Information about this *wh* element is passed up the tree in an GPSG-style way. The difference is that there is no trace but the extracted element is identical in phonological material with the filler. Interestingly, there is an HPSG variant of nonlocal dependencies that is very similar to what Torr & Stabler (2016) suggest and together with a modified Filler Head Schema the analyses are parallel: Hinrichs & Nakazawa (1994b) suggested that the linguistic objects that are involved in nonlocal dependencies are of type *sign* rather than *local*. This makes it possible to pass up information about a daughter including its phonological make up. If one assumes a version of HPSG permitting discontinuous constituents (Reape 1994; Müller 2019a, Chapter 10 of this volume) and a Filler Head Schema that requires that the phonology of the filler is identical to the phonology in the SLASH list and that does not insert the fronted element into the constituent order domain of the head (since it is in there already), we get an analysis of the type described in Torr & Stabler (2016). Figure 1 shows the analysis that was suggested by Torr & Stabler (2016) and Figure 2 the HPSG analog. Directional Minimalist Grammars use the '=' sign to indicate the direction in which an argument is required. =d means that a DP is required to the left of a head and d= encode the requirement of a DP to the right. This is like the '/' notation of Categorial Grammar (see Steedman 2000 and Kubota 2019, Chapter M of this volume). *likes* has the category d= =d v, which means that it is a verb requiring a d to its right (the object) and a d to its left (the subject). *who* is of category d and has a −wh feature, something that has to be checked for a derivation to be complete. *Jack* is the subject of *likes* and fulfills the =d requirement of *likes*. Items like [pres] and [int] are empty elements. [pres] has a +case feature and can make *Jack* move to its

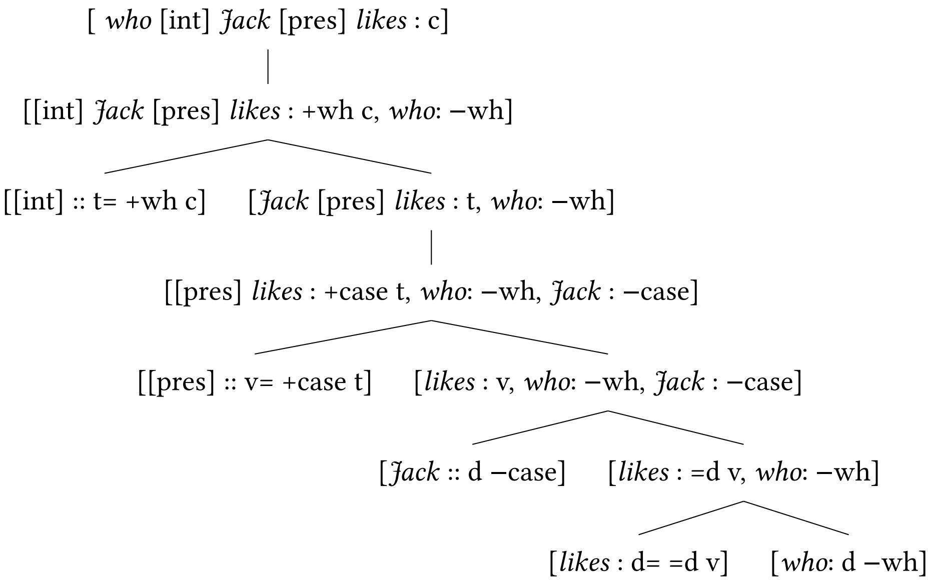


Figure 1: Derivation tree of *who Jack likes* in Directional Minimalist Grammar according to Torr & Stabler (2016)

specifier. The movement consumes the −case feature and puts *Jack* to the front of the string. This looks like a unary projection in the derivation tree. The empty interrogative head [int] selects for a t to its right. The result is a C projection that has a +wh feature. In the final step *who*, which is −wh moves to the left and the wh feature are removed. The important thing is that the information about the phonology of *who* and its wh feature is percolated up in the tree until it is finally bound off in the last derivation step.

Figure 2 shows the HPSG analog. The information about the local properties of the *wh* word including its phonology are passed up in the tree until they are bound off in a filler head configuration. The Filler-Head Schema binds off the nonlocal dependency and makes sure that the phonology of the filler is not realized twice (see Reape 1994, Müller 2019a: Section 6, Chapter 10 of this volume on linearization domains and Abeillé & Chaves 2019: Section 7, Chapter 17 of this volume on multi-dominance approaches in HPSG). An alternative to a binary branching Filler-Head Schema would be a unary branching rule that binds off the element in SLASH and adds the stored phonology to the phonology of the daughter. This would then be completely parallel to the unary branching assumed in Torr's Directional Minimalist Grammar.

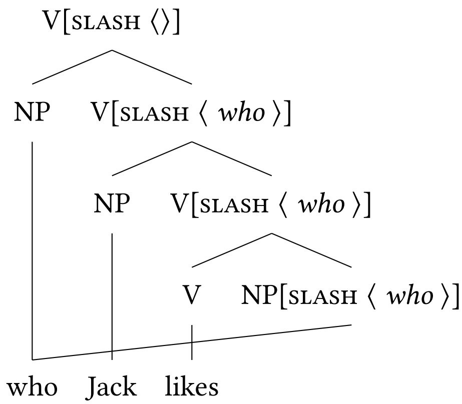


Figure 2: Possible HPSG analysis of *who Jack likes* using discontinuous constituents

Summing up: The fact that certain variants of Minimalism share properties with Categorial Grammar has been noticed early on (Berwick & Epstein 1995). Directional Minimalist Grammars were compared to CG and HPSG by Müller (2013). MGs were extended to include GPSG-style SLASH passing mechanisms by Kobele (2008) and continue to use them in the versions of Torr & Stabler (2016). We believe that this work is fruitful and well-formalized but formalization is insufficient for most of the work in Minimalism and idea from other frameworks are ignored more often than not.

2.3 Empirical quality

There are, then, issues about the quantity of data that is considered in Minimalist work. There are also issues about its quality (Schütze 2016). Research in HPSG is typically quite careful about data and often makes use of corpus and experimental data (see for example An & Abeillé 2017; Müller 1999b; 2002; Bildhauer & Cook 2010; Müller, Bildhauer & Cook 2012; Van Eynde 2015: Chapter 7; Abeillé et al. 2016; Shiraishi & Abeillé 2016 for examples of work with attested examples and for experimental work). This use of corpus data and attested examples is based on the insight that introspection alone is not sufficient, given that an enormous amount of time is spent to work out analyses and it would be unfortunate if these analyses were built on a shaky empirical basis. See Müller (2007a) and Meurers & Müller (2009) for the discussion of introspection vs. corpus data. Research in

Minimalism is often rather less careful.[6] In a review of a collection of Minimalist papers, Bender (2002: 434) comments that: "In these papers, the data appears to be collected in an off-hand, unsystematic way, with unconfirmed questionable judgments often used at crucial points in the argumentation". She goes on to suggest that the framework encourages "lack of concern for the data, above and beyond what is unfortunately already the norm in formal syntax, because the connection between analysis and data is allowed to be remote.". Similar things could be said about a variety of Minimalist work. Consider, for example, Aoun & Li (2003), who argue for quite different analyses of *that*-relatives and *wh*-relatives on the basis of the following (supposed) contrasts, which appear to represent nothing more than their own judgements (p. 110–112):

(5) a. The headway that Mel made was impressive.
 b. ?? The headway which Mel made was impressive.

(6) a. We admired the picture of himself that John painted in art class.
 b. * We admired the picture of himself which John painted in art class.

(7) a. The picture of himself that John painted in art class is impressive.
 b. *? The picture of himself which John painted in art class is impressive.

None of the native speakers we have consulted find significant contrasts here which could support different analyses. The example in (8a) with a *which* relative clause referring to *headway* can be found in Cole et al. (1982). Williams (1989: 437) and Falk (2010: 221) have examples with a reflexive coreferential with a noun in a relative clause introduced by *which* as in William's (8b) and corpus examples like (8c,d) can be found as well:

(8) a. The headway which we made was satisfactory.
 b. the picture of himself which John took
 c. The words had the effect of lending an additional clarity and firmness of outline to the picture of himself which Bill had already drawn in his mind–of a soulless creature sunk in hoggish slumber.[7]

[6] We hasten to say that we do not claim this to be true for all Minimalist work. There are researchers working with corpora or at least with attested examples (Wurmbrand 2003) and there is experimental work. Especially in Germany there were several large scale Collaborative Research Centers with a strong empirical focus which also fed back into theoretical work, including Minimalist work. The fact that we point out here is that there is work, including work by prominent Minimalists, that is rather sloppy as far as data is concerned.

[7] Wodehouse, P.G. 1917. *Uneasy Money*, London: Methuen & Co., p. 186, http://www.literaturepage.com/read.php?titleid=uneasymoney&abspage=186, 2018-09-18.

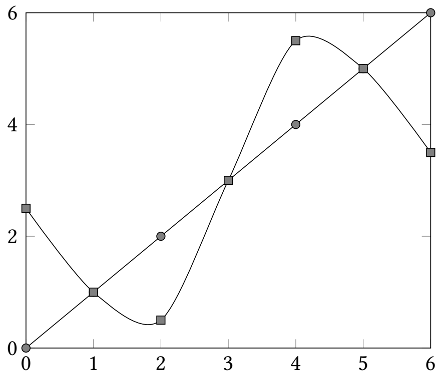
d. She refused to allow the picture of himself, which he had sent her, to be hung, and it was reported that she ordered all her portraits and busts of him to be put in the lumber attics.[8]

Given that it is relatively easy to come up with counterexamples it is surprising that authors do not do a quick check before working out rather complex analyses.

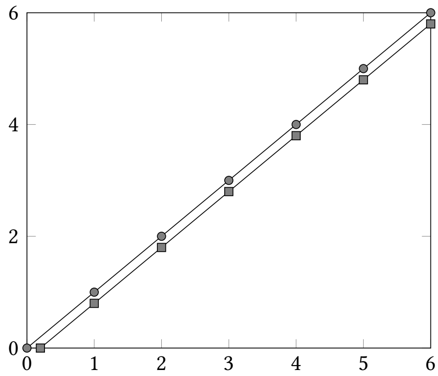
Note that we are not just taking one bad example of Minimalist work. It is probably the case that papers with dubious judgments can be found in any framework if only this is due to the repetitions of unwarranted claims made by others. The point is that Aoun & Li are influential (quoted by 455 other publications as of 2018-09-14). Others rely on these judgments or the analyses that were motivated by them. New conclusions are derived from analyses since theories make predictions. If this process continues for a while an elaborate theoretical edifice results that is not empirically supported. Note furthermore that the criticism raised here is not the sqabble of two authors working in an alternative framework. This criticism also comes from practitioners of Mainstream Generative Grammar. For example, Wolfgang Sternefeld and Hubert Haider, both very prominent figures in the German Generative Grammar school criticized the scientific standards in Minimalism heavily (Sternefeld & Richter 2012; Haider 2018). As Sternefeld & Richter (2012: 266–268) point out it is not just the case that Minimalist publications are based on empirically problematic foundations, it is even worse: researchers like Epstein & Seely (2006: Section 1.1) publish statements about research methodology that can only be understood as an immunization strategy. The authors discuss the two curves in Figure 3: one has a sine-like shape and hits three out of seven of the data points on a straight line and the other one is parallel to the line with the data points and hence hits none of the data points. The second approach gets none of the data right. Nevertheless the authors argue that one should prefer this theory since it was closer to the truth and more illuminating: "Clearly, Theory 1 is 'empirically preferable' by a 'winning score' of 3–0. The point […] is that Theory 2, despite getting none of the data correct, […] is 'closer to the truth' or 'more illuminating' than the empirically preferable Theory 1. Hence, we believe, Theory 2 is a better working or guiding hypothesis upon which to base future research." (Epstein & Seely 2006: 2)

As we will show in Section 3.4, Minimalist discussions of the important topic of labelling have also been marred by a failure to take account of relevant data.

[8]Jerrold, Clare. 1913. *The married life of Queen Victoria*, London: G. Bells & Sons, Ltd. https://archive.org/stream/marriedlifeofque00jerruoft/marriedlifeofque00jerruoft_djvu.txt, 2018-09-19.



(a) Theory 1



(b) Theory 2

Figure 3: Immunization in Minimalist work: According to Epstein & Seely (2006: Section 1.1) Theory 2, which gets none of the data points, is preferable to Theory 1 with three data points since Theory 2 is "more illuminating" and "closer to to the truth"

2.4 Argumentation for invisible entities and the assumption of innate linguistic knowledge

There are also differences in the kind of arguments that the two frameworks find acceptable. It is common within Minimalism to assume that some phenomenon which cannot be readily observed in some languages must be part of their grammatical system because it is clearly present in other languages. Notable examples would be case (Li 2008) or (object) agreement (Meinunger 2000: Chapter 4), which are assumed to play a role even though there are no visible manifestations within some languages (e.g., Mandarin Chinese and German, respectively). This stems from the longstanding Chomskyan assumption that language is the realization of a complex innate language faculty. From this perspective, there is much in any grammatical system that is a reflection of the language faculty and not in any simple way of the observable phenomena of the language in question. If some phenomenon plays an important role in many languages it is viewed as a reflection of the language faculty, and hence it must be a feature of all grammatical systems even those in which it is hard to see any evidence for it. An example – taken from a textbook on Minimalism (Hornstein, Nunes & Grohmann 2005: 124) – is an analysis of prepositional phrases in English. Figure 4 shows the analysis.[9]

[9]This analysis is actually a much simpler variant of the PP analysis which appeared in an earlier textbook by Radford (1997: 452). For discussion of this analysis see Sternefeld (2006: 549–550) and Müller (2016a: Section 4.6.1.2).

Due to theory internal assumptions the case requirement of the preposition can-

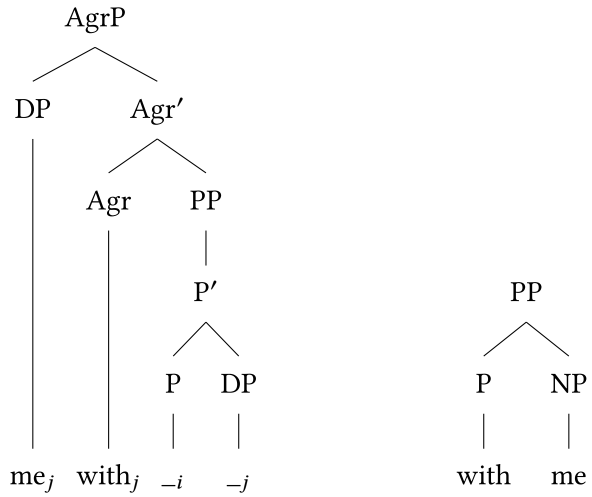


Figure 4: Minimalist analysis of a PP according to (Hornstein, Nunes & Grohmann 2005: 124) and the analysis assumed in HPSG and all other phrase structure-based frameworks

not be checked in the P-DP combination. According to the version of the theory adopted by the authors, case has to be checked in specifier positions. Therefore it was assumed that the preposition moves to an Agr head and the DP moves to the specifier position of this Agr head. The problem is of course that DP and P are in the wrong order now. However, the authors argue that this is the order that is manifested in Hungarian and that Hungarian is a language which has postpositions and these are agreeing with their nominal dependent. The authors assume that Hungarian postpositions are prepositions underlyingly and that the DP following the preposition moves to the left because of a movement process that is triggered by agreement. It is claimed that this movement exists both in Hungarian and in English but that the movement is covert (that is, invisible) in the latter language.

This line of argument would be reasonable if a complex innate language faculty were an established fact, but it isn't, and since Hauser, Chomsky & Fitch (2002), it seems to have been rejected within Minimalism. It follows that ideas about an innate language faculty should not be used to guide research on individual languages. Rather, as Müller (2015b: 25) puts it, "grammars should be motivated on a language-specific basis". Does this mean that other languages are irrelevant when investigating a specific language? Clearly not. As Müller also puts it, "In situations where more than one analysis would be compatible with a given dataset for language X, the evidence from language Y with similar constructs is most welcome and can be used as evidence in favor of one of the two analyses for language X." (2015b: 43). In practice, any linguist working on a new language will use apparently similar phenomena in other languages as a starting point. It is important, however, to recognize that apparently similar phenomena may turn out on careful investigation to be significantly different.[10]

3 Different views of grammar

We turn now to more substantive differences between HPSG and Minimalism, differences in their conceptions of grammar, especially syntax, and differences in their views of syntactic structure. As we will see, these differences are related. In this section we consider the former, and in the next we will look at the latter.

3.1 Declarative and constraint-based vs. derivational and generative-enumerative approaches

As is emphasized throughout this volume, HPSG assumes a declarative or constraint-based view of grammar. It also assumes that the grammar involves a complex systems of types and constraints. Finally, it assumes that syntactic analyses are complemented by separate semantic and morphological analyses. In each of these areas, Minimalism is different. It assumes a procedural view of grammar. It assumes that grammar involves just a few general operations. Finally, it assumes that semantics and morphology are simple reflections of syntax. We comment on each of these matters in the following subsections.

Whereas HPSG is a declarative or constraint-based approach, Minimalism seems to be firmly committed to a procedural approach. Chomsky (1995b: 219) remarks that: "We take L [a particular language] to be a generative procedure that constructs pairs (π, λ) that are interpreted at the articulatory-perceptual (A-P) and conceptual-intentional (C-I) interfaces, respectively, as 'instructions' to the performance systems.". Various arguments have been presented within HPSG for a declarative view, but no argument seems to be offered within Minimalism for a procedural view. Obviously, speakers and hearers do construct representations and must have procedures that enable them to do so, but this is a matter of

[10] Equally, of course, apparently rather different phenomena may turn out on careful investigation to be quite similar. For further discussion of HPSG and comparative syntax, see Borsley (2018).

performance, and there is no reason to think that the knowledge that is used in performance has a procedural character (see Section 5.1 on processing). Rather, the fact that it is used in both production and comprehension suggests that it should be neutral between the two and hence declarative. See also Wasow (2019: Section 3.1), Chapter G of this volume on this point.

Another difference between constraint-based and generative-enumerative approaches is that the first type of proposal provides a way to get graded acceptability into the picture (Pullum & Scholz 2001: Section 3.1). Since HPSG grammars are basically feature-value pairs with equality (or other relations) between values, it is possible to weigh constraints and it is possible to admit constraint violations and work with structures with violated constraints (see for example Sorace & Keller 2005 on cumulative constraint violation). So looking at the sentences in (9) we see that more and more constraints are violated:

(9) a. I am the chair of my department.
  b. * I are the chair of my department.
  c. * Me are the chair of my department.
  d. * Me are the chair of me's department.
  e. * Me are chair the of me's department.
  f. * Me are chair the me's department of.

In comparison to this a generative-enumerative grammar enumerates a set and a sequence either is in the set or it is not.[11]

For further discussion of the issues, see Section 5.1 of this paper and e. g. Pullum & Scholz (2001), Postal (2003), Sag & Wasow (2011; 2015), and Wasow (2019), Chapter G of this volume.

3.2 Underspecification

Another crucial difference between HPSG and Minimalism is that HPSG allows for the underspecification of information. In the absence of constraints, all principle options are possible. This is different in Minimalism. All structures that are derivable are predetermined by the numeration. Features have to be specified and they determine movement and properties of the derived objects. The general characterization of the frameworks is:

(10) a. Minimalism: Only what is explicitly ruled in works.

[11]For a discussion of Chomsky's (1964; 1975: Chapter 5) proposals to deal with different degrees of acceptability see Pullum & Scholz (2001: 29).

b. HPSG: Everything that is not ruled out works.

Let us consider some examples. The availability of type hierarchies makes it possible to underspecify part-of-speech information. For example, Sag (1997) assumes that complementizer (*comp*) and verb (*verb*) have a common supertype *verbal*. A head can then select for a complement with the category *verbal*. So rather than specifying two lexical items with different valence information or one with a disjunctive specification *verb* ∨ *comp*, one has just one lexical item selecting for *verbal*. Similarly, schemata (grammar rules) can contain underspecified types. A daughter in a dominance schema can have a value of a certain types that subsumes a number of other types. Let's say three. Without this underspecification one would need three schemata: one for every subtype of the more general type.

Quantifier scope can be underspecified as well (Copestake, Flickinger, Pollard & Sag 2005; Richter & Sailer 1999; Koenig & Richter 2019, Chapter C of this volume): constraints regarding which quantifier outscopes which other quantifier may be left unspecified. The absence of the respective constraints results in a situation where several scopings are possible. In transformational models it is usually assumed that quantifier elements move into certain positions covertly and scope relations are read off of the resulting tree (May 1985; Frey 1993; Sauerland & Elbourne 2002). This is unnecessary in HPSG.[12]

3.3 Types and constraints vs. general operations

The declarative-procedural contrast is an important one, but the contrast between the complex systems of types and constraints that are assumed within HPSG and the few general operations that form a Minimalist grammar is arguably more important.[13] Much work in Minimalism has three main operations Merge, Agree, and Move or Internal Merge. Merge combines two expressions, either words or phrases, to form a larger expression with the same label as one of the expressions (Chomsky 1995b: 244; 2008: 140). Its operation can be presented

[12] Apart from the possibility of a more compact specification of scope relations, HPSG analyses are actually superior on an empirical level. As Kiss (2005) and independently also Fanselow (2001) (working in a Minimalist setting) have shown, movement-based approaches cannot account for cases in which two quantified noun phrases are moved simultaneously while maintaining their relative order. Reconstruction predicts too many readings. Sauerland & Elbourne (2002) provided a solution to the problem but the analysis has an absurd complexity involving several movements, some of them invisible. For more discussion of scrambling and scope and the approaches mentioned here see Müller (2019c: 114–116).

[13] A procedural approach doesn't necessarily involve a very simple grammatical system. The Standard Theory of Transformational Grammar (Chomsky 1965) is procedural but has many different rules, both phrase structure rules and transformations.

as shown in Figure 5. In the case of English, the first alternative is represented by

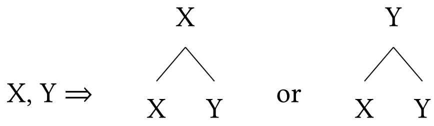


Figure 5: Merge

situations where a lexical head combines with a complement while the second is represented by situations where a specifier combines with a phrasal head. Chomsky (2008: 146) calls items merged with the first variant of Merge *first-merged* and those merged with the second variant *later-merged*.

Agree, as one might suppose, offers an approach to various kinds of agreement phenomena. It involves a probe, which is a feature or features of some kind on head, and a goal, which the head c-commands.[14] At least normally, the probe is a linguistic object with an uninterpretable feature or features with no value and the goal has a matching interpretable feature or features with appropriate values (Chomsky 2001: 3–5).[15] Agree values the uninterpretable feature or features and they are ultimately deleted, commonly after they have triggered some morphological effect. Agree can be represented as in Figure 6 (where the "u" prefix identifies a feature as uninterpretable, and we have just one uninterpretable feature on the probe and just one matching interpretable feature on the goal). Unsurprisingly subject-verb agreement is one manifestation of Agree, where X

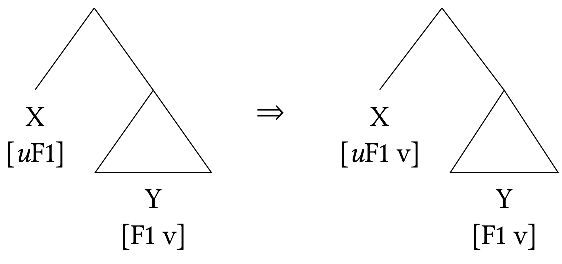


Figure 6: Agree

[14] It is not necessary to state the c-command restriction since it is assumed that linguistic objects are built bottom-up and when a head is combined with its dependent the head can "see" the dependent but not the context in which the resulting object is embedded. Information about dominating material will be available only later in the derivation. (Guido Mensching, p. c. 2018)

[15] Chomsky also assumes that the goal also has an uninterpretable feature of some kind to render it 'active'. In the case of subject-verb agreement, this is a Case feature on the subject.

is T(ense) and Y is a nominal phase, for Minimalism a DP, inside the complement of T. T presumably has two uninterpretable features, person and number, and the DP two matching interpretable features. Here, and elsewhere, Agree is a non-local relation involving elements which are not sisters. This contrasts with the situation in HPSG, in which subject-verb agreement is a consequence of a relation between the subject and its VP sister and a relation between the VP and the V that heads it.

Finally, Move, also called Internal Merge, is an operation which makes a copy of a constituent of some expression and merges it with the expression (Chomsky 1995b: Section 4.4; 2008: 140). The original element that is copied normally undergoes deletion. The process can be presented as in Figure 7.

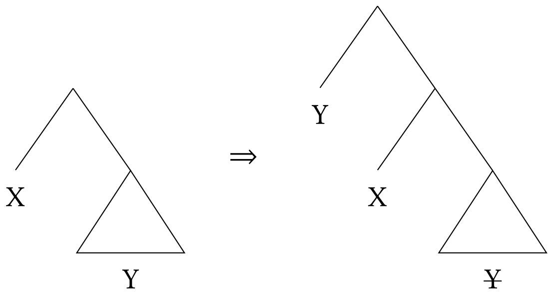


Figure 7: Move

This covers both the A′-movement process assumed for unbounded dependency constructions such as *wh*-interrogatives and the A-movement process assumed for raising sentences and passives. A question arises about so-called head-movement, where a head moves to a higher head position. This appears to mean that it must be possible for the copy to be merged with the head of the expression that contains it. However, this is incompatible with the widely assumed extension condition, which requires Merge to produce a larger structure. One response is the idea espoused in Chomsky (1995a: 368; 2001: 37) that head-movement takes place not in the syntax but in the PF component, which maps syntactic representations to phonetic representations. It seems that its status is currently rather unclear.

The three operations just outlined interact with lexical items to provide syntactic analyses. It follows that the properties of constructions must largely derive from the lexical items that they contain. Hence, the properties of lexical items are absolutely central to Minimalism. Oddly, the obvious implication – that the lexicon should be a major focus of research – seems to be ignored. As Newmeyer

(2005: 95, fn. 9) comments:

> […] in no framework ever proposed by Chomsky has the lexicon been as important as it is in the MP [Minimalist Program]. Yet in no framework proposed by Chomsky have the properties of the lexicon been as poorly investigated. (Newmeyer 2005: 95, fn. 9)

Sometimes it is difficult to derive the properties of constructions from the properties of visible lexical elements. But there is a simple solution: postulate an invisible element. The result is a large set of invisible functional heads. As we will see in Section 4.1.6, these heads do the work in Minimalism that is done by phrase types and the constraints on them in HPSG.

Although Minimalism is a procedural approach and HPSG a declarative approach, there are some similarities between Minimalism and early HPSG, the approach presented in Pollard & Sag (1987; 1994). In much the same way as Minimalism has just a few general mechanisms, early HPSG had just a few general phrase types. Research in HPSG in the 1990s led to the conclusion that this is too simple and that a more complex system of phrase types is needed to accommodate the full complexity of natural language syntax. Nothing like this happened within Minimalism, almost certainly because there was little attempt within this approach to deal with the full complexity of natural language syntax. As noted above, the approach has rarely been applied in detailed formal analyses. It looks too simple and it appears problematic in various ways. It is also a major source of the complexity that is characteristic of Minimalist syntactic structures, as we will see in Section 4.

3.4 Labelling

As we noted in the last section, Merge combines two expressions to form a larger expression with the same label as one of the expressions. But which of the original expressions provides the label for the larger expression? This issue has been discussed but not very satisfactorily. Chomsky defines which label is used in two different cases: the first case states that the label is the label of the head if the head is a lexical item and the second case states the label to be the label of the category from which something is extracted (Chomsky 2008: 145). As Chomsky notes, these rules are not unproblematic since the label is not uniquely determined in all cases. An example is the combination of two lexical elements since in such cases both elements can be the label of the resulting structure. Chomsky notices that this could result in deviant structures, but claims that this concern is unproblematic and ignores it. This means that rather fundamental notions in a grammar theory were ill-defined. A solution to this problem was provided in his 2013 paper (published five years later). However, this paper is inconsistent (Müller 2016a: Section 4.6.2) and even insiders find it incomprehensible. But this is not the point we want to focus on here. Rather we want to show one more time that empirical standards are not met. Chomsky uses underdetermination in his labeling rules to account for two possible structures in (11), an approach going back to Donati (2006):

(11) what [C [you wrote *t*]]

(11) can be an interrogative clause as in *I wonder what you wrote.* or a free relative clause as in *I will read what you wrote.*. According to the labeling rule that accounts for sentences from which an item is extracted, the label will be CP since the label is taken from the clause. However, since *what* is a lexical item, *what* can determine the label as well. If this labeling rule is applied *what you wrote* is assigned DP as a label and hence the clause can function as a DP argument of *read*.

Chomsky's proposal is interesting but it does not extend to cases involving free relative clauses with complex *wh*-phrases (so-called pied-piping) as they are attested in examples like (12):

(12) I'll read [whichever book] you give me.

The example in (12) is from one of the standard references on free relative clauses: Bresnan & Grimshaw (1978: 333), which is also cited in other mainstream generative work as for example (Groos & van Riemsdijk 1981).

Apart from the fact that complex *wh*-phrases are possible there is even more challenging data in the area of free relative clauses: the examples in (13) and (14) show that there are non-matching free relative clauses:

(13) Sie kocht, worauf sie Appetit hat.[16] (German)
she cooks where.on she appetite has
'She cooks what she feels like eating.'

(14) a. Worauf man sich mit einer Pro-form beziehen kann, [...] ist
where.upon one self with a Pro-form refer can is

[16] Bausewein (1990: 154).

eine Konstituente.[17]
a constituent

'If you can refer to something with a Pro-form, [...] it is a constituent.'

b. [Aus wem] noch etwas herausgequetscht werden kann, ist
out who yet something out.squeezed be can is
sozial dazu verpflichtet, es abzuliefern; ...[18]
socially there.to obliged it to.deliver

'Those who have not yet been bled dry are socially compelled to hand over their last drop.'

In (13) a relative clause with a PP relative phrase functions as an accusative object. In (14) the relative clauses function as subjects. (14b) is another example of a relative clause with a complex *wh* phrase. See Bausewein (1990) and Müller (1999a) for further discussion of free relative clauses and attested data.

According to Donati (2006: Section 5), pied piping does not exist in free relatives (see also Citko (2008: 930–932) for a rejection of this claim). Given how much attention the issue of Labeling has received and how central this is to Minimalist analyses this situation is quite surprising: an empirically false claim made in 2002/2003[19] is the basis for foundational work from 2002 until 2013 even though the facts are common knowledge in the field. Ott (2011) develops an analysis in which the category of the relative phrase is projected, but he does not have a solution for nonmatching free relative clauses, as he admits in a footnote on page 187. The same is true for Citko's analysis (2008), in which the extracted XP can provide the label. So, even though the data has been known for decades, it is ignored by authors and reviewers and foundational work is built on shaky empirical ground. See Müller (2016a: Section 4.6.2) for a more detailed discussion of labeling.

3.5 Feature deletion and "crashing at the interfaces"

In Section 3.3, we mentioned Case as an uninterpretable feature which renders a DP active. Like other uninterpretable features this is deleted as a result of Agree because it is not interpretable in LF. This means that Minimalism claims that a

[17] From the main text of: Günther Grewendorf, *Aspekte der deutschen Syntax. Eine Rektions-Bindungs-Analyse.* Studien zur deutschen Grammatik, number 33. Tübingen: Gunter Narr Verlag. 1988, p. 16, quoted from Müller (1999a: 61).

[18] Wiglaf Droste, taz, 01.08.97, p. 16, quoted from Müller (1999a: 61).

[19] Versions of Donati's paper were presented at *Going Romance* in 2002 and the most important Mainstream Generative Grammar conference in Europe GLOW in 2003.

case marked NP like *der Mann* 'the man' is not interpretable unless it is somehow stripped of its case information. So in Minimalism *der Mann* needs something on top of the DP that Agrees with and thereby consumes the case feature. While this seems cumbersome to most working outside Minimalism, there are actually deeper problems connected to the deletion of case features. There are situations in which you need case features more than once. An example of this is free relative clauses as the one in (15b):

(15) a. der Mann
the.NOM man

b. Ich treffe, wen ich treffen will.
I meet who.ACC I meet want.to
'I meet whoever I like to meet.'

wen is the accusative object in the relative clause. Since it is an object its case feature will be checked by the selecting verb *treffen* 'meet'. *wen* will then be a DP without any case information. However, the case of the relative phrase in free relative clauses is not arbitrary. It is important for the integration of the free relative clause in the matrix clause. The case of *wer* 'who' in a complete relative clause has to be known since it is important for the external distribution of the free relative clause, as the examples in (16) show:

(16) a. Wer mich treffen will, kann vorbeikommen.
who.NOM me meet wants.to may over.come
'He who wants to meet me may come over.'

b. * Ich treffe, wer mich treffen will.
I meet who.NOM me meet wants.to
'I meet whoever wants to meet me.'

HPSG also consumes resources in a way: items in valence representations are not projected up the tree once the requirement is saturated, but the difference is that objects with a certain structure and with certain features are not modified. A case-marked NP is not deprived of this case information. We think that this is the right way to deal with morphological markings and with feature specifications in general.

3.6 Some implications

We will look in detail at the implications for syntactic structure of this machinery in the next section. However, we will note some implications in the following paragraphs as a kind of preview of the next section.

First, the fact that Merge combines two expressions entails that syntactic structures are confined to binary branching and excludes various analyses that have been assumed within HPSG and other frameworks. Second, the assumption that expressions produced by Merge have the same label as one of the expressions that they consist of (Chomsky 2008: 145) is essentially the assumption that all complex expressions are headed. For HPSG, as for many other approaches, there are headed expressions and non-headed expressions, e.g., coordination and the NPN construction discussed in Sections 4.2.2 and 4.2.3, respectively.

As emphasized above, a further important feature of Minimalism is the view that semantics and morphology are simple reflections of syntax. The basic architiecture assumed in Minimalism is shown in Figure 8. Both phonology and semantics are read off the structures produced by syntax. The idea that seman-

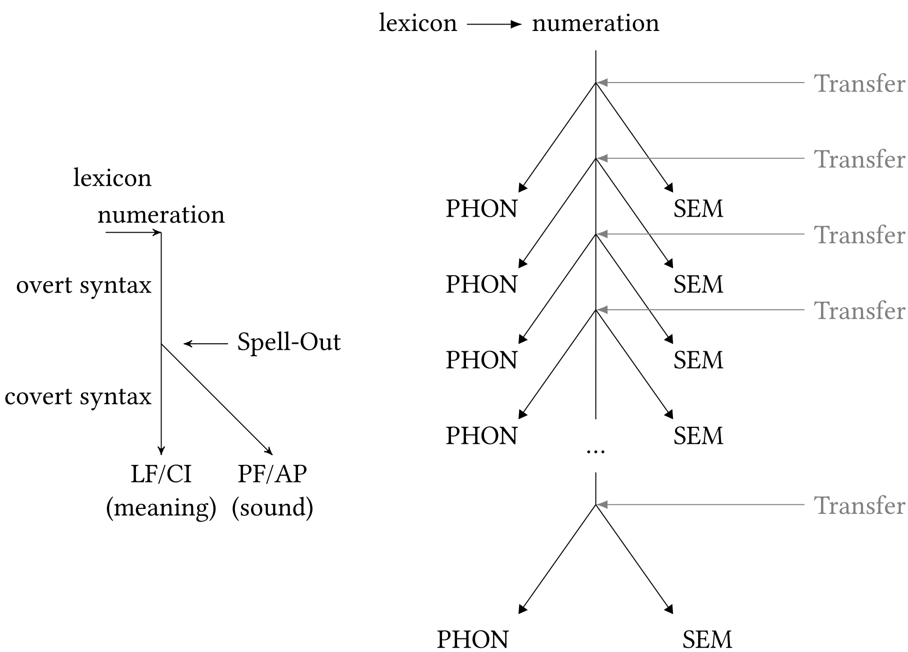


Figure 8: Syntax-centristic architecture in Minimalism before the Phase model (left) and in the Phase model (right) according to Richards (2015: 812, 830))

tics is a simple reflection of syntax goes back to the early years of Transformational Grammar. One aspect of this idea was formalized as the Uniform Theta

Assignment Hypothesis (UTAH) by Baker (1988: 46).

(17) Uniform Theta Assignment Hypothesis
Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

Minimalism abandoned the notion of D-structure, but within Minimalism the Hypothesis can be reformulated as follows:

(18) Uniform Theta Assignment Hypothesis (revised)
Identical thematic relationships between items are represented by identical structural relationships between those items when introduced into the structure.

We will look at some of the implications of this in the next section.

The idea that morphology is a simple reflection of syntax is also important. As we will discuss in the next section, it leads to abstract underlying structures and complex derivations and to functional heads corresponding to various suffixes. Again, we will say more about this in the next section.

4 Different views of syntactic structure

The very different views of grammar that are assumed in Minimalism and HPSG naturally lead to very different views of syntactic structure. The syntactic structures of Minimalism are both very complex and very simple. This sounds paradoxical but it isn't. They are very complex in that they involve much more structure than those assumed in HPSG and other approaches. But they are very simple in that they have just a single ingredient – they consist entirely of local trees in which there is a head and a single non-head. From the standpoint of HPSG, they are both too complex and too simple. We will consider the complexity in Section 4.1 and then turn to the simplicity in Section 4.2.

4.1 The complexity of Minimalist structures

For HPSG, as the chapters in this volume illustrate, linguistic expressions have a single relatively simple constituent structure with a minimum of phonologically empty elements.[20] For Minimalism, they have a complex structure containing a

[20]The relatively simple structures of HPSG are not an automatic consequence of its declarative nature. Postal's Metagraph Grammar framework (formerly known Arc Pair Grammar) is a declarative framework with structures that are similar in complexity to those of Minimalism (see Postal 2010).

variety of empty elements and with various constituents occupying more than one position in the course of the derivation. Thus the structures assumed within Minimalism are not at all minimalist. But this complexity is a more or less inevitable consequence of the Minimalist view of grammar outlined above.

4.1.1 Uniformity of structures due to semantic representation

There are a variety of sources of complexity, and some predate Minimalism.[21] This is true especially of the idea that semantics and morphology are simple reflections of syntax (on morphology see Section 4.1.3). For the syntax-semantics relation, UTAH, which we introduced on p. 960, is particularly important. It leads to a variety of abstract representations and movement processes. Consider, for example, the following:

(19) a. Who did Lee see?
     b. Lee saw who

Who bears the same thematic relation to the verb *see* in (19a) as in (19b). Assuming UTAH, it follows that *who* in (19a) should be introduced in the object position which it occupies in (19b) and then be moved to its superficial position. Consider next the following:

(20) a. Lee was seen by Kim.
     b. Kim saw Lee.

Here, *Lee* bears the same thematic relation to the verb *see* in (20a) as in (20b). Hence, it follows that *Lee* in (20a) should be introduced in the object position which it occupies in (20b) and then be moved to its superficial subject position. Finally, consider these examples:

(21) a. Lee seems to be ill.
     b. It seems that Lee is ill.

Here, *Lee* bears the same thematic relation to *ill* in (21a) as in (21b). Thus, it follows that *Lee* in (21a) should be introduced in the same position as *Lee* in (21b). The standard Minimalist approach assumes that *Lee* in both examples originates in a position adjacent to *ill* and is moved a short distance in (21b) but a longer distance in (21a).

[21] For interesting discussion of the historical development of the ideas that characterize Minimalism, see Culicover & Jackendoff (2005: Chapters 2 and 3).

These analyses are more or less inevitable if one accepts UTAH. But how sound is UTAH? Work in HPSG shows that it is quite possible to capture both the syntactic and the semantic properties of these sentence types without the assumption that the crucial constituents occupy more than one position. Thus, there is no reason to accept UTAH.

4.1.2 Lexical decomposition a la Generative Semantics

The idea that semantics is a simple reflection of syntax has led to other kinds of complexity. For example, it has led to revival of the idea once characteristic of Generative Semantics that lexical items may derive from complex expressions which in some sense represent their meanings.[22] Thus, Hale & Keyser (1993) argue that (22a) derives from a structure like that of (22b).

(22) a. Kim shelved the books.
b. Kim put the books on the shelf.

One problem with this proposal is that *shelve X* means more than just *put X on the shelf.* Thus, (23a) is not equivalent to (23b).

(23) a. Kim put his elbow on the shelf.
b. Kim shelved his elbow.

Moreover, as Culicover & Jackendoff (2005: 54–55) point out and as Hale & Keyser (1993: 105, Fn. 7) note themselves, denominal verbs can have many different interpretations.[23]

(24) a. Kim saddled the horse.
(Kim put the saddle on the horse.)
b. He microwaved the food.
(He put the food in the microwave and in addition he heated it.)
c. Lee chaired the meeting.
(Lee was the chairperson of the meeting.)
d. Sandy skinned the rabbit.
(Sandy removed the skin from the rabbit.)

[22] For typical Generative Semantics proposals of this kind, see McCawley (1968) and Postal (1970). Like Minimalism, Generative Semantics was characterized by extremely complex syntactic structures and for similar reasons. See Newmeyer (1986: Chapter 4) for discussion.

[23] The examples in (24c), (24g) and (24h) are taken from (Culicover & Jackendoff 2005: 54–55) or parallel to examples they discussed.

e. Kim pictured the scene.
(Kim constructed a mental picture of the scene.)

f. They stoned the criminal.
(They threw stones at the criminal.)

g. He fathered three children.
(He was the biological father of three children.)

h. He mothers his students.
(He treats his students the way a mother would.)

Denominal verbs need to be associated with the correct meanings, but there is no reason to think that syntax has a role in this.[24]

4.1.3 Complex structures and morphology

The idea that morphology is a simple reflection of syntax also leads to syntactic complexity. The fact that verbs in English and many other languages are marked for tense leads to the assumption that there is a T(ense) head at the heart of clause structure. Thus the sentence in (25) has the analysis in Figure 9.

(25) The cat chased the dog.

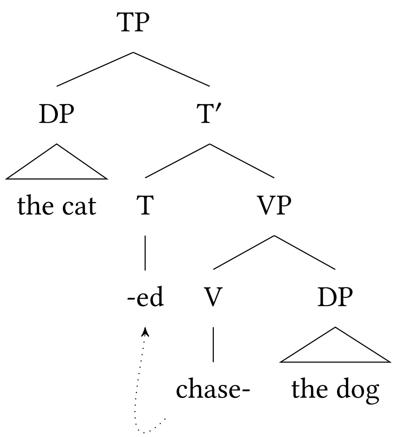


Figure 9: TP/VP analysis of simple English sentences

The verbal stem moves to the T head to pick up the *-ed* suffix.

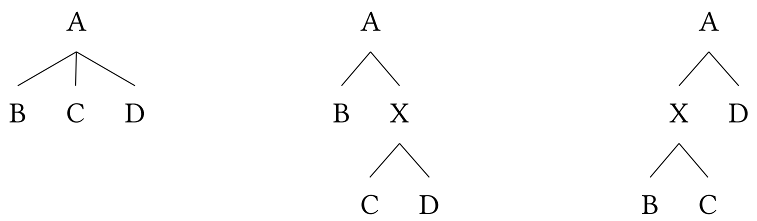
Similarly the fact that nouns in English and other languages are marked for number leads to the assumption that there is a Num(ber) head at the heart of noun

[24] See Culicover & Jackendoff (2005: 53–56) for further discussion.

phrase structure. These elements are not solely motivated by morphology. The assumption that verbs move to T and nouns to Num in some languages but not others provides a way of accounting for cross-linguistic word order differences (Pollock 1989). However, assumptions about morphology are an important part of the motivation. As discussed in Crysmann (2019), Chapter B of this volume, HPSG assumes a realizational approach to morphology, in which affixes are just bits of phonology realizing various properties of inflected words or derived lexemes. Hence, analyses like these are out of the question.

4.1.4 Binary branching

Another source of complexity which also predates Minimalism is the assumption that all structures are binary branching. As Culicover & Jackendoff (2005: 112–116) note, this idea goes back to the 1980s. It entails that there can be no structures of the form in Figure 10a. Rather all structure must take the form in Figure 10b or Figure 10c. As Culicover & Jackendoff discuss, the arguments for



(a) flat branching (b) binary branching (c) binary branching

Figure 10: Flat and binary branching

the binary branching restriction have never been very persuasive. Moreover, it is incompatible with various analyses which have been widely accepted in HPSG and other frameworks. We will return to this topic in Section 4.2.

4.1.5 Unbounded dependency constructions

As noted in Section 3, the simplicity of the Minimalist grammatical system means the properties of constructions must largely derive from the lexical items that they contain. Hence, the properties of lexical items are absolutely central to Minimalism and often this means the properties of phonologically empty items, especially empty functional heads. Thus, such elements are central feature of Minimalist syntactic structures. These elements do much the same work as phrase types and the associated constraints in HPSG.

The contrast between the two frameworks can be illustrated with unbounded dependency constructions. Detailed HPSG analyses of various unbounded dependency constructions are set out in Sag (1997; 2010) and Ginzburg & Sag (2000), involving a complex system of phrase types. For Minimalism, unbounded dependency constructions are headed by a phonologically empty complementizer (C) and have either an overt filler constituent or an invisible filler (an empty operator) in their specifier position. Essentially, then, they have the structure in Figure 11. All the properties of the construction must stem from the properties

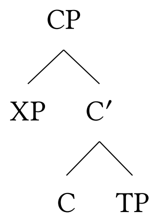


Figure 11: CP structures in Minimalism

of the C that heads it.

An important unbounded dependency construction is relative clauses. In English there are *wh*-relatives and non-*wh*-relatives and finite and non-finite relatives. *Wh*-relatives are illustrated by the following:

(26) a. someone [who you can rely on]
  b. someone [on whom you can rely]

(27) a. * someone [who to rely on]
  b. someone [on whom to rely]

These show that whereas finite *wh*-relatives allow either an NP or a PP as the filler, non-finite *wh*-relatives only allow a PP. In the HPSG analysis of Sag (1997), the facts are a consequence of constraints on two phrase types. A constraint on the type *fin-wh-fill-rel-cl* allows the first daughter to be an NP or a PP while a constraint on *inf-wh-fill-rel-cl* requires the first daughter to be a PP. For Minimalism, the facts must be attributed to the properties of the complementizer. There must be a complementizer which takes a finite TP complement and allows either an NP or a PP as its specifier and another complementizer which takes a non-finite TP complement (with an unexpressed subject) and only allows a PP as its specifier.

Non-*wh*-relatives require further phrase types within HPSG and further complementizers in Minimalism. However, rather than consider this, we will look at another unbounded dependency construction: *wh*-interrogatives. The basic data that needs to be accounted for is illustrated by the following:

(28) a. Who knows?
b. I wonder [who knows].
c. Who did Kim talk to?
d. I wonder [who Kim talked to].
e. I wonder [who to talk to].

Like *wh*-relatives, *wh*-interrogatives can be finite and non-finite. When they are finite their form depends on whether the *wh*-phrase is subject of the highest verb or something else. When it is subject of the highest verb, it is followed by what looks like a VP although it may be a clause with a gap in subject position. When the *wh*-phrase is something else, the following clause shows auxiliary-initial order if it is a main clause and subject-initial order if it is not. Non-finite *wh*-interrogatives are a simple matter, especially as the filler does not have to be restricted in the way that it does in non-finite *wh*-relatives. Ginzburg & Sag (2000) present an analysis which has two types for finite *wh*-interrogatives, one for subject-*wh*-interrogatives such as those in (28a) and (28b), and another for non-subject-*wh*-interrogatives such as those in (28c) and (28d). The latter is subject to a constraint requiring it to have the same value for the features IC (INDEPENDENT-CLAUSE) and INV (INVERTED). Main clauses are [IC +] and auxiliary-initial clauses are [INV +]. Hence the constraint ensures that a non-subject-*wh*-interrogative shows auxiliary-initial order just in case it is a main clause.

How can the facts be handled within Minimalism? As noted above, Minimalism analyses auxiliary-initial order as a result of movement of the auxiliary to C. It is triggered by some feature of C. Thus C must have this feature just in case (a) it heads a main clause and (b) the *wh*-phrase in its specifier position is not the subject of the highest verb. There are no doubt various ways in which this might be achieved, but the key point is the properties of a phonologically empty complementizer are crucial.

Borsley (2006b; 2017) discusses Minimalist analyses of relative clauses and *wh*-interrogatives and suggests that at least eight complementizers are necessary. One is optionally realized as *that*, and another is obligatorily realized as *for*. The other six are always phonologically empty. But it has been clear since Ross (1967a) and Chomsky (1977) that relative clauses and *wh*-interrogatives are not the only unbounded dependency constructions. Here are some others:

(29) a. What a fool he is! (*wh*-exclamative clause)
b. The bagels, I like. (topicalized clause)
c. Kim is more intelligent [than Lee is]. (comparative-clause)
d. Kim is hard [to talk to]. (*tough*-complement-clause)
e. Lee is too important [to talk to]. (*too*-complement-clause)
f. [The more people I met], [the happier I became]. (*the*-clauses)

Each of these constructions will require at least one empty complementizer. Thus, a comprehensive account of unbounded dependency constructions will require a large number of such elements. But with a large unstructured set of complementizers there can be no distinction between properties shared by some or all elements and properties restricted to a single element. There are a variety of shared properties. Many of the complementizers will take a finite complement, many others will take a non-finite complement, and some will take both. There will also be complementizers which take the same set of specifiers. Most will not attract an auxiliary, but some will, not only the compementizer in an example like (28c) but also the complementizers in the following, where the auxiliary is in italics:

(30) a. Only in Colchester *could* such a thing happen.
b. Kim is in Colchester, and so *is* Lee.
c. Such *is* life.
d. The more Bill smokes, the more *does* Susan hate him.

Thus, there are generalizations to be captured here. The obvious way to capture them is with the approach developed in the 1980s in HPSG work on the hierarchical lexicon (Flickinger, Pollard & Wasow 1985; Flickinger 1987), i.e. a detailed classification of complementizers which allows properties to be associated not just with individual complementizers but also with classes of complementizers. With this it should be possible for Minimalism not just to get the facts right but to capture the full set of generalizations. In many ways such an analysis would be mimicking the HPSG approach with its hierarchy of phrase types.[25] But in the present context the main point is the simplicity of the Minimalist grammatical system is another factor which leads to more complex syntactic structures than those of HPSG.

[25] For a fuller discussion of the issues see Borsley (2006b; 2017).

4.1.6 Syntactification of semantic categories

The left periphery of the clause is often much more complex than assumed in the last section as a result of the syntactification of semantic properties (Rizzi 2014), which is one aspect of the idea that semantics is a simple reflection of syntax. This is especially apparent in a sub-school that calls itself "cartographic". MGG comes with strong claims about the autonomy of syntax. There is a syntactic component and than there are the components of Phonological Form (PF) and Logical Form (LF), in more recent versions of the theory this is the articulatory-perceptual system (AP) and the conceptual-intentional system (CI). Figure 8 shows the early Minimalist architecture and the architecture assumed in the Phase-based models. Syntax was always regarded primary and PF and LF derived from syntactic representations. This is similar in Minimalism. The problem is that questions of intonation are connected to semantic and information structural properties (Halliday 1970: 36). A way around this is to stipulate syntactic features that can be interpreted by both PF and LF (Gussenhoven 1983). Another way of dealing with the data is to employ empty elements that are responsible for certain ordering of elements and that can be interpreted in the semantics. The accounts of Rizzi and Cinque are very prominent in this school of thought. For example, Rizzi (1997) suggests an analysis of the left periphery of clauses that incorporate special functional projections for topic and focus. His analysis is shown in Figure 12. In comparison no such projections exist in HPSG theories. HPSG grammars are surface oriented and the syntactic labels correspond for the most part to classical part of speech categorizations. So in examples with frontings like (31) the whole object is a verbal projection and not a Topic phrase, a Focus Phrase or a Force phrase.

(31) Bagels, I like.

Of course the fronted elements may be topics or foci but this is a property that is represented independently of syntactic information in parts of feature descriptions having to do with information structure. For treatment of information structure in HPSG see Engdahl & Vallduví (1996), De Kuthy (2000) and also Kuthy (2019), Chapter D of this volume. On determination of clause types see Ginzburg & Sag (2000) and Müller (2016b). For general discussion of the representation of information usually assigned to different linguistic "modules" and on "interfaces" between them in theories like LFG and HPSG see Kuhn (2007).

Cartographic approaches also assume a hierarchy of functional projections for the placement of adverbials. Some authors assume that all sentences in all languages have the same structure, which is supposed to explain orders of ad-

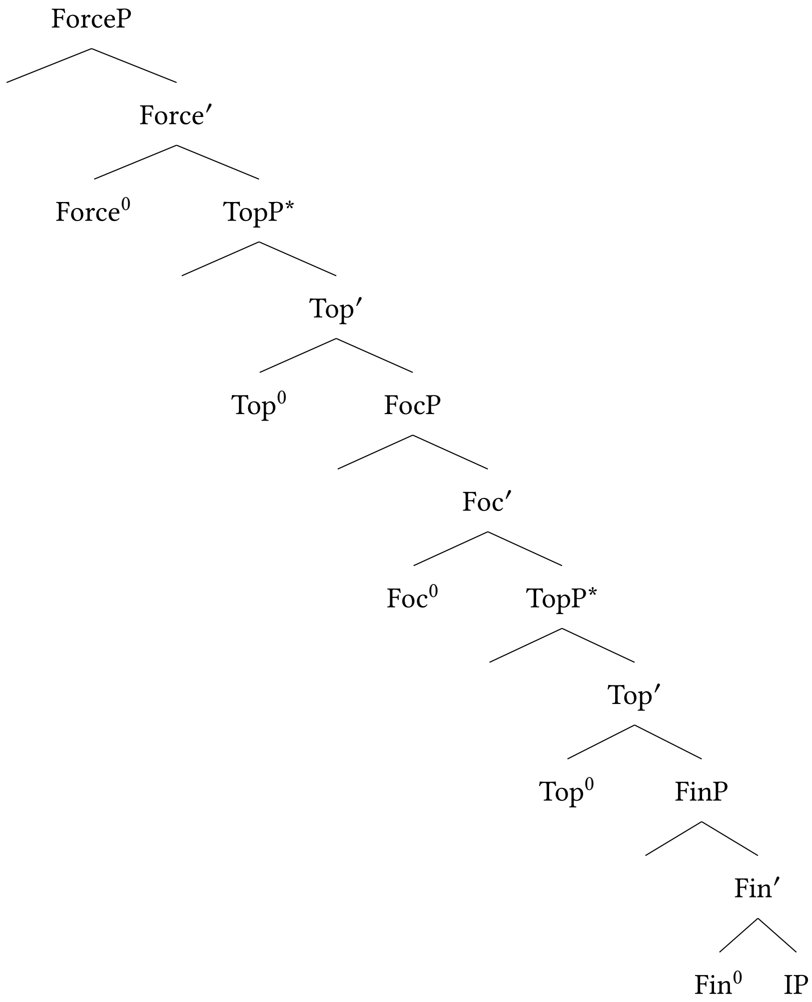


Figure 12: Syntactic structure of sentences following Rizzi (1997: 297)

verbials that seem to hold universally (e.g., Cinque 1999: 106 and Cinque & Rizzi 2010: 54–55). A functional head selects for another functional projection to establish this hierarchy of functional projections and the respective adverbial phrases can be placed in the specifier of the corresponding functional projection. Cinque (1999: 106) assumes 32 functional projections in the verbal domain. Cinque & Rizzi (2010: 57, 65) assume at least four hundred functional heads, which are – according to them – all part of a genetically determined UG.

In comparison, HPSG analyses assume that verbs project: a verb that is combined with an argument is a verbal projection. If an adverb attaches, a verbal projection with the same valence but augmented semantics results. Figure 13 shows the Cartographic and the HPSG structures. While the adverbs (Adv_1 and

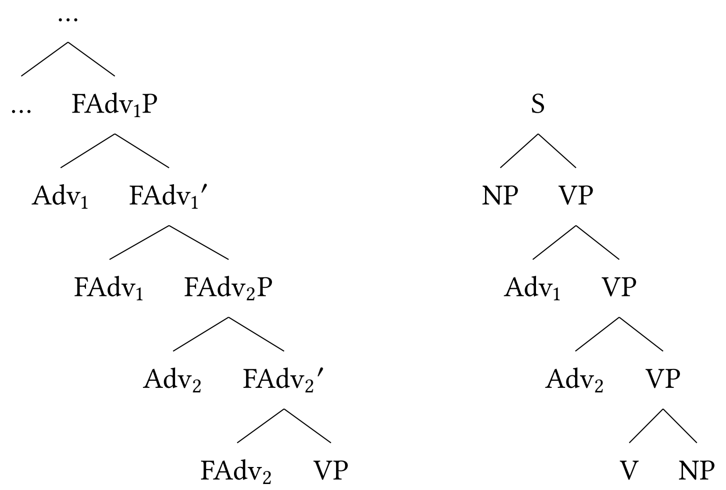


Figure 13: Treatment of adverbial phrases in Cartographic approaches and in HPSG

Adv_2 in the figure) attach to verbal projections in the HPSG analysis (S and VP are abbreviations standing for verbal projections with different valence requirements), the Cartographic approach assumes empty heads that select a clausal projection and provide a specifier position in which the adverbs can be realized. For the sake of exposition we called these heads $FAdv_1$ and $FAdv_2$. For example, $FAdv_2$ can combine with the VP and licences an Adv_2 in its specifier position. As is clear from the figure, the Cartographic approach is more complex since it involves two additional categories ($FAdv_1$ and $FAdv_2$) and nine nodes for the adverbial combination rather than five.

An interesting difference is that verbal properties are projected in the HPSG analysis. By doing this it is clear whether a VP contains an infinitive or a participle.

(32) a. Kim has met Sandy.
b. Kim will meet Sandy.

This property is important for the selection by a superordinate head, e.g., the auxiliary in the examples in (32). In a Cartographic approach one either has to assume that adverbial projections have features correlated with verbal morphology or one has to assume that superordinate heads may check properties of linguistic items that are deeply embedded.

If one believed in Universal Grammar (which researchers working in HPSG usually do not) and in innately specified constraints on adverb order, one would not assume that all languages contain the same structures, some of these structures being invisible. Rather one would assume linearization constraints (see Müller (2019a: Section 2), Chapter 10 of this volume) to hold crosslinguistically.[26] If adverbs of a certain type do not exist in a language, the linearization constraints would not do any harm. They just would never apply since there is nothing to apply to (Müller 2015b: 46).

4.1.7 Summary

Having discussed uniformity in theta role assignment, Generative Semantics-like approaches, branching, nonlocal dependencies and Cartographic approaches to the left periphery and adverb order within clauses, we conclude that a variety of features of Minimalism lead to structures that are much more complex than those of HPSG. HPSG shows that this complexity is unnecessary given a somewhat richer conception of grammar.

4.2 The simplicity of Minimalist structures

As we emphasized above, while minimalist structures are very complex, they are also simple in the sense that they have just a single ingredient, local trees

[26] Adjuncts are usually not siblings in local structures in HPSG (but see Kasper (1994) and Bouma & van Noord (1998: 62, 71)). There are nevertheless ways to impose order constraints on non-siblings. Engelkamp, Erbach & Uszkoreit (1992) discuss one approach, another approach would be to have Reape-style order domains (Reape 1994) in addition to the immediate dominance schemata for head-adjunct combination. See Müller (2019a), Chapter 10 of this volume for order domains.

consisting of a head and a single non-head. Most outsiders agree that this is too simple.

4.2.1 Binary branching, VPs, and verb-initial clauses

We look first at binary branching.[27] As we noted above, the assumption that all branching is binary is incompatible with various analyses which have been widely accepted in HPSG and other frameworks. For example, it means that the bracketed VP in (33), which contains two complements, cannot have the ternary branching structure in Figure 14, which is suggested in Pollard & Sag (1994: 36) and much other work.

(33) Kim [gave a book to Lee].

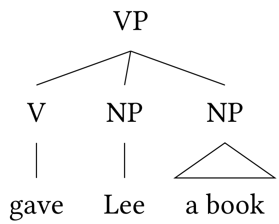


Figure 14: Flat structure for the VP *gave Lee a book*

Instead it has been assumed since Larson (1988) that the VP in examples like (33) has something like the structure in Figure 15. It is assumed that the verb originates in the lower VP and is moved into the higher VP. The higher V position to which the verb moves is commonly labelled *v* ("little *v*") and the higher phrase *v*P. The main argument for such an analysis appears to involve anaphora, especially contrasts like the following:

[27] In addition to structures with two or more branches, HPSG uses unary branching structures both in syntax and in the lexicon (lexical rules basically are unary branching structures, see Meurers (2001) and Davis & Koenig (2019: Section 5), Chapter 4 of this volume. For example, unary branching syntactic rules are used for semantic type shifting (Partee 1987). For respective HPSG analyses see Flickinger (2008: 91–92) and Müller (2009; 2012). The lack of unary branching structures in Minimalism is no problem since empty heads can be used instead. The empty head projects the properties that would be otherwise assigned to the mother node of the unary projection. See for example Ramchand (2005: 370). So, while the effects of unary projections can be modeled, the resulting structures are more complex. For a general discussion of empty elements and unary projections and lexical rules see Müller (2016a: Sections 19.2 and 19.5).

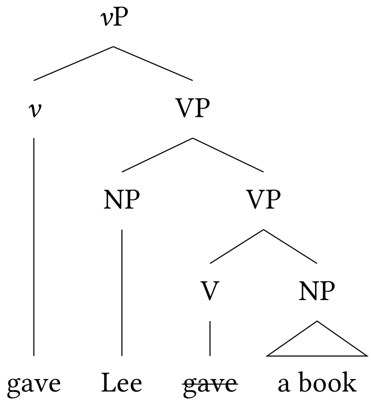


Figure 15: Larson-type analysis of VPs involving little *v*

(34) a. John showed Mary herself in the picture.
b. * John showed herself Mary in the picture.

The first complement can be the antecedent of a reflexive which is the second complement, but the reverse is not possible.

If constraints on anaphora refer to constituent structure as suggested by Chomsky (1981), the contrast suggests that the second NP should be lower in the structure than the first NP. But, as suggested by Pollard & Sag (1992), it is assumed in HPSG that constraints on anaphora refer not to constituent structure but to a list containing all arguments in order of obliquensss, in recent versions of HPSG the ARG-ST list (see also Branco (2019), Chapter 21 of this volume). On this view, anaphora can provide no argument for the complex structure in Figure 15. Therefore, both flat structures and binary branching structures with different branching directions as in Figure 16 are a viable option in HPSG. Müller (2015a: Section 2.4; 2019b) argues for such binary branching structures as a result of parametrizing the Head-Complement Schema for various variants of constituent order (head-initial and head-final languages with fixed constituent order and languages like German and Japanese with freer constituent order).

The fact that Merge combines two expressions also means that the auxiliary-initial clause in (35) cannot have a flat structure with both subjects and complement(s) as sisters of the verb, as in Figure 17.

(35) Will Kim be here?

It is standardly assumed in Minimalism that it has a structure of the form in Figure 18 or more complicated structures, as explained in Section 4.1.6. *Will* is

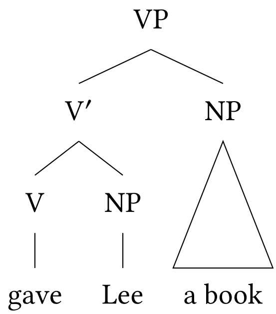


Figure 16: Possible analysis of VPs in HPSG with a branching direction differing from Larson-type structures

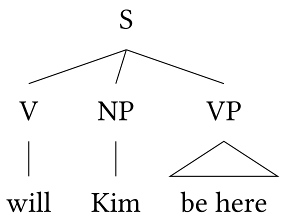


Figure 17: Flat structure for *Will Kim be there?*

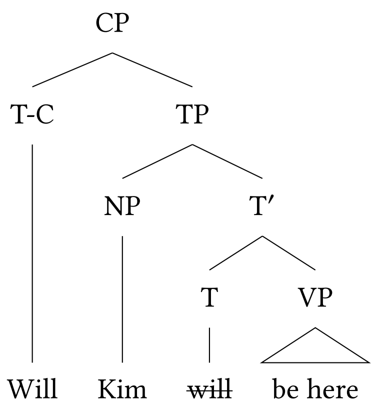


Figure 18: CP/TP structure for *Will Kim be there?*

analysed as a T(ense) element which moves to the C(omplementizer) position. A binary branching analysis of some kind is the only possibility within Minimalism provided the usual assumptions are made.

It is not just English auxiliary-initial clauses that cannot have a ternary branching analysis within Minimalism but verb-initial clauses in any language. A notable example is Welsh, which has verb-initial order in all types of finite clause. Here are some relevant examples:[28]

(36) a. Mi/Fe gerddith Emrys i 'r dre. (Welsh)
PRT walk.FUT.3SG Emrys to the town
'Emrys will walk to the town.'

b. Dywedodd Megan [cerddith Emrys i 'r dre].
say.PAST.3SG Megan walk.FUT.3SG Emrys to the town
'Megan said Emrys will walk to the town.'

A variety of transformational work, including work in Minimalism, has argued for an analysis like Figure 18 for Welsh finite clauses (see e.g., Jones & Thomas 1977, Sproat 1985, Sadler 1988, Rouveret 1994, and Roberts 2005). But Borsley (2006a) argues that there is no theory-neutral evidence for a structure of this kind. Hence, at least for Welsh, it seems that a simpler flat structure like Figure 17 is preferable.[29] Note, that we do not argue that structures like the one in Figure 18 are not appropriate for any language. The analog to head-movement analyses is standard among HPSG grammarians of German and there is data from apparent multiple frontings that seems to make a head-movement analysis unavoidable. See Müller (2017) for a book-length discussion of German clause structure. Müller (2019a: Section 4.1), Chapter 10 of this volume also discusses head-movement in HPSG.

4.2.2 Headedness and coordination

We turn now to the idea that all structures are headed. For HPSG, and many other approaches, there are headed structures and non-headed structure. Probably the most important example of the latter are coordinate structures such as those in (37) (see Sag 2003 and Abeillé & Chaves (2019), Chapter 17 of this volume for HPSG analyses).

[28] Positive main clause verbs are optionally preceded by a particle (*mi* or *fe*). We have included this in (36a) but not in (36b). When it appears it triggers so-called soft mutation. Hence (36a) has *gerddith* rather than the basic form *cerddith*, which is seen in (36b).

[29] Borsley (2016) argues for a similar flat structure for the Caucasian ergative SOV language Archi.

(37) [Kim and Lee] [wrote poems and painted pictures].

Much work in Minimalism assumes that coordinate structures are headed by the conjunction (Larson 1990: 596; Radford 1993: 89; Kayne 1994: Chapter 6; Johannessen 1998; Van Koppen 2005: 8; Bošković 2009: 474; Citko 2011: 27).[30] This suggests that both coordinate structures in (37) are conjunction phrases. This is highly problematic since the category of the phrases plays a role in accounting for their external distribution. So the VPs *wrote poems* and *painted pictures* have to be combined with a DP/NP to form a complete sentence. But according to the ConjP theory *Kim and Lee* is not a DP or NP it is a ConjP and hence incompatible with any requirements. Similarly, a T head in the analysis of (37) requires a VP argument but instead of a VP *wrote poems and painted pictures* there is only a ConjP.[31] It is fairly clear that conjunctions cannot be ordinary heads. Johannessen (1996: 669) suggests an analysis in which a coordinate structure has the features of the first conjunct. She depicts the analysis as in Figure 19. The prob-

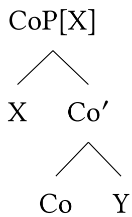


Figure 19: Analysis of coordination with projection of features from the first conjunct according to Johannessen (1996: 669)

lem is that it is unclear how this should be formalized: either the head category

[30] Kayne (1994: 57) differs from other proposals in not assuming the category Conj for the conjunction. Instead, he uses X^0 as the category in his structured examples. Since X is an underspecified variable his theory is underdetermined: while a ConjP is not compatible with any requirement by a governing head, an XP could appear as an argument of any dominating head. Kayne needs to work out a theory that determines the properties of the projected XP in relation to the coordinated items. We discuss this below.

[31] If one considers the part of speech labels only, one would expect the two ConjPs to be interchangeable, but of course they are not:

(i) * [Wrote poems and painted pictures] [Kim and Lee].

Of course the two ConjPs are not exactly of the same category since there may be further features that distinguish the two ConjPs. But how these features are distributed between the conjuncts and the mother is not worked out.

For a more detailed critique of the ConjP approach see Borsley (2005).

of the complete object is ConjP or it is X. Governing heads have to know where to look for the category. If they look at X, why is the part of speech information of Co projected? Why would governing heads not look at the category of other specifiers rather than their heads? Furthermore, coordinations are not equivalent to the first conjunct. There are cases where the coordination is a sum of the parts. For example, *Kim and Sandy* is a plural NP, as the agreement with the verb shows:

(38) Kim and Sandy laugh.

Johannessen's analysis seems to predict that the coordination of *Kim* and *Sandy* behaves like *Kim*, which is not the case. So, if one wants to assume an analysis with the conjunction as a head, one would have to assume that the head is a functor taking into account the properties of its specifier and complement, and projecting nominal information if they are nominal, verbal if they are verbal, etc (Steedman 1991). This would make them a unique type of a head with a unique relation to their specifier and complement. A problem for this approach is coordinate structures in which the conjuncts belong to different categories, e.g., the following:

(39) a. Hobbs is [a linguist and proud of it].
     b. Hobbs is [angry and in pain].

Such examples have led to HPSG analyses in which coordinate structures have whatever properties are common to the two conjuncts (Sag 2003). Within Minimalism, one might try to mimic such analyses by proposing that conjunctions have whatever properties are common to their specifier and complement. But a problem arises with an example like (40), where the conjuncts are not phrases but words.

(40) Kim [criticized and insulted] his boss.

To accommodate such examples, conjunctions would have to acquire not only part of speech information from the conjuncts but also selectional information. They would be heads which combine with a specifier and a complement to form an expression which, like a typical head, combines with a specifier and a complement. This would be a very strange situation and in fact it would make wrong predictions since the object *his boss* would be the third-merged item. It would hence be "later-merged" in the sense of Chomsky (2008: 146) and therefore treated as a specifier rather than a complement.[32]

[32] There have been attempts to argue that conjuncts are always phrases (Kayne 1994, Bruening 2018). But this position seems untenable (Abeillé 2006, Müller 2018b: Section 7).

4.2.3 Binary branching and headless structures: The NPN construction

Another problem for Minimalist theories is the NPN Construction discussed by Matsuyama (2004) and Jackendoff (2008). Examples are provided in (41):

(41) a. Student after student left the room.

b. Day after day after day went by, but I never found the courage to talk to her. (Bargmann 2015)

As Jackendoff argued it is not possible to identify one of the elements in the construction as the head. The construction has several peculiar properties and we share Jackendoff's view that these constructions are best treated by a phrasal configuration in which these highly idiosyncratic properties are handled. The construction is discussed in more detail in Müller (2019d), Chapter P of this volume and Bargmann's analysis within HPSG is provided. Bargmann's analysis also captures multiple repetitions of the PN sequence as they occur in (41b). Until now there is one proposal for NPN in the Minimalist framework: G. Müller (2011). G. Müller develops a reduplication account. He states that reduplication applies to words only and claims that German differs from English in not allowing adjective noun sequences in NPN constructions. He is aware of the possibility of these constructions in English (*miserable day after miserable day*) and states that his analysis is intended to account for the German data only. While this alone is a serious shortcoming of the analysis, the empirical claim does not hold water either as the following example from Müller (2019d), Chapter P of this volume shows:

(42)

Die beiden	tauchten	nämlich	geradewegs	wieder	aus	dem heimischen
the two	surfaced	namely	straightaway	again	from	the home

Legoland	auf,	wo	sie	im	Wohnzimmer,	schwarzen	Stein	um
Legoland	PART	where	they	in.the	living.room	black	brick	after

schwarzen	Stein,	vermeintliche	Schusswaffen	nachgebaut	hatten.[33]
black	brick	alledged	firearms	recreated	had

'The two surfaced straightaway from their home Legoland where they had recreated alledged firearms black brick after black brick.'

Apart from failing on the reduplication of adjective-noun combinations like *schwarzen Stein* 'black brick', the reduplication approach also fails on NPN patterns with several PN repetitions as in (41b): if the preposition is responsible for reduplicating content it is unclear how the first *after* is supposed to combine

[33] taz, 05.09.2018, p. 20, quoted from Müller (2019d: 1113).

with *day* and *day after day*. It is probably possible to design analyses of the NPN construction involving several empty heads but it is clear that these solutions would come at a rather high price.

4.2.4 Movement for more local phenomena like scrambling, passive and raising

We want now to consider the dependencies that Minimalism analyzes in terms of Move/Internal Merge. In the next section we look at unbounded dependencies, but first we consider local dependencies in passives, unaccusatives, raising sentences, and scrambling. The following illustrate the first three of these:

(43) a. Kim has been hit.
 b. Kim has disappeared.
 c. Kim seems to be clever.

These differ from unbounded dependency constructions in that whereas the gaps in the latter are positions in which overt NPs can appear, this is not true of the supposed gap positions in (43):

(44) a. * It has been hit Kim.
 b. * It has disappeared Kim.
 c. * It seems Kim to be clever.

This is a complication if they involve the same mechanism, but is unsurprising if they involve different mechanisms, as in HPSG and most other frameworks.

4.2.4.1 Passive

In the classical analysis of the passive in MGG, it is assumed that the morphology of the participle suppresses the agent role and removes the ability to assign accusative case. In order to receive case the underlying object has to move to the subject position, i.e. Spec,TP where it gets nominative (Chomsky 1981: 124).

(45) a. The mother gave [the girl] [a cookie].
 b. [The girl] was given [a cookie] (by the mother).

The analysis assumed in recent Minimalist work differs in detail but is movement-based like its predecessors. While movement-based approaches seem to work well for SVO languages like English, they are problematic for SOV languages like German. To see why consider the examples in (46):

(46) a. weil das Mädchen dem Jungen den Ball schenkte
because the.NOM girl the.DAT boy the.ACC ball gave
'because the girl gave the ball to the boy'

b. weil dem Jungen der Ball geschenkt wurde
because the.DAT boy the.NOM ball given was
'because the ball was given to the boy'

c. weil der Ball dem Jungen geschenkt wurde
because the.NOM ball the.DAT boy given was

In comparison to (46c), (46b) is the unmarked order (Höhle 1982). *der Ball* 'the ball' in (46b) occurs in the same position as *den Ball* in (46a), that is, no movement is necessary. Only the case differs. (46c) is, however, somewhat marked in comparison to (46b). So, if one assumed (46c) to be the normal order for passives and (46b) is derived from this by movement of *dem Jungen* 'the boy', (46b) should be more marked than (46c), contrary to the facts. To solve this problem, an analysis involving abstract movement has been proposed for cases such as (46b): the elements stay in their positions, but are connected to the subject position and receive their case information from there. Grewendorf (1993: 1311) assumes that there is an empty expletive pronoun in the subject position of sentences such as (46b) as well as in the subject position of sentences with an impersonal passive such as (47):[34]

(47) weil heute nicht gearbeitet wird
because today not worked is
'because there will be no work done today'

A silent expletive pronoun is something that one cannot see or hear and that does not carry any meaning. Such entities are not learnable from input and hence innate domain specific knowledge would be required and of course, approaches that do not have to assume very specific innate knowledge are preferable. For further discussion of language acquisition see Section 5.2 and Ginzburg (2019), Chapter F of this volume.

HPSG does not have this problem since passive is treated by lexical rules that map verbal stems onto participle forms with a reduced argument structure list. The first element (the subject in the active) is suppressed so that the second element (if there is any) becomes first. In SVO languages like English and Ice-

[34] See Koster (1986: 11–12) for a parallel analysis for Dutch as well as Lohnstein (2014) for a movement-based account of the passive that also involves an empty expletive for the analysis of the impersonal passive.

landic this element is realized before the verb: there is a valence feature for subjects/specifiers and items that are realized with the respective schema are serialized to the left of the verb. In SOV languages like German and Dutch the subject is treated like other arguments and hence it is not put in a designated position before the finite verb (Müller 2019a: Section 4, Chapter 10 of this volume). No movement is involved in this valence-based analysis of the passive. The problem of MGG analyses is that they mix two phenomena: passive and subject requirement. Since these two phenomena are kept separate in HPSG, problems like that discussed above can be avoided. See Müller (2016a: Section 3.4, Chapter 20) for further discussion.

4.2.4.2 Scrambling

Discussing passive, we already touched on problems related to local reordering of arguments, so-called *scrambling*. In what follows, we want to discuss scrambling in more detail. Languages like German have a freer constituent order than English. A sentence with a ditransitive verb allows for six permutations of the arguments, two of which are given in (48):

(48) a. [weil] der Mann der Frau das Buch gibt
because the.NOM man the.DAT woman the.ACC book gives
'because the man gives the book to the woman'

b. [weil] das Buch der Mann der Frau gibt
because the.ACC book the.NOM man the.DAT woman gives

It was long argued that scrambling should be handled as movement as well (Frey 1993). An argument that has often been used to support the movement-based analysis is the fact that scope ambiguities exist in sentences with reorderings which are not present in sentences in the base order. The explanation of such ambiguities comes from the assumption that the scope of quantifiers can be derived from their position in the superficial structure as well as their position in the underlying structure. If the position in both the surface and deep structure are the same, that is, when there has not been any movement, then there is only one reading possible. If movement has taken place, however, then there are two possible readings (Frey 1993: 185):

(49) a. Es ist nicht der Fall, daß er mindestens einem Verleger fast
it is not the case that he at.least one publisher almost

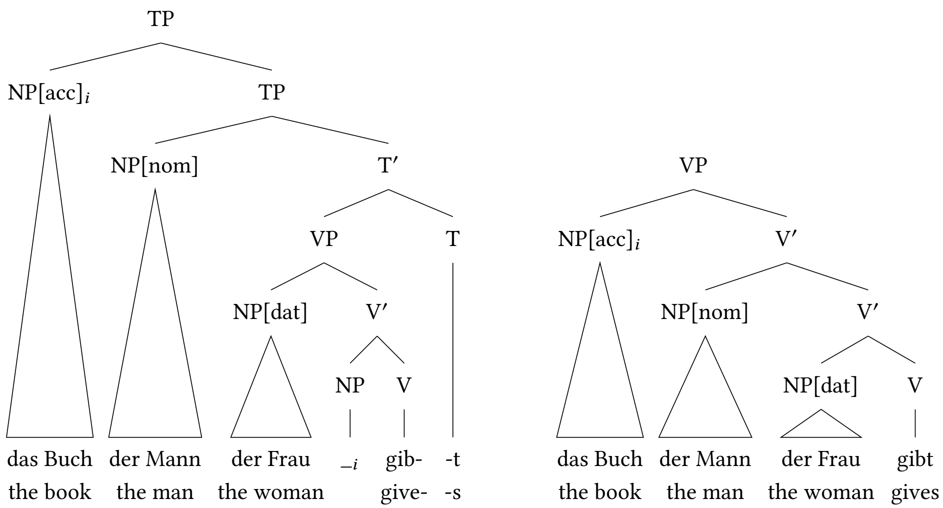


Figure 20: Analysis of local reordering as movement to Spec TP and "base-generation" analysis assumed in HPSG

jedes Gedicht anbot.
every poem offered

'It is not the case that he offered at least one publisher almost every poem.'

b. Es ist nicht der Fall, daß er fast jedes Gedicht$_i$ mindestens einem
it is not the case that he almost every poem at.least one
Verleger $\_{i}$ anbot.
publisher offered

'It is not the case that he offered almost every poem to at least one publisher.'

It turns out that approaches assuming traces run into problems as they predict certain readings which do not exist for sentences with multiple traces (see Kiss 2001: 146 and Fanselow 2001: Section 2.6). For instance in an example such as (50), it should be possible to interpret *mindestens einem Verleger* 'at least one publisher' at the position of $\_{i}$, which would lead to a reading where *fast jedes Gedicht* 'almost every poem' has scope over *mindestens einem Verleger* 'at least one publisher'. However, this reading does not exist.

(50) Ich glaube, dass mindestens einem Verleger$_i$ fast jedes Gedicht$_j$ nur
I believe that at.least one publisher almost every poem only

dieser Dichter $\_i$ $\_j$ angeboten hat.
this poet offered has

'I think that only this poet offered almost every poem to at least one publisher.'

The alternative to movement-based approaches are so-called "base-generation" approaches in which the respective orders are derived directly. Fanselow (2001), working within the Minimalist Program, suggests such an analysis in which arguments can be combined with their heads in any order. This is the HPSG analysis that was suggested by Gunji (1986) for Japanese and is standardly used in HPSG grammars of German (Hinrichs & Nakazawa 1994a; Kiss 1995; Meurers 1999; Müller 2003; 2017). See also Müller (2019a), Chapter 10 of this volume.

Sauerland & Elbourne (2002: 308) discuss analogous examples from Japanese, which they credit to Kazuko Yatsushiro. They develop an analysis where the first step is to move the accusative object in front of the subject. Then, the dative object is placed in front of that and then, in a third movement, the accusative is moved once more. The last movement can take place to construct either a structure that is later passed to LF or as a movement to construct the Phonological Form. In the latter case, this movement will not have any semantic effects. While this analysis can predict the correct available readings, it does require a number of additional movement operations with intermediate steps.

4.2.5 Nonlocal dependencies

Having dealt with phenomena treated via Move/Internal Merge in Minimalism but involving more local phenomena, we now turn to genuine nonlocal dependencies and compare the Move/Internal Merge approach to the HPSG approach to nonlocal dependencies.

4.2.5.1 Gaps without filler

The Move/Internal Merge approach seems quite plausible for typical examples of an unbounded dependency, but issues arise with less typical examples. Within this approach one expects to see a clause-initial filler-constituent and a gap somewhere in the following clause. This is what we commonly find, but there are unbounded dependency constructions in which there is a gap but no visible higher constituent matching it. Consider e.g., the following:

(51) a. the book [Kim bought _]
b. Lee is too important [for you to talk to _].

c. Lee is important enough [for you to talk to _].
d. Kim is easy [for anyone to talk to _].

Within Minimalist assumptions, it is more or less necessary to assume that such examples contain an invisible filler (a so-called empty operator). Unless there is some independent evidence for such invisible fillers, they are little more than an ad hoc device to maintain the Move/Internal Merge approach. Within the HPSG SLASH-based approach to unbounded dependencies, there is no assumption that there should always be a filler at the top of an unbounded dependency (Pollard & Sag 1994: Chapter 4, see also Borsley & Crysmann (2019), Chapter 14 of this volume). Hence, the examples in (51) are completely unproblematic.

4.2.5.2 Filler without gaps: Resumptive pronouns

There are also unbounded dependency constructions which seem to have not a gap but a resumptive pronoun (RP). Among many languages that are relevant here is Welsh, which has RPs in both *wh*-interrogatives and relative clauses, as the following illustrate:

(52) a. Pa ddyn werthodd Ieuan y ceffyl iddo *fo*?
which man sell.PAST.3SG Ieuan the horse to.3SGM he
'Which man did Ieuan sell the horse to?'

b. y dyn werthodd Ieuan y ceffyl iddo *fo*
the man sell.PAST.3SG Ieuan the horse to.3SGM he
'the man that Ieuan sold the horse to'

Willis (2011) and Borsley (2010; 2013) present evidence that Welsh RPs involve the same mechanism as gaps. Within Minimalism, this means that they must involve Move/Internal Merge. But one expects to see a gap where Move/Internal Merge has applied. One Minimalist response suggests that instead of being deleted, the copy left behind by Move/Internal Merge is somehow turned into a pronoun (see McCloskey 2006). A problem for this approach is that it makes it surprising that RPs universally look like ordinary pronouns (McCloskey 2002). Another approach exploits the complexity of Minimalist structures and proposes that there is a gap in the structure somewhere near the RP. Thus, for example, Willis (2011) proposes that examples like those in (52) with an RP in prepositional object position have a coindexed operator in the specifier position of PP, which undergoes movement. Similar approaches are outlined in Aoun et al. (2001) and Boeckx (2003). For detailed objections to both approaches, see Borsley (2013: Section 3). Within the SLASH-based approach of HPSG, there is no reason to think that there will always be a gap at the bottom of a dependency, and it is not difficult to accommodate RPs. See Vaillette (2001), Taghvaipour (2010), Borsley (2013) and Crysmann (2012; 2016) for slightly different approaches.[35] See also Borsley & Crysmann (2019), Chapter 14 of this volume for a more detailed discussion of nonlocal dependencies.

4.3 Conclusion

Thus, there are variety of phenomena which suggest that the Minimalist view of constituent structure is too simple. The restriction to binary branching, the assumption that all structures are headed, and Move/Internal Merge all seem problematic. It looks, then, as if the Minimalist view is both too complex and too simple.

5 Psycholinguistic issues

Although they differ in a variety of ways, HPSG and Minimalism agree that grammatical theory is concerned with linguistic knowledge. They focus first and foremost on the question: what form does linguistic knowledge take? But there are other questions that arise here, notably the following:

- How is linguistic knowledge put to use?
- How is linguistic knowledge acquired?

Both questions are central concerns for psycholinguistics. Thus, in considering the answers that HPSG and Minimalism can give we are considering their relevance to psycholinguistics. Chomskyan approaches, including Minimalism, have focused mainly on the second question and have paid little attention to the first. HPSG has had more to say about the first and has shown less interest in the second. However, there is a large body of work on acquisition in Construction

[35] Also relevant here are examples with more than one gap such as the following:

(i) a. Who does Kim like _ and Lee hate _?
b. Which book did you criticize _ without reading _?

There have been various attempts to accommodate such examples within the Move/Internal Merge approach, but it is not clear that any of them is satisfactory. In contrast such examples are expect within the SLASH-based approach (Levine & Sag 2003). See also (Pollard & Sag 1994: Section 4.6).

Grammar and since HPSG is a constructionist theory (Müller 2019d, Chapter P of this volume) all the insights carry over to HPSG. Clearly an adequate grammatical theory should be able to give satisfactory answers to both questions. In this section we will look briefly at the relation of the two theories to processing and then consider more fully their relation to acquisition.

5.1 Processing

We noted in Section 3 that whereas HPSG is a declarative or constraint-based approach to grammar, Minimalism has a procedural view of grammar. This contrast means that HPSG is much more suitable than Minimalism for incorporation into an account of the processes that are involved in linguistic performance.

The most obvious fact about linguistic performance is that it involves both production and comprehension. As noted in Section 3, this suggests that the knowledge that is used in production and comprehension should have a declarative character as in HPSG and not a procedural character as in Minimalism.

A second important feature of linguistic performance is that it involves different kinds of information utilized in any order that is necessary. Sag & Wasow (2011: 367–368) illustrate with the following examples:

(53) a. The sheep that was sleeping in the pen stood up.
  b. The sheep in the pen had been sleeping and were about to wake up.

In (53a), morphological information determines the number of sheep before non-linguistic information determines that pen means 'fenced enclosure' and not 'writing implement'. In (53b), on the other hand, non-linguistic information determines that pen means 'fenced enclosure' before morphological information determines the number of sheep. This is unproblematic for an approach like HPSG in which linguistic and non-linguistic knowledge takes the form of constraints which are not ordered in any way.[36] It is quite unclear how the facts can be accommodated within Minimalism given that linguistic knowledge with its procedural form is quite different from non-linguistic knowledge.

Other features of HPSG also make it attractive from a processing point of view. Firstly, there is the fact emphasized earlier that linguistic expressions have a single relatively simple constituent structure with a minimum of phonologically empty elements. Secondly there is the fact that all constraints are purely local and never affect anything larger than a local tree consisting of an expression and its

[36]See also Lücking (2019), Chapter K of this volume of this volume on the interaction of gesture and speech.

daughters. Both these properties make processing easier than it would otherwise be. Minimalism has neither property and hence again seems less satisfactory than HPSG in this area.

Someone might suppose that the fact that Minimalism treats linguistic knowledge as knowledge about how to construct syntactic structures means that it is well-suited for incorporation into accounts of linguistic performance. In fact this is not at all the case. The way standard Minimalism[37] constructs syntactic structures is quite unlike the way speakers and hearers construct them. Speakers begin with representations of meanings they want to communicate and gradually turn them into an appropriate sequence of sounds, constructing whatever syntactic structures are necessary to do this. Hearers in contrast begin with a sequence of sounds from which they attempt to work out what meanings are being communicated. To do this, they have to segment the sounds into words and determine what sorts of syntactic structures the words are involved in. Language processing is incremental and all channels are used in parallel (Marslen-Wilson 1975; Tanenhaus et al. 1995; 1996). Information about phonology, morpho-syntax, semantics, information structure and even world knowledge (as in the examples (53) above) are used as soon as they are available. Hence, parsing (54) is an incremental process: the hearer hears *Kim* first and as soon as the first sounds of *may* reach her the available information is integrated and hypothesis regarding further parts of the utterance are built.[38]

(54) Kim may go to London.

The construction of syntactic structures within Minimalism is a very different matter. It begins with a set of words, and they are gradually assembled into a syntactic structure, from which representations of sound and meaning can be derived either once a complete structure has been constructed or at the end of each phase if the derivation is broken up into phases. Moreover, the nature of English means that the construction of a syntactic structure essentially proceeds from right to left. Consider the analysis of (54): here, *go* can only be integrated into the structure after its complement to London has been constructed, and may

[37] For a discussion of non-standard versions like Phillips (2003) and Chesi (2015) see Sag & Wasow (2011) and Müller (2019c: 525).

[38] Note that the architecture in Figure 8 poses additional problems. A *numeration* is a selection of lexical items that is used in a derivation. Since a multitude of empty elements are assumed in Minimalist analyses it is unclear how such a numeration is constructed since it cannot be predicted at the lexical level which empty elements will be needed in the course of a derivation. Due to the empty elements, there are infinitely many possible numerations that might be appropriate for the analysis of a given input string. The problem would disappear if derivations could draw lexical items from the lexicon directly.

can only be integrated into the structure after the construction of its complement *go to London*, and only after that can *Kim* be integrated into the structure. This is quite different from construction of syntactic structures by speakers and hearers, which proceeds from left to right.

These issues have led researchers like Phillips (2003) and Chesi (2015) to propose rather different versions of Minimalism. However, they are still procedural approaches, and they have the problem that any system of procedures which resembles what speakers do will be very different from what hearers do and vice versa. The right response to the problems outlined above is not a different procedural version of Minimalism but a declarative version, neutral between production and comprehension. It would probably not be difficult to develop a declarative version of the framework. It would presumably have an external merge phrase type and an internal merge phrase type both subject to appropriate constraints. This would be better from a processing point of view than any procedural version of Minimalism. However, the complexity of its structures and the fact that its constraints are not purely local would still make it less satisfactory than HPSG in this area.

5.2 Acquisition

Acquisition has long been a central concern for Chomskyans and it has long been argued that acquisition is made possible by the existence of a complex innate language faculty (Chomsky 1965: Section I.8). Since the early 1980s the dominant view has been that the language faculty consists of a set of principles responsible for the properties which languages share and a set of parameters responsible for the ways in which they may differ (Chomsky 1981: 6). On this view acquiring a grammatical system is a matter of parameter-setting (Chomsky 2000: 8). Proponents of HPSG have always been sceptical about these ideas (see e.g., the remarks about parameters in Pollard & Sag (1994: 31) and have favoured accounts with "an extremely minimal initial ontology of abstract linguistic elements and relations" (Green 2011: 378). Thus, the two frameworks appear to be very different in this area. It is not clear, however, that this is really the case.

The idea that acquiring a grammatical system is a matter of parameter-setting is only as plausible as the idea of a language faculty with a set of parameters. It seems fair to say that this idea has not been as successful as was hoped when it was first introduced in the early 1980s. Outsiders have always been sceptical, but they have been joined in recent times by researchers sympathetic to many Chomskyan ideas. Thus, Newmeyer (2005: 75) writes as follows:

> […] empirical reality, as I see it, dictates that the hopeful vision of UG as providing a small number of principles each admitting of a small number of parameter settings is simply not workable. The variation that one finds among grammars is far too complex for such a vision to be realized.

At least some Minimalists have come to similar conclusions. Thus, Boeckx (2011: 206) suggests that:

> some of the most deeply-embedded tenets of the Principles-and-Parameters approach, and in particular the idea of Parameter, have outlived their usefulness. (Boeckx 2011: 206)

Much the same view is expressed in Hornstein (2009: 164–168).

A major reason for scepticism about parameters is that estimates of how many there are seem to have steadily increased. Fodor (2001: 734) considers that there might be just twenty parameters, so that acquiring a grammatical system is a matter of answering twenty yes/no-questions. Newmeyer (2005: 44) remarks that "I have never seen any estimate of the number of binary-valued parameters needed to capture all of the possibilities of core grammar that exceeded a few dozen". However, Roberts & Holmberg (2005) comment that "[n]early all estimates of the number of parameters in the literature judge the correct figure to be in the region of 50–100". Clearly, a hundred is a lot more than twenty. Newmeyer (2017: Section 6.3) speaks of "hundreds, if not thousands". This is worrying. As Newmeyer (2006: 6) observes, "it is an ABC of scientific investigation that if a theory is on the right track, then its overall complexity decreases with time as more and more problematic data fall within its scope. Just the opposite has happened with parametric theory. Year after year more new parameters are proposed, with no compensatory decrease in the number of previously proposed ones.".

The growing scepticism appears to tie in with the proposal by Hauser, Chomsky & Fitch (2002: 1573) that "FLN [the "Faculty of language–narrow sense"] comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces". On this view there seems to be no place for parameters within FLN. This conclusion is also suggested by Chomsky's remarks (2005) that "There is no longer a conceptual barrier to the hope that the UG [Universal Grammar] might be reduced to a much simpler form" (p. 8) and that "we need no longer assume that the means of generation of structured expressions are highly articulated and specific to language" (p. 9). It's hard to see how such remarks are compatible with the assumption that UG includes 50–100 parameters. But if parameters are not part of UG, it is not at all clear what their status might be.

It looks, then, as Chomskyans are gradually abandoning the idea of parameters. But if it is abandoned, grammar acquisition is not a matter of parameter-setting. Hence, it is not clear that Chomskyans can invoke any mechanisms that are not available to HPSG.

This might suggest that HPSG and Minimalism are essentially in the same boat where acquisition is concerned. However, this is not the case given the very different nature of grammatical systems in the two frameworks. The complex and abstract structures that are the hallmark of Minimalism and earlier Chomskyan frameworks pose major problems for acquisition. Furthermore the machinery that is assumed in addition to the basic operations Internal and External Merge are by no means trivial. There are numerations (subsets of the lexicon) that are assumed to play a role in a derivation, there is Agree and it has to be somehow acquired what the restriction on possible probe/goal relations is and which features are interpretable and which uninterpretable. Certain categories are Phase boundaries, others are not. There are complex conditions on Labeling. It is this that has led to the assumption that acquisition must be assisted by a complex language faculty. In contrast, HPSG structures are quite closely related to the observable data and so pose less of a problem for acquisition and hence create less need for some innate apparatus. Thus, HPSG probably has an advantage over Minimalism in this area too.

5.3 Restrictiveness

There is one further issue that we should discuss here. It appears to be quite widely assumed that one advantage that Minimalism has over alternatives like HPSG is that it is more "restrictive", in other words that it makes more claims about what is and is not possible in language. It looks then as if there might be an argument for Minimalism here. It is not clear, however, that this is really the case.

Minimalism would be a restrictive theory making interesting claims about language if it assumed a relatively small number of parameters. However, the idea that there is just a small number of parameters seems to have been abandoned, and at least some minimalists have abandoned the idea of parameters altogether (see Section 5.2). If there is either a large number of parameters or no parameters at all, Minimalism is not restrictive in the way that it once was. However, it does still embody some restrictions on grammatical systems. The assumption that syntactic structures are confined to binary branching is an important restriction, as is the assumption that expressions produced by Merge have the same label as one of the expressions that they consist of. But we have argued that both assumptions are quite dubious. It also seems to be assumed that case and agreement are features of all grammatical systems. This would be another important restriction, but this also seems dubious given that many languages show no clear evidence for one or both of these features. It looks to us, then, as if the restrictiveness of Minimalism is largely a matter of imposing certain dubious restrictions on grammatical systems.

Note also that there are problems with restrictiveness of a more formal nature. Earlier versions of MGG assumed $\overline{\mathrm{X}}$ theory and although this was not assumed initially it was quickly argued that the $\overline{\mathrm{X}}$ scheme is universal and that this is a restriction on grammatical systems that aids language acquisition (Haegeman 1994: 106). However, Kornai & Pullum (1990: 41, 47) show that $\overline{\mathrm{X}}$ theory is not restrictive at all as soon as empty elements are allowed in grammars: all languages that can be analyzed with a context-free grammar can be analyzed with an $\overline{\mathrm{X}}$ grammar with empty heads. Chomsky (1995b: Section 4.3) abandoned $\overline{\mathrm{X}}$ theory and replaced it by notions like first-merged and later-merged (Chomsky 1995b: 245; 2008) but the principled problem remains. Since as many empty heads as needed can be assumed in any position the predictions as far as restrictiveness is concerned are limited. See also Hornstein (2009: 165) and Starke (2014: 140) on heads and features and restrictiveness.

An example that is usually discussed when it comes to restrictiveness is question formation (Musso et al. 2003). Researchers in MGG state that certain ways of expressing things never occur although they may be imaginable. So the question is asked why questions are never formed by reversing the order of words in a string. So rather than (55b) the question that would correspond to (55a) would be (55c):

(55) a. Kim saw Sandy near the swimming pool.
 b. Did Kim see Sandy near the swimming pool?
 c. Pool swimming the near Sandy saw Kim?

Interestingly, such reorderings can be derived in systems that allow for so-called *remnant movement*, as Hubert Haider (p. c. 2018) pointed out. Remnant movement analyses are sometimes suggested for partial verb phrase fronting (G. Müller 1998). In the analysis of the following sentence, the object of *gelesen* 'read' is moved out of the VP and the VP remnant is then fronted:

(56) [$_{\mathrm{VP}}$ _$_j$ Gelesen]$_i$ hat [das Buch]$_j$ [keiner _$_i$].
 read has the book nobody
 'Nobody read the book.'

With such a system in place the reorderings can be derived as follows: the element 5 can move to the left of 4. The unit containing 4 and 5 can move to the left of 3 and [[5 [4 _]] [3 _]] can move to the left of 2 and so on.

(57) a. [1 [2 [3 4]]]

b. [3 4] → [4 [3 _]] → [2 [4 [3 _]]] → [[4 [3 _]] [2 _]] → [1 [[4 [3 _]] [2 _]]] → [[[4 [3 _]] [2 _]] [1 _]]

Of course there are reasons for the absence of certain imaginable constructions in the languages of the world. The reason for the absence of question formation like (55c) is simply short-term memory. Operations like those are ruled out due to performance constraints and hence should not be modeled in competence grammars. So it is unproblematic that remnant movement systems allow the derivation of strings with reverse order and it is unproblematic that one might develop HPSG analyses that reverse strings. Similarly certain other restrictions have been argued not to be part of the grammar proper. For instance Subjacency (Baltin 1981: 262; 2017; Rizzi 1982: 57; Chomsky 1986: 38–40) does not hold in the form stated in MGG (Müller 2004; 2016a: Section 13.1.5) and it is argued that several of the island constraints should not be modeled by hard constraints in competence grammars. See Chaves (2019), Chapter 16 of this volume for further discussion.

It is true that the basic formalism does not pose any strong restrictions on what could be said in an HPSG theory. As Pollard (1996) points out, this is the way it should be. The formalism should not be the constraining factor. It should be powerful enough to allow everything to be expressed in insightful ways and in fact, the basic formalism of HPSG has Turing-power, the highest power in the Chomsky hierarchy (Pollard 1999). This means that the general formalism is above the complexity that is usually assumed for natural languages: mildly context-sensitive. What is important though is that theories of individual languages are much more restrictive, getting the generative power down (Müller 2016a: Chapter 17).

These remarks should not be understood as a suggestion that languages vary without limit, as Joos (1958: 96) suggested. No doubt there are universal tendencies and variation is limited, but the question is whether this is due to innate linguistic constraints or a consequence of what we do with language and how our general cognitive capabilities are structured. While Minimalism starts out with claims about universal features about languages and tries to confirm these claims in language after language, researchers working in HPSG aim to develop fragments of languages that are motivated by facts from these languages and generalize over several internally motivated grammars. This leaves the option open that languages have very little in common as far as syntax is concerned. For example, Koenig & Michelson (2012) discuss the Northern Iroquoian language Oneida and argue that this language does not have syntactic valence. If they are correct, not even central concepts like valence and argument structure would be universal. The only remaining universal would be that we combine linguistic objects. This corresponds to Merge in Minimalism, without the restriction to binarity.

Acknowledgements

We know that this paper is very long. We nevertheless think that it is important to discuss the issues discussed in this paper and are very grateful to the editors for giving us this much space.

We thank Tom Wasow for discussion and comments on an earlier version of the paper. We thank David Adger and Andreas Pankau for discussion and Sebastian Nordhoff and Felix Kopecky for help with Figure 3.

We thank Shalom Lappin and Richard Sproat for discussion of implementation issues.

References

Abeillé, Anne. 2006. In defense of lexical coordination. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in formal syntax and semantics*, vol. 6, 7–36. Paris: CNRS. http://www.cssp.cnrs.fr/eiss6/, accessed 2018-2-25.

Abeillé, Anne & Rui P. Chaves. 2019. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 469–516. Berlin: Language Science Press. DOI:??

Abeillé, Anne, Berthold Crysmann & Aoi Shiraısh. 2016. Syntactic mismatch in french peripheral ellipsis. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, 1–30. Paris: CSSP.

An, Aixiu & Anne Abeillé. 2017. Agreement and interpretation of binominals in french. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky, Lexington*, 26–43. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-an-abeille.pdf.

Aoun, Joseph E. & Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar*. Cambridge, MA: MIT Press.

Aoun, Joseph, Lina Choueiri & Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32(3). 371–404.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function change*. Chicago: The University of Chicago Press.

Baltin, Mark. 1981. Strict bounding. In Carl Lee Baker & John J. McCarthy (eds.), *The logical problem of language acquisition*, 257–295. Cambridge, MA: MIT Press.

Baltin, Mark. 2017. Extraposition. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 2nd edn. (Blackwell Handbooks in Linguistics), 237–271. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9781118358733

Bargmann, Sascha. 2015. *Syntactically flexible VP-idioms and the N-after-N Construction*. Poster presentation at the 5th General Meeting of PARSEME, Iasi, 23–24 September 2015.

Bausewein, Karin. 1990. Haben kopflose Relativsätze tatsächlich keine Köpfe? In Gisbert Fanselow & Sascha W. Felix (eds.), *Strukturen und Merkmale syntaktischer Kategorien* (Studien zur deutschen Grammatik 39), 144–158. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford University dissertation. http://faculty.washington.edu/ebender/dissertation/, accessed 2018-2-25.

Bender, Emily M. 2002. Review of *Step by step: Essays on Minimalist Syntax in honor of Howard Lasnik* by Roger Martin, David Michaels, and Juan Uriagereka. *Journal of Linguistics* 38(2). 432–439.

Bender, Emily M. & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 797–840. Berlin: Language Science Press. DOI:??

Berwick, Robert C. & Samuel David Epstein. 1995. On the convergence of 'Minimalist' Syntax and Categorial Grammar. In Anton Nijholt, Giuseppe Scollo & Rene Steetskamp (eds.), *Algebraic methods in language processing*, 143–148. Enschede: University of Twente. http://eprints.eemcs.utwente.nl/9555/01/twlt10.pdf, accessed 2018-9-25.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the Stimulus revisited. *Cognitive Science* 35(7). 1207–1242.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topic-hood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79.

Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.
Boeckx, Cedric. 2003. *Islands and chains*. Amsterdam: John Benjamins Publishing Co.
Boeckx, Cedric. 2011. Approaching parameters from below. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 205–221. Oxford: Oxford University Press.
Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??
Borsley, Robert D. 2005. Against ConjP. *Lingua* 115(4). 461–482.
Borsley, Robert D. 2006a. On the nature of Welsh VSO clauses. *Lingua* 116(4). 462–490.
Borsley, Robert D. 2006b. *Syntactic and lexical approaches to unbounded dependencies*. Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.
Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language and Linguistic Theory* 27(2). 225–265.
Borsley, Robert D. 2010. An HPSG approach to Welsh unbounded dependencies. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 47–67. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.
Borsley, Robert D. 2011. Constructions, functional heads, and comparative correlatives. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics*, vol. 8, 7–20. Paris: CNRS. http://www.cssp.cnrs.fr/eiss8/borsley-eiss8.pdf, accessed 2018-8-23.
Borsley, Robert D. 2013. On the nature of Welsh unbounded dependencies. *Lingua* 133. 1–29. DOI:10.1016/j.lingua.2013.03.005
Borsley, Robert D. 2016. HPSG and the nature of agreement in Archi. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of agreement in cross-theoretical perspective*, 118–149. Oxford: Oxford University Press.
Borsley, Robert D. 2017. Explanations and 'engineering solutions'? Aspects of the relation between Minimalism and HPSG. In Stefan Müller (ed.), *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar,*

University of Kentucky, Lexington, 82–102. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2017/hpsg2017-borsley.pdf, accessed 2018-9-9.

Borsley, Robert D. 2018. Comparative syntax: An HPSG perspective. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Clausal architecture and its consequences: Synchronic and diachronic perspectives*, vol. 1. In Preparation.

Bošković, Željko. 2009. Unifying first and last conjunct agreement. *Natural Language and Linguistic Theory* 27(3). 455–496. DOI:10.1007/s11049-009-9072-6

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press.

Branco, António. 2019. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 613–665. Berlin: Language Science Press. DOI:??

Bresnan, Joan & Jane Grimshaw. 1978. The syntax of free relatives in English. *Linguistic Inquiry* 9. 331–392.

Bresnan, Joan & Ronald M. Kaplan. 1982. Introduction: grammars as mental representations of language. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), xvii–lii. Cambridge, MA/London: MIT Press.

Bruening, Benjamin. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language* 94(1). 1–42. DOI:10.1353/lan.2018.0000

Chaves, Rui. 2019. Island phenomena and related matters. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 415–468. Berlin: Language Science Press. DOI:??

Chesi, Cristiano. 2015. On directionality of phrase structure building. *Journal of Psycholinguistic Research* 44(1). 65–89. DOI:10.1007/s10936-014-9330-6

Chomsky, Noam. 1957. *Syntactic structures* (Janua Linguarum / Series Minor 4). The Hague/Paris: Mouton.

Chomsky, Noam. 1964. Degrees of grammaticalness. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The structure of language*, 384–389. Englewood Cliffs, NJ: Prentice-Hall.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press.

Chomsky, Noam. 1977. On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 1986. *Barriers* (Linguistic Inquiry Monographs 13). Cambridge, MA: MIT Press.

Chomsky, Noam. 1990. On formalization and formal linguistics. *Natural Language and Linguistic Theory* 8(1). 143–147.

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 1–52. Cambridge, MA/London: MIT Press.

Chomsky, Noam. 1995a. Bare phrase structure. In Hector Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, 51–109. Washington, DC: Georgetown U Press.

Chomsky, Noam. 1995b. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. *New horizons in the study of language and mind*. Cambridge, UK: Cambridge University Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2002. *On nature and language*. Cambridge, UK: Cambridge University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI:10.1016/j.lingua.2012.12.003

Chomsky, Noam & Howard Lasnik. 1995. The theory of principles and parameters. In Noam Chomsky (ed.), *The Minimalist Program*, chap. 1, 13–127. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York, Oxford: Oxford University Press.

Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 51–65. Oxford: Oxford University Press.

Citko, Barbara. 2008. Missing labels. *Lingua* 118(7). 907–944.

Citko, Barbara. 2011. *Symmetry in syntax: Merge, move and labels*. Vol. 129. Cambridge, UK: Cambridge University Press. DOI:10.1017/CBO9780511794278

Cole, Peter, Wayne Harbert & Gabriella Hermon. 1982. Headless relative clauses in Quechua. *International Journal of American Linguistics* 48(2). 113–124. http://www.jstor.org/stable/1264676, accessed 2018-9-13.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Crysmann, Berthold. 2012. Resumption and island-hood in Hausa. In Philippe de Groote & Mark-Jan Nederhof (eds.), *Formal Grammar. 15th and 16th International Conference on Formal Grammar, FG 2010 Copenhagen, Denmark, August 2010, FG 2011 Lubljana, Slovenia, August 2011*, vol. 7395 (Lecture Notes in Computer Science). Berlin: Springer Verlag.

Crysmann, Berthold. 2016. An underspecification approach to Hausa resumption. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 194–214. Stanford, CA: CSLI Publications.

Crysmann, Berthold. 2019. Morphology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 673–674. Berlin: Language Science Press. DOI:??

Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition. foundations of syntax*. Vol. 1 (Oxford Linguistics). Oxford: Oxford University Press.

Culicover, Peter W. & Ray Jackendoff. 1999. The view from the periphery: The English comparative correlative. *Linguistic Inquiry* 30(4). 543–571. DOI:10.1162/002438999554200

Culicover, Peter W. & Ray S. Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.

Davis, Anthony & Jean-Pierre Koenig. 2019. The nature and role of the lexicon in HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 47–92. Berlin: Language Science Press. DOI:??

De Kuthy, Kordula. 2000. *Discontinuous NPs in German — A case study of the interaction of syntax, semantics and pragmatics*. Saarbrücken: Universität des Saarlandes dissertation.

Donati, Caterina. 2006. On *wh*-head-movement. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), Wh*-movement: Moving on* (Current Studies in Linguistics 42), 21–46. Cambridge, MA: MIT Press.

Engdahl, Elisabet & Enric Vallduví. 1996. Information packaging in HPSG. In Claire Grover & Enric Vallduví (eds.), *Edinburgh Working Papers in Cognitive Science, vol. 12: Studies in HPSG*, chap. 1, 1–32. Edinburgh: Centre for Cognitive Science, University of Edinburgh. ftp://ftp.cogsci.ed.ac.uk/pub/CCS-WPs/wp-12.ps.gz, accessed 2018-3-20.

Engelkamp, Judith, Gregor Erbach & Hans Uszkoreit. 1992. Handling linear precedence constraints by unification. In Henry S. Thomson (ed.), *30th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 201–208. Newark, DE: Association for Computational Linguistics. auch als CLAUS-Report, Nr. 19, Universität des Saarlandes erschienen.

Epstein, Samuel David & T. Daniel Seely. 2006. *Derivations in Minimalism* (Cambridge Studies in Linguistics 111). Cambridge, UK: Cambridge University Press.

Falk, Yehuda N. 2010. An unmediated analysis of relative clauses. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2010 conference*, 207–227. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/15/, accessed 2018-3-20.

Fanselow, Gisbert. 2001. Features, θ-roles, and free constituent order. *Linguistic Inquiry* 32(3). 405–437.

Flickinger, Dan, Carl Pollard & Tom Wasow. 2019. The evolution of HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 7–43. Berlin: Language Science Press. DOI:??

Flickinger, Daniel P. 1987. *Lexical rules in the hierarchical lexicon*. Stanford University dissertation.

Flickinger, Daniel P. 2008. Transparent heads. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 87–94. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/abstr-flickinger.shtml, accessed 2018-2-25.

Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL.

Fodor, Janet Dean. 2001. Setting syntactic parameters. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 730–767. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9780470756416.ch23

Fong, Sandiway. 2014. Unification and efficient computation in the Minimalist Program. In L. Francis & L. Laurent (eds.), *Language and recursion*, 129–138. Berlin: Springer Verlag.

Fong, Sandiway & Jason Ginsburg. 2012. Computation with doubling constituents: Pronouns and antecedents in Phase Theory. In Anna Maria Di Sciullo (ed.), *Towards a Biolinguistic understanding of grammar: Essays on interfaces* (Linguistik Aktuell/Linguistics Today 194), 303–338. Amsterdam: John Benjamins Publishing Co.

Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus* (studia grammatica 35). Berlin: Akademie Verlag.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Ginzburg, Jonathan. 2019. Acquisition. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 773–774. Berlin: Language Science Press. DOI:??

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Green, Georgia M. 2011. Modeling grammar growth: Universal Grammar without innate principles or parameters. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 378–403. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd. DOI:10.1002/9781444395037.ch11

Grewendorf, Günther. 1993. German: A grammatical sketch. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax – Ein internationales Handbuch zeitgenössischer Forschung*, vol. 9.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1288–1319. Berlin: Walter de Gruyter Verlag. DOI:10.1515/9783110095869.1

Groos, Anneke & Henk van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar. In A. Belletti, L. Brandi & L. Rizzi (eds.), *Theory of markedness in Generative Grammar*, 171–216. Pisa: Scuola Normale Superiore.

Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, 1–21. Stanford, CA: CSLI Publications.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19. 377–419.

Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory*. 2nd edn. (Blackwell Textbooks in Linguistics 1). Oxford: Blackwell Publishers Ltd.

Haider, Hubert. 2018. *On Minimalist theorizing and scientific ideology in grammar theory*. Ms. Universität Salzburg.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 53–109. Cambridge, MA/London: MIT Press.

Halliday, Michael A. K. 1970. *A course in spoken English: intonation*. Oxford University Press.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579. DOI:10.1126/science.298.5598.1569

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994a. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–38. Stanford, CA: CSLI Publications.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994b. Partial-VP and split-NP topicalization in German—An HPSG analysis. In Erhard W. Hinrichs, Walt Detmar Meurers & Tsuneko Nakazawa (eds.), *Partial-VP and split-NP topicalization in German—An HPSG analysis and its implementation* (Arbeitspapiere des SFB 340 No. 58). Eberhard-Karls-Universität Tübingen.

Höhle, Tilman N. 1982. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 75–153. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag. Republished as Höhle (2018).

Höhle, Tilman N. 2018. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 107–191. Berlin: Language Science Press. DOI:10.5281/zenodo.1145680

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge University Press. DOI:10.1017/CBO9780511575129

Hornstein, Norbert, Jairo Nunes & Kleantes K. Grohmann. 2005. *Understanding Minimalism* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press. DOI:10.1017/CBO9780511840678

Huybregts, Riny & Henk van Riemsdijk. 1982. *Noam chomsky on the Generative Enterprise*. Dordrecht: Foris.

Jackendoff, Ray S. 2008. Construction after Construction and its theoretical challenges. *Language* 84(1). 8–28.

Johannessen, Janne Bondi. 1996. Partial agreement and coordination. *Linguistic Inquiry* 27(4). 661–676.

Johannessen, Janne Bondi. 1998. *Coordination* (Oxford Studies in Comparative Syntax). Oxford: Oxford University Press.

Jones, Morris & Alun R. Thomas. 1977. *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press.

Joos, Martin. 1958. Comment on *Phonemic overlapping* by Bernhard Bloch. In Martin Joos (ed.), *Readings in linguistics: The development of descriptive linguistics in America since 1925*. New York: American Council of Learned Societies. https://archive.org/details/readingsinlingui00joos.

Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 39–70. Stanford, CA: CSLI Publications.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.

Kiss, Tibor. 2001. Configurational and relational scope determination in German. In Walt Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 141–175. Stanford, CA: CSLI Publications.

Kiss, Tibor. 2005. Semantic constraints on relative clause extraposition. *Natural Language and Linguistic Theory* 23(2). 281–334. DOI:10.1007/s11049-003-1838-7

Kobele, Gregory M. 2008. Across-the-board extraction in Minimalist Grammars. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+9)*, 113–128.

Koenig, Jean-Pierre & Karin Michelson. 2012. The (non)universality of syntactic selection and functional application. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, vol. 9, 185–205. Paris: CNRS.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Kornai, András & Geoffrey K. Pullum. 1990. The X-bar Theory of phrase structure. *Language* 66(1). 24–50.

Koster, Jan. 1986. The relation between pro-drop, scrambling, and verb movements. *Groningen Papers in Theoretical and Applied Linguistics* 1. 1–43.

Kubota, Yusuke. 2019. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1013–1069. Berlin: Language Science Press. DOI:??

Kuhn, Jonas. 2007. Interfaces in constraint-based theories of grammar. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 613–650. Oxford: Oxford University Press. DOI:10.1093/oxfordhb/9780199247455.013.0020

Kuthy, Kordula De. 2019. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 709–743. Berlin: Language Science Press. DOI:??

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391.

Larson, Richard K. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* (21). 589–632.

Levine, Robert D. & Ivan A. Sag. 2003. Some empirical issues in the grammar of extraction. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 236–256. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-9-25.

Li, Yen-hui Audrey. 2008. Case, 20 years later. In Marjorie K.M. Chan & Hana Kang (eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL-20)*, 41–68. Columbus, Ohio: The Ohio State University.

Lin, Dekang. 1993. Principle-based parsing without overgeneration. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, 112–120. Columbus, Ohio, USA: Association for Computational Linguistics. DOI:10.3115/981574.981590

Lin, Dekang. 2003. Dependency-based evaluation of MINIPAR. In Anne Abeillé (ed.), *Treebanks: Building and Using Parsed Corpora* (Text, Speech and Language Technology 20), 317–329. Kluwer Academic Publishers. DOI:10.1007/978-94-010-0201-1_18

Lohnstein, Horst. 2014. Artenvielfalt in freier Wildbahn: Generative Grammatik. In Jörg Hagemann & Sven Staffeldt (eds.), *Syntaxtheorien: Analysen im Vergleich* (Stauffenburg Einführungen 28), 165–185. Tübingen: Stauffenburg Verlag.

Lücking, Andy. 2019. Gesture. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 887–934. Berlin: Language Science Press. DOI:??

Marslen-Wilson, William D. 1975. Sentence perception as an interactive parallel process. *Science* 189(4198). 226–228. DOI:10.1126/science.189.4198.226

Matsuyama, Tetsuya. 2004. The N *after* N Construction: A constructional idiom. *English Linguistics* 1(11). 55–84.

May, Robert. 1985. *Logical form: Its structure and derivation* (Linguistic Inquiry Monographs 12). Cambridge, MA/London: MIT Press.

McCawley, James D. 1968. Lexical insertion in a Transformational Grammar without Deep Structure. In *Fourth regional meeting of the Chicago Linguistic Society*, 71–81. University of Chicago.

McCloskey, James. 2002. Resumption, successive cyclicity and the locality of operations. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the Minimalist Program*, 184–226. Oxford: Blackwell Publishers Ltd. DOI:10.1002/9780470755662

McCloskey, James. 2006. Resumption. In Martin Everaert & Henk van Riemsdijk (eds.), *The blackwell companion to syntax*, 94–117. Oxford: Blackwell Publishers Ltd.

Meinunger, André. 2000. *Syntactic aspects of topic and comment* (Linguistik Aktuell/Linguistics Today 38). Amsterdam: John Benjamins Publishing Co.

Meurers, Walt Detmar. 1999. German partial-VP fronting revisited. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 129–144. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217.

Meurers, Walt Detmar & Stefan Müller. 2009. Corpora and syntax. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, vol. 29 (Handbücher zur Sprach- und Kommunikationswissenschaft), chap. 42, 920–933. Berlin: Mouton de Gruyter.

Michaelis, Laura A. & Knud Lambrecht. 1996. Toward a Construction-based theory of language function: The case of nominal extraposition. *Language* 72(2). 215–247.

Müller, Gereon. 1998. *Incomplete category fronting: A derivational approach to remnant movement in German* (Studies in Natural Language and Linguistic Theory 42). Dordrecht: Kluwer Academic Publishers.

Müller, Gereon. 2011. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. In Stefan Engelberg, Anke Holler & Kristel Proost (eds.), *Sprachliches Wissen zwischen Lexikon und Grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010), 211–249. Berlin: de Gruyter.

Müller, Stefan. 1999a. An HPSG-analysis for free relative clauses in German. *Grammars* 2(1). 53–105.

Müller, Stefan. 1999b. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2003. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI:10.1017/S0022226703002032

Müller, Stefan. 2004. Complex NPs, subjacency, and extraposition. *Snippets* 8. 10–11.

Müller, Stefan. 2007a. Qualitative Korpusanalyse für die Grammatiktheorie: Introspektion vs. Korpus. In Gisela Zifonun & Werner Kallmeyer (eds.), *Sprachkorpora – Datenmengen und Erkenntnisfortschritt* (Institut für Deutsche Sprache Jahrbuch 2006), 70–90. Berlin: Walter de Gruyter.

Müller, Stefan. 2007b. The Grammix CD Rom: A software collection for developing typed feature structure grammars. In Tracy Holloway King & Emily M. Bender (eds.), *Grammar Engineering across Frameworks 2007* (Studies in Computational Linguistics ONLINE). Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/GEAF/2007/, accessed 2018-2-25.

Müller, Stefan. 2009. On predication. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 213–233. Stanford, CA: CSLI Publications.

Müller, Stefan. 2012. *On the copula, specificational constructions and type shifting*. Ms. Freie Universität Berlin.

Müller, Stefan. 2013. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950. DOI:10.1353/lan.2013.0061

Müller, Stefan. 2014. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford, CA: CSLI Publications.

Müller, Stefan. 2015a. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook* (Handbooks of Linguistics and Communication Science 42.2), 937–973. Berlin: Walter de Gruyter. DOI:10.1515/9783110363708-004

Müller, Stefan. 2015b. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.91

Müller, Stefan. 2016a. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2016b. Satztypen: Lexikalisch oder/und phrasal. In Rita Finkbeiner & Jörg Meibauer (eds.), *Satztypen und Konstruktionen im Deutschen* (Linguistik – Impulse & Tendenzen 65), 72–105. Berlin: de Gruyter.

Müller, Stefan. 2017. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Revise and resubmit.

Müller, Stefan. 2018a. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 2nd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.5281/zenodo.1193241

Müller, Stefan. 2018b. The end of lexicalism as we know it? *Language* 94(1). e54–e66. DOI:10.1353/lan.2018.0014

Müller, Stefan. 2019a. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

Müller, Stefan. 2019b. *Germanic syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin. https://hpsg.hu-berlin.de/~stefan/Pub/germanic.html, accessed 2019-6-30.

Müller, Stefan. 2019c. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 3rd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.5281/zenodo.3364215

Müller, Stefan. 2019d. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

Müller, Stefan, Felix Bildhauer & Philippa Cook. 2012. Beschränkungen für die scheinbar mehrfache Vorfeldbesetzung im Deutschen. In Colette Cortès (ed.), *Satzeröffnung: Formen, Funktionen, Strategien* (Eurogermanistik 31), 113–128. Tübingen: Stauffenburg Verlag.

Musso, Mariacristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel & Cornelius Weiller. 2003. Broca's area and the language instinct. *Nature Neuroscience* 6(7). 774–781. DOI:10.1038/nn1077

Newmeyer, Frederick J. 1986. *Linguistic theory in America*. 2nd edn. New York, NY: Academic Press.

Newmeyer, Frederick J. 2005. *Possible and probable languages: A Generative perspective on linguistic typology*. Oxford: Oxford University Press.

Newmeyer, Frederick J. 2006. *A rejoinder to 'On the role of parameters in Universal Grammar: A reply to Newmeyer' by Ian Roberts and Anders Holmberg*. https://ling.auf.net/lingbuzz/000248, accessed 2018-9-25.

Newmeyer, Frederick J. 2017. Where, if anywhere, are parameters? A critical historical overview of parametric theory. In Claire Bowern, Laurence Horn & Raffaella Zanuttini (eds.), *On looking into words (and beyond): Structures, relations, analyses* (Empirically Oriented Theoretical Morphology and Syntax 3), 547–569. Language Science Press. DOI:10.5281/zenodo.495467

Ott, Dennis. 2011. A note on free relative clauses in the theory of Phases. *Linguistic Inquiry* 42(1). 183–192.

Partee, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen A. G. Groenendijk, Dick de Jongh & Martin J. B. Stokhof (eds.), *Studies in Discourse Representation Theory and the theory of generalized quantifiers* (Groningen-Amsterdam Studies in Semantics 8), 115–143. Dordrecht: Foris Publications.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34(1). 37–90. DOI:10.1162/002438903763255922

Pollard, Carl. 1997. *Lectures on the foundations of hpsg*. Ms. Ohio State University. https://www1.essex.ac.uk/linguistics/external/clmt/papers/hpsg/pollard-foundations.ps, accessed 2018-9-29.

Pollard, Carl J. 1996. *The nature of constraint-based grammar*. http://lingo.stanford.edu/sag/L221a/pollard-96.txt, accessed 2018-2-25. Paper presented at the Pacific Asia Conference on Language, Information, and Computation, Kyung Hee University, Seoul, Korea.

Pollard, Carl J. 1999. Strong generative capacity in HPSG. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects* *of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 281–298. Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Postal, Paul M. 1970. On the surface verb 'remind'. *Linguistic Inquiry* 1(1). 37–120.

Postal, Paul M. 2003. (Virtually) conceptually necessary. *Journal of Linguistics* 39(3). 599–620. DOI:10.1017/S0022226703002111

Postal, Paul M. 2004. *Skeptical linguistic essays*. Oxford: Oxford University Press.

Postal, Paul M. 2010. *Edge-based clausal syntax: A study of (mostly) English object structure*. Cambridge, MA: MIT Press.

Pullum, Geoffrey K. 1989. Formal linguistics meets the Boojum. *Natural Language and Linguistic Theory* 7(1). 137–143. DOI:10.1007/BF00141350

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between generative-enumerative and model-theoretic syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retor (eds.), *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Computer Science 2099), 17–43. Berlin: Springer Verlag.

Radford, Andrew. 1993. Head-hunting: On the trail of the nominal Janus. In Greville G. Corbett, Norman M. Fraser & Scott McGlashan (eds.), *Heads in grammatical theory*, 73–113. Cambridge University Press. DOI:10.1017/CBO9780511659454.005

Radford, Andrew. 1997. *Syntactic theory and the structure of English: A Minimalist approach* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.

Ramchand, Gillian. 2005. Post-Davidsonianism. *Theoretical Linguistics* 31(3). 359–373. DOI:10.1515/thli.2005.31.3.359

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Richards, Marc. 2015. Minimalism. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42.3 (Handbooks of

Linguistics and Communication Science), 803–839. Berlin: Mouton de Gruyter. DOI:10.1515/9783110363685

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 45–46. Berlin: Language Science Press. DOI:??

Richter, Frank & Manfred Sailer. 1999. Underspecified semantics in HPSG. In Harry Bunt & Reinhard Muskens (eds.), *Computing meaning*, vol. 1, 95–112. Dordrecht: Springer Netherlands. DOI:10.1007/978-94-011-4231-1_5

Rizzi, Luigi. 1982. Violations of the *wh* island constraint and the Subjacency Condition. In Luigi Rizzi (ed.), *Issues in Italian syntax* (Studies in Generative Grammar 11), 49–76. Dordrecht: Foris Publications.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.

Rizzi, Luigi. 2014. Syntactic Cartography and the Syntacticisation of Scope-Discourse Semantics. In Anne Reboul (ed.), *Mind, Values, and Metaphysics: Philosophical Essays in Honor of Kevin Mulligan*, vol. 2, 517–533. Cham: Springer International Publishing. DOI:10.1007/978-3-319-05146-8_30

Roberts, Ian. 2005. *Principles and Parameters in a VSO language: A case study in Welsh*. Oxford: Oxford University Press.

Roberts, Ian F. & Anders Holmberg. 2005. On the role of parameters in Universal Grammar: A reply to Newmeyer. In Hans Broekhuis, N. Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk* (Studies in Generative Grammar 86), 538–553. Berlin: Mouton de Gruyter. DOI:10.1515/9783110892994.538

Ross, John R. 1967a. *Constraints on variables in syntax*. MIT, Cambridge, Massachusetts. [Published in 1986 as *Infinite Syntax!* Norwood, NJ: Ablex Publishing] Ph.D. Dissertation.

Ross, John Robert. 1967b. *Constraints on variables in syntax*. MIT dissertation. http://files.eric.ed.gov/fulltext/ED016965.pdf, accessed 2018-2-25. Reproduced by the Indiana University Linguistics Club and later published as Ross (1986).

Ross, John Robert. 1986. *Infinite syntax!* (Language and Being 5). Norwood, New Jersey: Ablex Publishing Corporation.

Rouveret, Alain. 1994. *Syntaxe du gallois: principes généraux et typologie*. Paris: Editions du CNRS.

Sadler, Louisa. 1988. *Welsh syntax: a government binding approach*. London: Croom Helm.

Sag, Ivan A. 1983. On parasitic gaps. *Linguistics and Philosophy* 6(1). 35–45. DOI:10.1007/BF00868089

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.

Sag, Ivan A. 2003. Coordination and underspecification. In Jongbok Kim & Stephen Mark Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 267–291. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/3/, accessed 2018-9-25.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.

Sag, Ivan A. & Janet Dean Fodor. 1994. Extraction without traces. In Raul Aranovich, William Byrne, Susanne Preuss & Martha Senturia (eds.), *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, 365–384. Stanford, CA: CSLI Publications/SLA. http://lingo.stanford.edu/sag/papers/sag-fodor-wccfl.pdf, accessed 2018-2-25.

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Sag, Ivan A. & Thomas Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguistic Research* 44(1). 47–63.

Salkoff, Morris. 1988. Analysis by fusion. *Lingvisticæ Investigationes* 12(1). 49–84. DOI:10.1075/li.12.1.03sal

Sauerland, Uli & Paul Elbourne. 2002. Total reconstruction, PF movement, and derivational order. *Linguistic Inquiry* 33(2). 283–319.

Schütze, Carson T. 2016. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. 2nd edn. (Classics in Linguistics 2). Berlin: Language Science Press. DOI:10.17169/langsci.b89.100

Shiraishi, Aoi & Anne Abeillé. 2016. Peripheral ellipsis and verb mismatch. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway-King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 662–680. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2016/headlex2016-shiraishi-abeille.pdf.

Sorace, Antonella & Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115(11). 1497–1524. DOI:10.1016/j.lingua.2004.07.002

Sproat, Richard. 1985. Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3(2). 173–216.

Sproat, Richard & Shalom Lappin. 2005. *A challenge to the Minimalist community.* http://linguistlist.org/issues/16/16-1156.html, accessed 2018-3-21.

Stabler, Edward P. 2001. Minimalist Grammars and recognition. In Christian Rohrer, Antje Rossdeutscher & Hans Kamp (eds.), *Linguistic form and its computation* (Studies in Computational Linguistics 1), 327–352. Stanford, CA: CSLI Publications.

Starke, Michael. 2014. Towards elegant parameters: Language variation reduces to the size of lexically-stored trees. In M. Carme Picallo (ed.), *Linguistic variation in the Minimalist framework*, 140–152. Oxford: Oxford University Press.

Steedman, Mark J. 1991. Structure and intonation. *Language* 67(2). 260–296.

Steedman, Mark J. 2000. *The syntactic process* (Language, Speech, and Communication). Cambridge, MA: MIT Press.

Sternefeld, Wolfgang. 2006. *Syntax: Eine morphologisch motivierte generative Beschreibung des Deutschen* (Stauffenburg Linguistik 31). Tübingen: Stauffenburg Verlag.

Sternefeld, Wolfgang & Frank Richter. 2012. Wo stehen wir in der Grammatiktheorie? — Bemerkungen anläßlich eines Buchs von Stefan Müller. *Zeitschrift für Sprachwissenschaft* 31(2). 263–291. DOI:10.1515/zfs-2012-0010

Taghvaipour, Mehran A. 2010. *Persian relative clauses in HPSG: An Analysis of Persian relative clause constructions in Head-Driven Phrase Structure Grammar.* Saarbrücken: LAMBERT Academic Publishing.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634. DOI:10.1126/science.7777863

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1996. Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In Toshio Inui & James L. McClelland (eds.), *Information integration in perception and communication* (Attention and Performance XVI), 457–478. Cambridge, MA: MIT Press.

Torr, John & Edward P. Stabler. 2016. Coordination in Minimalist Grammars: Excorporation and across the board (head) movement. In.

Torr, John, Milos Stanojevic, Mark Steedman & Shay B. Cohen. 2019. Wide-coverage neural A* parsing for Minimalist Grammars. In Anna Korhonen, David Traum & Lluís Màrquez (eds.), *Proceedings of the 57th annual meeting* *of the association for computational linguistics*, 2486–2505. Florence, Italy: Association for Computational Linguistics. DOI:10.18653/v1/P19-1238

Vaillette, Nathan. 2001. Irish gaps and resumptive pronouns in HPSG. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 Conference, Norwegian University of Science and Technology, Trondheim*, 284–299. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2001/, accessed 2015-12-23.

Van Eynde, Frank. 2015. *Predicative constructions: From the Fregean to a Montagovian treatment* (Studies in Constraint-Based Lexicalism). Stanford, CA: CSLI Publications.

Van Koppen, Marjo. 2005. *One probe–two goals: Aspects of agreement in Dutch dialects*. Vol. 105 (LOT Dissertation Series). Utrecht Institute of Linguistics: Netherlands Graduate School of Linguistics.

Wasow, Tom. 2019. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 775–796. Berlin: Language Science Press. DOI:??

Williams, Edwin. 1989. The anaphoric nature of θ-roles. *Linguistic Inquiry* 20(3). 425–456.

Willis, David. 2011. The limits of resumption in Welsh *wh*-dependencies. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces*, 189–222. Amsterdam: John Benjamins Publishing Co.

Wurmbrand, Susanne. 2003. *Long passive (corpus search results)*. Ms. University of Connecticut.

Chapter M

HPSG and Categorial Grammar

Yusuke Kubota

National Institute for Japanese Language and Linguistics

This chapter aims to offer an up-to-date comparison of HPSG and categorial grammar (CG). Since the CG research itself consists of two major types of approaches with overlapping but distinct goals and research strategies, I start by giving an overview of these two variants of CG (section 2). This is followed by a comparison of HPSG and CG at a broad level, in terms of the general architecture of the theory (section 3), and then, by a more focused comparison of specific linguistic analyses of some selected phenomena (section 4). The chapter ends by briefly touching on issues related to computational implementation and human sentence processing (section 5). Thoughtout the discussion, I attempt to highlight both the similarities and differences between HPSG and CG research, in the hope of stimulating further research in the two research communities on their respective open questions, and so that the two communities can continue to learn from each other.

1 Introduction

The goal of this chapter is to provide a comparison between HPSG and CATEGORIAL GRAMMAR (CG). The two theories share certain important insights, mostly due to the fact that they are among the so-called 'lexicalist', 'non-transformational' theories of syntax that were proposed as major alternatives to the mainstream transformational syntax in the 1980s (see Borsley & Börjars (2011) for an overview of these theories). However, due to the differences in the main research goals in the respective communities in which these approaches have been developed, there are certain nontrivial differences between them as well. The present chapter assumes researchers working in HPSG or other non-CG theories of syntax as its main audience, and aims to inform them of key aspects of CG which

Yusuke Kubota. 2019. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1013–1069. Berlin: Language Science Press. DOI:??

make it distinct from other theories of syntax. While computational implementation and investigations of the formal properties of grammatical theory have been important in both HPSG and CG research, I will primarily focus on the linguistic aspects in the ensuing discussion, with pointers (where relevant) to literature on mathematical and computational issues. Throughout the discussion, I presuppose basic familiarity with HPSG (with pointers to relevant chapters in the handbook). The present handbook contains chapters that compare HPSG with other grammatical theories, including the present one. I encourage the reader to take a look at the other theory comparison chapters too (as well as other chapters dealing with specific aspects of HPSG in greater detail), in order to obtain a fuller picture of the theoretical landscape in current (non-transformational) generative syntax research.

2 Two varieties of CG

CG is actually not a monolithic theory, but is a family of related approaches (or, perhaps more accurately, it is much *less of* a monolithic theory than either HPSG or LFG is). For this reason, I will start my discussion by sketching some important features of two major varieties of CG, Combinatory Categorial Grammar (CCG) (Steedman 2000; Steedman 2012) and Type-Logical Categorial Grammar (TLCG; or 'Type-Logical Grammar') (Morrill 1994; Moortgat 2011; Kubota & Levine 2020).[1] After presenting the 'core' component of CG that is shared between the two approaches—which is commonly referred to as the 'AB grammar'—I introduce aspects of the respective approaches in which they diverge from each other.

2.1 Notation and presentation

Before getting started, some comments are in order as to the notation and the mode of presentation adopted. Two choices are made for the notation. First, CCG and TLCG traditionally adopt different notations of the slash. I stick to the TLCG notation throughout this chapter for notational consistency. Second, I present all the fragments below in the so-called labelled deduction notation of (Prawitz-style) natural deduction. In particular, I follow Oehrle (1994) and Morrill (1994) in the use of 'term labels' in labelled deduction to encode prosodic and semantic information of linguistic expressions. This involves writing linguistic expressions

[1]For more detailed introductions to these different variants of CG, see Steedman & Baldridge (2011) (on CCG) and Oehrle (2011) (on TLCG), both included in Borsley & Börjars (2011).

as *tripartite signs*, formally, tuples of prosodic form, semantic interpretation and syntactic category (or syntactic type). Researchers familiar with HPSG should find this notation easy to read and intuitive; the idea is essentially the same as how linguistic signs are conceived of in HPSG. In the CG literature, this notation has its roots in the conception of 'multidimensional' linguistic signs in earlier work by Dick Oehrle (Oehrle 1988). But the reader should be aware that this is *not* the standard notation in which either CCG or TLCG is typically presented. Also, logically savvy readers may find this notation somewhat confusing since it (unfortunately) obscures certain aspects of CG pertaining to its logical properties. In any event, it is important to keep in mind that different notations co-exist in the CG literature (and the logic literature behind it), and that, just as in mathematics in general, different notations can be adopted for the same formal system to highlight different aspects of it in different contexts. As noted in the introduction, for the mode of presentation, the emphasis is consistently on linguistic (rather than computational or logical) aspects. Moreover, I have taken the liberty to gloss over certain minor differences among different variants of CG for the sake of presentation. The reader is therefore encouraged to consult primary sources as well, especially when details matter.

2.2 The AB grammar

I start with a simple fragment of CG called the AB GRAMMAR, consisting of just two syntactic rules in (1):

(1) a. FORWARD SLASH ELIMINATION

$$\frac{\mathsf{a};\ A/B \qquad \mathsf{b};\ B}{\mathsf{a} \circ \mathsf{b};\ A}\ /\mathrm{E}$$

b. BACKWARD SLASH ELIMINATION

$$\frac{\mathsf{b};\ B \qquad \mathsf{a};\ B\backslash A}{\mathsf{b} \circ \mathsf{a};\ A}\ \backslash\mathrm{E}$$

With the somewhat minimal lexicon in (2), we can license the sentence *John loves Mary* as in (3). The two slashes / and \ are used to form 'complex' syntactic categories (more on this below) indicating valence information: the transitive verb *loves* is assigned the category (NP\S)/NP since it first combines with an NP to its right (i.e. the direct object) and then another NP to its right (i.e. the subject).

(2) a. john; NP

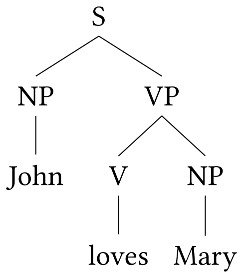
b. mary; NP

c. ran; NP\S

d. loves; (NP\S)/NP

(3)
$$\dfrac{\mathsf{john};\ \mathrm{NP} \qquad \dfrac{\mathsf{mary};\ \mathrm{NP} \quad \mathsf{loves};\ (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP}}{\mathsf{loves} \circ \mathsf{mary};\ \mathrm{NP}\backslash\mathrm{S}}{/\mathrm{E}}}{\mathsf{john} \circ \mathsf{loves} \circ \mathsf{mary};\ \mathrm{S}}{\backslash\mathrm{E}}$$

At this point, this is just like the familiar PSG analysis of the following form, except that the symbol VP is replaced by NP\S:

(4)


Things will start looking more interesting as we make the fragment more complex (and also by adding the semantics), but before doing so I first introduce some basic assumptions, first on syntactic categories (below) and then on semantics (next section).

Syntactic categories (or syntactic types) are defined recursively in CG. This can be concisely written using the so-called 'BNC notation' as follows:[2, 3]

(5) a. BaseType := { N, NP, PP, S }
    b. Type := BaseType | Type\Type | Type/Type

In words, anything that is a BaseType is a Type, and any complex expression of form A\ B or A/B where A and B are both Types is a Type. To give some examples, the following expressions are syntactic types according to the definition in (5):[4]

(6) a. S\S

[2] See section 3.3 below for the treatment of syntactic features (such as those used for agreement). I ignore this aspect for the fragment developed below for the sake of exposition. The treatment of syntactic features (or its analog) is a relatively underdeveloped aspect of CG syntax literature, as compared to HPSG research (where the whole linguistic theory is built on the basis of a theory/formalism of complex feature structures).

[3] Recognizing PP as a basic type is somewhat non-standard, although there does not seem to be any consensus on what should be regarded as a (reasonably complete) set of basic syntactic types for natural language syntax.

[4] I omit parentheses for a sequence of the same type of slash, for which disambiguation is obvious—for example, A\A\A is an abbreviation for (A\(A\A)).

b. (NP\S)/NP/NP

c. (S/(NP\S))\(S/NP)

d. ((NP\S)\(NP\S))\((NP\S)\(NP\S))

One important feature of CG is that, like HPSG, it lexicalizes the valence (or subcategorization) properties of linguistic expressions. Unlike HPSG, where this is done by a list (or set) valued syntactic feature, in CG, complex syntactic categories directly represent the combinatoric (i.e. valence) properties of lexical items. For example, lexical entries for intransitive and transitive verbs in English will look like the following (semantics is omitted here but will be supplied later):

(7) a. ran; NP\S

b. read; (NP\S)/NP

c. introduces; (NP\S)/PP/NP

(7a) says that the verb *ran* combines with its argument NP *to its left* to become an S. Likewise, (7c) says that *read* first combines with an NP *to its right* and then another NP to its left to become an S.

One point to keep in mind (though it may not seem to make much difference at this point) is that in CG, syntactic rules are thought of as logical rules and the derivations of sentences like (3) as *proofs* of the well-formedness of particular strings as sentences. From this 'logical' point of view, the two slashes should really be thought of as directional variants of implication (that is, both A/B and $B\backslash A$ essentially mean '*if* there is a B, *then* there is an A'), and the two rules of Slash Elimination introduced in (1) should be thought of as directional variants of MODUS PONENS ($B \rightarrow A, B \vdash A$). This analogy between natural language syntax and logic is emphasized in particular in the TLCG research.

2.3 Syntax-semantics interface in CG

One attractive property of CG as a theory of natural language syntax is its straightforward syntax-semantics interface. In particular, there is a functional mapping from syntactic categories to semantic types.[5] For the sake of exposition, I assume an extensional fragment of Montagovian model-theoretic semantics in what follows, but it should be noted that the CG syntax is mostly neutral to the choice

[5]Technically, this is ensured in TLCG by the homomorphism from the syntactic type logic to the semantic type logic (the latter of which is often implicit) and the so-called Curry-Howard correspondence between proofs and terms (van Benthem 1988).

of the specific variant of semantic theory to go with it.[6]

Assuming the standard recursive definition of semantic types as in (8) (with basic types e (individuals) and t (truth values)), we can define the function Sem that returns, for each syntactic category given as input, its semantic type, as in (9) and (10).

(8) a. BaseSemType := { e, t }

b. SemType := BaseSemType | SemType $\rightarrow$ SemType

(9) (Base Case)

a. Sem(NP) = Sem(PP) = e

b. Sem(N) = $e \rightarrow t$

c. Sem(S) = t

(10) (Recursive Clause)

For any complex syntactic category of the form A/B (or $B\backslash A$),

$\mathsf{Sem}(A/B)$ (= $\mathsf{Sem}(B\backslash A)$) = $\mathsf{Sem}(B) \rightarrow \mathsf{Sem}(A)$

For example, we have $\mathsf{Sem}(\mathrm{S}/(\mathrm{NP}\backslash\mathrm{S})) = (e \rightarrow t) \rightarrow t$ (for subject position quantifier in CCG).

Syntactic rules with semantics can then be written as in (11) (where the semantic effect of these rules is FUNCTION APPLICATION) and a sample derivation with semantic annotation is given in (12).

(11) a. FORWARD SLASH ELIMINATION

$$\frac{\mathsf{a};\ \mathcal{F};\ A/B \qquad \mathsf{b};\ \mathcal{G};\ B}{\mathsf{a} \circ \mathsf{b};\ \mathcal{F}(\mathcal{G});\ A}\ /\mathrm{E}$$

b. BACKWARD SLASH ELIMINATION

$$\frac{\mathsf{b};\ \mathcal{G};\ B \qquad \mathsf{a};\ \mathcal{F};\ B\backslash A}{\mathsf{b} \circ \mathsf{a};\ \mathcal{F}(\mathcal{G});\ A}\ \backslash\mathrm{E}$$

(12)

$$\cfrac{\mathsf{john};\ \mathbf{j};\ \mathrm{NP} \qquad \cfrac{\mathsf{loves};\ \mathbf{love};\ (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP} \qquad \mathsf{mary};\ \mathbf{m};\ \mathrm{NP}}{\mathsf{loves} \circ \mathsf{mary};\ \mathbf{love}(\mathbf{m});\ \mathrm{NP}\backslash\mathrm{S}}\ /\mathrm{E}}{\mathsf{john} \circ \mathsf{loves} \circ \mathsf{mary};\ \mathbf{love}(\mathbf{m})(\mathbf{j});\ \mathrm{S}}\ \backslash\mathrm{E}$$

A system of CG with only the Slash Elimination rules like the fragment above is called the AB GRAMMAR, so called because it corresponds to the earliest form of CG formulated by Ajdukiewicz (1935) and Bar-Hillel (1953).

[6] See for example Martin (2013) and Bekki & Mineshima (2017) for recent proposals on adopting compositional variants of (hyper)intensional dynamic semantics and proof theoretic semantics, respectively, for the semantic component of CG-based theories of natural language.

2.4 Combinatory Categorial Grammar

2.4.1 An 'ABC' fragment: AB grammar with order-preserving 'combinatory' rules

To do some interesting linguistic analysis, we need some more machinery. I now extend the AB fragment above by adding two types of rules: TYPE-RAISING and (Harmonic) FUNCTION COMPOSITION. These are a subset of rules typically entertained in CCG. I call the resultant system ABC GRAMMAR (AB + Function Composition).[7] Though it is an impoverished version of CCG, the ABC fragment already enables an interesting and elegant analysis of NONCONSTITUENT COORDINATION (NCC), originally due to Steedman (1985) and Dowty (1988), which is essentially identical to the analysis of NCC in the current versions of both CCG and TLCG. I will then discuss the rest of the rules constituting CCG in the next section. The reason for drawing a distinction between the 'ABC' fragment and (proper) CCG is just for the sake of exposition. The rules introduced in the present section have the property that they are all derivable as *theorems* in the (associative) Lambek calculus, the calculus that underlies most variants of TLCG. For this reason, separating the two sets of rules helps clarify the similarities and differences between CCG and TLCG.

The TYPE RAISING and FUNCTION COMPOSITION rules are defined as in (13) and (14), respectively.

(13) a. Forward Function Composition

$$\frac{\mathsf{a}; \mathcal{F}; A/B \qquad \mathsf{b}; \mathcal{G}; B/C}{\mathsf{a} \circ \mathsf{b}; \lambda x.\mathcal{F}(\mathcal{G}(x)); A/C}\text{FC}$$

b. Backward Function Composition

$$\frac{\mathsf{b}; \mathcal{G}; C\backslash B \qquad \mathsf{a}; \mathcal{F}; B\backslash A}{\mathsf{b} \circ \mathsf{a}; \lambda x.\mathcal{F}(\mathcal{G}(x)); C\backslash A}\text{FC}$$

(14) a. Forward Type Raising

$$\frac{\mathsf{a}; \mathcal{F}; A}{\mathsf{a}; \lambda v.v(\mathcal{F}); B/(A\backslash B)}\text{TR}$$

b. Backward Type Raising

$$\frac{\mathsf{a}; \mathcal{F}; A}{\mathsf{a}; \lambda v.v(\mathcal{F}); (B/A)\backslash B}\text{TR}$$

The Type-Raising rules are essentially rules of 'type lifting' familiar in the formal semantics literature, except that they specify the 'syntactic effect' of type lifting explicitly (such that the function-argument relation is reversed). Similarly Function Composition rules can be understand as function composition in the usual sense (as in mathematics and functional programming), except, again, that the syntactic effect is explicitly specified.

[7] This is not a standard terminology, but giving a name to this fragment is convenient for the purpose of the discussion below.

As noted by Steedman (1985), with Type Raising and Function Composition, we can analyze a string of words such as *John loves* as a constituent of type S/NP, that is, an expression that is looking for an NP to its right to become an S:

(15)
$$\dfrac{\dfrac{\textsf{john}; \mathbf{j}; \mathrm{NP}}{\textsf{john}; \lambda f.f(\mathbf{j}); \mathrm{S}/(\mathrm{NP}\backslash\mathrm{S})}\,\mathrm{TR} \qquad \textsf{loves}; \mathbf{love}; (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP}}{\textsf{john} \circ \textsf{loves}; \lambda x.\mathbf{love}(x)(\mathbf{j}); \mathrm{S}/\mathrm{NP}}\,\mathrm{FC}$$

Assuming generalized conjunction (with the standard definition for the generalized conjunction operator ⊓ *a la* Partee & Rooth (1983) and the polymorphic syntactic category $(X\backslash X)/X$ for *and*), the analysis for a RIGHT-NODE RAISING (RNR) sentence such as (16) is straightforward, as in (17).

(16) John loves, and Bill hates, Mary.

(17)
$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ \textsf{john} \circ \textsf{loves};\\ \lambda x.\mathbf{love}(x)(\mathbf{j}); \mathrm{S}/\mathrm{NP}\end{array} \qquad \dfrac{\begin{array}{c}\textsf{and};\\ \sqcap; (X\backslash X)/X\end{array} \qquad \begin{array}{c}\vdots\\ \textsf{bill} \circ \textsf{hates};\\ \lambda x.\mathbf{hate}(x)(\mathbf{b}); \mathrm{S}/\mathrm{NP}\end{array}}{\begin{array}{c}\textsf{and} \circ \textsf{bill} \circ \textsf{hates};\\ \sqcap(\lambda x.\mathbf{hate}(x)(\mathbf{b})); (\mathrm{S}/\mathrm{NP})\backslash(\mathrm{S}/\mathrm{NP})\end{array}}\,\mathrm{FA}}{\textsf{john} \circ \textsf{loves} \circ \textsf{and} \circ \textsf{bill} \circ \textsf{hates}; (\lambda x.\mathbf{love}(x)(\mathbf{j})) \sqcap (\lambda x.\mathbf{hate}(x)(\mathbf{b})); \mathrm{S}/\mathrm{NP}}\,\mathrm{FA} \qquad \begin{array}{c}\textsf{mary};\\ \mathbf{m}; \mathrm{NP}\end{array}}{\textsf{john} \circ \textsf{loves} \circ \textsf{and} \circ \textsf{bill} \circ \textsf{hates} \circ \textsf{mary}; \mathbf{love}(\mathbf{m})(\mathbf{j}) \wedge \mathbf{hate}(\mathbf{m})(\mathbf{b}); \mathrm{S}}\,\mathrm{FA}$$

Dowty (1988) showed that this analysis extends straightforwardly to the (slightly) more complex case of ARGUMENT CLUSTER COORDINATION (ACC), such as (18), as in (19) (here, VP, TV and DTV are abbreviations of NP\S, (NP\S)/NP, and (NP\S)/NP/NP, respectively).

(18) Mary gave Bill the book and John the record.

(19)
$$\dfrac{\begin{array}{c}\textsf{mary};\\ \mathbf{m};\\ \mathrm{NP}\end{array} \qquad \dfrac{\begin{array}{c}\textsf{gave};\\ \mathbf{give};\\ \mathrm{DTV}\end{array} \qquad \dfrac{\dfrac{\dfrac{\begin{array}{c}\textsf{bill};\\ \mathbf{b}; \mathrm{NP}\end{array}}{\begin{array}{c}\textsf{bill};\\ \lambda P.P(\mathbf{b});\\ \mathrm{DTV}\backslash\mathrm{TV}\end{array}}\,\mathrm{TR} \qquad \dfrac{\begin{array}{c}\textsf{the} \circ \textsf{book};\\ \iota(\mathbf{bk}); \mathrm{NP}\end{array}}{\begin{array}{c}\textsf{the} \circ \textsf{book};\\ \lambda Q.Q(\iota(\mathbf{bk}));\\ \mathrm{TV}\backslash\mathrm{VP}\end{array}}\,\mathrm{TR}}{\begin{array}{c}\textsf{bill} \circ \textsf{the} \circ \textsf{book};\\ \lambda R.R(\mathbf{b})(\iota(\mathbf{bk})); \mathrm{DTV}\backslash\mathrm{VP}\end{array}}\,\mathrm{FC} \qquad \dfrac{\begin{array}{c}\textsf{and};\\ \sqcap;\\ (X\backslash X)/X\end{array} \qquad \begin{array}{c}\vdots\\ \textsf{john} \circ \textsf{the} \circ \textsf{record};\\ \lambda R.R(\mathbf{j})(\iota(\mathbf{rc}));\\ \mathrm{DTV}\backslash\mathrm{VP}\end{array}}{\begin{array}{c}\textsf{and} \circ \textsf{john} \circ \textsf{the} \circ \textsf{record};\\ \sqcap(\lambda R.R(\mathbf{j})(\iota(\mathbf{rc})));\\ (\mathrm{DTV}\backslash\mathrm{VP})\backslash(\mathrm{DTV}\backslash\mathrm{VP})\end{array}}}{\begin{array}{c}\textsf{bill} \circ \textsf{the} \circ \textsf{book} \circ \textsf{and} \circ \textsf{john} \circ \textsf{the} \circ \textsf{record};\\ \lambda R.R(\mathbf{j})(\iota(\mathbf{rc})) \sqcap \lambda R.R(\mathbf{b})(\iota(\mathbf{bk})); \mathrm{DTV}\backslash\mathrm{VP}\end{array}}}{\begin{array}{c}\textsf{gave} \circ \textsf{bill} \circ \textsf{the} \circ \textsf{book} \circ \textsf{and} \circ \textsf{john} \circ \textsf{the} \circ \textsf{record};\\ \mathbf{give}(\mathbf{j})(\iota(\mathbf{rc})) \sqcap \mathbf{give}(\mathbf{b})(\iota(\mathbf{bk})); \mathrm{VP}\end{array}}}{\begin{array}{c}\textsf{mary} \circ \textsf{gave} \circ \textsf{bill} \circ \textsf{the} \circ \textsf{book} \circ \textsf{and} \circ \textsf{john} \circ \textsf{the} \circ \textsf{record};\\ \mathbf{give}(\mathbf{j})(\iota(\mathbf{rc}))(\mathbf{m}) \wedge \mathbf{give}(\mathbf{b})(\iota(\mathbf{bk}))(\mathbf{m}); \mathrm{S}\end{array}}$$

Here, by Type Raising, the indirect and direct objects become functions that can be combined via Function Composition, to form a non-standard constituent that can then be coordinated. After two such expressions are conjoined, the verb is fed as an argument to return a VP. Intuitively, the idea behind this analysis is that *Bill the book* is of type DTV\VP since if it were to combine with an actual ditransitive verb (such as *gave*), we would obtain a VP (*gave Bill a book*). Note that in both the RNR and ACC examples above, the right semantic interpretation for the whole sentence is assigned compositionally via the rules given above in (13) and (14).

2.4.2 From ABC to CCG

CCG is a version of CG developed by Mark Steedman since the 1980s with extensive linguistic application. The best sources for CCG are the three books by Steedman (Steedman 1997; 2000; Steedman 2012), which present treatments of major linguistic phenomena in CCG, and gives pointers to earlier literature. CCG is essentially a rule-based extension of the AB grammar. We have already seen in the previous section the two key components that constitute this extension: Type Raising and (Harmonic) Function Composition.[8] There are aspects of natural language syntax that cannot be handled adequately in this simple system, and in such situations, CCG makes (restricted) use of additional rules. This point can be illustrated nicely with two issues that arise in connection with the analysis of long-distance dependencies.

The basic idea behind the CCG analysis of long-distance dependencies, due originally to Ades & Steedman (1982), is very simple and is similar in spirit to the HPSG analysis in terms of SLASH feature percolation (see Chapter 14 for the treatment of long-distance dependencies in HPSG). Specifically, CCG analyzes extraction dependencies via a chain of function composition, as illustrated by the derivation for (20) in (21).

(20) This is the book that John thought that Mary read __.

[8] There is actually a subtle point about Type Raising rules. Recent versions of CCG (Steedman 2012) do not take them to be syntactic rules but rather assume that Type Raising is an operation in the lexicon. This choice seems to be motivated by parsing considerations (so as to eliminate as many unary rules as possible from the syntax). It is also worth noting in this connection that the CCG-based syntactic fragment that Jacobson (1999; 2000) assumes for her Variable-Free Semantics is actually a quite different system from Steedman's version of CCG in that it crucially assumes Geach rules, another type of unary rules likely to have similar computational consequences as Type-Raising rules, in the syntactic component.

(21)

mary; $\mathbf{m}$; NP
—— TR
mary; $\lambda P.P(\mathbf{m})$; $S/(NP\backslash S)$ read; $\mathbf{read}$; $(NP\backslash S)/NP$
—— FC
that; $\lambda p.p$; S'/S mary ∘ read; $\lambda x.\mathbf{read}(x)(\mathbf{m})$; S/NP
—— FC
thought; $\mathbf{think}$; $(NP\backslash S)/S'$ that ∘ mary ∘ read; $\lambda x.\mathbf{read}(x)(\mathbf{m})$; S'/S
—— FC
john; $\mathbf{j}$; NP
—— TR
john; $\lambda P.P(\mathbf{j})$; $S/(NP\backslash S)$ thought ∘ that ∘ mary ∘ read; $\lambda x.\mathbf{think}(\mathbf{read}(x)(\mathbf{m}))$; $(NP\backslash S)/NP$
—— FC
that; $\lambda P \lambda Q \lambda x. Q(x) \wedge P(x)$; $(N\backslash N)/(S/NP)$ john ∘ thought ∘ that ∘ mary ∘ read; $\lambda x.\mathbf{think}(\mathbf{read}(x)(\mathbf{m}))(\mathbf{j})$; S/NP
—— FA
that ∘ john ∘ thought ∘ that ∘ mary ∘ read; $\lambda Q \lambda x.Q(x) \wedge \mathbf{think}(\mathbf{read}(x)(\mathbf{m}))(\mathbf{j})$; $N\backslash N$

Like (many versions of) HPSG, CCG does not assume any empty expression at the gap site. Instead, the information that the subexpressions (constituting extraction pathway) such as *Mary read* and *thought that Mary read* are missing an NP on the right edge is encoded in the syntactic category of the linguistic expression. *Mary read* is assigned the type S/NP since it is a sentence missing an NP on its right edge. *thought that Mary read* is of type VP/NP since it is a VP missing an NP on its right edge, etc. Expressions that are not originally functions (such as the subject NPs in the higher and lower clauses inside the relative clause in (20)) are first type-raised. Then, function composition effectively 'delays' the saturation of the object NP argument of the embedded verb, until the whole relative clause meets the relative pronoun, which itself is a higher-order function that takes a sentence missing an NP (of type S/NP) as an argument.

The successive passing of the /NP specification to larger structures is essentially analogous to the treatment of extraction via the SLASH feature in HPSG. However, unlike HPSG, which has a dedicated feature that handles this information passing, CCG achieves the effect via the ordinary slash that is also used for local syntactic composition.[9]

This difference immediately raises some issues for the CCG analysis of extraction. First, in (20), the NP gap happens to be on the right edge of the sentence, but this is not always the case. Harmonic function composition alone cannot handle non-peripheral extraction of the sort found in examples such as the following:

[9] There are several complications that arise in the CCG analysis because of this architectural difference, which I discuss immediately below. But one interesting consequence of the CCG approach is that the quite complex set of specifications of SLASH feature percolation (including the *termination* of percolation) are all unnecessary in CCG.

(22) This is the book that John thought that [Mary read __ at school].

Assuming that *at school* is a VP modifier of type (NP\S)\(NP\S), what is needed here is a mechanism that assigns the type VP/NP to the string *read __ at school*, despite the fact that the missing NP is not on the right edge. CCG employs a special rule of 'crossed' function composition for this purpose, defined as follows:

(23) Crossed Function Composition

$$\frac{\mathsf{a};\ \mathcal{G};\ A/B \qquad \mathsf{b};\ \mathcal{F};\ A\backslash C}{\mathsf{a} \circ \mathsf{b};\ \lambda x.\mathcal{F}(\mathcal{G}(x));\ C/B}\,\mathrm{xFC}$$

Unlike its harmonic counterpart (in which a has the type $B\backslash A$), in (23) the directionality of the slash is different in the two premises, and the resultant category inherits the slash originally associated with the inherited argument (i.e. $/B$).

Once this non-order-preserving version of function composition is introduced in the grammar, the derivation for (22) is straightforward, as in (24):

(24)
$$\frac{\dfrac{\mathsf{mary};\ \mathbf{m};\ \mathrm{NP}}{\begin{array}{l}\mathsf{mary};\\ \lambda P.P(\mathbf{m});\\ \mathrm{S}/(\mathrm{NP}\backslash\mathrm{S})\end{array}}\,\mathrm{TR} \qquad \dfrac{\begin{array}{l}\mathsf{read};\\ \mathbf{read};\ (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP}\end{array} \qquad \begin{array}{l}\mathsf{at} \circ \mathsf{school};\\ \mathbf{at\text{-}school};\ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S})\end{array}}{\mathsf{read} \circ \mathsf{at} \circ \mathsf{school};\ \lambda x.\mathbf{at\text{-}school}(\mathbf{read}(x));\ (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP}}\,\mathrm{xFC}}{\mathsf{mary} \circ \mathsf{read} \circ \mathsf{at} \circ \mathsf{school};\ \lambda x.\mathbf{at\text{-}school}(\mathbf{read}(x))(\mathbf{m});\ \mathrm{S}/\mathrm{NP}}\,\mathrm{FC}$$

Unless appropriately constrained, the addition of the crossed composition rule leads to potential overgeneration, since non-extracted expressions cannot change word order so freely in English. For example, without additional restrictions, the simple CCG fragment above overgenerates examples such as the following:

(25) *Mary read at school it.

Here, I will not go into the technical details of how this issue is addressed in the CCG literature.[10] In contemporary versions of CCG, the application of special rules such as crossed composition in (23) is regulated by the notion of 'structural control' borrowed into CCG from the 'multi-modal' variant of TLCG (see Baldridge (2002) and Steedman & Baldridge (2011)).

Another issue that arises in connection to extraction is how to treat multiple gaps corresponding to a single filler. The simple fragment developed above cannot license examples involving PARASITIC GAPS such as the following:[11]

[10] The word order in (25) is of course fine in examples instantiating Heavy NP Shift.

[11] Multiple gaps in coordination (i.e. ATB extraction) is not an issue, since these cases can be handled straightforwardly via the polymorphic definition of generalized conjunction in CCG, in just the same way that unsaturated shared arguments in each conjunct are identified with one another.

(26) a. This is the article that I filed __ without reading __.

b. Peter is a guy who even the best friends of __ think __ should be closely watched.

Since neither type raising nor function composition changes the number of 'gaps' passed on to a larger expression, we need a new mechanism here. Steedman (1987) proposes the following rule to deal with this issue:

(27) Substitution

$$\frac{\mathsf{a};\ \mathcal{G};\ A/B \qquad \mathsf{b};\ \mathcal{F};\ (A\backslash C)/B}{\mathsf{a} \circ \mathsf{b};\ \lambda x.\mathcal{F}(x)(\mathcal{G}(x));\ C/B}\ \mathrm{S}$$

This rule has the effect of 'collapsing' the arguments of the two inputs into one, to be saturated by a single filler. The derivation for the adjunct parasitic gap example in (26a) then goes as follows (where VP is an abbreviation for NP\S):

(28)
$$\dfrac{\begin{matrix}\mathsf{filed;} \\ \mathbf{file};\mathrm{VP/NP}\end{matrix} \qquad \dfrac{\begin{matrix}\mathsf{without;} \\ \mathbf{wo};(\mathrm{VP}\backslash\mathrm{VP})/\mathrm{VP}\end{matrix} \qquad \begin{matrix}\mathsf{reading;} \\ \mathbf{read};\mathrm{VP/NP}\end{matrix}}{\begin{matrix}\mathsf{without} \circ \mathsf{reading;} \\ \lambda x.\mathbf{wo}(\mathbf{read}(x));(\mathrm{VP}\backslash\mathrm{VP})/\mathrm{NP}\end{matrix}}\mathrm{FC}}{\begin{matrix}\mathsf{filed} \circ \mathsf{without} \circ \mathsf{reading;} \\ \lambda x.\mathbf{wo}(\mathbf{read}(x))(\mathbf{file}(x));\mathrm{VP/NP}\end{matrix}}\mathrm{S}$$

Like the crossed composition rule, the availability of the substitution rule should be restricted to extraction environments. In earlier versions of CCG, this was done by a stipulation on the rule itself. Baldridge (2002) proposed an improvement of the organization of the CCG rule system in which the applicability of particular rules is governed by lexically specified 'modality' encodings. See Steedman & Baldridge (2011) for this relatively recent development in CCG.

2.5 Type-Logical Categorial Grammar

The 'rule-based' nature of CCG should be clear from the above exposition. TLCG takes a somewhat different perspective on the underlying architecture of the grammar of natural language. Specifically, in TLCG, the rule system of grammar is literally taken to be a kind of logic. Consequently, all (or almost all) grammar rules are logical inference rules reflecting the properties of (typically a small number of) logical connectives such as / and \ (which are, as noted in section 2.2, viewed as directional variants of implication). This conceptual shift can be best illustrated by first replacing the ABC grammar introduced in section 2.4.1

by the Lambek calculus, where all the rules posited as primitive rules in the former are derived as *theorems* (in the technical sense of the term) in the latter.

Before moving on, I should hasten to note that the TLCG literature is more varied than the CCG literature, consisting of several related but distinct lines of research. I choose to present one particular variant called Hybrid Type-Logical Categorial Grammar (Kubota & Levine 2020) in what follows, in line with the present chapter's linguistic emphasis (for a more in-depth discussion on the linguistic application of TLCG, see Carpenter (1998) and Kubota & Levine (2020)). A brief comparison with major alternatives is offered in section 2.5.3. Other variants of TLCG, most notably, the Categorial Type Logics (Moortgat 2011) and Displacement Calculus (Morrill 2010) emphasize the logical and computational aspects. Moot & Retoré (2012) is a good introduction to TLCG with emphasis on these latter aspects.

2.5.1 The Lambek calculus

In addition to the Slash Elimination rules (reproduced here as (29)), which are identical to the two rules in the AB grammar from section 2.2, the Lambek calculus posits the Slash Introduction rules, which can be written in the current labelled deduction format as in (30).[12]

(29) a. Forward Slash Elimination

$$\frac{\mathsf{a}; \mathcal{F}; A/B \qquad \mathsf{b}; \mathcal{G}; B}{\mathsf{a} \circ \mathsf{b}; \mathcal{F}(\mathcal{G}); A}/\mathrm{E}$$

b. Backward Slash Elimination

$$\frac{\mathsf{b}; \mathcal{G}; B \qquad \mathsf{a}; \mathcal{F}; B\backslash A}{\mathsf{b} \circ \mathsf{a}; \mathcal{F}(\mathcal{G}); A}\backslash\mathrm{E}$$

(30) a. Forward Slash Introduction

$$\cfrac{\cfrac{\begin{array}{ccc} \vdots & \underline{[\varphi; x; A]^n} & \vdots \\ \vdots & \vdots & \vdots \end{array}}{\mathsf{b} \circ \varphi; \mathcal{F}; B}}{\mathsf{b}; \lambda x.\mathcal{F}; B/A}/\mathrm{I}^n$$

b. Backward Slash Introduction

$$\cfrac{\cfrac{\begin{array}{ccc} \vdots & \underline{[\varphi; x; A]^n} & \vdots \\ \vdots & \vdots & \vdots \end{array}}{\varphi \circ \mathsf{b}; \mathcal{F}; B}}{\mathsf{b}; \lambda x.\mathcal{F}; A\backslash B}\backslash\mathrm{I}^n$$

I illustrate the workings of these rules in some examples first and then come back to some formal and notational issues. One consequence that immediately follows is that Type Raising and Function Composition (as well as other theorems; see, e.g., Jäger (2005: sec 2.2.5, pp. 46–)) are now derivable as theorems. As an illustration, the proofs for (14a) and (13a) are shown in (31) and (32), respectively.

[12] Morrill (1994) was the first to recast the Lambek calculus in this labelled deduction format.

(31) $$\dfrac{\dfrac{[\varphi; v; A\backslash B]^1 \qquad \mathsf{a}; \mathcal{F}; A}{\mathsf{a}\circ\varphi; v(\mathcal{F}); B}\,\backslash\mathrm{E}}{\mathsf{a}; \lambda v.v(\mathcal{F}); B/(A\backslash B)}\,/\mathrm{I}^1$$

(32) $$\dfrac{\dfrac{\mathsf{a}; \mathcal{F}; A/B \qquad \dfrac{[\varphi; x; C]^1 \qquad \mathsf{b}; \mathcal{G}; B/C}{\mathsf{b}\circ\varphi; \mathcal{G}(x); B}\,/\mathrm{E}}{\mathsf{a}\circ\mathsf{b}\circ\varphi; \mathcal{F}(\mathcal{G}(x)); A}\,/\mathrm{E}}{\mathsf{a}\circ\mathsf{b}; \lambda x.\mathcal{F}(\mathcal{G}(x)); A/C}\,/\mathrm{I}^1$$

These are formal theorems, but they intuitively make sense. For example, what's going on in (32) is simple. We first hypothetically assume the existence of some expression of type C and combine it with B/C. This gives us a larger expression of type B, which then can be fed as an argument to A/B. At that point, we withdraw the initial hypothesis and conclude that what we really had was just something that would become an A *if* there is a C to its right, namely, an expression of type A/C. Thus, a sequence of expression of types A/B and B/C is proven to be of type A/C. This type of proof is known as HYPOTHETICAL REASONING, since it involves a step of positing a hypothesis initially and withdrawing that hypothesis at a later point.

Getting back to some notational issues, there are two crucial things to keep in mind about the notational convention adopted here (which I implicitly assumed above). First, the connective $\circ$ in the prosodic component designates string concatenation and is associative in both directions (i.e. $(\varphi_1 \circ \varphi_2) \circ \varphi_3 \equiv \varphi_1 \circ (\varphi_2 \circ \varphi_3)$). In other words, hierarchical structure is irrelevant for the prosodic representation. Thus, the applicability condition on the Forward Slash Introduction rule (30a) is simply that the prosodic variable φ of the hypothesis appear as the rightmost element of the string prosody of the input expression (i.e. $\mathsf{b} \circ \varphi$). Since the penultimate step in (32) satisfies this condition, the rule is applicable here. Second, note in this connection that the application of the Introduction rules is conditioned on the position of the prosodic variable, and *not* on the position of the hypothesis itself in the proof tree (this latter convention is more standardly adopted when the Lambek calculus is presented in Prawitz-style natural deduction, though the two presentations are equivalent—see, e.g., Carpenter (1998) and Jäger (2005)).

Hypothetical reasoning with Slash Introduction makes it possible to recast the CCG analysis of nonconstituent coordination from section 2.4.1 within the logic of / and \. This reformulation fully retains the essential analytic ideas of the original CCG analysis but makes the underlying 'logic' of syntactic composition more transparent.

The following derivation illustrates how the 'reanalysis' of the string *Bill the book* as a derived constituent of type (VP/NP/NP)\VP (the same type as in (19)) can be obtained in the Lambek calculus:

(33)
$$\cfrac{\cfrac{\cfrac{[\varphi; f; \mathrm{VP/NP/NP}]^1 \quad \textsf{bill}; \mathbf{b}; \mathrm{NP}}{\varphi \circ \textsf{bill}; f(\mathbf{b}); \mathrm{VP/NP}}\,/\mathrm{E} \quad \textsf{the} \circ \textsf{book}; \iota(\mathbf{bk}); \mathrm{NP}}{\varphi \circ \textsf{bill} \circ \textsf{the} \circ \textsf{book}; f(\mathbf{b})(\iota(\mathbf{bk})); \mathrm{VP}}\,/\mathrm{E}}{\textsf{bill} \circ \textsf{the} \circ \textsf{book}; \lambda f.f(\mathbf{b})(\iota(\mathbf{bk})); (\mathrm{VP/NP/NP})\backslash\mathrm{VP}}\,\backslash\,\mathrm{I}^1$$

At this point, one may wonder what the relationship is between the analysis of nonconstituent coordination via Type Raising and Function Composition in the ABC grammar in section 2.4.1 and the hypothetical reasoning-based analysis in the Lambek calculus just presented. Intuitively, they seem to achieve the same effect in slightly different ways. The logic-based perspective of TLCG allows us to obtain a deeper understanding of the relationship between them. To facilitate comparison, I first recast the Type Raising + Function Composition analysis from section 2.4.1 in the Lambek calculus. The relevant part is the part that derives the 'noncanonical constituent' *Bill the book*:

(34)
$$\cfrac{\cfrac{\cfrac{[\varphi_3; R; \mathrm{DTV}]^3 \quad \cfrac{\cfrac{[\varphi_2; P; \mathrm{DTV}]^2 \quad \textsf{bill}; \mathbf{b}; \mathrm{NP}}{\varphi_2 \circ \textsf{bill}; P(\mathbf{b}); \mathrm{TV}}\,/\mathrm{E}}{\textsf{bill}; \lambda P.P(\mathbf{b}); \mathrm{DTV}\backslash\mathrm{TV}}\,\backslash\,\mathrm{I}^2}{\varphi_3 \circ \textsf{bill}; R(\mathbf{b}); \mathrm{TV}}\,\backslash\mathrm{E} \quad \cfrac{\cfrac{[\varphi_1; Q; \mathrm{TV}]^1 \quad \textsf{the} \circ \textsf{book}; \iota(\mathbf{bk}); \mathrm{NP}}{\varphi_1 \circ \textsf{the} \circ \textsf{book}; Q(\iota(\mathbf{bk})); \mathrm{VP}}\,/\mathrm{E}}{\textsf{the} \circ \textsf{book}; \lambda Q.Q(\iota(\mathbf{bk})); \mathrm{TV}\backslash\mathrm{VP}}\,\backslash\,\mathrm{I}^1}{\varphi_3 \circ \textsf{bill} \circ \textsf{the} \circ \textsf{book}; R(\mathbf{b})(\iota(\mathbf{bk})); \mathrm{VP}}\,\backslash\mathrm{E}}{\textsf{bill} \circ \textsf{the} \circ \textsf{book}; \lambda R.R(\mathbf{b})(\iota(\mathbf{bk})); \mathrm{DTV}\backslash\mathrm{VP}}\,\backslash\,\mathrm{I}^3$$

By comparing (34) and (33), we see that (34) contains some redundant steps. First, hypothesis 2 is introduced only to be replaced by hypothesis 3. This is completely redundant since we could have obtained exactly the same result by directly combining hypothesis 3 with the NP *Bill*. Similarly, hypothesis 1 can be eliminated by replacing it with the TV $\varphi_3 \circ$ book on the left-hand side of the third line from the bottom. By making these two simplifications, we obtain the derivation in (33).

The relationship between the more complex proof in (34) and the simpler one in (33) is parallel to the relationship between an unreduced lambda term (such as $\lambda R[\lambda Q[Q(\iota(\mathbf{bk}))](\lambda P[P(\mathbf{b})](R))]$) and its β-normal form (i.e. $\lambda R.R(\mathbf{b})(\iota(\mathbf{bk}))$). In fact, there is a formally precise one-to-one relationship between linear logic (of which the Lambek calculus is known to be a straightforward extension) and the typed lambda calculus known as the Curry-Howard Isomorphism (Howard 1969), according to which the lambda term that represents the proof (34) β-reduces to the term that represents the proof (33).[13] Technically, this is known as PROOF

[13] There is a close relationship between these lambda terms representing proofs and the lambda terms that we write to notate semantic translations, especially if we write the latter at each step of derivation *without* performing β-reduction. But it is important to keep in mind that

NORMALIZATION (Jäger (2005) contains a particularly useful discussion on this notion).

Thus, the logic-based architecture of the Lambek calculus (and various versions of TLCG, which are all extensions of the Lambek calculus) enables us to say, in a technically precise way, how (34) and (33) are the 'same' (or, more precisely, equivalent), by building on independently established results in mathematical logic and computer science. This is one big advantage of taking seriously the view, advocated by the TLCG research, that 'language *is* logic'.

2.5.2 Extending the Lambek calculus

Hypothetical reasoning is a very powerful (yet systematic) tool, but with forward and backward slashes, it is only good for analyzing expressions missing some material at the (right or left) periphery. This is problematic in the analyses of many linguistic phenomena, such as *wh*-extraction (where the 'gap' can be in a sentence-medial position—recall the discussion about crossed composition rules in CCG in section 2.4.2) and quantifier scope (where the quantifier needs to 'covertly' move from a sentence-medial position), as well as various kinds of discontinuous constituency phenomena (see for example Morrill et al. (2011), which contains analyses of various types of discontinuous constituency phenomena in a recent version of TLCG known as 'Displacement Calculus'). In what follows, I sketch one particular, relatively recent approach to this problem, known as HYBRID TYPE-LOGICAL CATEGORIAL GRAMMAR (Hybrid TLCG; Kubota 2010; 2015; Kubota & Levine 2015; 2020). This approach combines the Lambek calculus with Oehrle's (1994) term-labelled calculus, which deals with discontinuity by employing λ-binding in the prosodic component.

Hybrid TLCG extends the Lambek calculus with the Elimination and Introduction rules for the VERTICAL SLASH:

(35) a. VERTICAL SLASH INTRODUCTION

$$\dfrac{\dfrac{\begin{array}{ccc} \vdots & \underline{[\varphi; x; A]^n} & \vdots \\ \vdots & \vdots & \vdots \end{array}}{\mathsf{b}; \mathcal{F}; B}}{\lambda\varphi.\mathsf{b}; \lambda x.\mathcal{F}; B\upharpoonright A}\ {\upharpoonright}\mathrm{I}^n$$

b. VERTICAL SLASH ELIMINATION

$$\dfrac{\mathsf{a}; \mathcal{F}; A\upharpoonright B \quad \mathsf{b}; \mathcal{G}; B}{\mathsf{a}(\mathsf{b}); \mathcal{F}(\mathcal{G}); A}\ {\upharpoonright}\mathrm{E}$$

These rules allow us to model what (roughly) corresponds to syntactic movement

lambda terms representing syntactic proofs and lambda terms notating semantic translations are distinct things.

operations in derivational frameworks. This is illustrated in (36) for the $\forall > \exists$ reading for the sentence *Someone talked to everyone today*:

(36)

$$
\dfrac{
\begin{matrix}\lambda\sigma.\sigma(\text{everyone});\\ \boldsymbol{\forall}_{\textbf{person}};\\ S\upharpoonright(S\upharpoonright NP)\end{matrix}
\qquad
\dfrac{
\dfrac{
\begin{matrix}\lambda\sigma.\sigma(\text{someone});\\ \boldsymbol{\exists}_{\textbf{person}};\\ S\upharpoonright(S\upharpoonright NP)\end{matrix}
\quad
\text{②}\,
\dfrac{
\dfrac{
\dfrac{
\begin{bmatrix}\varphi_2;\\ x_2;\\ NP\end{bmatrix}^2
\quad
\dfrac{
\begin{matrix}\text{talked} \circ \text{to};\\ \textbf{talked-to};\\ (NP\backslash S)/NP\end{matrix}
\quad
\begin{bmatrix}\varphi_1;\\ x_1;\\ NP\end{bmatrix}^1
}{
\begin{matrix}\text{talked} \circ \text{to} \circ \varphi_1;\\ \textbf{talked-to}(x_1); NP\backslash S\end{matrix}
}\;/E
}{
\begin{matrix}\varphi_2 \circ \text{talked} \circ \text{to} \circ \varphi_1;\\ \textbf{talked-to}(x_1)(x_2); S\end{matrix}
}\;\backslash E
\quad
\begin{matrix}\text{today};\\ \textbf{tdy};\\ S\backslash S\end{matrix}
}{
\begin{matrix}\varphi_2 \circ \text{talked} \circ \text{to} \circ \varphi_1 \circ \text{today};\\ \textbf{tdy}(\textbf{talked-to}(x_1)(x_2)); S\end{matrix}
}\;\backslash E
}{
\begin{matrix}\lambda\varphi_2.\varphi_2 \circ \text{talked} \circ \text{to} \circ \varphi_1 \circ \text{today};\\ \lambda x_2.\textbf{tdy}(\textbf{talked-to}(x_1)(x_2)); S\upharpoonright NP\end{matrix}
}\;\upharpoonright I^2\ \text{①}
}{
\begin{matrix}\text{someone} \circ \text{talked} \circ \text{to} \circ \varphi_1 \circ \text{today};\\ \boldsymbol{\exists}_{\textbf{person}}(\lambda x_2.\textbf{tdy}(\textbf{talked-to}(x_1)(x_2))); S\end{matrix}
}\;\upharpoonright E
}{
\begin{matrix}\lambda\varphi_1.\text{Someone} \circ \text{talked} \circ \text{to} \circ \varphi_1 \circ \text{today};\\ \lambda x_1.\boldsymbol{\exists}_{\textbf{person}}(\lambda x_2.\textbf{tdy}(\textbf{talked-to}(x_1)(x_2))); S\upharpoonright NP\end{matrix}
}\;\upharpoonright I^1
}{
\begin{matrix}\text{someone} \circ \text{talked} \circ \text{to} \circ \text{everyone} \circ \text{today};\\ \boldsymbol{\forall}_{\textbf{person}}(\lambda x_1.\boldsymbol{\exists}_{\textbf{person}}(\lambda x_2.\textbf{tdy}(\textbf{talked-to}(x_1)(x_2)))); S\end{matrix}
}\;\upharpoonright E
$$

A quantifier has the ordinary GQ meaning ($\boldsymbol{\exists}_{\textbf{person}}$ and $\boldsymbol{\forall}_{\textbf{person}}$ abbreviate the terms $\lambda P.\exists x[\textbf{person}(x) \wedge P(x)]$ and $\lambda P.\forall x[\textbf{person}(x) \rightarrow P(x)]$, respectively), but its phonology is a function of type $(\textbf{st} \to \textbf{st}) \to \textbf{st}$ (where **st** is the type of strings). By abstracting over the position in which the quantifier 'lowers into' in an S via the Vertical Slash Introduction rule (35a), we can obtain an expression of type $S\upharpoonright NP$ (phonologically $\textbf{st} \to \textbf{st}$) (①), which can be given as an argument to the quantifier. Then, by function application via $\upharpoonright E$ (②), the subject quantifier *someone* semantically scopes over the sentence and lowers its phonology to the 'gap' position kept track of by λ-binding in phonology (note that this result obtains by function application and beta-reduction of the prosodic term). The same process takes place for the object quantifier *everyone* to complete the derivation. The scopal relation between multiple quantifiers depends on the order of application of this hypothetical reasoning. The surface scope reading is obtained by switching the order of the hypothetical reasoning for the two quantifiers (which results in the same string of words, but with the opposite scope relation).

This formalization of quantifying-in by Oehrle (1994) has later been extended by Barker (2007) for more complex types of scope-taking phenomena known as PARASITIC SCOPE in the analysis of symmetrical predicates (such as *same* and *different*).[14] Empirical application of parasitic scope include 'respective' readings (Kubota & Levine 2016b), 'split scope' of negative quantifiers (Kubota & Levine 2016a) and modified numerals such as *exactly N* (Pollard 2014).

Hypothetical reasoning with prosodic λ-binding enables a simple analysis of *wh* extraction too, as originally noted by Muskens (2003). The key idea is that sentences with medial gaps can be analyzed as expressions of type $\mathrm{S}\upharpoonright\mathrm{NP}$, as in the derivation for (37) in (38).

(37) Bagels$_i$, Kim gave t_i to Chris.

(38)

$$\dfrac{\begin{array}{l}\textsf{bagels};\\ \mathbf{b};\mathrm{NP}\end{array} \quad ② \; \dfrac{\begin{array}{l}\lambda\sigma\lambda\varphi.\varphi\circ\sigma(\epsilon);\\ \lambda\mathcal{F}.\mathcal{F};\\ (\mathrm{S}\upharpoonright X)\upharpoonright(\mathrm{S}\upharpoonright X)\end{array} \quad ① \; \dfrac{\dfrac{\begin{array}{l}\textsf{kim};\\ \mathbf{k};\mathrm{NP}\end{array} \quad \dfrac{\dfrac{\begin{array}{l}\textsf{gave};\\ \mathbf{gave};\\ \mathrm{VP/PP/NP}\end{array} \quad \begin{bmatrix}\varphi;\\ x;\\ \mathrm{NP}\end{bmatrix}^1}{\textsf{gave}\circ\varphi;\mathbf{gave}(x);\mathrm{VP/PP}}/\mathrm{E} \quad \begin{array}{l}\textsf{to}\circ\textsf{chris};\\ \mathbf{c};\mathrm{PP}\end{array}}{\textsf{gave}\circ\varphi\circ\textsf{to}\circ\textsf{chris};\mathbf{gave}(x)(\mathbf{c});\mathrm{VP}}/\mathrm{E}}{\textsf{kim}\circ\textsf{gave}\circ\varphi\circ\textsf{to}\circ\textsf{chris};\mathbf{gave}(x)(\mathbf{c})(\mathbf{k});\mathrm{S}}\backslash\mathrm{E}}{\begin{array}{l}\lambda\varphi.\textsf{kim}\circ\textsf{gave}\circ\varphi\circ\textsf{to}\circ\textsf{chris};\\ \lambda x.\mathbf{gave}(x)(\mathbf{c})(\mathbf{k});\mathrm{S}\upharpoonright\mathrm{NP}\end{array}}\upharpoonright\mathrm{I}^1}{\lambda\varphi.\varphi\circ\textsf{kim}\circ\textsf{gave}\circ\textsf{to}\circ\textsf{chris};\lambda x.\mathbf{gave}(x)(\mathbf{c})(\mathbf{k});\mathrm{S}\upharpoonright\mathrm{NP}}\upharpoonright\mathrm{E}}{\textsf{bagels}\circ\textsf{kim}\circ\textsf{gave}\circ\textsf{to}\circ\textsf{chris};\mathbf{gave}(\mathbf{b})(\mathbf{c})(\mathbf{k});\mathrm{S}}\upharpoonright\mathrm{E}$$

Here, after deriving an $\mathrm{S}\upharpoonright\mathrm{NP}$, which keeps track of the gap position via the λ-bound variable φ, the topicalization operator fills in the gap with an empty string and concatenates the topicalized NP to the left of the string thus obtained. This way, the difference between 'overt' and 'covert' movement reduces to a lexical difference in the prosodic specifications of the operators that induce them. A covert movement operator throws in some material in the gap position, whereas an overt movement operator 'closes off' the gap with an empty string.

As illustrated above, hypothetical reasoning for the Lambek slashes / and \ and for the vertical slash $\upharpoonright$ have important empirical motivations, but the real strength of a 'hybrid' system like Hybrid TLCG which recognizes both types of slashes is that it extends automatically to cases in which 'directional' and 'non-directional' phenomena interact. A case in point comes from the interaction of nonconstituent coordination and quantifier scope. Examples such as those in

[14] 'Parasitic scope' is a notion coined by Barker (2007) where, in transformational terms, some expression takes scope at LF by parasitizing on the scope created by a different scopal operator's LF movement. In versions of (TL)CG of the sort discussed here, this corresponds to double lambda-abstraction via the order-insensitive slash.

(39) allow for at least a reading in which the shared quantifier outscopes conjunction.[15]

(39) a. I gave a couple of books to Pat on Monday and to Sandy on Tuesday.
b. Terry said nothing to Robin on Thursday or to Leslie on Friday.

I now illustrate how this wide scope reading for the quantifier in NCC sentences like (39) is immediately predicted to be available in the fragment developed so far (Hybrid TLCG actually predicts both scopal relations for all NCC sentences; see Kubota & Levine (2015) for how the distributive scope is licensed). The derivation for (39b) is given in (40).

(40)

$$\dfrac{[\varphi_1; P; \mathrm{VP/PP/NP}]^1 \quad [\varphi_2; x; \mathrm{NP}]^2}{\varphi_1 \circ \varphi_2; P(x); \mathrm{VP/PP}}\,/\mathrm{E}$$

$$\dfrac{\varphi_1 \circ \varphi_2; P(x); \mathrm{VP/PP} \quad \mathsf{to} \circ \mathsf{robin}; \mathbf{r}; \mathrm{PP}}{\varphi_1 \circ \varphi_2 \circ \mathsf{to} \circ \mathsf{robin}; P(x)(\mathbf{r}); \mathrm{VP}}\,/\mathrm{E}$$

$$\dfrac{\varphi_1 \circ \varphi_2 \circ \mathsf{to} \circ \mathsf{robin}; P(x)(\mathbf{r}); \mathrm{VP} \quad \mathsf{on} \circ \mathsf{thursday}; \mathbf{onTh}; \mathrm{VP\backslash VP}}{\varphi_1 \circ \varphi_2 \circ \mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday}; \mathbf{onTh}(P(x)(\mathbf{r})); \mathrm{VP}}\,\backslash\mathrm{E}$$

$$\dfrac{\varphi_1 \circ \varphi_2 \circ \mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday}; \mathbf{onTh}(P(x)(\mathbf{r})); \mathrm{VP}}{\varphi_2 \circ \mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday}; \lambda P.\mathbf{onTh}(P(x)(\mathbf{r})); \mathrm{(VP/PP/NP)\backslash VP}}\,\backslash\mathrm{I}^1$$

$$\dfrac{\varphi_2 \circ \mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday}; \lambda P.\mathbf{onTh}(P(x)(\mathbf{r})); \mathrm{(VP/PP/NP)\backslash VP}}{\mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday}; \lambda x\lambda P.\mathbf{onTh}(P(x)(\mathbf{r})); \mathrm{NP\backslash(VP/PP/NP)\backslash VP}}\,\backslash\mathrm{I}^2$$

$$\dfrac{\begin{array}{l}\mathsf{or};\\ \lambda \mathcal{V}\lambda \mathcal{W}.\mathcal{W} \sqcup \mathcal{V};\\ (X\backslash X)/X\end{array} \quad \begin{array}{c}\vdots\\ \begin{array}{l}\mathsf{to} \circ \mathsf{leslie} \circ \mathsf{on} \circ \mathsf{friday};\\ \lambda x\lambda P.\mathbf{onFr}(P(x)(\mathbf{l}));\\ \mathrm{NP\backslash(VP/PP/NP)\backslash VP}\end{array}\end{array}}{\begin{array}{l}\mathsf{or} \circ \mathsf{to} \circ \mathsf{leslie} \circ \mathsf{on} \circ \mathsf{friday};\\ \lambda \mathcal{W}.\mathcal{W} \sqcup [\lambda x\lambda P.\mathbf{onFr}(P(x)(\mathbf{l}))];\\ \mathrm{(NP\backslash(VP/PP/NP)\backslash VP)\backslash(NP\backslash(VP/PP/NP)\backslash VP)}\end{array}}\,/\mathrm{E}$$

$$\dfrac{\begin{array}{c}\vdots\\ \begin{array}{l}\mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday};\\ \lambda x\lambda P.\mathbf{onTh}(P(x)(\mathbf{r}));\\ \mathrm{NP\backslash(VP/PP/NP)\backslash VP}\end{array}\end{array} \quad \begin{array}{l}\mathsf{or} \circ \mathsf{to} \circ \mathsf{leslie} \circ \mathsf{on} \circ \mathsf{friday};\\ \lambda \mathcal{W}.\mathcal{W} \sqcup [\lambda x\lambda P.\mathbf{onFr}(P(x)(\mathbf{l}))];\\ \mathrm{(NP\backslash(VP/PP/NP)\backslash VP)\backslash(NP\backslash(VP/PP/NP)\backslash VP)}\end{array}}{\begin{array}{l}\mathsf{to} \circ \mathsf{robin} \circ \mathsf{on} \circ \mathsf{thursday} \circ \mathsf{or} \circ \mathsf{to} \circ \mathsf{leslie} \circ \mathsf{on} \circ \mathsf{friday};\\ \lambda x\lambda P\lambda z.\mathbf{onTh}(P(x)(\mathbf{r}))(z) \vee \mathbf{onFr}(P(x)(\mathbf{l}))(z); \mathrm{NP\backslash(VP/PP/NP)\backslash VP}\end{array}}\,\backslash\mathrm{E}$$

[15] Whether the other scopal relation (one in which the quantifier meaning is 'distributed' to each conjunct, as in the paraphrase 'I gave a couple of books to Pat on Monday and I gave a couple of books to Sandy on Tuesday' for (39)) is possible seems to depend on various factors. With downward-entailing quantifiers such as (39b), this reading seems difficult to obtain without heavy contextualization and appropriate intonational cues. See Kubota & Levine (2015) for some discussion.

$$\vdots$$

$$\dfrac{\left[\begin{array}{l}\varphi_3;\\ x;\mathrm{NP}\end{array}\right]^3 \qquad \begin{array}{l}\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda x\lambda P\lambda z.\mathbf{onTh}(P(x)(\mathbf{r}))(z)\vee\\ \quad\mathbf{onFr}(P(x)(\mathbf{l}))(z);\\ \mathrm{NP}\backslash(\mathrm{VP}/\mathrm{PP}/\mathrm{NP})\backslash\mathrm{VP}\end{array}}{\begin{array}{l}\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda P\lambda z.\mathbf{onTh}(P(x)(\mathbf{r}))(z)\vee\\ \quad\mathbf{onFr}(P(x)(\mathbf{l}))(z);\\ (\mathrm{VP}/\mathrm{PP}/\mathrm{NP})\backslash\mathrm{VP}\end{array}}\ \backslash\mathrm{E}$$

$$\dfrac{\begin{array}{l}\textsf{said};\\ \mathbf{said};\\ \mathrm{VP}/\mathrm{NP}/\mathrm{PP}\end{array} \qquad \begin{array}{l}\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda P\lambda z.\mathbf{onTh}(P(x)(\mathbf{r}))(z)\vee\\ \quad\mathbf{onFr}(P(x)(\mathbf{l}))(z);\\ (\mathrm{VP}/\mathrm{PP}/\mathrm{NP})\backslash\mathrm{VP}\end{array}}{\begin{array}{l}\textsf{said}\circ\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda z.\mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(z)\vee\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(z);\mathrm{VP}\end{array}}\ \backslash\mathrm{E}$$

$$\dfrac{\begin{array}{l}\textsf{terry};\\ \mathbf{t};\mathrm{NP}\end{array} \qquad \begin{array}{l}\textsf{said}\circ\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda z.\mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(z)\vee\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(z);\mathrm{VP}\end{array}}{\begin{array}{l}\textsf{terry}\circ\textsf{said}\circ\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(\mathbf{t})\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(\mathbf{t});\mathrm{S}\end{array}}\ \backslash\mathrm{E}$$

$$\dfrac{\begin{array}{l}\textsf{terry}\circ\textsf{said}\circ\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(\mathbf{t})\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(\mathbf{t});\mathrm{S}\end{array}}{\begin{array}{l}\lambda\varphi_3.\textsf{terry}\circ\textsf{said}\circ\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda x.\mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(\mathbf{t})\vee\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(\mathbf{t});\mathrm{S}{\upharpoonright}\mathrm{NP}\end{array}}\ {\upharpoonright}\mathrm{I}^3$$

$$\dfrac{\begin{array}{l}\lambda\sigma.\sigma(\textsf{nothing});\\ \neg\exists_{\mathbf{thing}};\\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP})\end{array} \qquad \begin{array}{l}\lambda\varphi_3.\textsf{terry}\circ\textsf{said}\circ\varphi_3\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\\ \quad\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \lambda x.\mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(\mathbf{t})\vee\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(\mathbf{t});\mathrm{S}{\upharpoonright}\mathrm{NP}\end{array}}{\begin{array}{l}\textsf{terry}\circ\textsf{said}\circ\textsf{nothing}\circ\textsf{to}\circ\textsf{robin}\circ\textsf{on}\circ\textsf{thursday}\circ\textsf{or}\circ\textsf{to}\circ\textsf{leslie}\circ\textsf{on}\circ\textsf{friday};\\ \neg\exists_{\mathbf{thing}}(\lambda x.\mathbf{onTh}(\mathbf{said}(x)(\mathbf{r}))(\mathbf{t})\vee\mathbf{onFr}(\mathbf{said}(x)(\mathbf{l}))(\mathbf{t}));\mathrm{S}\end{array}}\ {\upharpoonright}\mathrm{E}$$

The key point in this derivation is that, via hypothetical reasoning, the string *to Robin on Thursday or to Leslie on Friday* forms a syntactic constituent with a full-fledged meaning assigned to it in the usual way. Then the quantifier takes scope above this whole coordinate structure, yielding the non-distributive, quantifier wide-scope reading.

Licensing the correct scopal relation between the quantifier and conjunction in the analysis of NCC remains a challenging problem in the HPSG literature. See section 4.2.1 for some discussion.

2.5.3 Notes on other variants of TLCG

2.5.3.1 Displacement Calculus and NL$_\lambda$ Among different variant of TLCG, Morrill's Displacement Calculus (Morrill et al. 2011; Morrill 2010) and Barker and Shan's continuation-based calculus NL$_\lambda$ (Barker & Shan 2015) are most closely related to Hybrid TLCG. Roughly speaking, Hybrid TLCG's vertical slash $\upharpoonright$ plays more or less the same role as the discontinuity connectives $\uparrow$ and $\downarrow$ in Displacement Calculus and the 'continuation' slashes $/\!/$ and $\backslash\!\backslash$ in NL$_\lambda$. Many empirical analyses of linguistic phenomena formulated in one of these variants of TLCG translate to the other two more or less straightforwardly (for example, the analyses of Gapping and symmetrical predicates in Kubota & Levine (2016a) and Kubota & Levine (2016b) build on Morrill's and Barker's analyses of the respective phenomena), though there are some important differences (some of which I point out below).

One major difference between the Displacement Calculus and Hybrid TLCG on the one hand and NL_λ on the other is that the latter takes **NL**, namely, the *Non*-associative Lambek Calculus, as the underlying calculus for the directional slashes / and \. Barker & Shan (2015) briefly comment on this property of their system, alluding to the possibility of controlling flexibility of constituency via the notion of 'structural control' (see section 2.5.3.2 below) in Multi-Modal Categorial Type Logics. This certainly is a viable view, but no explicit extension of NL_λ along these lines currently exists. Morrill's approach differs from Kubota & Levine's in certain important ways in the treatment of specific linguistic phenomena. The most substantial disagreement pertains to the treatment of island constraints. Unlike Kubota & Levine, who take (most) island constraints to arise from factors pertaining to on-line language processing, Morrill consistently holds the view that major island constraints are to be treated within the narrow syntax (Morrill 1994; Morrill 2010; 2017). See also a brief discussion about determiner gapping in section 4.2.2 for another point of disagreement.

2.5.3.2 Multi-Modal Categorial Type Logics There is a line of work in the TLCG literature most actively investigated back in the 90s and whose core architecture provides the theoretical underpinnings to Barker and Shan's NL_λ, called Multi-Modal Categorial Type Logics (MMCTL) (Moortgat & Oehrle 1994; Moortgat 2011; Bernardi 2002; Vermaat 2005) (but it should be kept in mind that NL_λ and MMCTL from the 90s differ from each other considerably in terms of empirical scope and orientation). One crucial difference between Hybrid TLCG and the Displacement Calculus on the one hand and MMCTL on the other is that, instead of recognizing a separate level of prosodic representation, MMCTL deals with various (somewhat heterogeneous) phenomena ranging from morpho-syntactic properties of verb clusters in Dutch cross-serial dependencies (Moortgat & Oehrle 1994) to technical difficulties with the Lambek Calculus in dealing with medial extraction (Bernardi 2002; Moortgat 2011) (see the discussion in sections 2.4.2 and 2.5.2 above) via the abstract notion of 'structural control', building on the technique of mixing different kinds of logic within a single deductive system originally developed in the literature of substructural logic (Restall 2000).

To see this point, it is instructive to take a look at the analysis of medial extraction in MMCTL, illustrated by the following derivation (adapted from Bernardi (2002)):

(41)

$$
\begin{prooftree}
\AxiomC{$(\mathrm{N}\backslash\mathrm{N})/(\mathrm{S}/_a\mathrm{NP}) \vdash (\mathrm{N}\backslash\mathrm{N})/(\mathrm{S}/_a\mathrm{NP})$}
\AxiomC{$\mathrm{NP} \vdash \mathrm{NP}$}
\AxiomC{$(\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \vdash (\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP}$}
\AxiomC{$[\ \mathrm{NP} \vdash \mathrm{NP}\ ]^1$}
\RightLabel{$/_a\mathrm{E}$}
\BinaryInfC{$(\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ_a \mathrm{NP} \vdash \mathrm{NP}\backslash\mathrm{S}$}
\AxiomC{$(\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S}) \vdash (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S})$}
\RightLabel{$\backslash\mathrm{E}$}
\BinaryInfC{$((\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ_a \mathrm{NP}) \circ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S}) \vdash \mathrm{NP}\backslash\mathrm{S}$}
\RightLabel{$\backslash\mathrm{E}$}
\BinaryInfC{$\mathrm{NP} \circ (((\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ_a \mathrm{NP}) \circ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S})) \vdash \mathrm{S}$}
\RightLabel{Diss}
\UnaryInfC{$\mathrm{NP} \circ (((\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S})) \circ_a \mathrm{NP}) \vdash \mathrm{S}$}
\RightLabel{MixAssoc}
\UnaryInfC{$(\mathrm{NP} \circ ((\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S}))) \circ_a \mathrm{NP} \vdash \mathrm{S}$}
\RightLabel{$/_a\mathrm{I}^1$}
\UnaryInfC{$\mathrm{NP} \circ ((\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S})) \vdash \mathrm{S}/_a\mathrm{NP}$}
\RightLabel{$/\mathrm{E}$}
\BinaryInfC{$(\mathrm{N}\backslash\mathrm{N})/(\mathrm{S}/_a\mathrm{NP}) \circ ((\mathrm{NP} \circ ((\mathrm{NP}\backslash\mathrm{S})/_a\mathrm{NP} \circ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S}))) \circ_a \mathrm{NP}) \vdash \mathrm{N}\backslash\mathrm{N}$}
\end{prooftree}
$$

Here, the derivation is written in the sequent style natural deduction (see Chapter 2). A sequent is a construct of the form $\Gamma \vdash A$ where the consequent A is a syntactic type and the antecedent Γ is a structured object consisting of possibly multiple syntactic types that are combined with one another via binary connectives (of several sorts). On the leaves of the trees, we find axioms of the form $A \vdash A$ and the sequent on the final line of the proof tree is obtained by applying the rules of logic (such as the Elimination and Introduction rules similar to the ones we have posited above) and 'structural rules', which manipulate the forms of the antecedents of the sequents that are derived.

Intuitively, what is going on in the above derivation is that an object NP that is hypothesized on the right of a transitive verb is 'moved' to the right edge via the structural rule of Mixed Associativity in the structured antecedent on the third line from the bottom. This licenses the $/_a$I rule, withdrawing the hypothesis to create the $\mathrm{S}/_a\mathrm{NP}$ category that is suitable as an argument to the relative pronoun. After lexical substitution, the derivation in (41) embodies, for example, a formal proof for the fact that the string 'which Sara wrote __ there' is a well-formed relative clause in English containing a medial gap.

There are a couple of important points of difference between MMCTL and other variants of TLCG that should already be clear from this rather cursory exposition. First, the analysis of extraction via structure rewriting operations in the abstract substructural component of the type logic is very different from the prosodic λ abstraction analysis presented above, due originally to Muskens (2003) and incorporated in many contemporary variants of TLCG (such as Hybrid TLCG and Pollard's Linear Categorial Grammar (Mihaliček & Pollard 2012; Pollard 2013)). Second, the exact ontological status of this 'multi-modal' substructural component is somewhat unclear. As noted above, the structure-changing operations have been put to use not only for extraction in English but also for dealing with complex morpho-syntactic phenomena such as verb clustering in Dutch cross-serial dependencies (Moortgat & Oehrle 1994) and clitic pronouns in

French (Kraak 1998). All this is done, not in a separate prosodic component, but inside a complex logic of syntactic types, where these syntactic types are taken as primitives that enter into binary composition operations of various sorts having different combinatorial possibilities. The different ontological assumptions in different variants of contemporary TLCG may have to do with the different research goals and research practices. When the emphasis is on linguistic application, a clear separation of ontologically distinct components is an important requirement, but when the emphasis is on studying the meta-logical properties of the formal calculus, building directly on the rich literature of substructural logic and formalizing the type logic for natural language syntax literally *as* a substructural logic may be a natural choice.

2.5.3.3 Linear Categorial Grammar There is a family of related approaches in the TLCG tradition (Oehrle 1994; de Groote 2001; Muskens 2003; Mihaliček & Pollard 2012; Pollard 2013) that are distinctly different from the variants of TLCG discussed above in *not* recognizing the directional slashes of the Lambek calculus as primitive logical connectives. I call this family of approaches Linear Categorial Grammar (LCG) in what follows.[16]

LCG is essentially a subsystem of Hybrid TLCG without the rules for the directional slashes. The original conceptual motivation for this architecture is similar to the motivation for linearization-based HPSG—the idea, originally due to Curry (1961), that separating the combinatoric component of grammar and surface word order realization leads to a cleaner theoretical architecture. To get a flavor of LCG, an analysis of extraction from section 2.5.2 can be reformulated in an LCG fragment as follows:

[16] Among these variants, Abstract Categorial Grammar (ACG) has mostly been studied in the formal grammar literature as a meta-framework for embedding different types of linguistic theories (rather than as a linguistic theory by itself), despite the fact that de Groote (2001), the very first paper that advocates it, motivated it by pointing out some empirical issues with standard variants of TLCG. For a highly readable and interesting recent work on using ACG as a meta-framework for embedding Tree Adjoining Grammar, see Pogodalla (2017).

(42)

$$
\dfrac{
\begin{array}{c}\textsf{bagels};\\ \mathbf{b}; \mathrm{NP}\end{array}
\quad
\dfrac{
\begin{array}{l}\lambda\sigma\lambda\varphi.\\ \quad \varphi \circ \sigma(\epsilon);\\ \lambda\mathcal{F}.\mathcal{F};\\ (\mathrm{S}\upharpoonright X)\upharpoonright(\mathrm{S}\upharpoonright X)\end{array}
\quad
\dfrac{
\dfrac{
\begin{array}{l}\textsf{kim};\\ \mathbf{k};\\ \mathrm{NP}\end{array}
\quad
\dfrac{
\dfrac{
\begin{array}{l}\lambda\varphi_1\lambda\varphi_2\lambda\varphi_3.\\ \quad \varphi_3 \circ \textsf{gave} \circ \varphi_1 \circ \varphi_2;\\ \mathbf{gave};\\ \mathrm{S}\upharpoonright\mathrm{NP}\upharpoonright\mathrm{PP}\upharpoonright\mathrm{NP}\end{array}
\quad
\begin{bmatrix}\varphi;\\ x;\\ \mathrm{NP}\end{bmatrix}^1
}{
\begin{array}{l}\lambda\varphi_2\lambda\varphi_3.\varphi_3 \circ \textsf{gave} \circ \varphi \circ \varphi_2;\\ \mathbf{gave}(x); \mathrm{S}\upharpoonright\mathrm{NP}\upharpoonright\mathrm{PP}\end{array}
}\upharpoonright\mathrm{E}
\quad
\begin{array}{l}\textsf{to} \circ \textsf{chris};\\ \mathbf{c}; \mathrm{PP}\end{array}
}{
\lambda\varphi_3.\varphi_3 \circ \textsf{gave} \circ \varphi \circ \textsf{to} \circ \textsf{chris}; \mathbf{gave}(x)(\mathbf{c}); \mathrm{S}\upharpoonright\mathrm{NP}
}\upharpoonright\mathrm{E}
}{
\textsf{kim} \circ \textsf{gave} \circ \varphi \circ \textsf{to} \circ \textsf{chris}; \mathbf{gave}(x)(\mathbf{c})(\mathbf{k}); \mathrm{S}
}\backslash\mathrm{E}
}{
\lambda\varphi.\textsf{kim} \circ \textsf{gave} \circ \varphi \circ \textsf{to} \circ \textsf{chris}; \lambda x.\mathbf{gave}(x)(\mathbf{c})(\mathbf{k}); \mathrm{S}\upharpoonright\mathrm{NP}
}\upharpoonright\mathrm{I}^1
}{
\lambda\varphi.\varphi \circ \textsf{kim} \circ \textsf{gave} \circ \textsf{to} \circ \textsf{chris}; \lambda x.\mathbf{gave}(x)(\mathbf{c})(\mathbf{k}); \mathrm{S}\upharpoonright\mathrm{NP}
}\upharpoonright\mathrm{E}
}{
\textsf{bagels} \circ \textsf{kim} \circ \textsf{gave} \circ \textsf{to} \circ \textsf{chris}; \mathbf{gave}(\mathbf{b})(\mathbf{c})(\mathbf{k}); \mathrm{S}
}\upharpoonright\mathrm{E}
$$

The key difference from the Hybrid TLCG derivation above in (38) is that here the syntactic categories of linguistic expressions do not specify word order. Note in particular that the ditransitive verb *gave* is syntactically of type $\mathrm{S}\upharpoonright\mathrm{NP}\upharpoonright\mathrm{PP}\upharpoonright\mathrm{NP}$, and specifies the order of the arguments relative to the verb itself in the prosodic representation, which is formally a function of type $\textbf{st}\multimap\textbf{st}\multimap\textbf{st}\multimap\textbf{st}$.

While the clear separation of the underlying combinatorics and surface word order is certainly appealing, the conceptual elegance comes at some nontrivial empirical cost. Specifically, the simple analysis of NCC as constituent coordination discussed in section 2.4.1, which is the hallmark of many variants of CG, does not straightforwardly carry over to this setup. This issue was already anticipated by Muskens (2001), and is discussed in detail in Kubota & Levine (2015). To see what's at issue, note that the distinction between S/NP and NP\S, which is crucial for distinguishing well-formed RNR sentence and ungrammatical strings in the CCG/Lambek calculus analysis of RNR is lost in LCG, since the two will be collapsed into the same type $\mathrm{S}\upharpoonright\mathrm{NP}$. But then, neither the syntactic type nor the prosodic type (which is uniformly $\textbf{st}\multimap\textbf{st}$) of the conjuncts provides enough information to determine conjoinability.

Worth (2016) presents the most comprehensive treatment of coordination in LCG addressing this problem (see also Kanazawa (2015)). The core idea of his proposal is to encode word order information via fine-grained subtypes of the prosodic types of linguistic expressions. Whether this extension of LCG offers the same (or better) empirical coverage as more traditional variants of TLCG (including Hybrid TLCG) is currently an open question.

3 Architectural similarities and differences

3.1 Broad architecture

One important property common to HPSG and CG is that they are both 'lexicalist' theories of syntax in the broader sense.[17] This is partly due to an explicit choice made at an early stage of the development of HPSG to encode valence information in the syntactic categories of linguistic expressions, following CG (see Chapters 2 and 4).[18] The two theories share many similarities in the analyses of specific linguistic phenomena due to this basic architectural similarity. For example, many phenomena that are treated by means of A movement operations (or via empty categories) in mainstream syntax, such as passivization, raising/control in English and 'complex predicate' phenomena in a typologically broad range of languages are generally treated by the sharing of valence information in the lexicon in these theories. For HPSG analyses of these phenomena, see Chapters 9, 12 and 13. Steedman & Baldridge (2011) contains a good summary of CG analyses of local dependencies (passivization, raising/control). Kubota (2014) contains a comparison of HPSG and CG analyses of complex predicates. The heavy reliance on 'lexicalist' analyses of local dependencies is perhaps the most important property that is shared in common in HPSG and various versions of CG.

the chapter label seems currently missing from the Lexicon chapter

But emphasizing this commonality too much may be a bit misleading, since the valence features of HPSG and the slash connectives in CG have very different ontological statuses in the respective theories. The valence features in HPSG are primarily specifications, closely tied to the specific phrase structure rules, that dictate the ways in which hierarchical representations are built. To be sure, the lexical specifications of the valence information play a key role in the movement-free analyses of local dependencies along the lines noted above, but still, there is

[17] I say 'broader sense' here since not all variants of either HPSG or CG subscribe to the so-called 'lexical integrity hypothesis' (see Chapter 4), which says that syntax and morphology are distinct components of grammar. For example, the treatments of verb clustering in Dutch by Moortgat & Oehrle (1994) and in Japanese by Kubota (2014) seem to go against the tenet of the lexical integrity hypothesis. See also Chapter 4 for some discussion on lexicalism.

[18] This point is explicitly noted by the original founders of HPSG in the following passage in Pollard & Sag (1987: 11):

> A third principle of universal grammar posited by HPSG, the Subcategorization Principle, is essentially a generalization of the 'argument cancellation' employed in categorial grammar

a rather tight connection between these valence specifications originating in the lexicon and the ways in which they are 'cancelled' in specific phrase structure rules.

Things are quite different in CG, especially in TLCG. As discussed in section 2, TLCG views the grammar of natural language *not* as a structure-building system, but as a logical deductive system. The two slashes / and \ are thus not 'features' that encode the subcategorization properties of words in the lexicon, but have a much more general and fundamental role within the basic architecture of grammar in TLCG. These connectives are literally implicational connectives within a logical calculus. Thus, in TLCG, 'derived' rules such as type-raising and function composition are *theorems*, in just the same way that the transitivity inference is a theorem in classical propositional logic. Note that this is not just a matter of high-level conceptual organization of the theory, since, as discussed in section 2, the ability to assign 'constituent' statuses to non-canonical constituents in the CG analyses of NCC directly exploits this property of the underlying calculus. The straightforward mapping from syntax to semantics discussed in section 2.3 is also a direct consequence of adopting this 'derivation as proof' perspective on syntax, building on the results of Curry-Howard correspondence (Howard 1969) in setting up the syntax-semantics interface.[19]

Another notable difference between (especially a recent variant of) HPSG and CG is that CG currently lacks a detailed theory of 'constructions', that is, patterns and (sub)regularities that are exhibited by linguistic expressions that cannot (at least according to the proponents of 'constructionist' approaches) be lexicalized easily. As discussed in Chapter P, recent Sign-Based Construction Grammar (SBCG) variants of HPSG (Sag et al. 2012) incorporate ideas from Construction Grammar (Goldberg 1995) and capture such generalizations via a set of constructional templates (or schemata), which are essentially a family of related phrase structure rules that are organized in a type inheritance hierarchy.

Such an architecture seems nearly impossible to implement literally in CG, except via empty operators or lexical operations corresponding to each such constructional schema. In particular, in TLCG, syntactic rules are logical inference rules, so, if we strictly adhere to its slogan 'language is logic', there is no option to freely add syntactic rules in the deductive system. However, the crucial tenet of Construction Grammar in its strongest form, namely, the claim that the

[19] Although CCG does not embody the idea of 'derivation as proof' as explicitly as TLCG does, it remains true to a large extent that the role of the slash connective within the overall theory is different from that of the valence features in HPSG, given that CCG and TLCG share many key ideas in the analyses of actual empirical phenomena.

grammar of natural language robustly exhibits patterns that cannot be perspicuously captured within the lexicon, seems still highly controversial even within the HPSG literature. Thus, if the relevant generalizations can all be lexicalized, along the lines suggested by Müller & Wechsler (2014) (see also Chapter P and Steedman & Baldridge (2011), the latter of which briefly discusses ways in which some of the empirical generalizations that Goldberg (1995) adduces to the notion of constructions can be lexicalized within CCG), then it may not pose any fundamental problem for the strongly lexicalist architecture of CG. That being said, the so-called 'peripheral' phenomena is largely an uncharted territory in CG research waiting for serious attention in future research.

3.2 Syntax-semantics interface

As should be clear from the exposition in section 2, both CCG and TLCG (at least in the simplest form) adopt a very rigid, one-to-one correspondence between syntax and semantics. Steedman's work on CCG has demonstrated that this simple and systematic mapping between syntax and semantics enables attractive analyses of a number of empirical phenomena at the syntax-semantics interface, including some notorious problems such as the scope parallelism issue in right-node raising known as the Geach paradigm (*Every boy loves, and every girl detests, some saxophonist*; Geach (1970)). Other important work on issues at the syntax-semantics interface includes Jacobson's (1999; 2000) work on pronominal anaphora in Variable-Free Semantics (covering a wide range of phenomena including the paycheck/Bach-Peters paradigms and binding parallelism in right-node raising), Barker & Shan's (2015) work on 'continuation-based' semantics (weak crossover, superiority effects and 'parasitic scope' treatments of symmetrical predicates and sluicing) and Kubota and Levine's (2015; 2017; 2020) Hybrid TLCG, dealing with interactions between coordination, ellipsis and scopal phenomena.

As discussed in Chapter C, recent HPSG work on complex empirical phenomena at the syntax-semantics interface makes heavy use of underspecification. For example, major analyses of nonconstituent coordination in recent HPSG use some version of underspecification framework to deal with complex interactions between coordination and scopal operators. (Yatabe 2001; Beavers & Sag 2004; Park et al. 2019; Yatabe & Tam 2019). In a sense, HPSG retains a rigid phrase structure-based syntax (modulo the flexibility entertained with the use of the linearization-based architecture) and deals with the complex mapping to semantics via the use of underspecification languages in the semantic component (such as Minimal Recursion Semantics (Copestake et al. 2005) and Lexical Resources Semantics (Richter & Sailer 2004)). CG, on the other hand, sticks to a tight mapping from syntax to semantics, but makes the syntactic component itself flexible.[20]

3.3 Morpho-syntax and word order

While there is relatively less detailed work on morphology and morpho-syntax interface in CG as compared to HPSG, there are several ideas that have been entertained in the CG literature that have either influenced some HPSG work or which are closely related to a certain line of work in in the HPSG literature. I review some of these in this section.

3.3.1 Linearization-based HPSG and the phenogrammar/tectogrammar distinction in CG

The idea of separating surface word order and the underlying combinatorics, embodied in the so-called LINEARIZATION-BASED version of HPSG (Reape 1994; Kathol 2000; cf. Chapter 10), has its origin in the work by the logician Haskell Curry (Curry 1961), in which he proposed the distinction between PHENOGRAMMAR (the component pertaining to surface word order) and TECTOGRAMMAR (underlying combinatorics). This same idea has influenced a certain line of work in the CG literature too (the most recent incarnation of which is the family of approaches identified as 'Linear Categorial Grammar' in section 2.5.3.3). Important early work was done by Dowty (1982a; 1996b) in a variant of CG which is essentially an AB grammar with 'syncategoremmatic' rules that directly manipulate string representations, of the sort utilized in Montague Grammar, for dealing with various sorts of discontinuous constituency.[21]

Dowty's early work has influenced two separate lines of work in the later development of CG. First, a more formally sophisticated implementation of an enriched theory of phenogrammatical component of the sort sketched in Dowty (1996b) was developed in the Multi-Modal CTL work in the 90s, by exploiting the notion of 'modal control' (see section 2.5.3.2 above; as already noted, this

[20] But note that in principle this architecture does not preclude the use of underspecification. Steedman's (2012) recent approach to quantification crucially employs a limited use of underspecification for the treatment of indefinites. Similarly, Bekki and Mineshima's Dependent Type Semantics (Bekki 2014; Bekki & Mineshima 2017), a version of compositional proof theoretic semantics that has a straightforward interface to CG syntax, makes use of underspecification for the treatment of anaphora.

[21] See also Chapter 2, section 1.2 for a discussion of the influence that early forms of CG (Bach 1979; 1980; Dowty 1982a,b) had on Head Grammar (Pollard 1984), a precursor of HPSG.

technique was later incorporated into CCG by Baldridge (2002)). Some empirical work in this line of research includes Moortgat & Oehrle (1994) (on Dutch cross-serial dependencies; see also Dowty (1996a) for an accessible exposition of this analysis), Kraak (1998) (French clitic climbing), Whitman (2009) ('right-node wrapping' in English) and Kubota (2010; 2014) (complex predicates in Japanese). Second, as already noted in section 2.5.3.3, the Curry/Dowty idea of pheno/tecto distinction has also been the core motivation for the underlying architecture of LCG—in fact, among different variants of CG, LCG can be thought of as an extremism approach in relegating word order completely from the combinatorics, by doing away with the distinction between the Lambek forward and backward slashes.

One issue that arises for approaches that distinguish between the levels of phenogrammar and tectogrammar, across the HPSG/CG divide, is how closely these two components interact with one another. Kubota (2014) discusses some data in the morpho-syntax of complex predicates in Japanese which (according to him) would call for an architecture of grammar in which the pheno and tecto components interact with one another closely, and which would thus be problematic for the simpler LCG-type architecture. It would be interesting to see whether/to what extent this same criticism would carry over to linearization-based HPSG, which is similar (at least in its simplest form) to LCG in maintaining a clear separation of the pheno/tecto components.[22]

3.3.2 Syntactic features and feature neutralization

As compared to HPSG, the status of syntactic features in CG is somewhat unclear, despite the fact that such 'features' are often used in linguistic analyses in the CG literature. One reason that a full-blown theory of syntactic features has not been developed in CG research to date seems to be that as compared to HPSG, syntactic features play a far less major role in linguistic analysis in CG. Another possible reason is that empirical work on complex linguistic phenomena (especially on languages other than English) are still very few in number in CG.

It is certainly conceivable to develop a theory of syntactic features and feature underspecification within CG by borrowing ideas from HPSG, for which there is already a rich tradition of foundational work on this issue. In fact, the work on

[22]But note also in this connection that linearization-based HPSG is by no means monolithic; for example, Yatabe & Tam (2019) (discussed below in section 4.2.1) propose a somewhat radical extension of the linearization-based approach in which semantic composition is done at the level of word order domains.

Unification-based Categorial Grammar (Calder et al. 1988) explored at the end of the 80s seems to have had precisely such a goal. Unfortunately, this approach remains largely isolated from other developments in the literature (of either CG or other grammatical theories/formalisms). Another possibility would be to pursue a more logic-based approach. For some ideas, see Bayer & Johnson (1995); Bayer (1996) and Morrill (1994). Morrill (1994) in particular briefly explores the idea of implementing syntactic features via the notion of dependent types. There is some renewed interest in linguistic application of ideas from Dependent Type Theory (Martin-Löf 1984) in the recent literature of CG and formal semantics (see, for example, Chatzikyriakidis & Luo (2017)), so, pursuing this latter type of approach in connection with this new line of work may lead to some interesting developments.

One issue that is worth noting in connection to syntactic features is the treatment of case syncretism and feature neutralization (cf. Chapter 7, Section 2). The work by Morrill (1994); Bayer (1996); Bayer & Johnson (1995) mentioned above proposed an approach to feature neutralization by positing meet and join connectives (which are like conjunction and disjunction in propositional logic) in CG. The key idea of this approach was recast in HPSG by means of inheritance hierarchies by Levy (2001), Levy & Pollard (2002) and Daniels (2002). See Chapter 7, Section 2 for an exposition of this HPSG work on feature neutralization.

3.3.3 Morphological theory in HPSG and CG

Recent work by Olivier Bonami and Berthold Crysmann (see Chapter B) on developing a morphological theory building on Greg Stump's Paradigm Function Morphology (PFM; Stump (2001)), with extensive use of inheritance hierarchies in HPSG, has enriched the morphological component of HPSG substantially. CG research currently lacks a comparably well developed theory of morphology. But given the similarity between HPSG and CG as lexicalist theories, adopting PFM or its extension (perhaps along the lines of Bonami and Crysmann) seems relatively straightforward. In fact, early work on morphology in CG such as Hoeksema (1984), Moortgat (1984), Hoeksema & Janda (1988) and Raffelsiefen (1992) view morphological processes not as structure-building processes but as functional mapping from one form to another, and in this sense, can be thought of as prefiguring the core ideas of PFM.

A somewhat different (and arguably more controversial) question is whether/to what extent the CG architecture is compatible with the view on morphology that is more in line with the Chomskian tradition, represented by Distributed Morphology (DM; Halle & Marantz (1993)). Given the 'surface-oriented' nature of

CCG, trying to combine CCG with DM is perhaps not a very fruitful idea to pursue, but things are arguably different with TLCG. Aside from the general lexicalist setup (as reflected, for example, in the explicit encoding of valance information in syntactic categories), TLCG is, in a sense, more similar to the mainstream syntax than to either CCG or HPSG. This should be clear from the analogy between hypothetical reasoning and syntactic movement in the exposition of (Hybrid) TLCG in section 2.5. Recent research in TLCG in other domains (primarily, syntax-semantics interface) has in fact shown that the logic-based setup of TLCG enables synthesis of ideas from both lexicalist and derivational approaches to syntax (Kubota 2014; Kubota & Levine 2016a,b; 2019). It remains to be seen whether a similar insight on the relationship between configurational and lexicalist approaches to morphology can be obtained by recasting both in the logic-based combinatorics of TLCG.

4 Specific empirical phenomena

Part II of the present handbook contains an excellent introduction to recent developments of HPSG research on major linguistic phenomena. I will therefore presuppose familiarity with such recent analyses, and my discussion below aims to highlight the differences between HPSG and CG in the analyses of a selected empirical phenomena. In order to make the ensuing discussion maximally informative, I focus on phenomena over which there is some ongoing major cross-theoretical debate, and those for which I believe one or the other theory would benefit from recent developments/rich research tradition in the other.

4.1 Long-distance dependencies

As noted in section 2.4, CCG treats long-distance dependencies via a sequence of function composition, which is similar to the SLASH percolation analysis in HPSG. CCG offers a treatment of major aspects of long-distance dependencies, including island effects (Steedman 2000) and parasitic gaps (Steedman 1987). Earlier versions of CCG involved a somewhat ad-hoc stipulation on the use of crossed composition rules (Steedman 1997). This was overcome in the more recent, 'multi-modal' variant of CCG (Baldridge 2002), which controls the application of such non-order-preserving rules via a fine-grained system of 'lexicalized modality'. The modality specifications in this new version of CCG enables one to relocate language-specific idiosyncrasies to the lexicon, in line with the general spirit of lexicalist theories of grammar.

The situation is somewhat different in TLCG. TLCG typically makes use of a movement-like operation for the treatment of extraction phenomena (via hypothetical reasoning), but the specific implementations differ considerably in different variants of TLCG. Major alternatives include the approach in terms of 'structural control' in Multi-Modal CTL (along the lines sketched in section 2.5.3.2, cf. Bernardi (2002); Moortgat (2011); see also Morrill (1994)), and the one involving prosodic λ-binding in LCG and related approaches (see section 2.5.2). In either approach, extraction phenomena are treated by means of some form of hypothetical reasoning, and this raises a major technical issue in the treatment of multiple gap phenomena. The underlying calculus of TLCG is a version of linear logic, and this means that the implication connective is resource sensitive. This is problematic in situations in which a single filler corresponds to multiple gaps, as in parasitic gaps and related phenomena. These cases of extraction require some sort of extension of the underlying logic or some special operator that is responsible for resource duplication. Currently, the most detailed treatment of extraction phenomena in the TLCG literature is Morrill (2017), which lays out in detail an analysis of long-distance dependencies capturing both major island constraints and parasitic gaps within the most recent version of Morrill's Displacement Calculus.

There are several complex issues that arise in relation to the linguistic analysis of extraction phenomena. One major open question is whether island constraints should be accounted for within narrow grammar. Both Steedman and Morrill follow the standard practice in generative grammar research in taking island effects to be syntactic, but this consensus has been challenged by a new body of research in the recent literature proposing various alternative explanations on different types of island constraints (see Chapter 16, Levine (2017) and Newmeyer (2016) for an overview of this line of work and pointers to the relevant literature). Recent syntactic analyses of long-distance dependencies in the HPSG literature explicitly avoid directly encoding major island constraints within the grammar (Sag 2010; Chaves 2012b). Unlike CCG and Displacement Calculus, Kubota & Levine's Hybrid TLCG opts for this latter type of view (that is, the one that is generally in line with recent HPSG work).

Another major empirical problem related to the analysis of long-distance dependencies is the so-called EXTRACTION PATHWAY MARKING phenomenon (McCloskey 1979; Zaenen 1983). While this issue has received considerable attention in the HPSG literature, through a series of work by Levine and Hukari (see Levine & Hukari 2006), there is currently no explicit treatment of this phenomenon in the CG literature. CCG can probably incorporate the HPSG analysis relatively easily, given the close similarity between the SLASH percolation mechanism and the step-by-step inheritance of the /NP specification in the function composition-based approach in CCG. Extraction pathway marking poses a much trickier challenge to TLCG, in which extraction is typically handled by a single-chain movement-like process by means of hypothetical reasoning (but see Kubota & Levine (2020: Chapter 7) for a sketch of a possible approach which mimics successive cyclic movement in the type-logical setup).

4.2 Coordination and ellipsis

Coordination and ellipsis are both major issues in contemporary syntactic theory. There are moreover some phenomena, such as Gapping and Stripping, which seem to lie at the boundary of the two empirical domains (see for example the recent overview by Johnson (2019)). There are some important similarities and differences between analytic ideas entertained in the HPSG and CG literature for problems in these empirical domains.

4.2.1 Analyses of nonconstituent coordination

CG is perhaps best known in the linguistics literature for its analysis of nonconstituent coordination. Steedman's work on CCG (Steedman 1997; 2000; Steedman 2012) in particular has shown how this analysis of coordination interacts smoothly with analyses of other major linguistic phenomena (such as long-distance dependencies, control and raising and quantification) to achieve a surface-oriented grammar that has wide empirical coverage and at the same time has attractive computational properties. Kubota & Levine (2015; 2020) offer an up-to-date TLCG analysis of coordination, and compare it with major alternatives in both the CCG and HPSG literature.

As compared to long-distance dependencies, coordination (in particular NCC) has received considerably less attention in the (H)PSG literature initially (Sag et al. (1985) is an important exception in the early literature). Things started to change somewhat around 2000, with a series of related proposals appearing one after another, including Yatabe (2001), Beavers & Sag (2004), Chaves (2007) and Crysmann (2008) (see Chapters 17 and 20). Here, I take up Beavers & Sag (2004) and Yatabe (2001) (updated in Yatabe & Tam (2019)) as two representative proposals in this line of work. The two proposals share some common assumptions and ideas, but they also differ in important respects.

Both Beavers & Sag and Yatabe (2001) adopt linearization-based HPSG, together with (a version of) Minimal Recursion Semantics for semantics. Of the two, Beavers & Sag's analysis is more in line with standard assumptions in HPSG. The basic idea of Beavers & Sag's analysis is indeed very simple: by exploiting the flexible mapping between the combinatoric component and the surface word order realization in linearization-based HPSG, they essentially propose a surface deletion-based analysis of NCC according to which NCC examples are analyzed as follows:

(43) [$_S$ Terry gave no man a book on Friday] or [$_S$ ~~Terry gave no man~~ a record on Saturday].

where the material in strike-out is underlyingly present but undergoes deletion in the prosodic representation.

In its simplest form, this analysis gets the scopal relation between the quantifier and coordination wrong in examples like (43) (a well-known problem for the conjunction reduction analysis from the 70s; cf. Partee (1970)). Beavers & Sag address this issue by introducing a condition called 'Optional Quantifier Merger':

(44) **Optional Quantifier Merger**: For any elided phrase denoting a generalized quantifier in the domain of either conjunct, the semantics of that phrase may optionally be identified with the semantics of its non-elided counterpart.

As noted by Levine (2011) and Kubota & Levine (2015), this condition does not follow from any general principle and is merely stipulated in Beavers & Sag's account.

Yatabe (2001) and Yatabe & Tam (2019) (the latter of which contains a much more accessible exposition of essentially the same proposal as the former) propose a somewhat different analysis. Unlike Beavers & Sag, who assume that semantic composition is carried out on the basis of the meanings of *signs* on each node (which is the standard assumption about semantic composition in HPSG), Yatabe shifts the locus of semantic composition to the list of domain objects, that is, the component that directly gets affected by the deletion operation that yields the surface string.

This crucially changes the default meaning predicted for examples such as (43). Specifically, on Yatabe's analysis, the surface string for (43) is obtained by the 'compaction' operation on word order domains that collapses two quantifiers originally contained in the two conjuncts into one. The semantics of the whole sentence is computed on the basis of this resultant word order domain representation, which contains only *one* instance of a domain object corresponding to the quantifier. The quantifier is then required to scope over the whole coordinate structure due to independently motivated principles of underspecification resolution. While this approach successfully yields the wide-scope reading for quantifiers, the distributive, narrow scope reading for quantifiers (which was trivial for Beavers & Sag) now becomes a challenge. Yatabe and Tam simply stipulate a complex disjunctive constraint on semantic interpretation tied to the 'compaction' operation that takes place in coordination so as to generate the two scopal readings.

Kubota & Levine (2015) note that, in addition to the quantifier scope issue noted above, Beavers & Sag's approach suffers from similar problems in the interpretations of symmetrical predicates (*same*, *different*, etc.), summative predicates (*a total of X*, *X in total*, etc.) and the so-called 'respective' readings of plural and conjoined expressions (see Chaves (2012a) for a lucid discussion of the empirical parallels between the three phenomena and how the basic cases can receive a uniform analysis within HPSG). Yatabe & Tam (2019) offer a response to Kubota & Levine, working out explicit analyses of these more complex phenomena in linearization-based HPSG. A major point of disagreement between Kubota & Levine on the other hand and Yatabe & Tam on the other seems to be whether/to what extent an analysis of a linguistic phenomenon should aim to explain (as opposed to merely account for) linguistic generalizations. There is no easy answer to this question, and it is understandable that different theories put different degrees of emphasis on this goal (see also Chapter L for discussion on a related point). Whatever conclusion one draws from this recent HPSG/CG debate on the treatment of nonconstituent coordination, one point seems relatively uncontroversial: coordination continues to constitute a challenging empirical domain for any grammatical theory, consisting of both highly regular patterns (such as systematic interactions with scopal operators (Kubota & Levine 2015; 2020)) and puzzling idiosyncrasies (such as the summative agreement facts (Postal 1998; Yatabe & Tam 2019) and extraposed relative clauses with split antecedents (Perlmutter & Ross 1970; Yatabe & Tam 2019)).

4.2.2 Gapping and Stripping

Descriptively, Gapping is a type of ellipsis phenomenon that occurs in coordination and which deletes some material including the main verb:[23]

[23] There is some disagreement as to whether Gapping is restricted to coordination. Kubota & Levine (2016a), following authors such as Johnson (2009), take Gapping to be restricted to coordination. Park et al. (2019) take a different view, and argue that Gapping should be viewed as a type of ellipsis phenomenon that is not restricted to coordination environments. See

(45) a. Leslie **bought** a CD, and Robin ∅ a book.
b. Terry **can go** with me, and Pat ∅ with you.
c. John **wants to try to begin to write** a novel, and Mary ∅ a play.

Gapping has invoked some theoretical controversy in the recent HPSG/CG literature for the 'scope anomaly' issue that it exhibits. The relevant data involving auxiliary verbs such as (46a) and (46b) have long been known in the literature since Oehrle (1971; 1987) and Siegel (1987). McCawley (1993) later pointed out similar examples involving downward-entailing determiners such as (46c).

(46) a. Mrs. J can't live in Boston and Mr. J ∅ in LA.
b. Kim didn't play bingo or Sandy ∅ sit at home all evening.
c. No dog eats Whiskas or ∅ cat ∅ Alpo.

The issue here is that (46a), for example, has a reading in which the modal *can't* scopes over conjunction ('it's not possible for Mrs. J to live in NY and Mr. J to live in LA at the same time'). This is puzzling, since such a reading wouldn't be predicted on the (initially plausible) assumption that Gapping sentences would be interpreted by simply supplying the meaning of the missing material in the right conjunct.

Kubota & Levine (2014; 2016a) note some difficulties for earlier accounts of Gapping in the (H)PSG literature (Sag et al. 1985; Abeillé et al. 2014) and argue for a constituent coordination analysis of Gapping in TLCG, building on earlier analyses of Gapping in CG (Steedman 1990; Hendriks 1995b; Morrill & Solias 1993). The key idea of Kubota & Levine's analysis involves taking Gapping as coordination of clauses missing a verb in the middle, which can be transparently represented as a function from strings to strings of category $\mathrm{S}{\upharpoonright}((\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP})$:

(47) $\lambda\varphi.\mathsf{leslie} \circ \varphi \circ \mathsf{a} \circ \mathsf{cd}; \lambda R.\exists x.\mathbf{cd}(x) \wedge R(x)(\mathbf{l}); \mathrm{S}{\upharpoonright}((\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP})$

A special type of conjunction entry (prosodically of type $(\mathbf{st}\rightarrow\mathbf{st})\rightarrow(\mathbf{st}\rightarrow\mathbf{st})\rightarrow(\mathbf{st}\rightarrow\mathbf{st})$) then conjoins two such expressions and returns a conjoined sentence missing the verb only in the first conjunct (on the prosodic representation). By feeding the verb to this resultant expression, a proper form-meaning pair is obtained for Gapping sentences like those in (45).

The apparently unexpected wide scope readings for auxiliaries and quantifiers in (46) turn out to be straightforward on this analysis. I refer the interested

Kubota & Levine (2020) for a response to Park et al. (2019).

reader to Kubota & Levine (2016a) for details, but the key idea is that the apparently 'anomalous' scope in such examples isn't really anomalous on this approach, since the auxiliary (which prosodically lowers into the first conjunct) takes the whole conjoined gapped clause as its argument in the combinatoric component underlying semantic interpretation. Thus, the existence of the wide scope reading is automatically predicted. Puthawala (2018) extends this approach to a similar 'scope anomaly' data found in Stripping, in examples such as the following:

(48) John didn't sleep, or Mary (either).

Just like the Gapping examples in (46), this sentence has both wide scope ('neither John nor Mary slept') and narrow scope ('John was the one who didn't sleep, or maybe that was Mary') interpretations for negation.

The determiner gapping example in (46c) requires a somewhat more elaborate treatment. Kubota & Levine (2016a) analyze determiner gapping via higher-order functions. Morrill & Valentín (2017) criticize this approach for a certain type of overgeneration problem regarding word order and propose an alternative analysis in Displacement Calculus.

Park et al. (2019) propose an analysis of Gapping in HPSG that overcomes the limitations of previous (H)PSG analyses of Gapping (Sag et al. 1985; Chaves 2005; Abeillé et al. 2014), couched in Lexical Resources Semantics. In their analysis, the lexical entries of the clause-level conjunction words *and* and *or* are underspecified as to the relative scope between the propositional operator contributed by the modal auxiliary in the first conjunct and the boolean conjunction or disjunction connective that is contributed by the conjunction word itself. Park et al. argue that this is sufficient for capturing the scope anomaly in the Oehrle/Siegel data such as (46a) and (46b). Extension to the determiner gapping case (46c) is left for future work.

Here again, instead of trying to settle the debate, I'd like to draw the reader's attention to the different perspectives on grammar that seem to be behind the HPSG and (Hybrid) TLCG approaches. Kubota & Levine's approach attains theoretical elegance at the cost of employing abstract higher-order operators (both in semantic and prosody). This makes the relationship between the competence grammar and the on-line human sentence processing model indirect, and relatedly, it is likely to make efficient computational implementation less straightforward (for a discussion on the relationship between competence grammar and a model of sentence processing, see Chapter G). Park et al.'s (2019) approach on the other hand is more in line with the usual practice (and the shared spirit) of HPSG research where the main emphasis is on writing an explicit grammar fragment that is constraint-based and surface-oriented. This type of tension is perhaps not easy to overcome, but it seems useful (for researchers working in different grammatical theories) to at least recognize (and appreciate) the existence of these different theoretical orientations tied to different approaches.

4.2.3 Ellipsis

Analyses of major ellipsis phenomena in HPSG and CG share the same essential idea that ellipsis is a form of anaphora, without any invisible hierarchically structured representations corresponding to the 'elided' expression. See Chapter 20 and Ginzburg & Miller (2019) for an overview of approaches to ellipsis in HPSG.

Recent analyses of ellipsis in HPSG (Ginzburg & Sag 2000; Miller 2014) make heavy use of the notion of 'constructions' adopted from Construction Grammar (this idea is even borrowed into some of the CG analyses of ellipsis such as Jacobson (2016)). Many ellipsis phenomena are known to exhibit some form of 'syntactic sensitivity' (Kennedy 2003; Chung 2013; Yoshida et al. 2015), and this fact has long been taken to provide strong evidence for the 'covert structure' analyses of ellipsis popular in the mainstream syntactic literature (Merchant 2019).

Some of the early works on ellipsis in CG include Hendriks (1995a) and Morrill & Merenciano (1996). Morrill & Merenciano (1996) in particular show how hypothetical reasoning in TLCG allows treatments of important properties of ellipsis phenomena such as strict/sloppy ambiguity and scope ambiguity of elided quantifiers in VP ellipsis. Jäger (2005) integrates these earlier works with a general theory of anaphora in TLCG, incorporating the key empirical analyses of pronominal anaphora by Jacobson (1999; 2000). Jacobson's 1998; 2008 analysis of Antecedent-Contained Ellipsis is also important. Antecedent-Contained Ellipsis is often taken to provide a strong piece of evidence for the representational analysis of ellipsis in mainstream generative syntax. Jacobson offers a counterproposal to this standard analysis that completely dispenses with covert structural representations. While the above works from the 90s have mostly focused on VP ellipsis, recent developments in the CG literature, including Barker (2013) on sluicing, Jacobson (2016) on fragment answers and Kubota & Levine (2017) on pseudogapping, considerably extended the empirical coverage of the same line of analysis.

The relationship between recent CG analyses of ellipsis and HPSG counterparts seems to be similar to the situation with competing analyses on coordination. Both Barker (2013) and Kubota & Levine (2017) exploit hypothetical reasoning to treat the antecedent of an elided material as a 'constituent' with full-fledged semantic interpretation at an abstract combinatoric component of syntax. The anaphoric mechanism can then refer to both the syntactic and semantic information of the antecedent expression to capture syntactic sensitivity observed in ellipsis phenomena, without the need to posit hierarchical representations at the ellipsis site. Due to its surface-oriented nature, HPSG is not equipped with an analogous abstract combinatoric component that assigns 'constituent' statuses to expressions that do not (in any obvious sense) correspond to constituents in the surface representation. In HPSG, the major work in restricting the possible form of ellipsis is instead taken over by constructional schemata, which can encode syntactic information of the antecedent to capture connectivity effects, as is done, for example, with the use of the SAL-UTT feature in Ginzburg & Sag's (2000) analysis of sluicing (cf. Chapter 20).

Kubota & Levine (2020: Chapter 8) extend Kubota & Levine's (2017) approach further to the treatment of interactions between VP ellipsis and extraction, which has often been invoked in the earlier literature (in particular, Kennedy (2003)) as providing crucial evidence for covert structure analysis of ellipsis phenomena (see also Jacobson (2019) for a related proposal, cast in a variant of CCG). At least some of the counterproposals that Kubota & Levine formulate in their argument against the covert structure analysis seem to be directly compatible with the HPSG approach to ellipsis, but (so far as I am aware) no concrete analysis of extraction/ellipsis interaction currently exists in the HPSG literature.

4.2.4 Mismatches in right-node raising

While RNR has mostly been discussed in connection to coordination in the literature, it is well-known that RNR is not necessarily restricted to coordination environments (see, for example, Wilder (2019) for a recent overview). Moreover, it has recently been pointed out by Abeillé et al. (2016) that RNR in French admit certain types of syntactic mismatch between the RNR'ed material and the selecting head in a non-adjacent conjunct. The current literature seems to agree that RNR is not a unitary phenomenon, and that at least some type of RNR should be treated via a mechanism of surface ellipsis, which could be modelled as deletion of syntactic (or prosodic) objects or via some sort of anaphoric mechanism (cf. Chapter 20, Chaves (2014), Kubota & Levine (2017)).

One point that is worth emphasizing in this connection is that while the 'NCC as constituent coordination' analysis of RNR in CG discussed in section 2.4.1 (major evidence for which comes from the interactions between various sorts of scopal operators and RNR as noted in section 4.2.1) is well-known, neither CCG nor TLCG is by any means committed to the idea that *all* instances of RNR should be analyzed this way. In fact, given the extensive evidence for the non-unitary nature of RNR reviewed in Chaves (2014) and the syntactic mismatch data from French offered by Abeillé et al. (2016), it seems that a comprehensive account of RNR in CG (or, for that matter, in any other theory) would need to recognize the non-unitary nature of the phenomenon, along lines similar to Chaves's (2014) recent proposal in HPSG. While there is currently no detailed comprehensive account of RNR along these lines in the CG literature, there does not seem to be any inherent obstacle for formulating such an account.

4.3 Binding

Empirical phenomena that have traditionally been analyzed by means of Binding Theory (both in the transformational and the nontransformational literature; cf. Chapter 21) potentially pose a major challenge to the 'non-representational' view of the syntax-semantics interface common to most variants of CG. The HPSG Binding Theory in Pollard & Sag (1994) captures Principles A and B at the level of argument structure, while Principle C makes reference to the configurational structure (i.e. the feature-structure encoding of the constituent geometry). The status of Principle C itself is controversial to begin with, but if this condition needs to be stated in the syntax, it would possibly constitute one of the the greatest challenges to CG-based theories of syntax, since, unlike phrase structure trees, the proof trees in CG are not objects that a principle of grammar can directly refer to.

While there seems to be no consensus in the current CG literature on how the standard binding theoretic facts are to be accounted for, there are some important ideas and proposals in the wider literature of CG-based syntax (broadly construed to include work in the Montague Grammar tradition). First, as for Principle A, there is a recurrent suggestion in the literature that these effects can (and should) be captured simply via strictly lexical properties of reflexive pronouns (e.g. Szabolcsi (1992); see Büring (2005) for a concise summary). For example, for a reflexive in the direct object position of a transitive verb bound by the subject NP, the following type assignment (where the reflexive pronoun first takes a transitive verb and then the subject NP as arguments) suffices to capture its bound status:

(49) himself; $\lambda R \lambda x.R(x)(x)$; $\mathrm{NP}\backslash((\mathrm{NP}\backslash \mathrm{S})/\mathrm{NP})\backslash \mathrm{S}$

This approach is attractively simple, but there are at least two things to keep in mind, in order to make it a complete analysis of Principle A in CG. First, while this lexical treatment of reflexive binding may at first sight appear to capture the locality of binding quite nicely, CG's flexible syntax potentially overgenerates unacceptable long-distance binding readings for (English) reflexives. Since right-node raising can take place across clause boundaries, it seems necessary to assume that hypothetical reasoning for the Lambek-slash (or a chain of function composition that has the same effect in CCG) can generally take place across clause boundaries. But then, expressions such as *thinks Bill hates* can be assigned the same syntactic type (i.e. (NP\S)/NP) as lexical transitive verbs, overgenerating non-local binding of a reflexive from a subject NP in the upstairs clause (**John$_i$ thinks Bill hates himself$_i$*).

In order to prevent this situation while still retaining the lexical analysis of reflexivization sketched above, some kind of restriction needs to be imposed as to the way in which reflexives combine with other linguistic expressions. One possibility would be to distinguish between lexical transitive verbs and derived transitive verb-like expressions by positing different 'modes of composition' in the two cases in a 'multi-modal' version of CG.

The other issue is that the lexical entry in (49) needs to be generalized to cover all cases in which a reflexive is bound by an argument that is higher in the obliqueness hierarchy. This amounts to positing a polymorphic lexical entry for the reflexive. The use of polymorphism is not itself a problem, since it is needed in other places in the grammar (such as coordination) anyway. But this account would amount to capturing the Principle A effects purely in terms of the specific lexical encoding for reflexive pronouns (unlike the treatment in HPSG which explicitly refers to the obliqueness hierarchy).

While Principle A effects are in principle amenable to a relatively simple lexical treatment along lines sketched above, Principle B turns out to be considerably more challenging for CG. To see this point, note that the lexical analysis of reflexives sketched above crucially relies on the fact that the constraint associated with reflexives corresponds to a straightforward semantic effect of variable binding. Pronouns instead require *disjointness* of reference from less oblique co-arguments, but such an effect cannot be captured by simply specifying some appropriate lambda term as the semantic translation for the pronoun.

To date, the most detailed treatment of Principle B effects in CG that explicitly addresses this difficulty is the proposal by Jacobson (2007), formulated in a version of CCG (Steedman (1997) proposes a different approach to binding, which will be briefly discussed at the end of this section). The key idea of Jacobson's account of Principle B effects is that NPs are divided by a binary-valued feature ±p, with pronouns marked NP[+p] and all other NPs NP[−p]. In all lexical entries of the form in (50), all NP (and PP) arguments in any realization of /\$ are specified as [−p].

(50) k; P; VP/\$

The effect of this restriction is to rule out pronouns from argument positions of verbs with ordinary semantic denotations. On this approach, the only way a lexically specified functional category can take [+p] arguments is via the application of the following irreflexive operator:[24]

(51) $\lambda\varphi.\varphi$; $\lambda f \lambda u \lambda v. f(u)(v)$, $u \neq v$; (VP/NP[+p])↾(VP/NP[-p])

The greyed-in part $u \neq v$ separated from the truth conditional meaning by a comma is a presupposition introduced by the pronoun-seeking variant of the predicate. It says that the subject and object arguments are forced to pick out different objects in the model. For the semantics of pronouns themselves, we can assume, following the standard practice, that free (i.e. unbound) pronouns are simply translated as arbitrary variables (cf. Cooper 1979).

Crucially, the operator in (51) is restricted in its domain of application to the set of signs which are specified in the lexicon. We notate this restriction by using the dashed line notion in what follows. Then (52) will be derived as in (53).

(52) John praises him.

[24]For expository purposes, I state the operator in (51) in its most restricted form, dealing with only the case where there is a single syntactic argument apart from the subject. A much broader coverage is of course necessary in order to handle cases like the following:

(i) a. *John$_i$ warned Mary about him$_i$.
  b. *John talked to Mary$_i$ about her$_i$.
  c. *John explained himself$_i$ to him$_i$.

What we need in effect is a schematic type specification that applies to a pronoun in any or all argument positions, i.e., stated on an input of the form VP/\$/XP[−p]/\$ to yield an output of the form VP/\$/XP[+p]/\$. To ensure the correct implementation of this extension, some version of the 'wrapping' analysis needs to be assumed (cf. Jacobson 2007), so that the order of the arguments in verbs' lexical entries is isomorphic to the obliqueness hierarchy (of the sort discussed by Pollard & Sag (1992)).

Cases such as the following also call for an extension (also a relatively straightforward one):

(ii) *John$_i$ is proud of him$_i$.

By assuming (following Jacobson 2007) that the [±p] feature percolates from NPs to PPs and by generalizing the irreflexive operator still further so that it applies not just to VP/XP[−p] but AP/XP[−p] as well, the ungrammaticality of (ii) follows straightforwardly.

(53)
$$\dfrac{\dfrac{\dfrac{\begin{array}{l}\lambda\varphi.\varphi;\\ \lambda f\lambda u\lambda v.f(u)(v), u \neq v;\\ (\mathrm{VP/NP[+p]})\upharpoonright(\mathrm{VP/NP[-p]})\end{array} \quad \begin{array}{l}\textsf{praises};\\ \mathbf{praise}; \mathrm{VP/NP[-p]}\end{array}}{\textsf{praises}; \lambda u\lambda v.\mathbf{praise}(u)(v), u \neq v; \mathrm{VP/NP[+p]}} \quad \textsf{him}; z; \mathrm{NP[+p]}}{\textsf{praises} \circ \textsf{him}; \lambda v.\mathbf{praise}(z)(v), z \neq v; \mathrm{VP}} \quad \textsf{john}; \mathbf{j}; \mathrm{NP[-p]}}{\textsf{john} \circ \textsf{praises} \circ \textsf{him}; \mathbf{praise}(z)(\mathbf{j}), z \neq \mathbf{j}; \mathrm{S}}$$

The presupposition $z \neq \mathbf{j}$ ensures that the referent of the pronoun is different from John.

Thus, Jacobson's approach captures the relevant conditions on the interpretation of pronouns essentially as a type of lexical presupposition tied to the denotation of the pronoun-taking verb, and the syntactic feature [±p] mediates the distributional correlation between the pronoun and the verb that subcategorizes for it. The idea is essentially the same as in the HPSG Binding Theory, except that the relevant condition is directly encoded as a restriction on the denotation itself since the standard CG syntax-semantics interface does not admit of syntactic indices of the sort assumed in HPSG.

Unlike Jacobson's proposal outlined above, Steedman's (1997) analysis of binding conditions in CCG recognizes the syntactic forms of the logical language that is used to write the denotations of linguistic expressions as the 'level' at which binding conditions are stated. This approach can be thought of as a 'compromise' which enabes a straightforward encoding of the HPSG-style Binding Conditions by (slightly) deviating from the CG doctrine of not admitting any representational object at the syntax-semantics interface (see Dowty (1996a) for a critique of Steedman's (1997) approach to binding discussing this issue clearly).

Steedman's approach can be best illustrated by taking a look at the analysis of (54).[25]

(54) Every student$_i$ praised him$_i$.

According to Steedman, pronouns receive translations of the form $\mathbf{pro}(x)$, where **pro** is effectively a term that marks the presence of (the translation of) a pronoun at some particular syntactic position in the logical formula that represents the meaning of the sentence.

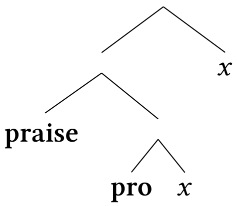
With this assumption, the translation for (54) that needs to be ruled out (via Principle B) is as follows:

[25] At the same time that he formulates an essentially syntactic account of Principle B via the term **pro** in the translation language, Steedman (1997) briefly speculates on the (somewhat radical) possibility of relegating Principle B entirely to the pragmatic component of pronominal anaphora resolution. However, the relevant discussion is rather sketchy, and the details of such a pragmatic alternative are not entirely clear.

(55) $\forall x[\mathbf{student}(x) \rightarrow \mathbf{praise}(\mathbf{pro}(x))(x)]$

And this is where the CCG Binding Theory kicks in. The relevant part of the structure of the logical formula in (55) can be more perspicuously written as a tree as in (56), which makes clear the hierarchical relation between sub-terms.

(56)



Principle B states that pronouns need to be locally free. (56) violates this condition since there is a locally c-commanding term x that binds $\mathbf{pro}(x)$ (where we say that a term α binds term β when they are semantically bound by the same operator).

Principles A and C are formulated similarly by making crucial reference to the structures of the terms that represent the semantic translations of sentences.

What we can see from the comparison of different approaches to binding in CG and the treatment of binding in HPSG is that although HPSG and CG are both lexicalist theories of syntax, and there is a general consensus that binding conditions are to be formulated lexically rather than configurationally, there are important differences in the actual implementations of the conditions between approaches that stick to the classical Montagovian tradition (embodying the tenet of 'direct compositionality' in Jacobson's terms) and those that make use of representational devices more liberally. This seems to be related to the tendency of HPSG and (Steedman's) CCG to make more liberal use of underspecified semantic representations as opposed to other variants of CG that generally eschew such intermediate representations.

Finally, some comments are in order regarding the status of Principle C, the part of Binding Theory that is supposed to rule out examples such as the following:

(57) a. *He$_i$ talked to John$_i$.
b. *He$_i$ talked to John's$_i$ brother.

The formulation of Principle C has always been a problem in lexicalist theories of syntax. While Principles A and B can be stated by just making reference to the local argument structure of a predicate in the lexicon, the global nature of Principle C seems to require looking at the whole configurational structure of the sentence in which the proper noun appears. In fact, Pollard & Sag (1994) opt for this solution, and their definition of the Principle C has a somewhat exceptional status within the whole theory (which otherwise adheres to strict locality conditions) in directly referring to the configurational structure.

Essentially the same problem arises in CG. Steedman's (1997) formulation of Principle C can be thought of as an analog of Pollard & Sag's (1994) proposal, where global reference to hierarchical structure is made not at the level of phrase structure, but instead at the level of 'logical structure', that is, in the syntactic structure of the logical language used for writing the meanings of natural language expressions. As already noted above, if one takes the Motagovian, or 'direct compositional' view of the syntax-semantics interface that is more traditional/standard in CG research, this option is unavailable.

Thus, Principle C has a somewhat cumbersome place within lexicalist theories in general. However, unlike Principles A and B, the status of Principle C in the grammar is still considerably unclear and controversial to begin with. If, as suggested by some authors (cf., e.g., Levinson (1987; 1991)), the effects of Principle C can be accounted for by pragmatic principles, that would remove one major sticking point in both HPSG and CG formulations of the Binding Theory.

5 A brief note on processing and implementation

The discussion above has mostly focused on linguistic analysis. In this final section, I will briefly comment on implications for psycholinguistics and computational linguistics research.

One attractive feature of CCG (but not CG in general), when viewed as an integrated model of the competence grammar and human sentence processing, is that it enables 'surface-oriented', incremental analyses of strings from left to right. This aspect was emphasized in the early literature of CCG (Ades & Steedman 1982; Crain & Steedman 1985), but it does not seem to have had much impact on psycholinguistic research in general since then. A notable exception is the work by Pickering & Barry (1991; 1993) in early 90s. There is also some work on the relationship between processing and TLCG (see Morrill (2010: Chapters 9 and 10), and references therein). In any event, a serious investigation of the relationship between competence grammar and human sentence processing from a CG perspective (either CCG or TLCG) seems to be a research topic that is waiting to be explored, much like the situation with HPSG (see Chapter G).

As for connections to computational linguistics/NLP research, like HPSG, large-scale computational implementation has been an important research agenda for CCG (see, e.g., Clark & Curran (2007)). I refer the reader to Steedman (2012: Chapter 13) for an excellent summary on this subject (this chapter contains a discussion of human sentence processing as well). Together with work on linguistically informed 'deep parsing' in HPSG, CCG parsers seem to be attracting some renewed interest in CL/NLP research recently, due to the new trend of combining the insights of statistical approaches and linguistically-informed approaches. In particular, the straightforward syntax-semantics interface of (C)CG seems to be an attractive feature in building CL/NLP systems that have an explicit logical representation of meaning. See Lewis & Steedman (2013) and Mineshima et al. (2016) for this type of work. TLCG research has traditionally been less directly related to CL/NLP research. But there are recent attempts at constructing large-scale treebanks (Moot 2015) and combining TLCG frameworks with more mainstream approaches in NLP research such as distributional semantics (Moot 2018).

6 Conclusion

As should be clear from the above discussion, HPSG and CG share many important similarities, mainly due to the fact that they are both variants of 'lexicalist' syntactic theories. This is particularly clear in the analyses of local dependencies in terms of lexically encoded argument structure information. Important differences emerge once we turn our attention to less canonical types of phenomena, such as noncanonical types of coordination (nonconstituent coordination, Gapping) and the treatment of 'constructional' patterns that are not easily lexicalizable. In general, HPSG has a richer and more comprehensive treatment of various empirical phenomena, whereas CG has a lot to offer to grammatical theory (perhaps somewhat paradoxically) due to the very fact that the potentials of the logic-based perspective it embodies has not been explored in full detail. It is more likely than not that the two will continue to develop as distinct theories of natural language syntax (and semantics). I hope that the discussion in the present chapter has made it clear that there are still many occasions for fruitful interactions between the two approaches both at the level of specific analytic ideas for specific phenomena and at the more general, foundational level pertaining to the core architecture of grammatical theory.

Acknowledgements

[acknowledgments to be added]

References

Abeillé, Anne, Gabriela Bîlbîie & Francois Mouret. 2014. A Romance perspective on Gapping constructions. In H. Boas & F. Gonzálvez García (eds.), *Romance in construction grammar*, 227–265. Amsterdam: John Benjamins.

Abeillé, Anne, Berthold Crysmann & Aoi Shiraıshi. 2016. Syntactic mismatch in french peripheral ellipsis. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, 1–30. Paris: CSSP.

Ades, Anthony E. & Mark J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4(4). 517–558.

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1. 1–27.

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 4(10). 515–531.

Bach, Emmon. 1980. In defense of passive. *Linguistics and Philosophy* (3). 297–342.

Baldridge, Jason. 2002. *Lexically specified derivational control in Combinatory Categorial Grammar*. University of Edinburgh dissertation.

Bar-Hillel, Yehoshua. 1953. A quasi-arithmetic notation for syntactic descriptions. *Language* 29. 47–58.

Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30(4). 407–444.

Barker, Chris. 2013. Scopability and sluicing. *Linguistics and Philosophy* 36(3). 187–223.

Barker, Chris & Chung-chieh Shan. 2015. *Continuations and natural language*. Oxford: Oxford University Press.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72(3). 579–616.

Bayer, Samuel & Mark Johnson. 1995. Features and agreement. In Hans Uszkoreit (ed.), *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 70–76. Cambridge, MA: Association for Computational Linguistics.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational*

Linguistics, Katholieke Universiteit Leuven, 48–69. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

Bekki, Daisuke. 2014. Representing anaphora with dependent types. In Nicholas Asher & Sergei Soloviev (eds.), *Logical Aspects of Computational Linguistics 2014*, 14–29. Heidelberg: Springer.

Bekki, Daisuke & Koji Mineshima. 2017. Context-passing and underspecification in dependent type semantics. In Stergios Chatzikyriakidis & Zhaohui Luo (eds.), *Modern perspectives in type-theoretical semantics*, 11–41. Heidelberg: Springer.

Bernardi, Raffaella. 2002. *Reasoning with polarity in Categorial Type Logic*. University of Utrecht dissertation.

Borsley, Robert D. & Kersti Börjars (eds.). 2011. *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*. Oxford, UK/ Cambridge, MA: Blackwell Publishers Ltd.

Büring, Daniel. 2005. *Binding theory*. Cambridge, England: Cambridge University Press.

Calder, Jonathan, Ewan Klein & Henk Zeevat. 1988. Unification categorial grammar: a concise, extendable grammar for natural language processing. In *Proceedings of the 12th conference on computational linguistics*, vol. 1, 83–86. Stroudsburg, PA, USA: Association for Computational Linguistics.

Carpenter, Bob. 1998. *Type-logical semantics*. Cambridge, MA: MIT Press.

Chatzikyriakidis, Stergios & Zhaohui Luo (eds.). 2017. *Modern perspectives in type-theoretical semantics*. Heidelberg: Springer.

Chaves, Rui. 2012a. Conjunction, cumulation and respectively readings. *Journal of Linguistics* 48(2). 297–344.

Chaves, Rui. 2012b. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30(2). 465–512.

Chaves, Rui P. 2014. On the disunity of right-node raising phenomena: Extraposition, ellipsis, and deletion. *Language* 4. 834–886. DOI:10.1353/lan.2014.0081

Chaves, Rui Pedro. 2005. A linearization-based approach to gapping. In G. Jäger, P. Monachesi, G. Penn & S. Wintner (eds.), *FG-MOL 2005: the 10th conference on Formal Grammar and the 9th meeting on mathematics of language*, 207–220. University of Edinburgh, Scotland.

Chaves, Rui Pedro. 2007. *Coordinate structures - constraint-based syntax-semantics processing*. Portugal: University of Lisbon dissertation. http://www.clul.ul.pt/clg/diss_rpc.pdf.

Chung, Sandra. 2013. Syntactic identity in sluicing: how much, and why. *Linguistic Inquiry* 44(1). 1–44.

Clark, Stephen & James Curran. 2007. Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics* 33(4). 493–552.

Cooper, Robin. 1979. The interpretation of pronouns. In Frank Heny & Helmut S. Schnelle (eds.), *Selections from the third groningen round table*, vol. 10 (Syntax and Semantics), 61–92. New York: Academic Press.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Crain, Stephen & Mark J. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language processing*, 320–358. Cambridge, UK: Cambridge University Press.

Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, 47–62. Stanford, CA: CSLI Publications.

Curry, Haskell B. 1961. Some logical aspects of grammatical structure. In Roman Jakobson (ed.), *Structure of language and its mathematical aspects*, vol. 12 (Symposia on Applied Mathematics), 56–68. Providence: American Mathematical Society.

Daniels, Michael W. 2002. On a type-based analysis of feature neutrality and the coordination of unlikes. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 137–147. Stanford, CA: CSLI Publications.

de Groote, Philippe. 2001. Towards abstract categorial grammars. In *Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter*, 148–155.

Dowty, David. 1996a. Non-constituent coordination, wrapping, and multimodal categorial grammars. Expanded draft of Aug 96 paper from International Congress of Logic, Methodology, and Philosophy of Science in M. L. Dalla Chiara et al. (eds.) *Structures and Norms in Science*, 347–368, 1997.

Dowty, David R. 1982a. Grammatical relations and Montague Grammar. In Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic representation* (Synthese Language Library 15). Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1982b. More on the categorial analysis of grammatical relations. In Annie Zaenen (ed.), *Subjects and other subjects: Proceedings of the Harvard Conference on the Representation of Grammatical Relations*, 115–153. Bloomington: Indiana University Linguistics Club.

Dowty, David R. 1988. Type raising, functional composition, and nonconstituent coordination. In Richard Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial Grammars and natural language structures*, 153–198. Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1996b. Toward a Minimalist theory of syntactic structure. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 11–62. Berlin: Mouton de Gruyter. Published version of a Ms. dated January 1990.

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.

Ginzburg, Jonathan & Philip Miller. 2019. Ellipsis in head-driven phrase structure grammar. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The oxford handbook of ellipsis*, 75–121. Oxford: OUP.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning, and use of English interrogatives* (CSLI Lecture Notes 123). Stanford, CA: CSLI Publications.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: The University of Chicago Press.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20*, 111–176. Cambridge, Mass.: MIT Press.

Hendriks, Petra. 1995a. *Comparatives and categorial grammar*. University of Groningen dissertation.

Hendriks, Petra. 1995b. Ellipsis and multimodal categorial type logic. In Glyn V. Morrill & Richard T. Oehrle (eds.), *Formal Grammar: Proceedings of the Conference of the European Summer School in Logic, Language and Information, Barcelona, 1995*, 107–122.

Hoeksema, Jack. 1984. *Categorial morphology*. Rijksuniversiteit Groningen dissertation.

Hoeksema, Jack & Richard D. Janda. 1988. Implications of process-morphology for categorial grammar. In Richard T. Oehrle, Emmon Bach & Deirdre Wheeler (eds.), *Categorial grammars and natural language structures*, 199–247. Dordrecht: D. Reidel Publishing Company.

Howard, William A. 1969. The formulae-as-types notion of construction. Published in Jonathan P. Seldin and J. Roger Hindley (eds.) 1980. *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism*, pp. 479–490. Academic Press.

Jacobson, Pauline. 1998. Antecedent contained deletion and pied-piping: evidence for a variable-free semantics. In Devon Strolovitch & Aaron Lawson (eds.), *Pro-*

ceedings from Semantics and Linguistic Theory VIII, 74–91. Ithaca, New York: Cornell University.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22(2). 117–184.

Jacobson, Pauline. 2000. Paycheck pronouns, Bach-Peters sentences, and variable-free semantics. *Natural Language Semantics* 8(2). 77–155.

Jacobson, Pauline. 2007. Direct compositionality and variable-free semantics: the case of "Principle B" effects. In Chris Barker & Pauline Jacobson (eds.), *Direct compositionality*, 191–236. Oxford: Oxford University Press.

Jacobson, Pauline. 2008. Direct compositionality and variable-free semantics: the case of Antecedent Contained Deletion. In Kyle Johnson (ed.), *Topics in ellipsis*, 30–68. Cambridge, England: Cambridge University Press.

Jacobson, Pauline. 2016. The short answer: implications for direct compositionality (and vice versa). *Language* 92(2). 331–375.

Jacobson, Pauline. 2019. Ellipsis in categorial grammar. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The oxford handbook of ellipsis*, 122–141. Oxford: OUP.

Jäger, Gerhard. 2005. *Anaphora and Type-Logical Grammar*. Berlin: Springer.

Johnson, Kyle. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry* 40(2). 289–328.

Johnson, Kyle. 2019. Gapping and stripping. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 562–604. Oxford: OUP.

Kanazawa, Makoto. 2015. Syntactic features for regular constraints and an approximation of directional slashes in Abstract Categorial Grammars. In Yusuke Kubota & Robert Levine (eds.), *Proceedings for ESSLLI 2015 Workshop 'Empirical Advances in Categorial Grammar'*, 34–70. University of Tsukuba & Ohio State University.

Kathol, Andreas. 2000. *Linear syntax*. New York, Oxford: Oxford University Press.

Kennedy, Christopher. 2003. Ellipsis and syntactic representation. In Kerstin Schwabe & Susanne Winkler (eds.), *The interfaces: deriving and interpreting omitted structures*, 29–53. Amsterdam: John Benjamins.

Kraak, Esther. 1998. A deductive account of French object clitics. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax*, vol. 30 (Syntax and Semantics), 271–312. San Diego: Academic Press.

Kubota, Yusuke. 2010. *(In)flexibility of constituency in Japanese in Multi-Modal Categorial Grammar with Structured Phonology*. Ohio State University dissertation.

Kubota, Yusuke. 2014. The logic of complex predicates. English. *Natural Language and Linguistic Theory* 32(4). 1145–1204. DOI:10.1007/s11049-014-9246-8

Kubota, Yusuke. 2015. Nonconstituent coordination in Japanese as constituent coordination: an analysis in Hybrid Type-Logical Categorial Grammar. *Linguistic Inquiry* 46(1). 1–42.

Kubota, Yusuke & Robert Levine. 2014. Scope anomaly of Gapping. In *Proceedings of NELS 44*, 247–260. GLSA.

Kubota, Yusuke & Robert Levine. 2015. Against ellipsis: arguments for the direct licensing of 'non-canonical' coordinations. *Linguistics and Philosophy* 38(6). 521–576.

Kubota, Yusuke & Robert Levine. 2016a. Gapping as hypothetical reasoning. *Natural Language and Linguistic Theory* 34(1). 107–156.

Kubota, Yusuke & Robert Levine. 2016b. The syntax-semantics interface of 'respective' predication: a unified analysis in Hybrid Type-Logical Categorial Grammar. *Natural Language and Linguistic Theory* 34(3). 911–973.

Kubota, Yusuke & Robert Levine. 2017. Pseudogapping as pseudo-VP ellipsis. *Linguistic Inquiry* 48(2). 213–257.

Kubota, Yusuke & Robert Levine. 2019. Modal auxiliaries and negation: a type-logical account. In *WoLLIC 2019: logic, language, information, and computation*, 415–432. Heidelberg: Springer.

Kubota, Yusuke & Robert Levine. 2020. *Coordination, ellipsis and Type-Logical Categorial Grammar*. Cambridge, Mass.: MIT Press.

Levine, Robert D. 2011. Linearization and its discontents. Presented at the 18th International Conference on HPSG, Seattle, Washington. USA.

Levine, Robert D. 2017. *Syntactic analysis: an HPSG-based approach*. Cambridge, England: Cambridge University Press.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions* (CSLI Lecture Notes 166). Stanford, CA: CSLI Publications.

Levinson, Stephen C. 1987. Pragmatics and the grammar of anaphora. *Journal of Linguistics* 23. 379–434.

Levinson, Stephen C. 1991. Pragmatic reduction of the Binding Conditions revisited. *Journal of Linguistics* 27. 107–161.

Levy, Roger. 2001. Feature indeterminacy and the coordination of unlikes in a totally well-typed HPSG. http://www.mit.edu/~rplevy/papers/feature-indet.pdf, accessed 2018-7-29. Unpublished manuscript.

Levy, Roger & Carl Pollard. 2002. Coordination and neutralization in HPSG. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 221–234. Stanford, CA: CSLI Publications.

Lewis, Mike & Mark Steedman. 2013. Combined distributional and logical semantics. *Transactions of the Association for Computational Linguistics* 1. 179–192.

Martin, Scott. 2013. *The dynamics of sense and implicature*. Ohio State University dissertation.

Martin-Löf, Per. 1984. *Intuitionistic type theory*. Bibliopolis.

McCawley, James D. 1993. Gapping with shared operators. In David A. Peterson (ed.), *Berkeley Linguistics Society*, 245–253. Berkeley, California.

McCloskey, James. 1979. *Transformational syntax and model-theoretic semantics*. Dordrecht: Reidel.

Merchant, Jason. 2019. Ellipsis: a survey of analytical approaches. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 19–45. Oxford: OUP.

Mihaliček, Vedrana & Carl Pollard. 2012. Distinguishing phenogrammar from tectogrammar simplifies the analysis of interrogatives. In Philippe de Groote & Mark-Jan Nederhof (eds.), *Formal Grammar 2010/2011*, 130–145. Heidelberg: Springer.

Miller, Philip. 2014. A corpus study of pseudogapping and its theoretical consequences. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, vol. 10, 73–90.

Mineshima, Koji, Ribeka Tanaka, Pascual Martínez-Gómez, Yusuke Miyao & Daisuke Bekki. 2016. Building compositional semantics and higher-order inference system for a wide-coverage Japanese CCG parser. In *Proceedings of EMNLP 2016*, 2236–2242. Austin, Texas: Association for Computational Linguistics.

Moortgat, Michael. 1984. A Fregean restriction on metarules. In *Proceedings of NELS 14*, 306–325. GLSA.

Moortgat, Michael. 2011. Categorial type logics. In Johan F. A. K. van Benthem & G. B. Alice ter Meulen (eds.), *Handbook of logic and language*, 2nd edn., 95–179. Amsterdam: Elsevier.

Moortgat, Michael & Richard T. Oehrle. 1994. Adjacency, dependence, and order. In Paul Dekker & Martin Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium*, 447–466. Universiteit van Amsterdam: Instituut voor Taal, Logica, en Informatica.

Moot, Richard. 2015. A type-logical treebank for French. *Journal of Language Modelling* 3(1). 229–264.

Moot, Richard. 2018. Combining logical and distributional methods in type-logical grammars. *Journal of Language Modelling* 6(2). 287–317.

Moot, Richard & Christian Retoré. 2012. *The logic of categorial grammars: a deductive account of natural language syntax and semantics*. Heidelberg: Springer.

Morrill, Glyn. 2010. *Categorial grammar: logical syntax, semantics, and processing*. Oxford: Oxford University Press.

Morrill, Glyn. 2017. Grammar logicised: relativisation. *Linguistics and Philosophy* 40(2). 119–163.

Morrill, Glyn V. 1994. *Type Logical Grammars: Categorial logic of signs*. Dordrecht: Kluwer Academic Publishers.

Morrill, Glyn & Josep-Maria Merenciano. 1996. Generalizing discontinuity. *Traitement Automatique des Langues* 27(2). 119–143.

Morrill, Glyn & Teresa Solias. 1993. Tuples, discontinuity, and gapping in categorial grammar. In *Proceedings of the Sixth Conference on European Chapter of the Association for Computational Linguistics*, 287–296. Morristown, NJ, USA: Association for Computational Linguistics.

Morrill, Glyn & Oriol Valentín. 2017. A reply to Kubota and Levine on gapping. *Natural Language and Linguistic Theory* 35(1). 257–270.

Morrill, Glyn, Oriol Valentín & Mario Fadda. 2011. The displacement calculus. *Journal of Logic, Language and Information* 20(1). 1–48.

Müller, Stefan & Stephen Mark Wechsler. 2014. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Muskens, Reinhard. 2001. Categorial Grammar and Lexical-Functional Grammar. In Miriam Butt & Tracy Holloway King (eds.), *The Proceedings of the LFG '01 Conference*. University of Hong Kong. http://csli-publications.stanford.edu/LFG/6/lfg01.html.

Muskens, Reinhard. 2003. Language, lambdas, and logic. In Geert-Jan Kruijff & Richard Oehrle (eds.), *Resource sensitivity in binding and anaphora*, 23–54. Dordrecht: Kluwer.

Newmeyer, Frederick J. 2016. Nonsyntactic explanations of island constraints. *Annual Review of Linguistics* 2. 187–210.

Oehrle, R. T. 1988. Multi-dimensional compositional functions as a basis for grammatical analysis. In R. T. Oehrle, E. Bach & D. Wheeler (eds.), *Categorial grammars and natural language structures*, 349–389. Dordrecht: Reidel.

Oehrle, Richard T. 1971. On Gapping. MS., MIT.

Oehrle, Richard T. 1987. Boolean properties in the analysis of gapping. In Geoffrey J. Huck & Almerindo E. Ojeda (eds.), *Syntax and semantics: discontinuous constituency*, vol. 20, 203–240. Academic Press.

Oehrle, Richard T. 1994. Term-labeled categorial type systems. *Linguistics and Philosophy* 17(6). 633–678.

Oehrle, Richard T. 2011. Multi-Modal Type-Logical Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 225–267. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Park, Sang-Hee, Jean-Pierre Koenig & Rui P. Chaves. 2019. A semantic underspecification-based analysis of scope ambiguities in Gapping. In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti & Louise McNally (eds.), *Proceedings of Sinn und Bedeutung 23*, vol. 2, 237–251.

Partee, Barbara Hall. 1970. Negation, conjunction, and quantifiers: syntax vs. semantics. *Foundations of Language* 6. 153–165.

Partee, Barbara & Mats Rooth. 1983. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*, 361–383. Berlin: Walter de Gruyter.

Perlmutter, David M. & John Robert Ross. 1970. Relative clauses with split antecedents. *Linguistic Inquiry*. 350.

Pickering, Martin & Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6. 229–264.

Pickering, Martin & Guy Barry. 1993. Dependency Categorial Grammar and coordination. *Linguistics* 31(5). 855–902.

Pogodalla, Sylvain. 2017. A syntax-semantics interface for Tree-Adjoining Grammars through Abstract Categorial Grammars. *Journal of Language Modelling* 5(3). 527–605.

Pollard, Carl. 2013. Linear categorial grammar. MS., Lecture Notes at ESSLLI 2013.

Pollard, Carl. 2014. What numerical determiners mean: a non-ambiguity analysis. Talk presented at the Workshop on Semantics of Cardinals, Ohio State University, March 6, 2014.

Pollard, Carl J. 1984. *Generalized Phrase Structure Grammars, Head Grammars, and natural language*. Stanford University dissertation.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1992. Anaphors in English and the scope of Binding Theory. *Linguistic Inquiry* 23(2). 261–303.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Postal, Paul M. 1998. *Three investigations of extraction.* Cambridge, Mass.: MIT Press.

Puthawala, Daniel. 2018. Stripping isn't so mysterious, or anomalous scope, either. In Annie Foret, Greg Kobele & Sylvain Pogodalla (eds.), *Formal Grammar 2018*, 102–120.

Raffelsiefen, Renate. 1992. A nonconfigurational approach to morphology. In Mark Aronoff (ed.), *Morphology now*, 133–162. State University of New York Press.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Restall, Greg. 2000. *An introduction to substructural logics*. London: Routledge.

Richter, Frank & Manfred Sailer. 2004. Basic concepts of lexical resource semantics. In Arnold Beckmann & Norbert Preining (eds.), *ESSLLI 2003 – Course material I* (Collegium Logicum 5), 87–143. Wien: Kurt Gödel Society.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing Sign-Based Construction Grammar. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 1–29. Stanford, CA: CSLI Publications.

Sag, Ivan A., Gerald Gazdar, Thomas Wasow & Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language and Linguistic Theory* 3(2). 117–171.

Siegel, Muffy A. 1987. Compositionality, case, and the scope of auxiliaries. *Linguistics and Philosophy* 10(1). 53–75.

Steedman, Mark. 2012. *Taking scope*. Cambridge, Mass.: MIT Press.

Steedman, Mark J. 1985. Dependency and coordination in the grammar of Dutch and English. *Language* 61(3). 523–568.

Steedman, Mark J. 1987. Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory* 5(3). 403–439. DOI:10.1007/BF00134555

Steedman, Mark J. 1990. Gapping as constituent coordination. *Linguistics and Philosophy* 13(2). 207–265.

Steedman, Mark J. 1997. *Surface structure and interpretation* (Linguistic Inquiry Monographs 30). Cambridge, MA: MIT Press.

Steedman, Mark J. 2000. *The syntactic process* (Language, Speech, and Communication). Cambridge, MA: MIT Press.

Steedman, Mark & Jason Baldridge. 2011. Combinatory Categorial Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 181–224. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure* (Cambridge Studies in Linguistics 93). Cambridge University Press.

Szabolcsi, Anna. 1992. Combinatory grammar and projection from the lexicon. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters* (CSLI Lecture Notes 24), 241–268. Stanford: CSLI Publications.

van Benthem, Johan. 1988. The semantics of variety in categorial grammar. In Wojciech Buszkowski, Witold Marciszewski & Johan van Benthem (eds.), *Categorial grammar*, 37–55. Amsterdam: John Benjamins.

Vermaat, Willemijn. 2005. *The logic of variation. a cross-linguistic account of wh-question formation*. Utrecht University dissertation.

Whitman, Neal. 2009. Right-node wrapping: multimodal categorial grammar and the 'friends in low places' coordination. In Erhard Hinrichs & John Nerbonne (eds.), *Theory and evidence in semantics*, 235–256. Stanford: CSLI Publications.

Wilder, Chris. 2019. Conjunction reduction and right-node raising. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford handbook of ellipsis*, 681–718. Oxford: OUP.

Worth, Christopher. 2016. *English coordination in Linear Categorial Grammar*. Ohio State University dissertation.

Yatabe, Shûichi. 2001. The syntax and semantics of Left-Node Raising in Japanese. In Daniel P. Flickinger & Andreas Kathol (eds.), *Proceedings of the HPSG-2000 Conference, University of California, Berkeley*, 325–344. CSLI Publications. http://csli-publications.stanford.edu/HPSG/1/, accessed 2018-2-25.

Yatabe, Shûichi & Wai Lok Tam. 2019. In defense of an HPSG-based theory of non-constituent coordination: a reply to Kubota and Levine. MS., available at http://ling.auf.net/lingbuzz/003152.

Yoshida, Masaya, Tim Hunter & Michael Frazier. 2015. Parasitic gaps licensed by elided syntactic structure. *Natural Language and Linguistic Theory* 33(4). 1439–1471.

Zaenen, Annie. 1983. On syntactic binding. *Linguistic Inquiry* 14(3). 469–504.

Chapter N

HPSG and Lexical Functional Grammar

Doug Arnold
University of Essex

Change epigram in chapters/03.tex
or remove it there

Change the abstract in chapters/03.tex Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse

Doug Arnold. 2019. HPSG and Lexical Functional Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1071–1072. Berlin: Language Science Press. DOI:??

rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter O

HPSG and Dependency Grammar

Dick Hudson
London

Put abstract here with \abstract.

1 Two centuries of syntactic theory

In the early 19th century, European grammar was still dominated by the Latin grammar of Priscian which focused on individual words, their morphosyntactic properties and their relations (controlled especially by government and agreement); grammars and grammatical theory were mainly focused on school pedagogy, where the dominant model was the parsing of individual words. But these ideas, and especially government, defined 'dependency' relations holding most words together. The exception was the relation between the verb and its subject, which was still described in terms of the dominant classical logic based on the subject-predicate split. Putting these two traditions together, grammarians produced a mixed theory of sentence structure and a number of diagramming systems to represent such structures – most famously, the diagramming system invented in the USA by Reed & Kellogg (1877) (and still taught in the 21st century in some American schools). This is also the theory that Bloomfield brought back to the USA from Germany, and which he developed into Immediate Constituent analysis (which later turned into phrase-structure analysis); as in the earlier theory, the subject and predicate were equal, in contrast with other 'endocentric' constructions. Bloomfield combined this mixed theory with Wundt's theory of cognition, with the sentence as the 'whole' which defines its parts (and the word no longer in prime position), which allowed a consistent geometry, but phrase-structure trees did not appear till the middle of the 20th century. Meanwhile, however, both Humboldt and Grimm had suggested that the verb was

Dick Hudson. 2019. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1073–1076. Berlin: Language Science Press. DOI:??

the sole head of the sentence, with the subject as one of its dependents, and by the 1860s and 1870s, grammarians in Hungary, Russia and Germany (apparently working independently) were arguing for this view, half a century before it was formalised by Tesniere and named 'dependency analysis'. The first 'stemma' diagram appeared (in Hungary) in 1873. Another 19th-century reaction against classical logic was the logical tradition started (in Germany) by Frege, who may have learned to draw stemmas at school; this tradition gave rise (in Poland) to categorial grammar, which some (including Chomsky) see as a version of dependency analysis. One outcome of this history was the present-day geographical split between American phrase structure (PS) and European dependency structure (DS). Variations on the dependency theme Unsurprisingly, therefore, dependency theory has had more impact on Europeans than on Americans. The general idea of word-word dependencies was built into a number of different theoretical packages which combined it with other ideas, notably multiple levels (the Russian Mel'cuk) and information structure (the Czechs Sgall and Hajicova). However, dependency structure has also been popular internationally in natural-language processing (represented perhaps most notably by the Stanford Parser). 'Plain-vanilla' versions of DS and PS are very similar and are weakly equivalent, but as with phrase structure, such theories need to be supplemented, giving rise to theories in which structures are much richer. One such theory is Word Grammar (WG), which is probably closer to HPSG than any of the other DS theories. In WG, a word is allowed to depend on more than one other word (like re-entrance in HPSG) and dependencies are combined with extra mechanisms for coordination and for word order. This theory will be the main point of comparison with HPSG in the rest of the chapter.

2 Signs, constructions and levels

The contrast between PS and DS is orthogonal to choices about the number of levels (syntax, morphology, etc) and how they are related, but of course these choices are essential for any theoretical package. As in PS theories, different DS theories assume different answers, but Word Grammar takes a rather conservative position in which syntax is distinct both from morphology and from semantics. This view is hard to reconcile with the claim that language consists of 'constructions' or 'signs', both of which assume a direct link between 'form' and 'meaning'. In this view, units of phonological 'form' are only indirectly linked to units of meaning. Approaches which evoke 'signs' or 'constructions' can also be challenged for their conservative assumptions about plain-vanilla surface PS.

Arguably, DS is a better basis for capturing the fine detail of idiosyncratic constructions since these always involve individual lexical items linked by dependencies, and typically focus on just one dependent of a given lexeme rather than on entire multi-dependent phrases. Networks WG takes the whole of language (not just the lexicon) to be a gigantic network, which is a step further than HPSG (where PS rules are outside the network); the network is also not assumed to be a DAG because mutual dependency is allowed. One of the characteristics of network analyses is the central role of relation types (i.e. HPSG attributes). According to WG, but not HPSG, these types form a typed hierarchy which parallels the typed hierarchy of non-relational 'entities' such as words, phonemes and so on; and in both hierarchies, properties are inherited by (a special formalisation of) default inheritance. One of the consequences of this treatment of relations is that, just like entities, they can freely be created and learned as required, so there is no need to assume a universal hard-wired reservoir of relations. This is particularly helpful in DS, where dependencies are typed but different languages require different classifications and distinctions. Word order Another similarity between WG and HPSG is in the treatment of word order. In both theories, dominance (i.e. daughterhood in HPSG and dependency in WG) is separated from linear precedence. In WG, a word's position is treated as one of the word's property's linked to a second property ('landmark'), the word from which it takes its position; the word's landmark is normally the word on which it depends, but exceptions are allowed in cases such as extraction and pied piping. The landmark relation allows a treatment of pied piping which avoids the feature-percolation of HPSG.

3 Words, nodes and semantic phrases

The final topic is the Achille's heel of DS: the completely flat structures where a word has two or more dependents. This is problematic in DS (but not, of course, in HPSG) in examples such as typical French house, meaning 'typical for a French house', because there is no syntactic node that could carry the meaning 'French house'. Current WG provides a solution which moves WG in the direction of PS by distinguishing types from tokens, and then distinguishing 'sub-tokens' of tokens. In this analysis, the token house is distinct not only from the type HOUSE, but also from the sub-token house' which is modified by the dependent French, which in turn is distinct from house'' modified by typical. Sub-tokens are very similar in function to the phrases of HPSG but arguably not quite equivalent.

Abbreviations

Acknowledgements

References

Höhle, Tilman N. 1997. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Christa Dürscheid, Karl Heinz Ramers & Monika Schwarz (eds.), *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, 107–120. Tübingen: Max Niemeyer Verlag. Republished as Höhle (2018).

Höhle, Tilman N. 2018. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 417–433. Berlin: Language Science Press. First published as Höhle (1997). DOI:10.5281/zenodo.1145680

Reed, Alonzo & Brainerd Kellogg. 1877. *Higher lessons in English: A work on English grammar and composition*. Projekt Gutenberg. http://www.gutenberg.org/ebooks/7188, accessed 2018-3-23.

Chapter P

HPSG and Construction Grammar

Stefan Müller

Humboldt-Universität zu Berlin

This chapter discusses the main tenets of Construction Grammar (CxG) and shows that HPSG adheres to them. This discussion includes surface orientation, language acquisition without UG, and inheritance networks and shows how HPSG (and other frameworks) are positioned along these dimensions. Formal variants of CxG will be briefly discussed and their relation to HPSG will be pointed out. It is argued that lexical representations of valence are more appropriate than phrasal approaches, which are assumed in most variants of CxG. Other areas of grammar seem to require headless phrasal constructions (e.g., the NPN construction and certain extraction constructions) and it is shown how HPSG handles these. Derivational morphology is discussed as a further example of an early constructionist analysis in HPSG.

This chapter deals with Construction Grammar (CxG) and its relation to Head-Driven Phrase Structure Grammar (HPSG). The short version of the message is: HPSG is a Construction Grammar.[1] It had constructional properties right from the beginning and over the years – due to influence by Construction Grammarians like Fillmore and Kay – certain aspects were adapted, making it possible to better capture generalizations over phrasal patterns. In what follows I will first say what Construction Grammars are (Section 1), and I will explain why HPSG as developed in Pollard & Sag (1987; 1994) was a Construction Grammar and how it was changed to become even more Constructive (Section 1.2.3). Section 2 deals with so-called argument structure constructions, which are usually dealt with by

[1]This does not mean that HPSG is not a lot of other things at the same time. For instance, it is also a Generative Grammar in the sense of Chomsky (1965: 4), that is, it is explicit and formalized. HPSG is also very similar to Categorial Grammar (Müller 2013; Kubota 2019, Chapter M of this volume). Somewhat ironically, Head-Driven Phrase Structure Grammar is not entirely head-driven anymore (see Section 4.1), nor is it a phrase structure grammar (Richter 2019, Chapter 3 of this volume).

Stefan Müller. 2019. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1077–1128. Berlin: Language Science Press. DOI:??

assuming phrasal constructions in CxG, and explains why this is problematic and why lexical approaches are more appropriate. Section 3 explains Construction Morphology, Section 4 shows how cases that should be treated phrasally can be handled in HPSG, and Section 5 sums up the paper.

1 What is Construction Grammar?

Construction Grammar was developed as a theory that can account for non-regular phenomena as observed in many idioms (Fillmore, Kay & O'Connor 1988). It clearly set itself apart from theories like Government & Binding (Chomsky 1981), which assumes very abstract schemata for the combination of lexical items ($\overline{\text{X}}$ rules). The argument was that grammatical constructions are needed to capture irregular phenomena and their interaction with more regular ones. In contrast, Chomsky (1981: 7) considered rules for passive or relative clauses as epiphenomenological; everything was supposed to follow from general principles.[2] According to Chomsky, grammar consisted of a set of general combinatorical rules and some principles. The Minimalist Program (Chomsky 1995) is even more radical, since only two combinatorial rules are left (External and Internal Merge). Various forms of CxG object to this view and state that several very specific phrasal constructions are needed in order to account for language in its entirety and full complexity. Phenomena for which this is true will be discussed in Section 4. However, the case is not as clear in general, since one of the authors of Fillmore, Kay & O'Connor (1988) codeveloped a head-driven, lexical theory of idioms that is entirely compatible with the abstract rules of Minimalism (**KM2017a**; Sag 2007; Kay, Sag & Flickinger 2015). This theory will be discussed in Section 1.4.2.1. Of course, the more recent lexical theory of idioms is a Constructional theory as well. So the first question to answer in a chapter like this is: what is a construction in the sense of Construction Grammar? What is Construction Grammar? While it is relatively clear what a Construction is, the answer to the question regarding Construction Grammar is less straight-forward (see also Fillmore 1988: 35 on this). Section 1.1 provides the definition for the term *Construction* and Section 1.2 states the tenets of CxG and discusses to what extent the main frameworks cur-

[2] The passive in GB is assumed to follow from suppression of case assignment and the Case Filter, which triggers movement of the object to SpecIP. The important part of the analysis is the combination of the verb stem with the passive morphology. This is where suppression of case assignment takes place. This morphological part of the analysis corresponds to the passive construction in theories like HPSG and SBCG: a lexical rule. So in a sense there is a Passive Construction in GB as well.

rently on the market adhere to them.

1.1 The notion Construction

Fillmore, Kay & O'Connor (1988) discuss sentences like (1) and notice that they pose puzzles for standard accounts of the syntax and the syntax/semantics interface.

(1) a. The more carefully you do your work, the easier it will get.
    b. I wouldn't pay five dollars for it, let alone ten dollars.

The *the -er the -er* Construction is remarkable, since it combines aspects of normal syntax (clause structure and extraction) with idiosyncratic aspects like the special use of *the*. In (1a) the adverb phrase *more carefully* does not appear to the left of *work* but is fronted and *the* appears without a noun. The second clause in (1a) is structured in a parallel way. There have to be two of these *the* clauses to form the respective construction. Fillmore, Kay & O'Connor (1988) extensively discuss the properties of *let alone*, which are interesting for syntactic reasons (the fragments following *let alone*) and for information structural reasons. I will not repeat the discussion here but refer the reader to the paper.

In later papers, examples like (2) were discussed:

(2) a. What is this scratch doing on the table? (Kay & Fillmore 1999: 3)
    b. Frank dug his way out of prison. (Goldberg 1995: 199)

Again, the semantics of the complete sentences is not in an obvious relation to the material involved. The question in (2a) is not about actions of a scratch, but rather the question is why there is a scratch. Similarly, (2b) is special in that there is a directional PP that does not normally go together with verbs like *dug*. It is licensed by *way* in combination with a possessive pronoun.

Fillmore et al. (1988), Goldberg (1995), Kay & Fillmore (1999) and Construction Grammarians in general argue that the notion of Construction is needed for adequate models of grammar, that is, for models of grammar that are capable of analyzing the examples above. Fillmore et al. (1988: 501) define Construction as follows:

> Constructions on our view are much like the nuclear family (mother plus daughters) subtrees admitted by phrase structure rules, EXCEPT that (1) constructions need not be limited to a mother and her daughters, but may span wider ranges of the sentential tree; (2) constructions may specify, not only

> syntactic, but also lexical, semantic, and pragmatic information; (3) lexical items, being mentionable in syntactic constructions, may be viewed, in many cases at least, as constructions themselves; and (4) constructions may be idiomatic in the sense that a large construction may specify a semantics (and/or pragmatics) that is distinct from what might be calculated from the associated semantics of the set of smaller constructions that could be used to build the same morphosyntactic object. (Fillmore et al. 1988: 501)

A similar definition can be found in Goldberg's work. Goldberg (2006: 5) defines Construction as follows:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (Goldberg 2006: 5)

The difference between this definition and earlier definitions by her and others is that patterns that are stored because of their frequencies are included. This addition is motivated by psycholinguistic findings that show that forms may be stored even though they are fully regular and predictable (Bybee 1995; Pinker & Jackendoff 2005: 228).

Goldberg provides Table 1 with examples for Constructions. In addition to

Table 1: Examples of constructions, varying in size and complexity according to Goldberg (2009)

Word	e.g., *tentacle*, *gangster*, *the*
Word (partially filled)	e.g., *post*-N, V-*ing*
Complex word	e.g., *textbook*, *drive-in*
Idiom (filled)	e.g., *like a bat out of hell*
Idiom (partially filled)	e.g., *believe* <one's> *ears/eyes*
Covariational Conditional	The Xer the Yer (e.g., *The more you watch the less you know*)
Ditransitive	Subj V Obj1 Obj2 (e.g., *She gave him a kiss*; *He fixed her some fish tacos*.)
Passive	Subj aux VPpp (PPby) (e.g., *The cell phone tower was struck by lightning*.)

such constructions with a clear syntax-semantics or syntax-function relation, Goldberg (2013: 453) assumes a rather abstract VP construction specifying “statistical constraints on the ordering of postverbal complements, dependent on weight and information structure”.

If one just looks at Goldberg’s definition of Construction, all theories currently on the market could be regarded as Construction Grammars. As Peter Staudacher pointed out in the discussion after a talk by Knud Lambrecht in May 2006 in Potsdam, lexical items are form-meaning pairs and the rules of phrase structure grammars come with specific semantic components as well, even if it is just functional application. So, Categorial Grammar, GB-style theories paired with semantics (Heim & Kratzer 1998), GPSG, TAG, LFG, HPSG, and even Minimalism would be Construction Grammars. If one looks at the examples of Constructions in Table 1, things change a bit. Idioms are generally not the focus of work in Mainstream Generative Grammar (MGG).[3] MGG is usually concerned with explorations of the so-called Core Grammar as opposed to the Periphery, to which the idioms are assigned. The Core Grammar is the part of the grammar that is supposed to be acquired with help of innate domain specific knowledge, something whose existence Construction Grammar denies. But if one takes Hauser, Chomsky & Fitch (2002) seriously and assumes that only the ability to form complex linguistic objects out of less complex linguistic objects (Merge) is part of this innate knowledge, then the core/periphery distinction does not have much content and after all, Minimalists could adopt a version of Sag’s local, selection-based analysis of idioms (Sag 2007; Kay et al. 2015**KM2017a**). However, as is discussed in the next subsection, there are other aspects that really set Construction Grammar apart from MGG.

1.2 Basic tenets of Construction Grammar

Goldberg (2003) names the following tenets as core assumptions standardly made in CxG:

Tenet 1 All levels of description are understood to involve pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully abstract phrasal patterns. (See Table 1.)

[3] The term *Mainstream Generative Grammar* is used to refer to work in Transformational Grammar, for example Government & Binding (Chomsky 1981) and Minimalism (Chomsky 1995). Some authors working in Construction Grammar see themselves in the tradition of Generative Grammar in a wider sense, see for example Fillmore, Kay & O’Connor (1988: 501) and Fillmore (1988: 36).

Tenet 2 An emphasis is placed on subtle aspects of the way we conceive of events and states of affairs.

Tenet 3 A "what you see is what you get" approach to syntactic form is adopted: no underlying levels of syntax or any phonologically empty elements are posited.

Tenet 4 Constructions are understood to be learned on the basis of the input and general cognitive mechanisms (they are constructed), and are expected to vary cross-linguistically.

Tenet 5 Cross-linguistic generalizations are explained by appeal to general cognitive constraints together with the functions of the constructions involved.

Tenet 6 Language-specific generalizations across constructions are captured via inheritance networks much like those that have long been posited to capture our non-linguistic knowledge.

Tenet 7 The totality of our knowledge of language is captured by a network of constructions: a "constructicon".

I already commented on Tenet 1 above. Tenet 2 concerns semantics and the syntax-semantics interface, which are part of most HPSG analyses. In what follows I want to look in more detail at the other tenets. Something that is not mentioned in Goldberg's tenets but is part of the definition of Construction by Fillmore et al. (1988: 501) is the non-locality of Constructions. I will comment on this in a separate subsection.

1.2.1 Surface orientation and empty elements

Tenet 3 requires a surface-oriented approach. Underlying levels and phonologically empty elements are ruled out. This excludes derivational models of transformational syntax assuming an underlying structure (the so-called D-structure) and some derived structure or more recent derivational variants of Minimalism. There was a time where representational models of Government & Binding (GB, Chomsky 1981) did not assume a D-structure but just one structure with traces (Koster 1978: 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38; Veenstra 1998: 58). Some of these analyses are rather similar to HPSG analyses as they are assumed today (Kiss 1995; Bouma & van Noord 1998; Meurers 2000; Müller 2005;

2017a; 2019b). Chomsky's Minimalist work (Chomsky 1995) assumes a derivational model and comes with a rhetoric of building structure in a bottom-up way and sending complete phases to the interfaces for pronunciation and interpretation. This is incompatible with Tenet 3, but in principle, Minimalist approaches are very similar to Categorial Grammar, so there could be representational approaches adhering to Tenet 3.[4]

A comment on empty elements is in order: all articles introducing Construction Grammar state that CxG does not assume empty elements. Most of the alternative theories do use empty elements: see König (1999) on Categorial Grammar, Gazdar, Klein, Pullum & Sag (1985: 143) on GPSG, Bresnan (2001: 67) on LFG, Bender (2000) and Sag, Wasow & Bender (2003: 464) on HPSG/Sign-Based Construction Grammar. There are results from the 60s that show that phrase structure grammars containing empty elements can be translated into grammars that do not contain empty elements (Bar-Hillel, Perles & Shamir 1961: 153, Lemma 4.1) and sure enough there are versions of GPSG (Uszkoreit 1987: 76–77), LFG (Kaplan & Zaenen 1989; Dalrymple et al. 2001), and HPSG (Bouma et al. 2001; Sag 2010: 508) that do not use empty elements. Grammars with empty elements often are more compact than those without empty elements and express generalizations more directly. See for example Bender (2000) for copulaless sentences in African American Vernacular English and Müller (2014) on nounless NPs in German. The argument against empty elements usually refers to language acquisition: it is argued that empty elements cannot be learned since they are not detectable in the input. However, if the empty elements alternate with visible material, it can be argued that what is learned is the fact that a certain element can be left out. What is true, though, is that things like empty expletives cannot be learned since these empty elements are neither visible nor do they contribute to meaning. Their only purpose in grammars is to keep uniformity. For example, Grewendorf (1993) working in GB suggests an analysis of the passive in German that is parallel to the movement-based analysis of English passives (Chomsky 1981: 124). In order to account for the fact that the subject does not move to initial position in German, he suggests an empty expletive pronoun that takes the subject position and that is connected to the original non-moved subject. Such elements cannot be acquired without innate knowledge about the IP/VP system

[4] There is a variant of Minimalist Grammars (Stabler 2011), namely Top-down Phase-based Minimalist Grammar (TPMG) as developed by Chesi (2012; 2007) and Bianchi & Chesi (2006; 2012). There is no movement in TPMG. Rather, *wh*-phrases are linked to their "in situ" positions with the aid of a short-term memory buffer that functions like a stack. See also Hunter (2010; 2018) for a related account where the information about the presence of a *wh*-phrase is percolated in the syntax tree, like in GPSG/HPSG.

and constraints about the obligatory presence of subjects. The CxG criticism is justified here.

A frequent argumentation for empty elements in MGG is based on the fact that there are overt realizations of an element in other languages (e.g., object agreement in Basque and focus markers in Gungbe). But since there is no language-internal evidence for these empty elements, they cannot be learned and one would have to assume that they are innate. This kind of empty element is rightly rejected (by proponents of CxG and others).

Summing up, it can be said that all grammars can be turned into grammars without empty elements and hence fulfill Tenet 3. It was argued that the reason for assuming Tenet 3 (problems in language acquisition) should be reconsidered and that a weaker form of Tenet 3 should be assumed: empty elements are forbidden unless there is language-internal evidence for them. This revised version of Tenet 3 would allow one to count empty element versions of CG, GPSG, LFG, and HPSG among Construction Grammars.

1.2.2 Language acquisition without the assumption of UG

Tenets 4 and 5 are basically what everybody should assume in MGG if Hauser, Chomsky & Fitch (2002) are taken seriously. Of course, this is not what is done in large parts of the field. The most extreme variant is Cinque & Rizzi (2010), who assume at least 400 functional heads being part of Universal Grammar (UG) and being present in all grammars of all languages, although sometimes invisibly. Such assumptions beg the question why the genera of Bantu languages should be part of our genome and how they got there. Researchers working on language acquisition realized that the Principles & Parameters approach (Meisel 1995) makes wrong predictions. They now talk about Micro-Cues instead of parameters (Westergaard 2014) and these Micro-Cues are just features that can be learned. However, Westergaard still assumes that the features are determined by UG, a dubious assumption seen from a CxG perspective (and from the perspective of Hauser, Chomsky, Fitch and genetics in general; Bishop 2002).

Note that even those versions of Minimalism that do not follow the Rizzi-style Cartographic approaches are far from being minimalist in their assumptions. Some distinguish between strong and weak features, some assume enumerations of lexical items from which a particular derivation draws its input, and some assume that all movement has to be feature-driven. Still others assume that derivations work in so-called phases and that a phase, once completed, is "shipped to the interfaces". Construction of phases is bottom-up, which is incompatible with psycholinguistic results (see also Borsley & Müller 2019: Section 5.1, Chapter L in this volume). None of these assumptions is a natural one to make from a language acquisition point of view. Most of these assumptions do not have any empirical motivation; the only motivation usually given is that they result in "restrictive theories". But if there is no motivation for them, this means that the respective architectural assumptions have to be part of our innate domain-specific knowledge, which is implausible according to Hauser, Chomsky & Fitch (2002).

As research in computational linguistics shows, our input is rich enough to form classes, to determine the part of speech of lexical items, and even to infer syntactic structure thought to be underdetermined by the input. For instance, Bod (2009) shows that the classical auxiliary inversion examples that Chomsky still uses in his Poverty of the Stimulus arguments (Chomsky 1971: 29–33; Berwick, Pietroski, Yankama & Chomsky 2011) can also be learned from language input available to children. See also Freudenthal et al. (2006; 2007) on input-based language acquisition.

HPSG does not make any assumptions about complicated mechanisms like feature-driven movement and so on. HPSG states properties of linguistic objects like part of speech, case, gender, etc., and states relations between features like agreement and government. In this respect it is like other Construction Grammars and hence experimental results regarding theories of language acquisition can be carried over to HPSG. See also Ginzburg (2019), Chapter F of this volume on language acquisition.

1.2.3 Inheritance networks

This leaves us with Tenets 6 and 7, that is, *inheritance networks* and the constructicon. Inheritance is something that is used in the classification of knowledge. For example, the word *animal* is very general and refers to entities with certain properties. There are subtypes of this kind of entity: *mammal* and further subtypes like *mouse*. In inheritance hierarchies, the knowledge of superconcepts is not restated at the subconcepts but instead, the superconcept is referred to. This is like Wikipedia: the Wikipedia entry of *mouse* states that mice are mammals without listing all the information that comes with the concept of mammal. Such inheritance hierarchies can be used in linguistics as well. They can be used to classify roots, words, and phrases. An example of such a hierarchy used for the classification of adjectives and adjectival derivation is discussed in Section 3. See also Davis & Koenig (2019: Section 4), Chapter 4 of this volume on inheritance in the lexicon.

MGG does not make reference to inheritance hierarchies. HPSG did this right from the beginning in 1985 (Flickinger et al. 1985) for lexical items and since

1995 also for phrasal constructions (Sag 1997). LFG rejected the use of types but used macros in computer implementations. The macros were abbreviatory devices and did not play any role in theoretical work. This changed in 2004, when macros were suggested in theoretical work (Dalrymple, Kaplan & King 2004). And although any connection to constructionist work is vehemently denied by some of the authors, recent work in LFG has a decidedly constructional flavor (Asudeh, Dalrymple & Toivonen 2008; 2014).[5] LFG differs from frameworks like HPSG, though, in assuming a separate level of c-structure. c-structure rules are basically context-free phrase structure rules and they are not modeled by feature value pairs (although they could be; Kaplan 1995). This means that it is not possible to capture a generalization regarding lexical items, lexical rules, and phrasal schemata, or any two-element subset of these three kinds of objects. While HPSG describes all of these elements with the same inventory and hence can use common supertypes in the description of all three, this is not possible in LFG (Müller 2018b: Section 23.1).[6] For example, Höhle (1997) argued that complementizers and finite verbs in initial position in German form a natural class. HPSG can capture this since complementizers (lexical elements) and finite verbs in initial position (results of lexical rule applications or a phrasal schema) can have a common supertype. TAG is also using inheritance in the Meta Grammar (Lichte & Kallmeyer 2017).

Since HPSG's lexical entries, lexical rules, and phrasal schemata are all described by typed feature descriptions, one could call the set of these descriptions the constructicon. Therefore, Tenet 7 is also adhered to.

1.2.4 Non-locality

Fillmore, Kay & O'Connor (1988: 501) stated in their definition of Constructions that Constructions may involve more than mothers and immediate daughters (see p. 1079 above).[7] That is, daughters of daughters can be specified as well. A straightforward example of such a specification is given in Figure 1, which shows the TAG analysis of the idiom *take into account* following Abeillé & Schabes (1989: 7). The fixed parts of the idiom are just stated in the tree. NP↓ stands for an open slot into which an NP has to be inserted. The subscript NA says that adjunction to the respectively marked nodes is forbidden. Theories like Constructional HPSG

[5] See Toivonen (2013: 516) for an explicit reference to construction-specific phrase structure rules in the sense of Construction Grammar.

[6] One could use templates (Dalrymple et al. 2004; Asudeh et al. 2013) to specify properties of lexical items and of mother nodes in c-structure rules, but usually c-structure rules specify the syntactic categories of mothers and daughters, so this information has a special status within

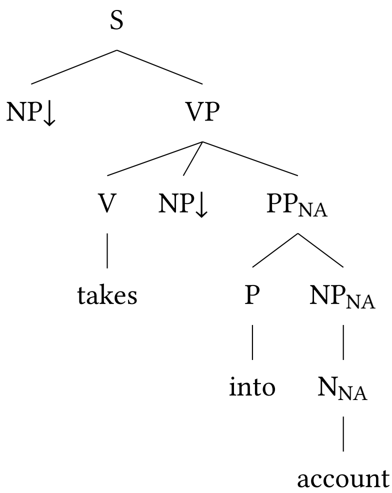


Figure 1: TAG tree for *take into account* by Abeillé & Schabes (1989: 7)

can state such complex tree structures like TAG can. Dominance relationships are modeled by feature structures in HPSG and it is possible to have a description that corresponds to Figure 1. The NP slots would just be left underspecified and can be filled in models that are total (see Richter 2007 and Richter 2019, Chapter 3 of this volume for formal foundations of HPSG).

It does not come without some irony that the theoretical approach that was developed out of Berkeley Construction Grammar and Constructional HPSG, namely Sign-Based Construction Grammar (Sag, Boas & Kay 2012; Sag 2012), is strongly local: it is made rather difficult to access daughters of daughters (Sag 2007). So, if one would stick to the early definition, this would rule out SBCG as a Construction Grammar. Fortunately, this is not justified. First, there are ways to establish nonlocal selection (Müller 2016) and second, there are ways to analyze idioms locally. Sag (2007), Kay, Sag & Flickinger (2015), and **KM2017a** develop a theory of idioms that is entirely based on local selection. For example, for *take into account*, one can state that *take* selects two NPs and a PP with the fixed lexical material *into* and *account*. The right form of the PP is enforced by means of the feature LEXICAL IDENTIFIER (LID). A special word *into* with the LID value *into* is specified as selecting a special word *account*. What is done in TAG via direct

the c-structure rules.

[7] This subsection is based on a much more thorough discussion of locality and SBCG in Müller (2016: Section 10.6.2.1.1 and Section 18.2).

specification is done in SBCG via a series of local selections of specialized lexical items. The interesting (intermediate) conclusion is: if SBCG can account for idioms via local selection, then theories like Categorial Grammar and Minimalism can do so as well. So, they cannot be excluded from Construction Grammars on the basis of arguments concerning idioms and non-locality of selection.

However, there may be cases of idioms that cannot be handled via local selection. For example, Richter & Sailer (2009) discuss the following idiom:

(3) glauben, X_Acc tritt ein Pferd
believe X kicks a horse
'be utterly surprised'

The X-constituent has to be a pronoun that refers to the subject of the matrix clause. If this is not the case, the sentence becomes ungrammatical or loses its idiomatic meaning.

(4) a. Ich glaube, mich / # dich tritt ein Pferd.
I believe me.ACC you.ACC kicks a horse

b. Jonas glaubt, ihn tritt ein Pferd.[8]
Jonas believes him kicks a horse
'Jonas is utterly surprised.'

c. # Jonas glaubt, dich tritt ein Pferd.
Jonas believes you kicks a horse
'Jonas believes that a horse kicks you.'

Richter & Sailer (2009: 313) argue that the idiomatic reading is only available if the accusative pronoun is fronted and the embedded clause is V2. The examples in (5) do not have the idiomatic reading:

(5) a. Ich glaube, dass mich ein Pferd tritt.
I believe that me a horse kicks
'I believe that a horse kicks me.'

b. Ich glaube, ein Pferd tritt mich.
I believe a horse kicks me
'I believe that a horse kicks me.'

They develop an analysis with a partly fixed configuration and some open slots, similar in spirit to the TAG analysis in Figure 1. However, their restrictions on

[8] http://www.machandel-verlag.de/der-katzenschatz.html, 2015-07-06.

Pferd clauses are too strict since there are variants of the idiom that do not have the accusative pronoun in the *Vorfeld*:

(6) ich glaub es tritt mich ein Pferd wenn ich einen derartigen
I believe EXPL kicks me a horse when I a such
Unsinn lese.[9]
nonsense read
'I am utterly surprised when I read such nonsense.'

So it might be the case that the organization of the embedded clause can be stated clause-internally, and hence it is an open question whether there are idioms that make nonlocal Constructions necessary.

What is not an open empirical question, though, is whether humans store chunks with complex internal structure or not. It is clear that we do, and much Construction Grammar literature emphasizes this. Constructional HPSG can represent such chunks directly in the theory, but SBCG cannot, since linguistic signs do not have daughters. So here, Constructional HPSG and TAG are the theories that can represent complex chunks of linguistic material with its internal structure, while other theories like GB, Minimalism, CG, LFG, SBCG, and DG cannot.

1.2.5 Summary

If all these points are taken together, it is clear that most variants of MGG are not Construction Grammars. However, CxG had considerable influence on other frameworks so that there are constructionist variants of LFG, HPSG, and TAG. HPSG in the version of Sag (1997) (also called Constructional HPSG) and the HPSG dialect Sign-Based Construction Grammar are Construction Grammars that follow all the tenets mentioned above.

1.3 Variants of Construction Grammar

The previous section discussed the tenets of CxG and to what degree other frameworks adhere to them. This section deals with frameworks that have Construction Grammar explicitly in their name. The following variants are usually named:

- Berkeley Construction Grammar (Fillmore 1985b; 1988; Kay & Fillmore 1999; Fried 2015)

[9] http://www.welt.de/wirtschaft/article116297208/Die-verlogene-Kritik-an-den-Steuerparadiesen.html, commentary section, 2018-02-20.

- Cognitive Construction Grammar (Lakoff 1987; Goldberg 1995; 2006)
- Cognitive Grammar (Langacker 1987; 2000; 2008; Dąbrowska 2004)
- Radical Construction Grammar (Croft 2001)
- Embodied Construction Grammar (Bergen & Chang 2005)
- Fluid Construction Grammar (Steels & De Beule 2006; Steels 2011)
- Sign-Based Construction Grammar (Sag 2010; 2012)

Berkeley Construction Grammar, Embodied Construction Grammar, Fluid Construction Grammar, and Sign-Based Construction Grammar are the ones that are more formal. All of these variants use feature value pairs and are constraint-based. They are sometimes also referred to as unification-based approaches. Berkeley Construction Grammar never had a consistent formalization. The variant of unification assumed by Kay & Fillmore (1999) was formally inconsistent (Müller 2006a: Section 2.4) and the computation of construction-like objects (CLOs) suggested by Kay (2002) did not work either (Müller 2006a: Section 3). Berkeley Construction Grammar was dropped by the authors, who joined forces with Ivan Sag and Laura Michaelis and eventually came up with an HPSG variant named Sign-Based Construction Grammar (Sag 2012). The differences between Constructional HPSG (Sag 1997) and SBCG are to some extent cosmetic: semantic relations got the suffix *-fr* for *frame* (*like-rel* became *like-fr*), phrases were called constructions (*hd-subj-ph* became *subj-head-cxt*), and lexical rules were called *derivational constructions*.[10] While this renaming would not have changed anything in terms of expressiveness of theories, there was another change that was not motivated by any of the tenets of Construction Grammar but rather by the wish to get a more restrictive theory: Sag, Wasow & Bender (2003) and Sag (2007) changed the feature geometry of phrasal signs in such a way that signs do not contain daughters. The information about mother-daughter relations is contained in lexical rules and phrasal schemata (Constructions) only. The phrasal schemata are more like GPSG immediate dominance schemata (phrase structure rules without constraints on the order of the daughters) in licensing a mother node when certain daughters are present, but without the daughters being represented as part of the AVM that stands for the mother node, as was common

[10] This renaming trick was so successful that it even confused some of the co-editors of the volume about SBCG (Boas & Sag 2012). See for example Boas (2014) and the reply in Müller & Wechsler (2014b).

in HPSG from 1985 till Sag, Wasow & Bender (2003).[11] This differs quite dramatically from what was done in Berkeley Construction Grammar, since BCxG explicitly favored a non-local approach (Fillmore 1988: 37; Fillmore et al. 1988: 501). Arguments were not canceled but passed up to the mother node. Adjuncts were passed up as well, so that the complete internal structure of an expression is available at the top-most node (Kay & Fillmore 1999: 9). The advantage of BCxG and Constructional HPSG (Sag 1997) is that complex expressions (e.g., idioms and other more transparent expressions with high frequency) can be stored as chunks containing the internal structure. This is not possible with SBCG, since phrasal signs never contain internal structures. For a detailed discussion of Sign-Based Construction Grammar see Section 1.4.2 and Müller (2016: Section 10.6.2).

Embodied Construction Grammar (Bergen & Chang 2005) uses typed feature descriptions for the description of linguistic objects and allows for discontinuous constituents. As argued by Müller (2016: Section 10.6.3), it is a notational variant of Reape-style HPSG (Reape 1994) (see Müller 2019a: Section 6, Chapter 10 of this volume for discontinuous constituents in HPSG).

Fluid Construction Grammar is also rather similar to HPSG. An important difference is that FCG attaches weights to constraints, something that is usually not done in HPSG. But in principle, there is nothing that forbids adding weights to HPSG as well, and in fact it has been done (Brew 1995; Briscoe & Copestake 1999; Miyao & Tsujii 2008), and it should be done to a larger extent (Miller 2013). Van Trijp (2013) tried to show that Fluid Construction Grammar is fundamentally different from SBCG, but I think he failed in every single respect. See Müller (2017b) for a detailed discussion, which cannot be repeated here for space reasons.

1.4 Constructional HPSG and formal variants of Construction Grammar

1.4.1 Constructional HPSG

As is discussed in other chapters in more detail, HPSG uses feature value pairs to model linguistic objects. One important tool is structure sharing. For example, determiner, adjective, and noun agree with respect to certain features in languages like German. The identity of properties is modeled by identity of features and this identity is established by identifying the values in descriptions. Now, it is obvious that certain features are always shared together. In order to facilitate the statement of respective constraints, feature value pairs are put into groups.

[11] The two approaches will be discussed in more detail in Section 1.4.1 and Section 1.4.2.

This is why HPSG feature descriptions are very complex. Information about syntax and semantics is represented under SYNTAX-SEMANTICS (SYNSEM), information about syntax under CATEGORY (CAT), and information that is projected along the head path of a projection is represented under HEAD. All feature structures have to have a type. The type may be omitted in the description, but there has to be one in the model. Types are organized in hierarchies. They are written in italics. (7) shows an example lexical item for the word *ate*:[12]

(7) Lexical item for the word *ate*:

$$
\begin{bmatrix}
\textit{word}\\
\textsc{phonology}\ \langle\ \textit{ate}\ \rangle\\
\textsc{syntax-semantics}\ \ldots \begin{bmatrix} \textit{local} \\ \textsc{category} \begin{bmatrix} \textit{category}\\ \textsc{head} \begin{bmatrix}\textit{verb}\\ \textsc{vform}\ \ \textit{fin}\end{bmatrix}\\ \textsc{spr}\ \ \left\langle \text{NP}[\textit{nom}]_{\boxed{1}} \right\rangle \\ \textsc{comps}\ \ \left\langle \text{NP}[\textit{acc}]_{\boxed{2}}\right\rangle \end{bmatrix}\\ \textsc{content}\ \ \ldots \begin{bmatrix}\textit{eat}\\ \textsc{agent}\ \ \boxed{1}\\ \textsc{theme}\ \ \boxed{2}\end{bmatrix}\end{bmatrix}
\end{bmatrix}
$$

The information about part of speech and finiteness is bundled under HEAD. The selection of a subject is represented under SPR (sometimes the feature SUBJ is used for subjects) and the non-subject arguments are represented as part of a list under COMPS. The semantic indices $\boxed{1}$ and $\boxed{2}$ are linked to thematic roles in the semantic representation.

Dominance structures can also be represented with feature value pairs. While Pollard & Sag (1987) and Pollard & Sag (1994) had a DAUGHTERS feature and then certain phrasal types constraining the daughters within the DTRS feature, Sag (1997) represented the daughters and constraints upon them at the top level of the sign[13]. This move made it possible to have subtypes of the type *phrase*, e.g.,

[12] The '…' stands for the feature LOCAL, which is irrelevant in the present discussion. It plays a role in the treatment of nonlocal dependencies (Borsley & Crysmann 2019, Chapter 14 of this volume).

[13] The top level is the outermost level. So in (7), PHONOLOGY and SYNTAX-SEMANTICS are on the

filler-head-phrase, *specifier-head-phrase*, and *head-complement-phrase*. Generalizations over these types could be captured within the type hierarchy. (8) shows a constraint on the type *head-complement-phrase*:

(8) *head-complement-phrase* ⇒

$$\begin{bmatrix} \text{SYNSEM|LOC|CAT|COMPS } \boxed{1} \\ \text{HEAD-DTR|SYNSEM|LOC|CAT|COMPS } \langle \boxed{2} \rangle \oplus \boxed{1} \\ \text{NON-HEAD-DTRS } \langle [\text{ SYNSEM } \boxed{2}] \rangle \end{bmatrix}$$

The constraint says that feature structures of type *head-complement-phrase* have to have a SYNSEM value, a HEAD-DTR feature, and a list-valued NON-HEAD-DTRS feature. The list has to contain a singleton element whose SYNSEM value is identical to the first element of the COMPS list of the head daughter ($\boxed{2}$), and the remainder of the COMPS list ($\boxed{1}$) is identical to the COMPS value of the phrase licensed by the schema (represented under SYNSEM|LOC|CAT|COMPS). '⊕' stands for the append relation concatenating two lists. For further details see Borsley & Abeillé (2019), Chapter 1 of this volume.

Dominance schemata (corresponding to grammar rules in phrase structure grammars) refer to such phrasal types. (9) shows how the lexical item in (7) can be used in a head-complement configuration:

top level.

(9) Analysis of *ate a pizza* in Constructional HPSG:

$$\left[\begin{array}{ll}
\multicolumn{2}{l}{\textit{head-complement-phrase}} \\
\textsc{phon} & \langle \textit{ate, a, pizza} \rangle \\
\textsc{synsem}|\textsc{loc} & \left[\begin{array}{ll} \textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \boxed{1} \\ \textsc{spr} & \boxed{2} \\ \textsc{comps} & \langle\rangle \end{array}\right] \\ \textsc{cont} & \ldots \end{array}\right] \\
\textsc{head-dtr} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{word}} \\ \textsc{phon} & \langle \textit{ate} \rangle \\ \textsc{synsem}|\textsc{loc} & \left[\begin{array}{ll} \textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \boxed{1} \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{verb}} \\ \textsc{vform} & \textit{fin} \end{array}\right] \\ \textsc{spr} & \boxed{2} \langle \mathrm{NP}[\textit{nom}] \rangle \\ \textsc{comps} & \langle \boxed{3}\, \mathrm{NP}[\textit{acc}] \rangle \end{array}\right] \\ \textsc{cont} & \ldots \end{array}\right] \end{array}\right] \\
\textsc{non-head-dtrs} & \left\langle \left[\begin{array}{ll} \textsc{phon} & \langle \textit{a, pizza} \rangle \\ \textsc{synsem} & \boxed{3} \end{array}\right] \right\rangle
\end{array}\right]$$

The description in the COMPS list of the head is identified with the SYNSEM value of the non-head daughter ($\boxed{3}$). The information about the missing specifier is represented at the mother node ($\boxed{2}$). Head information is also shared between head daughter and mother node. The respective structure sharings are enforced by principles: the Subcategorization Principle or Valence Principles make sure that all valents of the head daughter that are not realized in a certain configuration are still present at the mother node. The Head Feature Principle ensures that the head information of a head daughter in headed structures is identical to the head information on the mother node, that is, HEAD features are shared.

This is a very brief sketch of Constructional HPSG and is by no means intended to be a full-blown introduction to HPSG, but it provides a description of properties that can be used to compare Constructional HPSG to Sign-Based Construction Grammar in the next subsection.

1.4.2 Sign-Based Construction Grammar

Having discussed some aspects of Constructional HPSG, I now turn to SBCG. SBCG is an HPSG variant, so it shares most properties of HPSG but there are some interesting properties that are discussed in this section. Locality constraints are discussed in the next subsection, and changes in feature geometry in the subsections to follow. Subsection 1.4.2.6 discusses Frame Semantics.

1.4.2.1 Locality constraints

As mentioned in Section 1.4.2.1, SBCG assumes a strong version of locality: phrasal signs do not have daughters. This is due to the fact that phrasal schemata (= phrasal constructions) are defined as in (10):

(10) Head-Complement Construction following Sag et al. (2003: 481):
head-comp-cx ⇒

$$\begin{bmatrix} \textsc{mother}|\textsc{syn}|\textsc{val}|\textsc{comps}\ \langle\rangle \\ \textsc{head-dtr}\ \boxed{0} \begin{bmatrix} \textit{word} \\ \textsc{syn}|\textsc{val}|\textsc{comps}\ \boxed{A} \end{bmatrix} \\ \textsc{dtrs}\ \langle\ \boxed{0}\ \rangle \oplus \boxed{A}\ \textit{nelist} \end{bmatrix}$$

Rather than specifying syntactic and semantic properties of the complete linguistic object at the top level (as earlier versions of HPSG did), these properties are specified as properties under MOTHER. Hence a construction licenses a sign (a phrase or a complex word), but the sign does not include daughters. The daughters live on the level of the construction only. While earlier versions of HPSG licensed signs directly, SBCG needs a statement saying that all objects under MOTHER are objects licensed by the grammar (Sag, Wasow & Bender 2003: 478):[14]

(11) Φ is a Well-Formed Structure according to a grammar *G* if and only if:
1. there is a construction *C* in *G*, and
2. there is a feature structure *I* that is an instantiation of *C*, such that Φ is the value of the MOTHER feature of *I*.

The idea behind this change in feature geometry is that heads cannot select for daughters of their valents and hence the formal setting is more restrictive and hence affecting computational complexity (Ivan Sag, p.c. 2011). However, this restriction can be circumvented by just structure sharing an element of the daughters list with some value within MOTHER. The XARG feature making one argument available at the top level of a projection (Bender & Flickinger 1999) is such

[14] A less formal version of this constraint is given as the Sign Principle by Sag (2012: 105): "Every sign must be listemically or constructionally licensed, where: a sign is listemically licensed only if it satisfies some listeme, and a sign is constructionally licensed if it is the mother of some well-formed construct."

a feature. So, on the formal level, the MOTHER feature alone does not result in restrictions on complexity. One would have to forbid such structure sharings in addition, but then one could keep MOTHER out of the business and state the restriction for earlier variants of HPSG (Müller 2018b: Section 10.6.2.1.3).

Sag et al. (2003) differentiated between specifiers and complements, but this distinction was given up in later work on SBCG. Sag (2012) has just one valence list that includes both subjects and non-subjects. This is a return to the valence representations of Pollard & Sag (1987). An argument for this was never given, despite arguments for a separation of valence information by Borsley (1987). With one single valence feature, a VP would be an unsaturated projection and generalizations concerning phrases cannot be captured. For example, a generalization concerning extraposition (in German) is that maximal projections (that is projections with an empty COMPS list) can be extraposed (Müller 1999: Section 13.1.2). It is impossible to state this generalization in SBCG in a straightforward way (Müller 2018b: Section 10.6.2.3).

1.4.2.2 The Head Feature Principle

There have been some other developments as well. Sag (2012) got rid of the Head Feature Principle and stated identity of information explicitly within constructions. Structure sharing is not stated with boxed numbers but with capital letters instead. An exclamation mark can be used to specify information that is not shared (Sag 2012: 125). While the use of letters instead of numbers is just a presentational variant, the exclamation mark is a non-trivial extension. (12) provides an example: the constraints on the type *pred-hd-comp-cxt*:

(12) Predicational Head-Complement Construction following Sag (2012: 152):
pred-hd-comp-cxt ⇒

$$\begin{bmatrix} \text{MOTHER}|\text{SYN X ! [VAL } \langle \text{ Y } \rangle] \\ \text{HEAD-DTR Z: } \begin{bmatrix} \textit{word} \\ \text{SYN X: } \begin{bmatrix} \text{VAL } \langle \text{ Y } \rangle \oplus \text{L} \end{bmatrix} \end{bmatrix} \\ \text{DTRS } \langle \text{ Z } \rangle \oplus \text{L:}\textit{nelist} \end{bmatrix}$$

The X stands for all syntactic properties of the head daughter. These are identified with the value of SYN of the mother with the exception of the VAL value, which is specified to be a list with the element Y. It is interesting to note that the !-notation is not without problems: Sag (2012: 145) states that the version of SBCG that he presents is "purely monotonic (non-default)", but if the SYN value of the mother is not identical + overwriting of VAL, it is unclear how the type of SYN can be constrained. ! can be understood as explicitly sharing all features that are not mentioned after the !. Note, though, that the type has to be shared as well. This is not trivial, since structure sharing cannot be applied here, since structure sharing the type would also identify all features belonging to the respective value. So one would need a relation that singles out a type of a structure and identifies this type with the value of another structure. Note also that information from features behind the ! can make the type of the complete structure more specific. Does this affect the shared structure (e.g., HEAD-DTR|SYN in (12))? What if the type of the complete structure is incompatible with the features in this structure? What seems to be a harmless notational device in fact involves some non-trivial machinery in the background. Keeping the Head Feature Principle makes this additional machinery unnecessary.

1.4.2.3 Feature geometry and the FORM feature

The phrasal sign for *ate a pizza* in Constructional HPSG was given in (9). (13) is the Predicational Head Complement Construction with daughters and mother filled in.

(13)
$$\begin{bmatrix} \textit{head-comp-cx} & \\ \textsc{mother} & \begin{bmatrix} \textit{phrase} & \\ \textsc{form} & \langle \textit{ate, a, pizza} \rangle \\ \textsc{syn} & \begin{bmatrix} \textsc{cat} & \boxed{1} \\ \textsc{val} & \langle \text{NP}[\textit{nom}] \rangle \end{bmatrix} \\ \textsc{sem} & \ldots \end{bmatrix} \\ \textsc{head-dtr}\ \boxed{2} & \begin{bmatrix} \textit{word} & \\ \textsc{form} & \langle \textit{ate} \rangle \\ \textsc{syn} & \begin{bmatrix} \textsc{cat}\ \boxed{1} & \begin{bmatrix} \textit{verb} \\ \textsc{vf}\ \textit{fin} \end{bmatrix} \\ \textsc{val} & \langle \text{NP}[\textit{nom}], \boxed{3}\,\text{NP}[\textit{acc}] \rangle \end{bmatrix} \\ \textsc{sem} & \ldots \end{bmatrix} \\ \textsc{dtrs} & \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix}$$

As was explained in the previous subsection, Constructional HPSG groups all selectable information under SYNSEM and then differentiates into CAT and CONT.

SBCG goes back to Pollard & Sag (1987) and uses SYN and SEM. The idea behind SYNSEM was to exclude the selection of phonological information and daughters (Pollard & Sag 1994: 23). Since daughters are outside of the definition of *synsem*, they cannot be accessed from within valency lists. Now, SBCG pushes this idea one step further and also restricts the access to daughters in phrasal schemata (Constructions in SBCG terminology): since signs do not have daughters, constructions may not refer to the daughters of their parts. But obviously signs need to have a form part, since signs are per definition form-meaning pairs. It follows that the form part of signs is selectable in SBCG. This will be discussed in more detail in the following subsection. Subsection 1.4.2.5 discusses the omission of the LOCAL feature.

1.4.2.4 Selection of PHON and FORM values

The feature geometry of constructional HPSG has the PHON value outside of SYNSEM. Therefore verbs can select for syntactic and semantic properties of their arguments but not for their phonology. For example, they can require that an object has accusative case but not that it starts with a vowel. SBCG allows for the selection of phonological information (the feature is called FORM here) and one example of such a selection is the indefinite article in English, which has to be either *a* or *an* depending on whether the noun or nominal projection it is combined with starts with a vowel or not (Flickinger, Mail to the HPSG mailing list, 01.03.2016):

(14) a. an institute
b. a house

The distinction can be modeled by assuming a selection feature for determiners.[15] An alternative would be, of course, to capture all phonological phenomena by formulating constraints on phonology on the phrasal level (see Bird & Klein 1994 and Walther 1999 for phonology in HPSG).

Note also that the treatment of raising in SBCG admits nonlocal selection of phonology values, since the analysis of raising in SBCG assumes that the element on the valence list of the embedded verb is identical to an element in the ARG-ST list of the matrix verb (Sag 2012: 159). Hence, both verbs in (15) can see the phonology of the subject:

[15] In Standard HPSG there is mutual selection between the determiner and the noun. The noun selects the determiner via SPR and the determiner selects the noun via a feature called SPECIFIED (Pollard & Sag 1994: 45–54).

(15) Kim can eat apples.

In principle, there could be languages in which the form of the downstairs verb depends on the presence of an initial consonant in phonology of the subject. English allows for long chains of raising verbs and one could imagine languages in which all the verbs on the way are sensitive to the phonology of the subject. Such languages probably do not exist.

Now, is this a problem? Not really, but if one develops a general setup in a way to exclude everything that is not attested in the languages of the world (as for instance the selection of arguments of arguments of arguments), then it is a problem that heads can see the phonology of elements that are far away.

There are two possible conclusions for practitioners of SBCG: either the MOTHER feature could be given up, since one agrees that theories that do not make wrong predictions are sufficiently constrained and one does not have to explicitly state what cannot occur in languages, or one would have to address the problem with nonlocally selected phonology values and therefore assume a SYNSEM or LOCAL feature that bundles information that is relevant in raising and does not include the phonology. In the latter case, the feature geometry of SBCG would get more complicated. This additional complication is further evidence against MOTHER, adding to the argument I made about MOTHER in Subsection 1.4.2.1.

1.4.2.5 The LOCAL feature and information shared in nonlocal dependencies

Similarly, elements of the ARG-ST list contain information about FORM. In nonlocal dependencies, this information is shared in the GAP list (SLASH set or list in other versions of HPSG) and is available all the way to the filler. In other versions of HPSG, only LOCAL information is shared and elements in valence lists do not have a PHON feature. If the sign that is contained in the GAP list were identified with the filler, the information about phonological properties of the filler would be available at the extraction side and SBCG could be used to model languages in which the phonology of a filler is relevant for a head from which it is extracted. So for instance, *likes* could see the phonology of *bagels* in (16):

(16) Bagels, I think that Peter likes.

It would be possible to state constraints saying that the filler has to contain a vowel or two vowels or that it ends with a consonant. In addition, all elements on the extraction path (*that* and *think*) can see the phonology of the filler as well. While there are languages that mark the extraction path, I doubt that there are languages that have phonological effects across long distances. This problem can be and has been solved by assuming that the filler is not shared with the information in the GAP list, but parts of the filler are shared with parts in the GAP list: Sag (2012: 166) assumes that SYN, SEM, and STORE information is identified individually. Originally, the feature geometry of HPSG was motivated by the wish to structure share information. Everything within LOCAL was shared between filler and extraction side. This kind of motivation is given up in SBCG.

Note, also, that not sharing the complete filler with the gap means that the FORM value of the element in the ARG-ST list at the extraction side is not constrained. Without any constraints, the theory would be compatible with infinitely many models, since the FORM value could be anything. For example, the FORM value of an extracted adjective could be ⟨ *Donald Duck* ⟩ or ⟨ *Dunald Dock* ⟩ or any arbitrary chaotic sequence of letters/phonemes. To exclude this, one can stipulate the FORM values of extracted elements to be the empty list, but this leaves one with the unintuitive situation that the element in GAP has an empty FORM list while the corresponding filler has a different, filled one.

1.4.2.6 Frame Semantics

Another difference between SBCG and other variants of HPSG is the use of Frame Semantics (Fillmore 1982; Fillmore 1985a). The actual representations in SBCG are based on MRS (Minimal Recursion Semantics, Copestake et al. 2005, see also Koenig & Richter 2019, Chapter C of this volume) and the change seems rather cosmetic (relations have the suffix *-fr* for frame rather than *-rel* for relation and the feature is called FRAMES rather than RELATIONS), but there is one crucial difference: the labels of semantic roles are more specific then what is usually used in other variants of HPSG.[16] Sag (2012: 89) provides the following representation for the meaning contribution of the verb *eat*:

(17)
$$\begin{bmatrix} \textit{sem-obj} & \\ \text{INDEX} & \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{eating-fr} & \\ \text{LABEL} & l \\ \text{SIT} & s \\ \text{INGESTOR} & i \\ \text{INGESTIBLE} & j \end{bmatrix} \right\rangle \end{bmatrix}$$

[16] Pollard & Sag (1987: 95) and Pollard & Sag (1994) use role labels like KISSER and KISSEE that are predicate-specific. Generalizations over these feature names are impossible within the standard formal setting of HPSG (Pollard & Sag 1994: Section 8.5.3; Müller 1999: 24, Fn. 1).

While some generalizations over verbs of a certain type can be captured with role labels like INGESTOR and INGESTIBLE, this is limited to verbs of ingestion. More general role labels like AGENT and PATIENT (or PROTO-AGENT and PROTO-PATIENT, Dowty 1991) allow for more generalizations of broader classes of verbs (see Davis & Koenig 2000 and Wechsler, Koenig & Davis 2019, Chapter 9 of this volume).

1.4.3 Fluid Construction Grammar

One thing that makes SBCG different from other Construction Grammars is that SBCG assumes a strongly lexicalist stance (Sag & Wasow 2011): argument structure is encoded lexically. A ditransitive verb is a ditransitive verb since it selects for three NP arguments. This selection is encoded in valence features of lexical items. It is not assumed that phrasal configurations can license additional arguments as it is in Radical Construction Grammar, Embodied Construction Grammar, and Fluid Construction grammar. The next section discusses phrasal CxG approaches in more detail. Section 4 then discusses patterns that should be analyzed phrasally and which are problematic for entirely head-driven (or rather functor-driven) theories like Categorial Grammar, Dependency Grammar, and Minimalism.

2 Valence vs. phrasal patterns

Much work in Construction Grammar starts from the observation that children acquire patterns and, in later acquisition stages, abstract from these patterns to schemata containing open slots to be filled by variable material, for example subjects and objects (Tomasello 2003). The conclusion that is drawn from this is that language should be described with reference to phrasal patterns. Most Construction Grammar variants assume a phrasal approach to argument structure constructions (Goldberg 1995; 2006; Goldberg & Jackendoff 2004), with Constructional HPSG (Sag 1997), Boas's (2003) work, and SBCG (Sag et al. 2012; Sag 2012) being the three exceptions. So, for examples like the resultative construction in (18), Goldberg (1995) assumes that there is a phrasal construction [Sbj Verb Obj Obl] into which material is inserted and which contributes the resultative semantics as a whole.

(18) She fished the pond empty.

HPSG follows the lexical approach and assumes that *fish-* is inserted into a lexical construction (lexical rule), which licenses the combination with other parts of the resultative construction (Müller 2002).

I argued in several publications that the language acquisition facts can be explained in lexical models as well (Müller 2010: Section 6.3; Müller & Wechsler 2014a: Section 9). While a pattern-based approach claims that (19) is analyzed by inserting *Kim*, *loves*, and *Sandy* into a phrasal schema stating that NP[nom] verb NP[acc] or subject verb object are possible sequences in English, a lexical approach would state that there is a verb *loves* selecting for an NP[nom] and an NP[acc] (or for a subject and an object).

(19) Kim loves Sandy.

Since objects follow the verb in English (modulo extraction) and subjects precede the verb, the same sequence is licensed in the lexical approach. The lexical approach does not have any problems accounting for patterns in which the sequence of subject, verb, and object is discontinuous. For example, an adverb may intervene between subject and verb:

(20) Kim really loves Sandy.

In a lexical approach it is assumed that verb and object may form a unit (a VP). The adverb attaches to this VP and the resulting VP is combined with the subject. The phrasal approach has to assume either that adverbs are part of phrasal schemata licensing cases like (20) (see Uszkoreit 1987: Section 6.3.2 for such a proposal in a GPSG approach to German) or that the phrasal construction may license discontinuous patterns. Bergen & Chang (2005: 170) follow the latter approach and assume that subject and verb may be discontinuous but verb and object(s) have to be adjacent. While this accounts for adverbs like the one in (20), it does not solve the general problem, since there are other examples showing that verb and object(s) may appear discontinuously as well:

(21) Mary tossed me a juice and Peter a water.

Even though *tossed* and *Peter a water* are discontinuous in (21), they are an instance of the ditransitive construction. The conclusion is that what has to be acquired is not a phrasal pattern but rather the fact that there are dependencies between certain elements in phrases (see also Behrens 2009 for a similar view from a language acquisition perspective). I return to ditransitive constructions in Section 2.3.

I discussed several phrasal approaches to argument structure and showed where they fail (Müller 2006a,b; 2007; 2010; Müller & Wechsler 2014a,b; Müller 2018a). Of course, the discussion cannot be reproduced here, but I want to repeat four points showing that lexical valence representation is necessary and that effects that are the highlight of phrasal approaches can be achieved in lexical proposals as well. The first two are problems that were around in GPSG times and basically were solved by abandoning the framework and adopting a new framework which was a fusion of GPSG and Categorial Grammar: HPSG.[17]

2.1 Derivational morphology and valence

The first argument (Müller 2016: Section 5.5.1) is that certain patterns in derivational morphology refer to valence. For example, the -*bar* 'able' derivation productively applies to transitive verbs only, that is, to verbs that govern an accusative.

(22) a. unterstützbar
supportable

b. * helfbar
helpable

c. * schlafbar
sleepable

Note that the -*bar* 'able' derivation is like the passive in that it suppresses the subject and promotes the accusative object: the accusative object is the element adjectives derived with the -*bar* 'able' derivation predicate over. There is no argument realized with the adjective *unterstützbaren* 'supportable' attaching to *Arbeitsprozessen* 'work.processes' in *unterstützbaren Arbeitsprozessen*.[18] Hence one could not claim that the stem enters a phrasal construction with arguments and -*bar* attaches to this phrase. It follows that information about valency has to be present at the stem.

Note also that the resultative construction interacts with the -*bar* 'able' derivation. (23) shows an example of the resultative construction in German in which the accusative object is introduced by the construction: it is the subject of *leer* 'empty' but not a semantic argument of the verb *fischt* 'fishes'.

[17] For further criticism of GPSG see Jacobson (1987). A detailed discussion of reasons for abandoning GPSG can be found in Müller (2016: Section 5.5).

[18] Adjectives realize their arguments preverbally in German:

(i) der [seiner Frau treue] Mann
the his wife faithful man
'the man who is faithful to his wife'

unterstützbaren 'supportable' does not take an argument; it is a complete adjectival projection like *seiner Frau treue*.

(23) Sie fischt den Teich leer.
she fishes the pond empty

So even though the accusative object is not a semantic argument of the verb, the -*bar* 'able' derivation is possible and an adjective like *leerfischbar* 'empty.fishable' meaning 'can be fished empty' can be derived. This is explained by lexical analyses of the -*bar* 'able' derivation and the resultative construction, since if one assumes that there is a lexical item for the verb *fisch-* selecting an accusative object and a result predicate, then this item may function as the input for the -*bar* 'able' derivation. See Section 3 for further discussion of -*bar* 'able' derivation and Verspoor (1997), Wechsler (1997), Wechsler & Noh (2001), and Müller (2002: Chapter 5) for lexical analyses of the resultative construction in the framework of HPSG.

2.2 Partial verb phrase fronting

The second argument concerns partial verb phrase fronting (Müller 2016: Section 5.5.2). (24) gives some examples: in (24a) the bare verb is fronted and its arguments are realized in the middle field, in (24b) one of the objects is fronted together with the verb, and in (24c) both objects are fronted with the verb.

(24) a. Erzählen wird er seiner Tochter ein Märchen können.
tell will he his daughter a fairy.tale can

b. Ein Märchen erzählen wird er seiner Tochter können.
a fairy.tale tell will he his daughter can

c. Seiner Tochter ein Märchen erzählen wird er können.
his daughter a fairy.tale tell will he can

'He will be able to tell his daughter a fairy tale.'

The problem with sentences such as those in (24) is that the valence requirements of the verb *erzählen* 'to tell' are realized in various positions in the sentence. For fronted constituents, one requires a rule which allows a ditransitive to be realized without its arguments or with one or two objects. This basically destroys the idea of a fixed phrasal configuration for the ditransitive construction and points again in the direction of dependencies.

Furthermore, it has to be ensured that the arguments that are missing in the prefield are realized in the remainder of the clause. It is not legitimate to omit obligatory arguments or realize arguments with other properties like a different case, as the examples in (25) show:

(25) a. Verschlungen hat er es nicht.
devoured has he.NOM it.ACC not
'He did not devour it.'

b. * Verschlungen hat er nicht.
devoured has he.NOM not

c. * Verschlungen hat er ihm nicht.
devoured has he.NOM him.DAT not

The obvious generalization is that the fronted and unfronted arguments must add up to the total set of arguments selected by the verb. This is scarcely possible with the rule-based representation of valence in GPSG (Nerbonne 1986; Johnson 1986). In theories such as Categorial Grammar, it is possible to formulate elegant analyses of (25) (Geach 1970). Nerbonne (1986) and Johnson (1986) both suggest analyses for sentences such as (25) in the framework of GPSG which ultimately amount to changing the representation of valence information in the direction of Categorial Grammar. With a switch to CG-like valence representations in HPSG, the phenomenon of partial verb phrase fronting found elegant solutions (Höhle 2018a: Section 4; Müller 1996; Meurers 1999).

2.3 Coercion

An important observation in constructionist work is that, in certain cases, verbs can be used in constructions that differ from the constructions they are normally used in. For example, verbs that are usually used with one or two arguments may be used in the ditransitive construction:

(26) a. She smiled.
b. She smiled herself an upgrade.[19]
c. He baked a cake.
d. He baked her a cake.

The usual explanation for sentences like (26b) and (26d) is that there is a phrasal pattern with three arguments into which intransitive and strictly transitive verbs may be entered. It is assumed that the phrasal patterns are associated with a certain meaning (Goldberg 1996; Goldberg & Jackendoff 2004). For example, the benefactive meaning of (26d) is contributed by the phrasal pattern (Goldberg 1996: Section 6; Asudeh, Giorgolo & Toivonen 2014: 81).

[19] Douglas Adams. 1979. *The Hitchhiker's Guide to the Galaxy*, Harmony Books. Quoted from Goldberg (2003: 220).

The insight that a verb is used in the ditransitive pattern and thereby contributes a certain meaning is of course also captured in lexical approaches. Briscoe & Copestake (1999) suggested a lexical rule-based analysis mapping a transitive version of verbs like *bake* onto a ditransitive one and adding the benefactive semantics. This is parallel to the phrasal approach in that it says: three-place *bake* behaves like other three-place verbs (e.g., *give*) in taking three arguments and by doing so, it comes with a certain meaning (see Müller 2018a for a lexical rule-based analysis of the benefactive constructions that works for both English and German, despite the surface differences of the respective languages). The lexical rule is a form-meaning pair and hence a construction. As Croft put it 15 years ago: Lexical rule vs. phrasal schema is a false dichotomy (Croft 2003). But see Müller (2018a; 2006a; 2013) and Müller & Wechsler (2014a) for differences between the approaches.

Briscoe & Copestake (1999) paired their lexical rules with probabilities to be able to explain differences in productivity. This corresponds to the association strength that van Trijp (2011: 141) used to relate lexical items to phrasal constructions of various kinds.

2.4 Non-predictability of valence

The last subsection discussed phrasal models of coercion that assume that verbs can be inserted into constructions that are compatible with the semantic contribution of the verb. Müller & Wechsler (2014a: Section 7.4) pointed out that this is not sufficiently constrained. Müller & Wechsler discussed the examples in (27), among others:

(27) a. John depends on Mary. (*counts, relies,* etc.)
b. John trusts (*on) Mary.

While *depends* can be combined with a *on*-PP, this is impossible for *trusts*. Also the form of the preposition of prepositional objects is not always predictable from semantic properties of the verb. So there has to be a way to state that certain verbs go together with certain kinds of arguments and others do not. A lexical specification of valence information is the most direct way to do this. Phrasal approaches sometimes assume other means to establish connections between lexical items and phrasal constructions. For instance, Goldberg (1995: 50) assumes that verbs are "conventionally associated with constructions". The more technical work in Fluid CxG assumes that every lexical item is connected to various phrasal constructions via coapplication links (van Trijp 2011: 141). This is very similar to Lexicalized Tree Adjoining Grammar (LTAG; Schabes, Abeillé & Joshi 1988), where a rich syntactic structure is associated to a lexical anchor. So, phrasal approaches that link syntactic structure to lexical items are actually lexical approaches as well. As in GPSG, they include means to ensure that lexical items enter into correct constructions. In GPSG, this was taken care of by a number. I already discussed the GPSG shortcomings in previous subsections.

Concluding this section, it can be said that there has to be a connection between lexical items and their arguments and that a lexical representation of argument structure is the best way to establish such a relation.

3 Construction Morphology

The first publications in Construction Morphology were the master's thesis of Riehemann (1993), which later appeared as Riehemann (1998), and Koenig's WCCFL paper and thesis (**Koenig93a**; **Koenig94a-u**; Koenig 1999). Riehemann called her framework *Type-Based Derivational Morphology*, since it was written before influential work like Goldberg (1995) appeared and before the term *Construction Morphology* (Booij 2005) was used. Riehemann did a careful corpus study on adjective derivations with the suffix *-bar* '-able'. She noticed that there is a productive pattern that can be analyzed by a lexical rule relating a verbal stem to the adjective suffixed with *-bar*.[20] The productive pattern applies to verbs governing an accusative as in (28a) but is incompatible with verbs taking a dative as in (28b):

(28) a. unterstützbar
supportable

b. * helfbar
helpable

c. * schlafbar
sleepable

Intransitive verbs are also excluded, as (28c) shows. Riehemann suggests the schema in (29):

[20] She did not call her rule a lexical rule, but the difference between her template and the formalization of lexical rules by Meurers (2001) is the naming of the feature MORPH-B vs. DTR.

(29) Schema for productive adjective derivations with the suffix -*bar* in German according to Riehemann (1998: 68):

$$\begin{bmatrix} \textit{reg-bar-adj} \\ \text{PHONOLOGY } \boxed{1} + \textit{bar} \\ \text{MORPH-B } \left\langle \begin{bmatrix} \textit{trans-verb} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM|LOC } \begin{bmatrix} \text{CAT|VAL|COMPS } \langle \text{NP}[\textit{acc}]:\boxed{2} \rangle \oplus \boxed{3} \\ \text{CONT|NUC } \boxed{4} \begin{bmatrix} \text{ACT} \\ \text{UND } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{SYNSEM|LOCAL } \begin{bmatrix} \text{CATEGORY } \begin{bmatrix} \text{HEAD } & \textit{adj} \\ \text{VALENCE } & \begin{bmatrix} \text{SUBJ } & \langle \text{NP}:\boxed{2} \rangle \\ \text{COMPS } & \boxed{3} \end{bmatrix} \end{bmatrix} \\ \text{CONTENT|NUCLEUS } \begin{bmatrix} \text{RELN } & \lozenge \\ \text{ARG1 } & \boxed{2} \\ \text{ARG2 } & \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

MORPH-B is a list that contains a description of a transitive verb (something that governs an accusative object which is linked to the undergoer role ($\boxed{2}$) and has an actor).[21] The phonology of this element ($\boxed{1}$) is combined with the suffix -*bar* and forms the phonology of the complete lexical item. The resulting object is of category *adj* and the semantics of the accusative object of the input verb ($\boxed{2}$) is identified with the one of the subject of the resulting adjective. The semantics of the input verb ($\boxed{4}$) is embedded under a modal operator in the semantics of the adjective.

While the description of the -*bar* 'able' derivation given so far captures the situation quite well, there are niches and isolated items that are exceptions. According to Riehemann (1998: 5), this was the case for 7% of the adjectives she

[21]Note that the specification of the type *trans-verb* in the list under MORPH-B is redundant, since it is stated that there has to be an accusative object and that there is an actor and an undergoer in the semantics. Depending on further properties of the grammar, the specification of the type is actually wrong: productively derived particle verbs may be input to the -*bar* 'able' derivation, and these are not a subtype of *trans-verb*, since the respective particle verb rule derives both transitive (*anlachen* 'laugh at somebody') and intransitive verbs (*loslachen* 'start to laugh') (Müller 2003: 296). *Anlachen* does not have an undergoer in the semantic representation suggested by Stiebels (1996). See Müller (2003: 308) for a version of the -*bar* 'able' derivation schema that is compatible with particle verb formations as input.

looked at in her corpus study. Examples are verbs ending in -*ig* like *entschuldigen* 'to excuse'. The -*ig* is dropped in the derivation:

(30) entschuldbar
excuseable

Other cases are lexicalized forms like *essbar* 'safely edible', which have a special lexicalized meaning. Exceptions of the accusative requirement are verbs selecting a dative (31a), a prepositional object (31b), reflexive verbs (31c), and even intransitive, mono-valent verbs (31d):

(31) a. unentrinnbar
inescapable

b. verfügbar
available

c. regenerierbar
regenerable

d. brennbar
inflammable

To capture generalizations about productive, semi-productive and fixed patterns/items, Riehemann suggests a type hierarchy, parts of which are provided in Figure 2. The type *bar-adj* stands for all -*bar* adjectives and comes with the

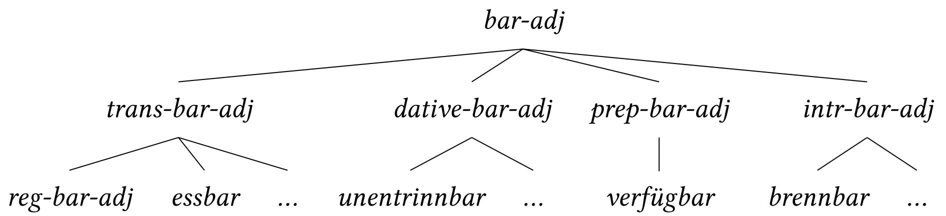


Figure 2: Parts of the type hierarchy for -*bar* 'able' derivation adapted from Riehemann (1998: 15)

constraints that apply to all of them. One subtype of this general type is *trans-bar-adj*, which subsumes all adjectives that are derived from transitive verbs. This includes all regularly derived -*bar*-adjectives, which are of the type *reg-bar-adj* but also *essbar* 'edible' and *sichtbar* 'visible'.

As this recapitulation of Riehemann's proposal shows, the analysis is a typical CxG analysis: V-*bar* is a partially filled word (see Goldberg's examples in Table 1).

The schema in (29) is a form-meaning pair. Exceptions and subregularities are represented in an inheritance network.

4 Phrasal patterns

Section 2 discussed the claim that Constructions in the sense of CxG have to be phrasal. I showed that this is not true and that in fact lexical approaches to valence have to be preferred under the assumptions usually made in non-transformational theories. However, there are other areas of grammar that give exclusively head-driven approaches like Categorial Grammar, Minimalism, and Dependency Grammar a hard time. In what follows I discuss the NPN construction and various forms of filler gap constructions.

4.1 The NPN Construction

Matsuyama (2004) and Jackendoff (2008) discuss the NPN Construction, examples of which are provided in (32):

(32) a. Student after student left the room.
b. Day after day after day went by, but I never found the courage to talk to her. (Bargmann 2015)

The properties of the NPN construction (with *after*) are summarized by Bargmann (2015) in a concise way and I will repeat his examples and summarization below to motivate his analysis in (40).

The examples in (32) show that the N-after-N Construction has the *distribution of NPs*.

As (33) shows, the construction is *partially lexically fixed*: *after* cannot be replaced by any other word (Matsuyama 2004: 73).

(33) Alex asked me question { after / * following / * succeeding } question.

The construction is *partially lexically flexible*: the choice of Ns is free, except for the fact that the Ns must be identical (34a), the Ns must be count nouns (34b), Ns must be in the singular (34c), and the Ns must be bare (34d).

(34)	a.	* bus after car	(N1 ≠ N2)
	b.	* water after water	(Ns = mass nouns)
	c.	* books after books	(Ns = plurals)
	d.	* a day after a day	(Ns have determiners)

The construction is *syntactically fixed*: N-after-N cannot be split by syntactic operations as the contrast in (35) shows (Matsuyama 2004):

(35) a. Man after man passed by.
b. * Man passed by after man.

If extraposition of the *after*-N constituent were possible, (35b) with an extraposed *after man* should be fine but it is not, so NPN seems to be a fixed configuration.

There is a syntax-semantics mismatch: while N-after-N is syntactically singular, as (36) shows, it is plural semantically, as (37) shows:

(36) Study after study { reveals / *reveal } the dangers of lightly trafficked streets.

(37) a. John ate { apple after apple / apples / *an apple } for an hour.
b. John ate { *apple after apple / *apples / an apple } in an hour.

Furthermore there is an aspect of semantic sequentiality: N-after-N conveys a temporal or spatial sequence: as Bargmann (2015) states, the meaning of (38a) is something like (38b).

(38) a. Man after man passed by.
b. First one man passed by, then another(, then another(, then another(, then …))).

The Ns in the construction do not refer to one individual each; rather, they contribute to a holistic meaning.

The NPN construction allows adjectives to be combined with the nouns, but this is restricted. N1 can only be preceded by an adjective if N2 is preceded by the same adjective:

(39) a. bad day after bad day (N1 and N2 are preceded by the same adjective.)
b. * bad day after awful day (N1 and N2 are preceded by different adjectives.)
c. * bad day after day (Only N1 is preceded by an adjective.)
d. day after bad day (Only N is preceded by an adjective.)

Finally, *after* N may be iterated to emphasize the fact that there are several referents of N, as the example in (32b) shows.

This empirical description is covered by the following phrasal construction, which is adapted from Bargmann (2015):[22]

(40) NPN Construction as formalized by Bargmann (2015):

$$\begin{bmatrix}
\textsc{phon}\ \langle \textit{… N …, after, … N …} \rangle \\
\textsc{ss|loc|cat} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textit{noun} \\ \textsc{count}\ - \\ \textsc{agr}\ \ \textit{3rdsing} \end{bmatrix} \\ \textsc{val} \begin{bmatrix} \textsc{spr}\ \ \langle\rangle \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \end{bmatrix} \\
\textsc{sr}\ \lambda P.\exists X.|X| > 1\ \&\ \forall x \in X{:}\ N'(x)\ \&\ \exists R^{order} \subseteq X^2\ \&\ P(x) \\
\textsc{dtrs} \left\langle \begin{bmatrix} \textsc{phon}\ \langle \textit{… N …} \rangle \\ \textsc{ss|l|c} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textit{noun} \\ \textsc{count}\ + \\ \textsc{agr}\ \textit{3rdsing} \end{bmatrix} \\ \textsc{val} \begin{bmatrix} \textsc{spr}\ \ \langle \text{Det} \rangle \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{sr}\ \ldots\ \lambda x.N'(x)\ \ldots \end{bmatrix}, \left(\begin{bmatrix} \textsc{phon}\ \langle \textit{after} \rangle \\ \ldots\ \textsc{head}\ \textit{prep} \\ \textsc{sr}\ \exists R^{order} \subseteq X^2 \end{bmatrix}, \begin{bmatrix} \textsc{phon}\ \langle \textit{… N …} \rangle \\ \textsc{ss|l|c} \begin{bmatrix} \textsc{head} \begin{bmatrix} \textit{noun} \\ \textsc{count}\ + \\ \textsc{agr}\ \textit{3rdsing} \end{bmatrix} \\ \textsc{val} \begin{bmatrix} \textsc{spr}\ \ \langle \text{Det} \rangle \\ \textsc{comps}\ \langle\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{sr}\ \ldots\ \lambda x.N'(x)\ \ldots \end{bmatrix} \right)^{+} \right\rangle
\end{bmatrix}$$

There is a list of daughters consisting of a first daughter and an arbitrarily long list of *after* N pairs. The '+' means that there has to be at least one *after* N pair. The nominal daughters select for a determiner via SPR, so they can be either bare nouns or nouns modified by adjectives. The semantic representation, non-standardly represented as the value of SR, says that there have to be several objects in a set X ($\exists X.|X| > 1$) and for all of them, the meaning of the $\overline{\text{N}}$ has to hold ($\forall x \in X : N'(x)$). Furthermore there is an order between the elements of X as stated by $\exists R^{order} \subseteq X^2$.

From looking at this construction, it is clear that it cannot be accounted for by standard $\overline{\text{X}}$ rules. Even without requiring $\overline{\text{X}}$ syntactic rules, there seems to be no

[22] Jackendoff and Bargmann assume that the result of combining N, P, and N is an NP. However this is potentially problematic, as Matsuyama's example in (40) shows (Matsuyama 2004: 71):

(i) All ranks joined in hearty cheer after cheer for every member of the royal family …

As Matsuyama points out, the reading of such examples is like the reading of *old men and women* in which *old* scopes over both *men* and *women*. This is accounted for in structures like the one indicated in (ii):

(ii) hearty [cheer after cheer]

Since adjectives attach to $\overline{\text{N}}$s and not to NPs, this means that NPN constructions should be $\overline{\text{N}}$s. Of course (ii) cannot be combined with a determiner, so one would have to assume that NPN constructions select for a determiner that has to be dropped obligatorily. Determiners are also dropped in noun phrases with mass nouns with a certain reading.

way to capture these constructions in head-based approaches like Minimalism, Categorial Grammar, or Dependency Grammar. For simple NPN constructions, one could claim that *after* is the head. *after* would be categorized as a third-person singular mass noun and select for two $\overline{\text{N}}$s. It would (non-compositionally) contribute the semantics stated above. But it is unclear how the general schema with arbitrarily many repetitions of *after* N could be accounted for. If one assumes that *day after day* forms a constituent, then the first *after* in (41) would have to combine an N with an NPN sequence.

(41) day after [day [after day]]

This means that we would have to assume two different items for *after*: one for the combination of $\overline{\text{N}}$s and another one for the combination of $\overline{\text{N}}$ with NPN combinations. Note that an analysis of the type in (41) would have to project information about the $\overline{\text{N}}$s contained in the NPN constructio,n since this information has to be matched with the single $\overline{\text{N}}$ at the beginning. In any case, a lexical analysis would require several highly idiosyncratic lexical items (prepositions projecting nominal information and selecting items they usually do not select). It is clear that a reduplication account of the NPN construction as suggested by G. Müller (2011) does not work, since patterns with several repetitions of PN as in (41) cannot be accounted for as reduplication. G. Müller (p. 241) stated that reduplication works for word-size elements only (in German) and hence his account does not extend to the English examples given above. (42) shows an attested German example containing adjectives, which means that G. Müller's approach is not appropriate for German either.

(42) Die beiden tauchten nämlich geradewegs wieder aus dem heimischen
the two surfaced namely straightaway again from the home
Legoland auf, wo sie im Wohnzimmer, schwarzen Stein um
Legoland PART where they in.the living.room black brick after
schwarzen Stein, vermeintliche Schusswaffen nachgebaut hatten.[23]
black brick alleged firearms recreated had
'The two surfaced straightaway from their home Legoland where they had recreated alleged firearms black brick after black brick.'

This subsection showed how a special phrasal pattern can be analyzed within HPSG. The next section will discuss filler-gap constructions, which were analyzed as instances of a single schema by Pollard & Sag (1994) but which were

[23] taz, 05.09.2018, p. 20

later reconsidered and analyzed as a family of subconstructions by Sag (1997; 2010).

4.2 Specialized sub-constructions

HPSG took over the treatment of nonlocal dependencies from GPSG (Gazdar 1981) (see also Flickinger, Pollard & Wasow 2019, Chapter 2 of this volume on the history of HPSG and Borsley & Crysmann 2019, Chapter 14 of this volume on unbounded dependencies). Pollard & Sag (1994: Chapters 4 and 5) had an analysis of topicalization constructions like (43) and an analysis of relative clauses. However, more careful examination revealed that more fine-grained distinctions have to be made. Sag (2010: 491) looked at the following examples:

(43) a. [My bagels,] she likes. (topicalized clause)
b. [*What* books] do they like? (*wh*-interrogative)
c. (the person) [*who* (*se* book)] they like (*wh*-relative)
d. [*What a* play] he wrote! (*wh*-exclamative)
e. [*the more* books] they read … (the-clause)

As Sag shows, the fronted element is specific to the construction at hand:

(44) a. * [*Which* bagels] / [*Who*], she likes. (topicalized clause)
b. * [*What a* book] do they like? (*wh*-interrogative)
c. % the thing [[*what*] they like] (*wh*-relative)
d. * [*Which* bagels] / [*What*] she likes! (*wh*-exclamative)
e. * [*which* books] they read, the more they learn. (the-clause)

A topicalized clause should not contain a *wh* item (44a), a *wh*-interrogative should not contain a *what a* sequence as known from *wh*-exclamatives (44b), and so on.

Furthermore, some of these constructions allow non-finite clauses and others do not:

(45) a. * Bagels, (for us) to like. (topicalized clause)
b. * It's amazing [what a dunce (for them) to talk to]. (*wh*-exclamative)
c. * The harder (for them) to come, the harder (for them) to fall. (the-clause)
d. I know how much time (* for them) to take. (*wh*-interrogative)
e. The time in which (*for them) to finish. (*wh*-relative)

So there are differences as far as fillers and sentences from which something is extracted are concerned. Sag discussed further differences like inversion/non-inversion in the clauses out of which something is extracted. I do not repeat the full discussion here but refer the reader to the original paper.

In principle, there are several ways to model the phenomena. One could assume empty heads as Pollard & Sag (1994: Chapter 5) suggested for the treatment of relative clauses. Or one could assume empty heads as they are assumed in Minimalism: certain so-called operators have features that have to be checked and cause items with the respective properties to move. Borsley (2006) discussed potential analyses of relative clauses involving empty heads and showed that one would need a large number of such empty heads, and since there is no theory of the lexicon in Minimalism, generalizations are missed (see also Borsley & Müller 2019, Chapter L of this volume). The alternative suggested by Sag (2010) is to assume a general Filler-Head Schema of the kind assumed in Pollard & Sag (1994) and then define more specific sub-constructions. To take an example, the *wh*-exclamative is a filler-head structure, so it inherits everything from the more general construction, but in addition, it specifies that the filler daughter must contain a *what a* part and states the semantics that is contributed by the exclamative construction.

5 Summary

This paper summarized the properties of Construction Grammar, or rather Construction Grammars, and showed that HPSG can be seen as a Construction Grammar, since it fulfills all the tenets assumed in CxG: it is surface-based, grammatical constraints pair form and function/meaning, the grammars do not rely on innate domain-specific knowledge, and the grammatical knowledge is represented in inheritance hierarchies. This sets HPSG and CxG apart from other generative theories that either assume innate language-specific knowledge (Minimalism, e.g., Chomsky 2013; Kayne 1994; Cinque & Rizzi 2010) or do not assume inheritance hierarchies for all linguistic levels (e.g., LFG).

I showed why lexical analyses of argument structure should be preferred over phrasal ones and that there are other areas in grammar where phrasal analyses are superior to lexical ones. I showed that they can be covered in HPSG, while they are problematic for proposals assuming that all structures have to have a head.

Acknowledgements

I thank Anne Abeillé, Bob Borsley, Rui Chaves, and Jean-Pierre Koenig for comments on earlier versions of this chapter and for discussion in general. I thank Frank Richter for discussion of formal properties of SBCG.

References

Abeillé, Anne & Yves Schabes. 1989. Parsing idioms in Lexicalized TAG. In Harold Somers & Mary McGee Wood (eds.), *Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics*, 1–9. Manchester, England: Association for Computational Linguistics.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2008. Constructions with lexical integrity: Templates as the lexicon-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2008 conference*. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/13/, accessed 2018-2-25.

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2013. Constructions with lexical integrity. *Journal of Language Modelling* 1(1). 1–54.

Asudeh, Ash, Gianluca Giorgolo & Ida Toivonen. 2014. Meaning and valency. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2014 conference*, 68–88. Stanford, CA: CSLI Publications.

Bargmann, Sascha. 2015. *Syntactically flexible VP-idioms and the N-after-N Construction*. Poster presentation at the 5th General Meeting of PARSEME, Iasi, 23–24 September 2015.

Bar-Hillel, Yehoshua, Micha A. Perles & Eliahu Shamir. 1961. On formal properties of simple phrase-structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 14(2). 143–172.

Behrens, Heike. 2009. Konstruktionen im Spracherwerb. *Zeitschrift für Germanistische Linguistik* 37(3). 427–444.

Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford University dissertation. http://faculty.washington.edu/ebender/dissertation/, accessed 2018-2-25.

Bender, Emily & Daniel P. Flickinger. 1999. Peripheral constructions and core phenomena: Agreement in tag questions. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 199–214. Stanford, CA: CSLI Publications.

Bergen, Benjamin K. & Nancy Chang. 2005. Embodied Construction Grammar in simulation-based language understanding. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction Grammars: Cognitive grounding and theoretical extensions*, 147–190. Amsterdam: John Benjamins Publishing Co.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the Stimulus revisited. *Cognitive Science* 35(7). 1207–1242.

Bianchi, Valentina & Cristiano Chesi. 2006. Phases, left branch islands, and computational nesting. In *Proceedings of the 29th Penn Linguistic Colloquium* (Upenn Working Papers in Linguistics 12.1), 15–28.

Bianchi, Valentina & Cristiano Chesi. 2012. Subject islands and the subject criterion. In Valentina Bianchi & Chesi Cristiano (eds.), *Enjoy linguistics! papers offered to luigi rizzi on the occasion of his 60th birthday*. Siena: CISCL Press.

Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491.

Bishop, Dorothy V. M. 2002. Putting language genes in perspective. *TRENDS in Genetics* 18(2). 57–59.

Boas, Hans C. 2003. *A Constructional approach to resultatives* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Boas, Hans C. 2014. Lexical approaches to argument structure: Two sides of the same coin. *Theoretical Linguistics* 40(1–2). 89–112.

Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based Construction Grammar* (CSLI Lecture Notes 193). Stanford, CA: CSLI Publications.

Bod, Rens. 2009. From exemplar to grammar: Integrating analogy and probability in language learning. *Cognitive Science* 33(4). 752–793. DOI:10.1111/j.1551-6709.2009.01031.x

Booij, Geert E. 2005. Construction-Dependent Morphology. *Lingue e linguaggio* 4. 31–46.

Borsley, Bob & Anne Abeillé. 2019. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 3–5. Berlin: Language Science Press. DOI:??

Borsley, Bob & Berthold Crysmann. 2019. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 353–353. Berlin: Language Science Press. DOI:??

Borsley, Robert D. 1987. *Subjects and complements in HPSG*. Report No. CSLI-87-107. Stanford, CA: Center for the Study of Language & Information.

Borsley, Robert D. 2006. *Syntactic and lexical approaches to unbounded dependencies*. Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.

Borsley, Robert D. & Stefan Müller. 2019. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 937–1012. Berlin: Language Science Press. DOI:??

Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1). 1–65.

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell Publishers Ltd.

Brew, Chris. 1995. Stochastic HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, 83–89. Dublin: Association for Computational Linguistics.

Briscoe, Ted J. & Ann Copestake. 1999. Lexical rules in constraint-based grammar. *Computational Linguistics* 25(4). 487–526. http://www.aclweb.org/anthology/J99-4002, accessed 2018-10-7.

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425–455.

Chesi, Cristiano. 2007. An introduction to phase-based Minimalist Grammars: Why *move* is top-down from left to right. In.

Chesi, Cristiano. 2012. *Phases and Cartography in linguistic computation: Toward a cognitively motivated computational model of linguistic competence*. University of Siena dissertation.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1971. *Problems of knowledge and freedom*. London: Fontana.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris Publications.

Chomsky, Noam. 1995. *The Minimalist Program* (Current Studies in Linguistics 28). Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI:10.1016/j.lingua.2012.12.003

Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 51–65. Oxford: Oxford University Press.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 3(2–3). 281–332. DOI:10.1007/s11168-006-6327-9

Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2003. Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in language: Studies in honour of Günter Radden* (Current Issues in Linguistic Theory 243), 49–68. Amsterdam: John Benjamins Publishing Co. DOI:10.1075/cilt.243.07cro

Dąbrowska, Ewa. 2004. *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Washington, D.C.: Georgetown University Press.

Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2001. Weak crossover and the absence of traces. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2001 conference*, 66–82. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/6/, accessed 2018-2-25.

Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2004. Linguistic generalizations over descriptions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2004 conference*, 199–208. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/9/, accessed 2018-2-25.

Davis, Anthony R. & Jean-Pierre Koenig. 2000. Linking as constraints on word classes in a hierarchical lexicon. *Language* 76(1). 56–91.

Davis, Anthony & Jean-Pierre Koenig. 2019. The nature and role of the lexicon in HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 47–92. Berlin: Language Science Press. DOI:??

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Fillmore, Charles. 1982. Frame Semantics. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 111–137. Seoul: Hanshin Publishing Co.

Fillmore, Charles J. 1985a. Frames and the semantics of understanding. *Quaderni di Semantica* 6(2). 22–53.

Fillmore, Charles J. 1985b. Syntactic intrusions and the notion of grammatical constructions. *Berkeley Linguistic Society* 11. 73–86.

Fillmore, Charles J. 1988. The mechanisms of "Construction Grammar". In Shelley Axmaker, Annie Jaisser & Helen Singmaster (eds.), *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, 35–55. Berkeley, CA: Berkeley Linguistics Society.

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538.

Flickinger, Dan, Carl Pollard & Tom Wasow. 2019. The evolution of HPSG. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 7–43. Berlin: Language Science Press. DOI:??

Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL.

Fordham, Andrew & Matthew Walter Crocker. 1994. Parsing with principles and probabilities. In Judith L. Klavans Philip Resnik (ed.), *The balancing act: combining symbolic and statistical approaches to language*. Las Cruces, New Mexico, USA: Association for Computational Linguistics.

Freudenthal, Daniel, Julian M. Pine, Javier Aguado-Orea & Fernand Gobet. 2007. Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science* 31(2). 311–341. DOI:10.1080/15326900701221454

Freudenthal, Daniel, Julian M. Pine & Fernand Gobet. 2006. Modeling the development of children's use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science* 30(2). 277–310. DOI:10.1207/s15516709cog0000_47

Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus* (studia grammatica 35). Berlin: Akademie Verlag.

Fried, Mirjam. 2015. Construction Grammar. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42.3 (Handbooks of Linguistics and Communication Science), 974–1003. Berlin: Mouton de Gruyter. DOI:10.1515/9783110363685

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12. 155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag. 1985. *Generalized Phrase Structure Grammar*. Cambridge, MA: Harvard University Press.

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.

Ginzburg, Jonathan. 2019. Acquisition. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 773–774. Berlin: Language Science Press. DOI:??

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: The University of Chicago Press.

Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. *Berkeley Linguistic Society* 22. 132–146. http://roa.rutgers.edu/files/415-0900/roa-415-goldberg-2.pdf, accessed 2018-3-20.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford: Oxford University Press.

Goldberg, Adele E. 2009. The nature of generalization in language. *Cognitive Linguistics* 20(1). 93–127.

Goldberg, Adele E. 2013. Argument structure Constructions vs. lexical rules or derivational verb templates. *Mind and Language* 28(4). 435–465. DOI:10.1111/mila.12026

Goldberg, Adele E. & Ray S. Jackendoff. 2004. The English resultative as a family of Constructions. *Language* 80(3). 532–568.

Grewendorf, Günther. 1993. German: A grammatical sketch. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax – Ein internationales Handbuch zeitgenössischer Forschung*, vol. 9.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1288–1319. Berlin: Walter de Gruyter Verlag. DOI:10.1515/9783110095869.1

Haider, Hubert. 1993. *Deutsche Syntax – generativ: Vorstudien zur Theorie einer projektiven Grammatik* (Tübinger Beiträge zur Linguistik 325). Tübingen: Gunter Narr Verlag.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298(5598). 1569–1579. DOI:10.1126/science.298.5598.1569

Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Vol. 13 (Blackwell Textbooks in Linguistics). Oxford: Blackwell Publishers Ltd.

Höhle, Tilman N. 1997. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Christa Dürscheid, Karl Heinz Ramers & Monika Schwarz (eds.), *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, 107–120. Tübingen: Max Niemeyer Verlag. Republished as Höhle (2018b).

Höhle, Tilman N. 2018a. Spuren in HPSG. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 491–498. Berlin: Language Science Press. DOI:10.5281/zenodo.1145680

Höhle, Tilman N. 2018b. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 417–433. Berlin: Language Science Press. First published as Höhle (1997). DOI:10.5281/zenodo.1145680

Hunter, Tim. 2010. Unifying adjunct islands and freezing effects in Minimalist Grammars. In Srinivas Bangalore, Robert Frank & Maribel Romero (eds.), *Proceedings of the 10th International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+ 10)*, 17–24. Yale: Linguistic Department, Yale University. http://aclweb.org/anthology/W10-4403, accessed 2019-3-22.

Hunter, Tim. 2018. Left-corner parsing of Minimalist Grammars. In Bob Berwick & Edward Stabler (eds.). Oxford: Oxford University Press.

Jackendoff, Ray S. 2008. Construction after Construction and its theoretical challenges. *Language* 84(1). 8–28.

Jacobson, Pauline. 1987. Review of Generalized Phrase Structure Grammar. *Linguistics and Philosophy* 10(3). 389–426.

Johnson, Mark. 1986. A GPSG account of VP structure in German. *Linguistics* 24(5). 871–882.

Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.), *Formal issues in Lexical-Functional Grammar* (CSLI Lecture Notes 47), 7–27. Stanford, CA: CSLI Publications.

Kaplan, Ronald M. & Annie Zaenen. 1989. Long-distance dependencies, constituent structure and functional uncertainty. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*, 17–42. Chicago/London: The University of Chicago Press.

Kay, Paul. 2002. An informal sketch of a formal architecture for Construction Grammar. *Grammars* 5(1). 1–19.

Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The What's X Doing Y? Construction. *Language* 75(1). 1–33.

Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. *A lexical theory of phrasal idioms*. Ms. CSLI Stanford.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.

Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

Koenig, Jean-Pierre & Frank Richter. 2019. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 675–708. Berlin: Language Science Press. DOI:??

Kolb, Hans-Peter & Craig L. Thiersch. 1991. Levels and empty categories in a Principles and Parameters based approach to parsing. In Hubert Haider & Klaus Netter (eds.), *Representation and derivation in the theory of grammar* (Studies in Natural Language and Linguistic Theory 22), 251–301. Dordrecht: Kluwer Academic Publishers.

König, Esther. 1999. LexGram: A practical Categorial Grammar formalism. *Journal of Language and Computation* 1(1). 33–52.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris Publications.

Koster, Jan. 1987. *Domains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris Publications.

Kubota, Yusuke. 2019. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 1013–1069. Berlin: Language Science Press. DOI:??

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. Vol. 1. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford, CA: CSLI Publications.

Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

Lichte, Timm & Laura Kallmeyer. 2017. Tree-Adjoining Grammar: A tree-based constructionist grammar framework for natural language understanding. In *Proceedings of the AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding* (Technical Report SS-17-02). Association for the Advancement of Artificial Intelligence.

Lohnstein, Horst. 1993. *Projektion und Linking: Ein prinzipienbasierter Parser fürs Deutsche* (Linguistische Arbeiten 287). Tübingen: Max Niemeyer Verlag.

Matsuyama, Tetsuya. 2004. The N *after* N Construction: A constructional idiom. *English Linguistics* 1(11). 55–84.

Meisel, Jürgen. 1995. Parameters in acquisition. In Paul Fletcher & Brian MacWhinny (eds.), *The handbook of child language*, 10–35. Oxford: Blackwell Publishers Ltd. DOI:10.1111/b.9780631203124.1996.00002.x

Meurers, Walt Detmar. 1999. German partial-VP fronting revisited. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 129–144. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar. 2000. *Lexical generalizations in the syntax of German non-finite constructions*. Arbeitspapiere des SFB 340 No. 145. Tübingen: Eberhard-Karls-Universität. http://www.sfs.uni-tuebingen.de/~dm/papers/diss.html, accessed 2018-2-25.

Meurers, Walt Detmar. 2001. On expressing lexical generalizations in HPSG. *Nordic Journal of Linguistics* 24(2). 161–217.

Miller, Philip. 2013. Usage preferences: The case of the English verbal anaphor *do so*. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 121–139. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2013/miller.pdf, accessed 2018-8-17.

Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.

Müller, Gereon. 2011. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. In Stefan Engelberg, Anke Holler & Kristel Proost (eds.), *Sprachliches Wissen zwischen Lexikon und Grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010), 211–249. Berlin: de Gruyter.

Müller, Stefan. 1996. Yet another paper about partial verb phrase fronting in German. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96: 16th International Conference on Computational Linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*, 800–805. Copenhagen, Denmark: Association for Computational Linguistics.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche* (Linguistische Arbeiten 394). Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2003. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI:10.1017/S0022226703002032

Müller, Stefan. 2005. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201. 3–39.

Müller, Stefan. 2006a. Phrasal or lexical Constructions? *Language* 82(4). 850–883. DOI:10.1353/lan.2006.0213

Müller, Stefan. 2006b. Resultativkonstruktionen, Partikelverben und syntaktische vs. lexikonbasierte Konstruktionen. In Kerstin Fischer & Anatol Stefanowitsch (eds.), *Konstruktionsgrammatik: Von der Anwendung zur Theorie* (Stauffenburg Linguistik 40), 177–202. Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2007. Phrasal or lexical Constructions: Some comments on underspecification of constituent order, compositionality, and control. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 373–393. Stanford, CA: CSLI Publications.

Müller, Stefan. 2010. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3). 601–655. DOI:10.1017/S0022226709990284

Müller, Stefan. 2013. Unifying everything: Some remarks on Simpler Syntax, Construction Grammar, Minimalism and HPSG. *Language* 89(4). 920–950. DOI:10.1353/lan.2013.0061

Müller, Stefan. 2014. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford, CA: CSLI Publications.

Müller, Stefan. 2016. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2017a. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Revise and resubmit.

Müller, Stefan. 2017b. Head-Driven Phrase Structure Grammar, Sign-Based Construction Grammar, and Fluid Construction Grammar: Commonalities and differences. *Constructions and Frames* 9(1). 139–174. DOI:10.1075/cf.9.1.05mul

Müller, Stefan. 2018a. *A lexicalist account of argument structure: Template-based phrasal LFG approaches and a lexical HPSG alternative* (Conceptual Foundations of Language Science 2). Berlin: Language Science Press. DOI:10.5281/zenodo.1441351

Müller, Stefan. 2018b. *Grammatical theory: From Transformational Grammar to constraint-based approaches*. 2nd edn. (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.5281/zenodo.1193241

Müller, Stefan. 2019a. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 267–311. Berlin: Language Science Press. DOI:??

Müller, Stefan. 2019b. *Germanic syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin. https://hpsg.hu-berlin.de/~stefan/Pub/germanic.html, accessed 2019-6-30.

Müller, Stefan & Stephen Mark Wechsler. 2014a. Lexical approaches to argument structure. *Theoretical Linguistics* 40(1–2). 1–76. DOI:10.1515/tl-2014-0001

Müller, Stefan & Stephen Mark Wechsler. 2014b. Two sides of the same slim Boojum: Further arguments for a lexical approach to argument structure. *Theoretical Linguistics* 40(1–2). 187–224. DOI:10.1515/tl-2014-0009

Nerbonne, John. 1986. 'Phantoms' and German fronting: Poltergeist constituents? *Linguistics* 24(5). 857–870. DOI:10.1515/ling.1986.24.5.857

Pinker, Steven & Ray S. Jackendoff. 2005. The faculty of language: What's special about it? *Cognition* 95(2). 201–236.

Pollard, Carl J. & Ivan A. Sag. 1987. *Information-based syntax and semantics* (CSLI Lecture Notes 13). Stanford, CA: CSLI Publications.

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar* (Studies in Contemporary Linguistics). Chicago: The University of Chicago Press.

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Richter, Frank. 2007. Closer to the truth: A new model theory for HPSG. In James Rogers & Stephan Kepser (eds.), *Model-theoretic syntax at 10 – Proceedings of the ESSLLI 2007 MTS@10 Workshop, August 13–17*, 101–110. Dublin: Trinity College Dublin.

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 45–46. Berlin: Language Science Press. DOI:??

Richter, Frank & Manfred Sailer. 2009. Phraseological clauses as Constructions in HPSG. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 297–317. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2009/, accessed 2018-2-25.

Riehemann, Susanne. 1993. *Word formation in lexical type hierarchies: A case study of* bar-*adjectives in German*. Eberhard-Karls-Universität Tübingen MA thesis.

Also published as SfS-Report-02-93, Seminar für Sprachwissenschaft, University of Tübingen.
Riehemann, Susanne Z. 1998. Type-based derivational morphology. *Journal of Comparative Germanic Linguistics* 2(1). 49–77. DOI:10.1023/A:1009746617055
Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2). 431–484.
Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/, accessed 2018-2-25.
Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.
Sag, Ivan A. 2012. Sign-Based Construction Grammar: An informal synopsis. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 69–202. Stanford, CA: CSLI Publications.
Sag, Ivan A., Hans C. Boas & Paul Kay. 2012. Introducing Sign-Based Construction Grammar. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar* (CSLI Lecture Notes 193), 1–29. Stanford, CA: CSLI Publications.
Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/Cambridge, MA: Blackwell Publishers Ltd.
Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic theory: A formal introduction*. 2nd edn. (CSLI Lecture Notes 152). Stanford, CA: CSLI Publications.
Schabes, Yves, Anne Abeillé & Aravind K. Joshi. 1988. *Parsing strategies with 'lexicalized' grammars: Application to Tree Adjoining Grammars*. Technical Report MS-CIS-88-65. University of Pennsylvania Department of Computer & Information Science.
Stabler, Edward P. 2011. Computational perspectives on Minimalism. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic Minimalism* (Oxford Handbooks in Linguistics), chap. 27, 616–641. Oxford: Oxford University Press.
Steels, Luc (ed.). 2011. *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11). Amsterdam: John Benjamins Publishing Co.
Steels, Luc & Joachim De Beule. 2006. *A (very) brief introduction to Fluid Construction Grammar*. Paper presented at the Third International Workshop on Scalable Natural Language Understanding (ScaNaLU 2006) June 8, 2006, following HLT/NAACL, New York City.

Stiebels, Barbara. 1996. *Lexikalische Argumente und Adjunkte: Zum semantischen Beitrag verbaler Präfixe und Partikeln* (studia grammatica 39). Berlin: Akademie Verlag.

Toivonen, Ida. 2013. English benefactive NPs. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2013 conference*, 503–523. Stanford, CA: CSLI Publications.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Luc Steels (ed.), *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11), 115–145. Amsterdam: John Benjamins Publishing Co. DOI:10.1075/cal.11.07tri

van Trijp, Remi. 2013. A comparison between Fluid Construction Grammar and Sign-Based Construction Grammar. *Constructions and Frames* 5(1). 88–116.

Veenstra, Mettina Jolanda Arnoldina. 1998. *Formalizing the Minimalist Program*. Rijksuniversiteit Groningen Ph.D. thesis.

Verspoor, Cornelia Maria. 1997. *Contextually-dependent lexical semantics*. University of Edinburgh dissertation.

Walther, Markus. 1999. *Deklarative prosodische Morphologie: Constraint-basierte Analysen und Computermodelle zum Finnischen und Tigrinya* (Linguistische Arbeiten 399). Tübingen: Max Niemeyer Verlag.

Wechsler, Stephen Mark. 1997. Resultative predicates and control. In Ralph C. Blight & Michelle J. Moosally (eds.), *Texas Linguistic Forum 38: The syntax and semantics of predication: Proceedings of the 1997 Texas Linguistics Society Conference*, 307–321. Austin, Texas: University of Texas, Department of Linguistics.

Wechsler, Stephen Mark & Bokyung Noh. 2001. On resultative predicates and clauses: Parallels between Korean and English. *Language Sciences* 23(4). 391–423.

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar*, 225–265. Berlin: Language Science Press. DOI:??

Westergaard, Marit. 2014. Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation* 14(1). 26–45.

Name index

Aamot, Elias, 26
Abbott, Miriam, 97
Abeillé, Ann, 374, 379–382
Abeillé, Anne, 21, 93, 97, 132, 189, 190, 197, 215–218, 255, 274, 275, 278, 282, 283, 287, 301, 317, 318, 320, 322, 325, 326, 328, 329, 386, 420, 422, 423, 432, 439, 447, 474, 478, 479, 481, 482, 491–493, 495, 501, 505, 553, 554, 559, 562, 627, 759, 943–945, 975, 977, 1048, 1049, 1051, 1052, 1086, 1087, 1093, 1107
Abner, Natasha, 887
Abney, Stephen, 187
Abrusán, Márta, 416
Abzianidze, Lasha, 26
Ackerman, Farrell, 65, 69, 70, 235, 236, 250
Ades, Anthony E., 1021, 1057
Adger, David, 106
Adolphs, Peter, 809
Aissen, Judith, 255, 322
Ajdukiewicz, Kazimierz, 274, 1018
Akmajian, Adrian, 427, 537
Alahverdzhieva, Katya, 887, 892, 898, 899, 906, 907, 910–912
Alexopoulou, Theodora, 731
Alibali, Martha W., 918
Allegranza, Valerio, 187, 195, 201
Allen, Cynthia, 748
Allwood, Jens, 423, 916
Almog, Lior, 536
Alotaibi, Mansour, 174, 374
Alqurashi, Abdulrahman, 373, 376, 377, 401
Altmann, Gerry, 779, 781
Altmann, Gerry T. M., 426
Ambridge, Ben, 711
Amy, Gérard, 437
An, Aixiu, 491, 945
Anderson, Stephen, 53, 54, 257
Andrews, Avery D., 108, 157
Antonov, Anton, 105
Aoun, Joseph, 984
Aoun, Joseph E., 946
Arad Greshler, Tali, 25
Arai, Manabu, 426
Aranovich, Raúl, 95
Argyle, Michael, 888
Ariel, Mira, 149
Arka, I Wayan, 233
Arka, I. Wayan, 97
Arnold, Doug, 207, 218, 389–391, 395, 450, 493, 848
Arnold, Jennifer E., 779
Arnon, Inbal, 444, 787
Aronoff, Mark, 518
Artawa, Ketut, 233
Artstein, Ron, 503
Asher, Nicholas, 615, 870

Asudeh, Ash, 670, 685, 691, 898, 1086, 1105
Auer, Peter, 919
Augustinus, Liesbeth, 759
Austin, John L., 841, 845, 863
Avgustinova, Tania, 373

Baayen, R Harald, 810
Bach, Emmon, 10, 14, 17, 18, 268, 675, 1040
Bachrach, Asaf, 424
Baker, C. L., 689
Baker, Curtis L., 553, 554, 556, 558
Baker, Mark, 231, 242
Baker, Mark C., 749, 960
Baldridge, Jason, 63, 800, 812, 1014, 1023, 1024, 1037, 1039, 1041, 1043
Baldwin, Timothy, 517–519, 521, 798, 814
Ball, Douglas L., 95, 108, 111, 112, 114
Baltin, Mark, 427
Banarescu, Laura, 815
Bangalore, Srinivas, 803
Bangerter, Adrian, 903
Bar-Hillel, Yehoshua, 18, 1018, 1083
Bargmann, Sascha, 523, 534, 978, 1110–1112
Barker, Chris, 1029, 1030, 1032, 1033, 1039, 1050
Barlow, Michael, 133
Barry, Guy, 1057
Barwise, John, 696
Barwise, Jon, 847, 859, 863
Bauer, Angelika, 919
Bausewein, Karin, 956, 957
Bavelas, Janet B., 893, 896, 914
Bayer, Samuel, 170, 496, 497, 1042
Beavers, John, 173, 249, 250, 420, 474, 477, 494, 497, 499, 500, 502, 1039, 1045
Bech, Gunnar, 330, 759
Beermann, Dorothee, 814
Behrens, Heike, 1102
Beißwenger, Michael, 843
Bekki, Daisuke, 1018, 1040
Belletti, Adriana, 552, 572
Bender, Emily, 553, 558, 570, 571
Bender, Emily M, 814
Bender, Emily M., 24–26, 28, 94, 104, 117, 174, 299, 473–475, 477, 528, 530, 531, 710, 713, 714, 723, 724, 803, 808, 812, 813, 815, 817, 899, 939, 940, 946, 1083, 1090, 1091, 1095
Beneš, Eduard, 296
Bergen, Benjamin K., 1090, 1091, 1102
Berman, Judith, 206
Bernardi, Raffaella, 1033, 1044
Berwick, Robert, 422
Berwick, Robert C., 28, 942, 945, 1085
Bever, Thomas, 780, 789
Bever, Thomas G., 435
Bhatt, Rajesh, 148
Bianchi, Valentina, 1083
Biberauer, Theresa, 749, 750, 768
Bickel, Balthasar, 104
Bierman, Arthur K., 908
Bierwisch, Manfred, 268
Bildhauer, Felix, 296–298, 710, 714, 734–738, 945
Bird, Steven, 669, 670, 815, 1098
Birdwhistell, Ray L., 887
Bishop, Dorothy V. M., 1084
Bjerre, Anne, 401
Bjerre, Tavs, 281

Blake, Barry J., 162
Blakemore, Diane, 390
Blanco, Eduardo, 818
Blank, Malou, 842
Blevins, James P., 287
Bloomfield, Leonard, 10, 472
Blunsom, Phil, 809, 819
Boas, Hans C., 175, 367, 1087, 1090, 1101
Bochner, Harry, 50, 85
Bod, Rens, 450, 1085
Boeckx, Cedric, 984, 989
Boguraev, Bran, 798
Bolc, Leonard, 373
Bolinger, Dwight, 431, 443
Bolt, Richard A., 897
Bonami, Olivier, 52, 82, 83, 287, 318, 320, 326, 333, 341, 343, 482, 569, 670, 691, 847
Bond, Francis, 803, 808, 815, 819
Booij, Geert E., 1107
Booij, Gert, 53
Bopp, Franz, 138
Borer, Hagit, 62
Borgonovo, Claudia, 440
Borsley, Bob, 21, 97, 189, 190, 267, 268, 282, 292, 358, 416, 418, 479, 482–484, 488, 527, 942, 984, 985, 1092, 1093, 1114
Borsley, Robert, 472, 474, 553, 566, 573, 574, 577, 578
Borsley, Robert D, 365
Borsley, Robert D., 21, 24, 105, 111, 112, 150, 174, 277, 278, 282, 283, 285, 286, 373, 374, 376, 377, 390, 391, 400, 401, 472, 481, 482, 492, 522, 537, 848, 939, 941, 950, 966, 967, 975, 976, 984, 985, 1013, 1014, 1084, 1096, 1115
Bosque, Ignacio, 494
Bouma, Gosse, 25, 26, 74, 80, 86, 166, 195, 230, 274, 275, 281, 283, 317, 330, 340, 393, 418, 427, 449, 759, 801, 810, 811, 815, 971, 1082, 1083
Boyé, Gilles, 83
Bošković, Željko, 976
Brachman, Ronald, 63
Branco, António, 25, 93, 114, 277–279, 614, 615, 623, 624, 626–628, 630, 631, 652, 846, 847, 973
Branigan, Holly, 781–786
Bratt, Elizabeth Owen, 162, 569
Breckinridge Church, Ruth, 918
Bresnan, Joan, 9, 50, 52, 101, 139, 187, 260, 448, 501, 566, 776, 777, 781, 937, 956, 1083
Brew, Chris, 1091
Briscoe, Ted, 71, 76, 78, 1091, 1106
Broadwell, George Aaron, 53, 55
Brown, Ralf, 615
Bruening, Benjamin, 48, 51, 52, 148, 493, 977
Bruland, Tore, 26
Bruneau, Thomas J., 888
Bruner, Jerome S., 902
Buitelaar, Paul, 503
Burger, Harald, 518, 521, 539
Burks, Arthur W., 908
Butt, Miriam, 155, 165, 316, 769, 812, 814
Butterworth, George, 895
Buys, Jan, 809, 819
Bybee, Joan, 1080
Bîlbîie, Gabriela, 480

Börjars, Kersti, 750, 1013, 1014
Bühler, Karl, 903
Büring, Daniel, 727, 1052
Bîlbîie, Gabriela, 389, 478, 480

Cairns, Helen Smith, 437
Calder, Jonathan, 1042
Callmeier, Ulrich, 23, 808, 809
Carbonell, Jaime, 615
Carletta, Jean, 916
Carlson, Greg, 244
Carpenter, Bob, 25, 63, 81, 857, 1025, 1026
Carreiras, Manuel, 429
Carroll, John, 799, 801, 812
Carroll, John Andrew, 801
Carter, David, 804, 815
Cassell, Justin, 917
Cattell, Ray, 440
Caute, Anna, 919
Cha, Jong-Yul, 398
Chan, Chung, 397, 398
Chang, Nancy, 1090, 1091, 1102
Chatzikyriakidis, Stergios, 1042
Chaves, Rui, 53, 479, 788, 992, 1044, 1047
Chaves, Rui P., 132, 173, 275, 287, 419–422, 424, 428, 431–434, 438, 439, 441, 446, 472, 474, 477, 488, 494, 497–500, 502, 503, 505, 787, 943, 944, 975, 1051, 1052
Chaves, Rui Pedro, 1045, 1049
Chen, Yufei, 809, 818, 819
Chesi, Cristiano, 987, 988, 1083
Chierchia, Gennaro, 628
Choi, Incheol, 318, 341
Chomsky, Noam, 3, 8–10, 49, 99, 106, 120, 132, 156, 186, 195, 235, 284, 372, 431, 432, 434, 438, 440, 443, 472, 473, 522, 552, 555, 621, 631, 749, 775, 787, 798, 843, 846, 937, 938, 940, 942, 949, 950, 952–955, 959, 966, 973, 977, 979, 988, 989, 991, 992, 1077, 1078, 1081–1085, 1115
Christiansen, Morten H., 788
Chung, Chan, 317, 330, 333, 335, 337, 340, 568
Chung, Daeho, 569
Chung, Sandra, 423, 425, 1050
Cienki, Alan J., 897
Cinque, Guglielmo, 282, 423, 440, 970, 1084, 1115
Citko, Barbara, 957, 976
Clark, Herbert H., 437, 843, 851, 853, 894, 903
Clark, Stephen, 803, 820, 1058
Clausen, David R., 437, 438
Clemens, Lauren Eby, 110
Clifton, Charles E. Jr., 426, 429
Cohen, Philip R., 850, 899
Cohen, Shay B., 942
Cole, Peter, 946
Collins, Chris, 27
Colston, Herbert L., 518, 520
Comrie, Bernard, 18, 104, 105, 147, 232, 277, 788
Cook, Jeanette, 26
Cook, Mark, 888
Cook, Philippa, 296, 297, 714, 734–736, 945
Cook, Susan Wagner, 918
Cooper, Robin, 679, 696, 853, 858–860, 862, 863, 870, 893, 910, 915, 1054

Cooperrider, Kensy, 894, 895
Copestake, Ann, 22, 23, 25, 26, 51, 71, 72, 76, 78, 240, 369, 393, 424, 446, 477, 528, 530, 676, 691, 692, 694–696, 712, 804–806, 809, 820, 858, 907, 911, 952, 1039, 1091, 1100, 1106
Copestake, Ann A., 26, 28
Corbett, Greville, 140, 490, 491
Corbett, Greville G., 97
Correa, Nelson, 631
Corver, Norbert, 520
Costa, Francisco, 25, 93, 847
Coulson, Seanna, 892
Cowper, Elizabeth A., 435
Crain, Stephen, 781, 1057
Crawford, Jean, 438
Creary, Lewis G., 15
Creel, Sarah C., 426
Cresti, Diana, 446
Crocker, Matthew Walter, 1082
Croft, William, 1090, 1106
Crouch, Richard, 685, 691
Crowgey, Berthold, 553
Crowgey, Joshua, 94, 565, 570, 571
Crysmann, Berthold, 25, 52, 54, 55, 82, 83, 93, 170, 173, 174, 267, 268, 292, 358, 373, 377, 393, 394, 416, 418, 474, 479, 488, 490, 497, 499, 527, 691, 808, 942, 964, 984, 985, 1045, 1092, 1114
Cubelli, Roberto, 919
Culicover, Peter, 440, 450, 480, 481, 501, 678, 711
Culicover, Peter W., 28, 424, 432, 439, 520, 939–941, 961–964
Culy, Chris, 397
Culy, Christopher, 11
Curran, James, 1058
Curran, James R., 803
Curry, Haskell B., 1035, 1040
Cutler, Anne, 535

Dahl, Östen, 551
Dalrymple, Mary, 138, 172, 173, 618, 631, 1083, 1086
Daniels, Michael W., 168, 170, 172, 173, 497, 1042
Danon, Gabi, 138
Davies, William D., 416, 436
Davis, Anthony, 66, 78, 97, 105, 109, 134, 237–242, 246–252, 255, 257, 278, 279, 373, 524, 685, 813, 855, 972, 1085, 1101
Davis, Anthony R., 97, 108, 1101
De Beer, Carola, 918
De Beule, Joachim, 1090
de Groote, Philippe, 1035
De Kok, Daniël, 811
De Kuthy, Kordula, 317, 330, 331, 361, 710, 711, 713–716, 719–721, 725, 727–729, 732–735, 737, 968
de Ruiter, Jan P., 918
de Ruiter, Jan Peter, 888, 918
de Saussure, Ferdinand, 844
De Swart, Henriette, 553, 576, 577
De Vicenzi, Marica, 429
Deane, Paul, 441
Deane, Paul D., 786
Deane, Paul D., 424, 438, 440
Delahunty, Gerald P., 436
Delais-Roussarie, Elisabeth, 670
Dellert, Johannes, 808
DeLong, Katherine A., 426
Demberg, Vera, 869
Demske, Ulrike, 761, 763, 764

Den Dikken, Marcel, 480
Denison, David, 747, 748
Dery, Jeruen E., 432, 438, 439
Desmet, Timothy, 429
Devlin, Keith, 847
Dhongde, Ramesh Vaman, 358
Di Sciullo, Anna-Maria, 47
Dimitriadis, Alexis, 618
DiRusso, Alyssa A., 918
Dixon, R.M.W., 104, 108
Dobrovol’skij, Dmitrij, 537
Domaneschi, Filippo, 888
Donati, Caterina, 956, 957
Donnellan, Keith S., 862
Donohue, Cathryn, 94, 269, 290, 291, 299
Donohue, Mark, 104
Dorna, Michael, 25
Dost, Ascander, 54, 55
Dowty, David, 9, 10, 14, 17, 18, 63, 237–239, 244, 250, 287, 714, 1019, 1020, 1040, 1041, 1055, 1101
Dowty, David R., 245, 501
Drellishak, Scott, 26, 105, 174, 473–475, 477, 808
Dridan, Rebecca, 804, 810
Drummond, Alex, 433
Dryer, Matthew S, 551
Dryer, Matthew S., 109
Dubenion-Smith, Shannon A., 758
Dubinsky, Stanley, 416, 436
Dukes, Michael, 94
Duncan, Susan D., 918
Dutilh, MWF, 810
Dyła, Stefan, 166
Dörre, Jochen, 25
Dąbkowski, Maksymilian, 94
Dąbrowska, Ewa, 1090

Eady, Stephen J., 437
Eberhard, Kathleen M., 780
Ebert, Cornelia, 887, 898, 913
Ebert, Robert Peter, 753, 758, 765
Eco, Umberto, 908
Edlund, Jens, 843
Egg, Markus, 685
Eisele, Andreas, 808
Eisenberg, Peter, 279
Ejerhed, Eva, 787
Ekman, Paul, 888, 892, 894
Elbourne, Paul, 952, 983
Elbourne, Paul D., 432
Embick, David, 52
Emele, Martin C., 25
Emerson, Guy, 26, 78, 530, 820, 899, 940
Emonds, Joseph E., 10
Enfield, Nick J., 895
Engdahl, Elisabet, 423, 440, 710, 712, 713, 715, 717, 718, 725, 730, 731, 737, 848, 906, 908, 968
Engdahl, Elisabeth, 787
Engelkamp, Judith, 277, 971
Enger, Hans-Olav, 145, 146
Epstein, Samuel David, 28, 945, 947, 948
Erbach, Gregor, 25, 277, 525–529, 538, 971
Erk, Katrin, 820
Ernst, Thomas, 527, 553, 554
Erteschik-Shir, Nomi, 423, 424, 438, 440, 787
Eschenbach, Carole, 631
Eshghi, Arash, 869
Espinal, Teresa, 535
Evans, Gareth, 642

Evans, Nicholas, 257
Everaert, Martin, 533
Eynde, Frank Van, 278

Faarlund, Jan Terje, 145, 146
Faghiri, Pegah, 318, 538
Falk, Yehuda N., 946
Faltz, Leonard, 683
Fan, Zhenzhen, 25
Fang, Yimai, 818
Fanselow, Gisbert, 297, 952, 982, 983
Farkas, Richárd, 818
Fast, Jakub, 477
Federmeier, Kara D., 426
Fedorenko, Evelina, 444
Feldhaus, Anke, 283
Fellbaum, Christiane, 523
Ferdinand de Saussure, 844[1]
Fernando, Tim, 853
Fernández, Eva M., 429
Fernández, Raquel, 843, 844, 869
Ferreira, Fernanda, 430, 436, 437
Fillmore, Charles, 237, 238, 243, 244, 689, 910, 1100
Fillmore, Charles J., 28, 29, 284, 519, 521, 524, 532, 537, 1078–1082, 1086, 1089–1091, 1100
Findlay, Jamie Y., 535
Firth, John Rupert, 820
Fitch, W. Tecumseh, 949, 989, 1081, 1084, 1085
Fleischer, Wolfgang, 518
Flickinger, Dan, 15–17, 22–26, 28, 63, 93, 450, 530, 778, 798, 801, 802, 804, 808, 810, 816, 817, 819, 821, 858, 937, 952, 967, 972, 1078, 1085, 1087, 1095, 1114
Flickinger, Daniel, 17, 25, 26, 63, 82, 817
Fodor, Janet A., 435
Fodor, Janet D., 437
Fodor, Janet Dean, 416–418, 422, 429–431, 435, 437, 786, 939, 989
Fodor, Jerry A., 775
Fokkens, Antske, 26, 95
Fokkens, Antske Sibelle, 809, 822
Foley, William, 237
Fong, Sandiway, 631, 941
Ford, Marilyn, 781
Fordham, Andrew, 1082
Fox, Danny, 416, 424
Frampton, John, 106
Francis, Elaine, 777
Francis, Elaine J., 401
Frank, Anette, 283, 632, 635, 656
Franz, Alex, 25
Fraser, Bruce, 519
Frazier, Lyn, 429
Frazier, Lynn, 435
Frege, Gottlob, 859
Freidin, Robert, 434
Freudenthal, Daniel, 1085
Frey, Werner, 952, 981, 1082
Fricke, Ellen, 903, 905
Fried, Mirjam, 1089
Friedman, Joyce, 812
Friesen, Wallace V., 888, 892, 894
Fritz-Huechante, Paola, 160
Fujii, Mamoru, 11
Fukuda, Shin, 431
Fuss, Eric, 149

Gabriel, Aguila-Multner, 490
Galantucci, Bruno, 914
Gallin, Daniel, 704, 714
Garnsey, Susan, 437

Garrett, M. F., 435
Garrod, Simon, 914
Gawron, Jean Mark, 13, 241, 421, 628
Gazdar, Gerald, 8–11, 14, 15, 18, 186, 269, 282, 296, 420, 428, 429, 470, 472, 496, 522, 523, 525, 527, 542, 789, 799, 937, 942, 943, 1083, 1114
Geach, Peter Thomas, 1039, 1105
Gennari, Silvia P., 788
Georgi, Ryan, 814
Gerbl, Niko, 402
Gerdts, Donna, 243
Gerwing, Jennifer, 896
Ghayoomi, Masood, 807
Ghodke, Sumukh, 815
Gibbon, Dafydd, 916
Gibbs, Raymond W., 518
Gibbs, Raymond W. Jr., 518, 520
Gibson, Edward, 425, 426, 429, 435, 444, 788, 789
Ginsburg, Jason, 941
Ginzburg, Jonathan, 22, 59, 97, 98, 113, 187, 188, 190–193, 205, 206, 208, 209, 212, 216, 273, 274, 284, 320, 328, 360, 363, 370, 371, 374, 391, 418, 425, 443–445, 449, 450, 676, 686, 687, 689, 690, 843, 844, 851–855, 857–859, 862–864, 866, 868, 869, 873, 892, 893, 895, 910, 915, 919, 939, 940, 965, 966, 968, 980, 1050, 1051, 1085
Giorgolo, Gianluca, 898, 1086, 1105
Givón, Talmy, 138, 149
Glaser, Elvira, 752
Glucksberg, Sam, 518, 521
Godard, Danièle, 93, 255, 274, 275, 283, 287, 301, 317, 318, 322, 325, 326, 328, 329, 374, 379–382, 416, 493, 495, 553, 554, 559, 562, 627, 759
Goldberg, Adele, 711
Goldberg, Adele E., 48, 62, 80, 85, 240, 246, 252, 344, 424, 438, 440, 441, 1038, 1039, 1079–1081, 1090, 1101, 1105–1107
Golde, Karin, 615, 624, 625
Goldin-Meadow, Susan, 918
Goldsmith, John, 420, 472, 487
Gollan, Tamar H., 776
Goodall, Grant, 167, 431, 438
Goodard, Danièle, 495, 501
Goodluck, Helen, 423
Goodman, Michael Wayne, 819
Goodman, Nelson, 908
Goodwin, Charles, 842, 894, 905
Grano, Thomas, 487
Green, Georgia M, 849, 851
Green, Georgia M., 988
Greenberg, Joseph H., 138, 140, 232
Grewendorf, Günther, 163, 749, 980, 1083
Grimshaw, Jane, 106, 956
Grinevald, Colette, 140
Grishman, Ralph, 782
Groenendijk, Jeroen, 687
Groenendijk, Marius, 25
Grohmann, Kleantes K., 948, 949
Groos, Anneke, 167, 956
Gross, Maurice, 519
Grosu, A., 786
Grosu, Alexander, 392, 417, 429, 441, 501
Grover, Claire, 162
Guilfoyle, Eithne, 235

Gunji, Takao, 277, 382, 383, 387, 983
Gussenhoven, Carlos, 720, 968
Gutmann, Sam, 106
Guéron, Jacqueline, 433
Götz, Thilo, 25

Hackl, Martin, 416
Haddar, Kais, 377
Hadjikhani, Nouchine, 895
Haegeman, Liliane, 150, 552, 991
Hagstrom, Paul, 567
Hahn, Florian, 907
Hahn, Michael, 377, 401
Haider, Hubert, 156, 163, 281, 732, 753, 947, 1082
Haji-Abdolhosseini, Mohammad, 849
Hale, Kenneth, 260, 317, 962
Hall, Edward T., 888
Halle, Morris, 50, 52, 1042
Halliday, Michael A. K., 847, 968
Hamblin, Charles Leonard, 687, 853
Hammarström, Harald, 94
Han, Chung-hye, 567
Harris, Alice, 54
Harris, Zellig Sabbetai, 820
Hashimoto, Chikara, 799, 809
Haspelmath, Martin, 51–53, 96, 101, 106, 316, 317, 473, 538
Haugereid, Petter, 25, 93
Hauser, Marc D., 949, 989, 1081, 1084, 1085
Hawkins, John A., 422, 749, 776, 788, 790
Head, Brian F., 147
Healey, Patrick G. T., 869
Hegarty, Michael, 440
Heim, Irene, 629, 675, 848, 1081
Heinz, Wolfgang, 106, 156, 159–161
Heldner, Mattias, 843
Hellan, Lars, 25, 26, 93, 627, 814
Henderson, John M., 430, 436
Hendriks, Petra, 1048, 1050
Hepple, Mark, 631
Herbelot, Aurélie, 26
Herring, Joshua, 798
Hershcovich, Daniel, 806
Herzig Sheinfux, Livnat, 25
Higginbotham, James, 299, 630
Higgins, Francis Roger, 522
Hillyard, Steven A., 426
Hilpert, Martin, 566
Himmelmann, Nikolaus P., 105
Hinrichs, Erhard, 255, 759, 760
Hinrichs, Erhard W., 277, 287, 317, 330, 331, 361, 376, 401, 943, 983
Hiramatsu, Kazuko, 438
Ho, Jia Qian, 536
Hobbs, Jerry R., 782
Hoberg, Ursula, 297
Hockett, Charles F., 10, 11
Hoeksema, Jack, 1042
Hofmeister, Philip, 425, 426, 430, 433, 438, 787
Holler, Anke, 389, 897
Holmberg, Anders, 989
Horn, George M., 519
Horn, Laurence, 416
Hornstein, Norbert, 618, 678, 948, 949, 989, 991
Horvat, Matic, 819
Hough, Julian, 869, 916
Howard, William A., 1027, 1038
Hsiao, Franny, 789
Huang, C.-T. James, 627
Huang, Cheng-Teh James, 431, 434, 440

Huck, Geoffrey J., 431
Huddleston, Rodney, 395
Huddleston, Rodney D., 432, 433, 472, 486, 496, 502, 505
Hudson, Dick, 804, 848
Hudson, Richard, 187, 498, 504
Hukari, Thomas, 416
Hukari, Thomas E., 167, 275, 418, 422, 425, 426, 432, 434, 439, 441, 442, 445, 503, 1044
Hunter, Julie, 870
Hunter, Tim, 1083
Hutchins, Sean, 430, 436
Huybregts, Riny, 940
Härd, John Evert, 758
Höhle, Tilman N., 268, 305, 662, 669, 671, 980, 1001, 1076, 1086, 1105, 1121, 1122

Iatridou, Sabine, 618
Ingria, Robert J. P., 168
Iordăchioaia, Gianina, 691, 704
Itakura, Shoji, 895
Ivanova, Angelina, 805, 806

Jackendoff, Ray, 50, 51, 56, 85, 186, 244, 480, 481, 553, 678, 848, 941
Jackendoff, Ray S., 17, 28, 450, 520, 625, 711, 715, 939–941, 961–964, 978, 1080, 1101, 1105, 1110
Jacobs, Joachim, 719
Jacobson, Pauline, 282, 450, 497, 647, 1021, 1050, 1051, 1053, 1054, 1103
Jacques, Guillaume, 105
Jaeger, T. Florian, 787
Jakob, Hanna, 918
Janda, Richard D., 1042
Jannedy, Stefanie, 891
Jelinek, Eloise, 102
Jiménez–Fernández, Ángel, 433, 438
Job, Remo, 429
Johannessen, Janne Bondi, 474, 976
Johansson, Gunnar, 908, 909
Johnson, David E., 472, 848, 870
Johnson, Kyle, 1045, 1047
Johnson, Mark, 170, 628, 800, 1042, 1105
Johnston, Michael, 899, 917
Jones, Bob, 472, 474, 553, 574, 577, 578
Jones, Morris, 975
Joos, Martin, 992
Josefsson, Gunlög, 145, 146
Joshi, Aravind K., 13, 800, 803, 1107
Jurafsky, Daniel, 50, 81
Jurka, Johannes, 431
Just, Marcel, 790
Jäger, Gerhard, 631, 1025, 1026, 1028, 1050

Kallmeyer, Laura, 1086
Kamide, Yuki, 426, 779
Kamp, Hans, 629, 656, 677, 847, 848
Kanazawa, Makoto, 1036
Kandybowicz, Jason, 449
Kaplan, David, 890
Kaplan, Ronald M., 173, 937, 1083, 1086
Karttunen, Lauri, 503, 629
Kasami, Tadao, 11, 800
Kasper, Robert, 676, 683, 684
Kasper, Robert T., 20, 275, 971
Kasper, Walter, 22, 25, 808
Kathol, Andreas, 12, 55, 138, 146, 147, 150, 173, 281, 286, 287, 292–296, 317, 330, 332, 393, 401, 573, 759, 847, 850, 1040

Kato, Yasuhiko, 564
Katzir, Roni, 424
Kauffman, Lynn E., 888
Kay, Martin, 132, 691, 692, 800, 812
Kay, Paul, 85, 207, 533, 534, 537, 1078, 1079, 1081, 1086, 1087, 1089–1091
Kayne, Richard, 359, 417, 427, 428, 431, 443
Kayne, Richard S., 372, 474, 493, 783, 976, 977, 1115
Keenan, Edward, 232, 683, 704
Keenan, Edward L., 18, 105, 115, 277, 788
Kehler, Andrew, 420, 421, 472, 488
Kelepir, Meltem, 564
Keller, Frank, 292, 393, 426, 427, 951
Kellogg, Brainerd, 1073
Kelly, Spencer D., 892
Kemper, Susan, 437
Kempson, Ruth, 843, 844, 857
Kendon, Adam, 843, 887, 888, 895, 897, 914
Kennedy, Christopher, 1050, 1051
Keyser, Samuel, 245
Keyser, Samuel Jay, 260, 317, 962
Kiefer, Bernd, 26
Kikuta, Chiharu Uda, 396, 397
Kim, Jin-Dong, 817
Kim, Jong-Bok, 24, 25, 78, 93, 207, 218, 255, 283, 287, 317, 318, 330, 367, 382–385, 397–399, 401, 402, 427, 450, 494, 499, 505, 552–556, 558, 559, 561, 562, 564, 565, 568–570, 572–574
Kim, Su Nam, 517–519, 521
Kim, Yong-Beom, 397
Kimball, John, 435
King, Jeffrey C., 856
King, Jonathan, 790
King, Paul, 20, 82, 298, 524
King, Paul J., 855
King, Tracy Halloway, 138
King, Tracy Holloway, 165, 769, 812, 814, 821, 1086
Kinn, Kari, 747
Kipp, Michael, 915, 916
Kiss, Tibor, 164, 279, 283, 292, 317, 330, 368, 393, 615, 623, 762, 952, 982, 983, 1082
Kita, Sotaro, 888, 897, 919
Klein, Ewan, 8, 15, 269, 628, 669, 670, 675, 849, 906, 1083, 1098
Klein, Wolfgang, 842, 844, 905
Klima, Edward S, 556, 558
Klimov, Georgij A., 104
Kluender, Robert, 423–426, 432, 435, 438, 440, 441, 790
Kobele, Gregory M., 943, 945
Koch, Peter, 842
Koenig, Jean-Pierre, 50, 51, 66, 77, 81, 83–85, 94, 97, 101, 102, 105, 108, 109, 117, 134, 238–240, 245–252, 255, 257, 278, 279, 373, 396, 477, 524, 528, 615, 685, 712, 805, 813, 855, 856, 907, 952, 972, 993, 1085, 1100, 1101, 1107
Kohrt, Annika, 441
Kolarova, Zornitza, 896
Kolb, Hans-Peter, 1082
Kolliakou, Dimitra, 731
Kong, Anthony Pak-Hin, 918
Konieczny, Lars, 775
Konietzko, Andreas, 711

Koning, Ekkehard, 651
Kopp, Stefan, 900
Kordoni, Valia, 25, 754
Kornai, András, 991
Koster, Jan, 618, 980
Kothari, Anubha, 416
Kouylekov, Milen, 815
Kraak, Esther, 1035, 1041
Krahmer, Emiel, 903
Kranstedt, Alfred, 896, 903, 904, 917
Kratzer, Angelika, 675, 846, 1081
Krauss, Michael, 95
Krauss, Robert M., 888, 918
Krenn, Brigitte, 525–529, 538
Krer, Mohamed, 566
Krieger, Hans-Ulrich, 809, 810
Krifka, Manfred, 719, 847, 848
Kripke, Saul A., 847
Kroch, Anthony, 415, 426, 446
Kroeger, Paul R., 113
Kropp Dakubu, Mary Esther, 95
Kruijff-Korbayová, Ivana, 709
Kruyt, Johanna G, 810
Kubota, Yusuke, 173, 401, 421, 800, 848, 943, 1014, 1025, 1028, 1030–1032, 1036, 1037, 1041, 1043, 1045–1051, 1077
Kuhn, Jonas, 968
Kuhnle, Alexander, 26, 820
Kuiper, Koenraad, 539
Kuno, Susumu, 416, 421, 423, 424, 434, 435, 438, 440, 625
Kupffer, Manfred, 890
Kupść, Anna, 25, 553, 565, 575, 577
Kutas, Marta, 426, 437, 790
Kuthy, Kordula De, 813, 848, 906, 908, 968
Kynette, Donna, 437
König, Esther, 1083
Kühnlein, Peter, 904

Ladd, Robert D., 444
Ladusaw, William, 244, 678
Lahm, David, 76, 704
Laka, Itziar, 100
Lakoff, George, 49, 420, 421, 472, 487, 917, 1090
Lambek, Joachim, 18
Lambrecht, Knud, 396, 537, 941
Landman, Fred, 244, 392
Langacker, Ronald W., 1090
Laporte, Éric, 519
Lappin, Shalom, 423, 424, 472, 615, 848, 870
Larson, Richard K., 972, 976
Larsson, Staffan, 848
Lascarides, Alex, 71, 72, 553, 558, 870, 898, 904, 905
Lasnik, Howard, 9, 431, 432, 437, 552, 940
Lau, Ellen, 426
Laubsch, Joachim, 24
Laurens, Frédéric, 389
Lausberg, Hedda, 916
Le, Quoc Anh, 917
Leass, Herbert, 615
Lee, Jungmee, 421
Lee, Juwon, 95
Lee, Sun-Hee, 318, 340, 341
Lee-Goldman, Russell, 391, 392, 851
Lees, Robert B, 372
Lehmann, Christian, 751
Lenz, Alexandra, 752, 753
LeSourd, Philip S., 148
Letcher, Ned, 799
Levelt, 918
Levesque, Hector J., 850

Levin, Beth, 225, 237, 242, 247, 251, 252, 256
Levin, Nancy, 420, 472, 487
Levine, Robert, 173, 416, 1014, 1025, 1028, 1030–1032, 1036, 1043, 1045–1051
Levine, Robert D., 24, 167, 168, 173, 275, 418, 420, 422, 425, 426, 432–434, 439, 441, 442, 445, 472, 488, 498, 503, 856, 985, 1044, 1046
Levinson, Stephen C., 843, 903, 1057
Levy, Roger, 169, 170, 172, 426, 429, 1042
Lewis, David, 851, 898
Lewis, M. Paul, 95
Lewis, Mike, 1058
Lewis, Richard L., 436
Lewis, William D., 814
Li, Yen-hui Audrey, 946, 948
Liang, L., 26
Lichte, Timm, 1086
Lightfoot, David, 748, 749
Lin, Chien-Jer Charles, 789
Lin, Dekang, 942
Linadarki, Evita, 777
Linardaki, Evita, 450
Linell, Per, 843
Link, Godehard, 292, 368, 631
Lipenkova, Janna, 26, 93, 95, 807
Loehr, Daniel, 887, 891, 911
Lohmann, Arne, 487
Lohnstein, Horst, 980, 1082
Louden, Mark, 750
Lucas, Christopher, 395
Luo, Zhaohui, 1042
LuperFoy, Susann, 615
Lødrup, Helge, 539
Lønning, Jan Tore, 23, 804
Löbner, Sebastian, 651, 652
Lücking, Andy, 843, 849, 851, 859, 860, 888–891, 893, 898, 900, 903, 904, 907, 908, 910, 914–917, 919, 986

MacDonald, John, 781
MacDonald, Maryellen C., 776, 781, 784, 785, 788
Machicao y Priemer, Antonio, 26, 160, 364, 373
Macken, Elizabeth, 26
MacKinlay, Andrew, 817
Maekawa, Takafumi, 217
Mak, Willem M., 426
Malchukov, Andrej, 155
Maling, Joan, 157, 165, 166
Malouf, Robert, 25, 26, 73, 75, 95, 164, 211, 212, 801, 810
Manandhar, Suresh, 25, 711, 714, 731
Manning, Chris, 615, 626
Manning, Christopher, 232, 235
Manning, Christopher D., 26, 97
Manzini, Rita, 618
Marandin, Jean-Marie, 287
Marantz, Alec, 48, 50–52, 62, 1042
Maratsos, Michael, 788, 790
Marciniak, Malgorzata, 25
Marcus, Mitchell P., 811, 815
Marcus, R. Barcan, 847
Marimon, Montserrat, 25, 93, 810, 815
Markantonatou, Stella, 536
Marrafa, Palmira, 615, 624, 626
Marshall, Catherine R., 851
Marslen-Wilson, William, 779
Marslen-Wilson, William D., 987
Martell, Craig, 916
Martin, Scott, 1018

Martin-Löf, Per, 859, 1042
Masataka, Nobuo, 903
Mateu, Jaume, 535
Matiasek, Johannes, 106, 156, 159–161
Matsuki, Kazunaga, 781
Matsuyama, Tetsuya, 978, 1110–1112
Matsuzaki, Takuya, 803, 810, 811
Matthews, Danielle, 903
Mauner, Gail, 245
May, Robert, 424, 433, 678, 952
McCawley, James D., 53, 423, 501, 528, 962, 1048
McClave, Evelyn, 891
McCloskey, James, 143, 423, 984, 1044
McGinn, Colin, 903
McGurk, Harry, 781
Mchombo, Sam, 50, 52, 139
Mchombo, Sam A., 566
McMurray, Bob, 780
McNeill, David, 843, 887, 888, 894–897, 901, 903, 913, 916, 918
Megerdoomian, Karine, 341
Mehler, Alexander, 903, 907, 914
Meier, Jane, 3
Meinunger, André, 948
Meisel, Jürgen, 1084
Melnik, Nurit, 25, 377, 388, 807, 809
Mendoza-Denton, Norma, 891
Merchant, Jason, 449, 1050
Merenciano, Josep-Maria, 1050
Meurers, Detmar, 25, 76, 317, 330
Meurers, Walt Detmar, 25, 26, 104, 163, 164, 283, 299, 317, 330, 331, 710, 719–721, 727–729, 732–735, 807, 856, 945, 972, 983, 1082, 1105, 1107
Michaelis, Laura A., 941
Michel, Daniel, 426
Michelson, Karin, 94, 101, 102, 117, 257, 993
Mihaliček, Vedrana, 1034, 1035
Miller, Philip, 57, 80, 81, 231, 627, 1050, 1091
Miller, Philip H., 97, 162, 320, 562, 669
Milward, David, 503
Mineshima, Koji, 1018, 1040, 1058
Minsky, Marvin, 17
Mitchell, Erika, 567
Mithun, Marianne, 104
Mitkov, Ruslan, 615
Miyamoto, Edson T, 789
Miyao, Yusuke, 25, 777, 811, 817, 1091
Moeljadi, David, 25
Momma, Shota, 776
Monachesi, Paola, 79, 103, 162, 316, 317, 324, 572
Montague, Richard, 9, 686, 859
Moore, John, 250
Moortgat, Michael, 19, 1014, 1025, 1033, 1034, 1037, 1041, 1042, 1044
Moosally, Michelle J, 492
Moot, Richard, 1025, 1058
Morante, Roser, 818
Morgado da Costa, Luis, 817
Morril, Glyn, 631
Morrill, Glyn, 19, 1014, 1025, 1028, 1032, 1033, 1042, 1044, 1048–1050, 1057
Morrill, Glynn, 497
Moshier, M. Drew, 20
Mouret, Francois, 499, 505
Mouret, François, 476, 478, 480, 499, 501, 505
Muansuwan, Nuttanart, 95
Mulder, Jean Gail, 114

Mullen, Tony, 187
Munn, Alan, 421
Muskens, Reinhard, 1030, 1034–1036
Musso, Mariacristina, 991
Muszyńska, Ewa, 810
Mykowiecka, Agnieszka, 25, 373
Müller, Cornelia, 888, 895, 897
Müller, Gereon, 978, 991, 1113
Müller, Stefan, 18, 22–29, 51, 77, 78, 85, 93–95, 104, 110, 111, 114, 117, 150, 156, 163, 164, 180, 204, 226, 254, 255, 258, 277–281, 283–287, 292, 294, 296–299, 317, 318, 330–333, 341, 344, 356, 361, 362, 364, 367, 373, 388, 391, 393, 394, 399, 401, 427, 428, 496, 499, 522, 523, 528, 538, 554, 735, 755, 759, 798–800, 807, 808, 811–813, 821, 848, 939, 940, 942–945, 948, 949, 952, 956, 957, 968, 971, 972, 975, 977, 978, 981, 983, 986, 987, 992, 1039, 1077, 1082–1084, 1086, 1087, 1090, 1091, 1096, 1100–1106, 1108, 1115

Na, Younghee, 431
Nakamura, Michiko, 789
Nakazawa, Tsuneko, 255, 277, 287, 317, 330, 331, 361, 376, 401, 759, 760, 943, 983
Neeleman, Ad, 440
Nerbonne, John, 24, 187, 393, 858, 1105
Netter, Klaus, 187, 201–203, 283
Neu, Julia, 25
Neumann, Günter, 450
Newmeyer, Frederick J, 962
Newmeyer, Frederick J., 9, 522, 954, 955, 988, 989, 1044
Newmyer, Fredrick, 553
Ni, Weijia, 426
Nichols, Johanna, 104
Nishigauchi, Taisuke, 423
Nivre, Joakim, 811, 815
Nobe, Shuichi, 891
Noh, Bokyung, 1104
Noziet, Georges, 437
Nunberg, Geoffrey, 344, 518, 520, 522, 523, 525, 530, 531, 534, 538, 843, 850, 905
Nunes, Jairo, 948, 949
Nykiel, Joann, 450
Nykiel, Joanna, 287, 450, 494, 499, 505
Nöth, Winfried, 888
Núñez, Rafael, 895

O'Connor, Mary Catherine, 1078, 1079, 1081, 1086
Oehrle, R. T., 1015
Oehrle, Richard, 19
Oehrle, Richard T., 1014, 1028, 1029, 1033–1035, 1037, 1041, 1048
Oepen, Stephan, 25, 26, 28, 801, 804, 808–810, 815, 816, 818, 819
Oesterreicher, Wulf, 842
Oliva, Karel, 283
Oostdijk, Nelleke, 811
Oppenheimer, Daniel M., 903
Orgun, Cemil Orhan, 670
Osenova, Petya, 25
Oshima, David Y., 415
Ott, Dennis, 957

Packard, Woodley, 26, 28, 809, 818

Paggio, Patrizia, 710, 712, 714, 721, 722, 725
Park, Byung-Soo, 401
Park, Myung-Kwan, 431, 432
Park, Sang-Hee, 477, 505, 506, 1039, 1047–1049
Partee, Barabara, 425
Partee, Barbara, 678, 1020
Partee, Barbara H., 972
Partee, Barbara Hall, 847, 1046
Patejuk, Agnieszka, 173
Payne, John R, 551, 567
Peacocke, Christopher, 908
Penn, Gerald, 25, 26, 807, 855–857
Perles, Micha A., 1083
Perlmutter, David, 322
Perlmutter, David M., 1047
Perlmutter, Donald M., 448
Perry, John, 847, 859, 863
Pesetsky, David, 106, 425, 443, 472
Peters, Stanley, 10, 16, 628
Petrick, Stanley Roy, 812
Petten, Cyma Van, 437
Petzell, Erik M., 750
Pfeiffer, Thies, 888, 917
Phillips, Colin, 427, 437, 439, 501, 776, 787, 987, 988
Pickering, Martin, 782–786, 1057
Pietroski, Paul, 942, 1085
Piirainen, Elisabeth, 537
Pinker, Steven, 240, 247, 251, 1080
Plank, Frans, 104
Poesio, Massimo, 641, 843, 844, 869, 870, 893, 898, 915
Pogodalla, Sylvain, 1035
Poletto, Cecilia, 749
Polinsky, Maria, 110, 148, 431
Pollard, Carl, 60, 62, 63, 79, 81, 85, 170, 172, 418, 420, 423, 424, 432, 440, 474, 475, 483, 489, 494, 615, 616, 619, 621, 624, 625, 630, 632, 633, 647, 676–679, 681, 682, 695, 763, 766, 768, 778, 844, 845, 848, 855, 857, 858, 862, 863, 937, 939, 1030, 1034, 1035, 1042, 1114
Pollard, Carl J., 8, 11, 15–22, 24, 113, 134, 136–138, 144, 146, 156, 157, 159–162, 187, 188, 190, 192, 193, 197, 205, 216, 227, 228, 242, 274, 277, 279, 287, 292, 293, 333, 356, 358, 360, 361, 363–367, 369, 372, 374, 377, 378, 385, 386, 389, 393–395, 397, 523–526, 669, 676, 679–681, 683, 712, 714, 717, 759, 783, 790, 800, 855, 937, 952, 955, 967, 972, 973, 984, 985, 988, 992, 1037, 1040, 1052, 1054, 1057, 1077, 1092, 1096, 1098, 1100, 1113–1115
Pollock, Jean-Yves, 282, 552, 555, 562, 964
Posner, Roland, 888, 914
Postal, Paul M., 421, 423, 428, 432, 504, 938, 940, 951, 960, 962, 1047
Potsdam, Eric, 148
Potts, Christopher, 392
Poulson, Laurie, 26
Poyatos, Fernando, 888
Prince, Ellen, 420, 472, 487
Pritchett, Bradley L., 417
Proudian, Derek, 15, 24
Przepiórkowski, Adam, 25, 59, 97, 156, 162–166, 169, 173, 299, 553,

565, 575–577, 682
Przepiórkowski, Adam, 105, 109, 278, 279
Pullum, Geoffrey K., 8–10, 15, 167, 269, 277, 395, 471, 799, 940, 951, 991, 1083
Pulman, Stephen G., 527
Purver, Matthew, 854, 855, 858, 859, 862
Puthawala, Daniel, 1049

Quine, Willard Van Orman, 905
Quirk, Randolph, 210

Radford, Andrew, 948, 976
Raffelsiefen, Renate, 1042
Ramchand, Gillian, 972
Ranta, Aarne, 812, 814, 859
Rappaport Hovav, Malka, 225, 237, 251, 252, 256
Rayner, K., 435
Reali, Florencia, 788
Reape, Michael, 195, 729
Reape, Mike, 12, 55, 150, 173, 269, 274, 281, 319, 332, 391, 798, 943, 944, 971, 1040, 1091
Reed, Alonzo, 1073
Reid, Nicholas, 140
Reinhart, Tanya, 628, 642
Reintges, Chris H., 377
Reiplinger, Melanie, 817
Reis, Marga, 228, 268, 297, 752
Rentier, Gerrit, 317, 330
Restall, Greg, 1033
Retoré, Christian, 28, 1025
Reuland, Eric, 618, 628
Reyle, Uwe, 629, 632, 635, 656, 677, 847, 848
Rich, Elaine, 615
Richards, Marc, 959
Richards, Norvin, 535
Richter, Frank, 81, 82, 298, 369, 393, 477, 523, 524, 528, 531–533, 615, 676, 691, 703–705, 712, 805, 855, 856, 858, 907, 940, 947, 952, 1040, 1077, 1087, 1088, 1100
Riehemann, Susanne, 528, 1107
Riehemann, Susanne Z., 77, 78, 528–531, 1107–1109
Riemsdijk, Henk van, 940
Rieser, Hannes, 869, 870, 893, 898, 904, 907, 915, 919
Ristad, Eric Sven, 16
Ritchart, Amanda, 449
Ritchie, R. W., 10
Rizzi, Luigi, 285, 322, 325, 968–970, 992, 1084, 1115
Roach, Kelly, 11
Roberts, Ian, 749, 754, 975, 989
Roberts, Taylor, 54
Rochemont, Michael S., 427
Roeper, T., 245
Rohde, Douglas LT, 815
Roland, Douglas, 426, 788
Rooth, Mats, 727, 1020
Rose, Miranda L., 918, 919
Ross, John R., 417, 420, 423, 427, 430–432, 434, 440, 446, 447, 786, 787, 966
Ross, John Robert, 358, 394, 472, 475, 514, 943, 1009, 1047
Ross, Malcolm, 105
Rounds, William C., 20
Roussou, Anna, 754
Rouveret, Alain, 975
Runner, Jeffrey T., 95

Ruwet, Nicolas, 519, 522, 523, 538
Ruys, Eddy, 424
Ryu, Byong-Rae, 175, 318
Röpke, Insa, 918

Saah, Kofi, 423
Sabbagh, James, 424, 501
Sacks, Harvey, 843, 868
Sadler, Louisa, 207, 493, 975
Sag, Ivan, 57, 60, 62, 63, 78–81, 85, 187, 188, 190–193, 204–209, 212, 213, 216, 231, 255, 553, 576, 577, 615, 616, 619, 621, 624–627, 630, 632, 633, 647, 675–679, 695
Sag, Ivan A, 235, 566
Sag, Ivan A., 542
Sag, Ivan A., 8, 11, 15, 20–24, 29, 52, 59, 94, 97, 98, 113, 134, 136–138, 144, 146, 149, 156, 157, 159, 162, 170, 172, 173, 175, 187, 188, 190, 192, 193, 197, 205, 216, 227, 228, 242, 269, 273, 274, 283, 284, 290, 291, 299, 320, 328, 333, 334, 344, 356, 358, 360–372, 374, 375, 377–379, 383, 385–391, 394, 395, 397, 418, 420, 423–427, 432, 433, 438–440, 442–447, 449–451, 474, 475, 477, 482–484, 486, 489, 490, 494, 497, 499, 500, 502, 519, 523–526, 528, 532, 553–556, 558, 632, 669, 676, 679–681, 683, 686, 687, 689, 690, 712, 714, 717, 762, 763, 766, 768, 775, 783, 787, 789, 790, 800, 801, 844, 845, 855, 857, 858, 862, 863, 870, 873, 937, 939–941, 943, 951, 952, 955, 965, 966, 968, 972, 973, 975, 977, 984–988, 1037–1039, 1044, 1045, 1048–1052, 1054, 1057, 1077, 1078, 1081, 1083, 1086, 1087, 1089–1092, 1095, 1096, 1098, 1100, 1101, 1113–1115
Sailer, Manfred, 369, 393, 523, 531, 532, 534, 536, 676, 691, 703, 705, 714, 715, 858, 952, 1040, 1088
Saito, Mamoru, 431, 437
Saleem, Safiyyah, 26, 96
Salem, Maha, 917
Salkoff, Morris, 941
Salomon, Ralf, 917
Salvi, Giampaolo, 325
Samvelian, Pollet, 318, 320, 341, 343, 344, 495, 538
Sandfeld, Kristian, 396
Sandøy, Mads H., 26
Santorini, Beatrice, 432
Sapp, Christopher D., 758
Sauerland, Uli, 432, 952, 983
Sauerland, Uli., 372
Schabes, Yves, 800, 1086, 1087, 1107
Schachter, Paul, 105, 372
Schegloff, Emanuel A., 868
Schein, Barry, 244
Schenk, André, 520, 538
Schiel, Florian, 916
Schlenker, Philippe, 898, 907, 913, 914
Schmerling, Susan, 420
Schmidt, Thomas, 915
Schmolze, James, 63
Scholz, Barbara C., 940, 951
Schuster, Sebastian, 818
Schwarze, Christioph, 757

Schwarzschild, Roger, 727
Schäfer, Ulrich, 26, 808, 810, 817
Schütze, Carson T., 945
Scott, Dana S., 20
Searle, John R., 687
Sedivy, Julie C., 869
Sedivy, Kathleen M. Eberhard & Julie C., 779
Seeley, Daniel, 617
Seely, T. Daniel, 947, 948
Seifart, Frank, 140
Seki, Hiroyuki, 11
Sekine, Kazuki, 918
Selkirk, Elizabeth, 711, 720, 733
Sells, Peter, 207, 218, 317, 330, 367, 561
Seuren, Peter, 629
Shabes, Yves, 11
Shamir, Eliahu, 1083
Shan, Chung-chieh, 1032, 1033, 1039
Sheinfux, Livnat Herzig, 538
Shieber, Stuart M., 10
Shiraishi, Aoi, 945
Shiraïshi, Aoi, 502
Siegel, Melanie, 23, 25, 93, 803, 808
Siegel, Muffy A., 1048
Siewierska, Anna, 96, 97, 105
Simons, Gary F., 95
Simpkins, Neil, 25
Sirai, Hidetoshi, 382, 383, 387
Skwarski, Filip, 670
Slama-Cazacu, Tatiana, 888
Slayden, Glenn C., 809
Sloetjes, Han, 915, 916
Smessaert, Hans, 651
Snider, Neal, 428, 429, 787
Soehn, Jan-Philipp, 533
Solberg, Lars Jørgen, 816
Solias, Teresa, 1048
Sonesson, Göran, 908
Song, Sanghoun, 710, 713, 714, 716, 717, 723, 724, 726, 739
Sorace, Antonella, 951
Sowa, Timo, 908
Spector, Benjamin, 416
Spencer, Andrew, 155, 218
Spevack, Samuel C., 780, 781
Spivey-Knowlton, Michael J., 779, 780
Sproat, Richard, 975
Sprouse, Jon, 426, 431, 438, 787
Spurk, Christian, 26
Srivastav, Veneeta, 358
Stabler, Edward, 27, 28, 798, 941–945, 1083
Stalnaker, Robert, 628, 851
Stanojevic, Milos, 942
Starke, Michael, 991
Staub, Adrian, 426, 429
Steedman, Mark, 501, 502, 781, 800, 942, 1014, 1021, 1023, 1024, 1037, 1039, 1040, 1045, 1058
Steedman, Mark J., 11, 13, 274, 709, 781, 943, 977, 1014, 1019–1021, 1024, 1043, 1045, 1048, 1053, 1055, 1057
Steels, Luc, 903, 1090
Steffen, Jörg, 26
Steinbach, Markus, 897
Stepanov, Arthur, 431
Sternefeld, Wolfgang, 947, 948
Stiebels, Barbara, 1108
Stokhof, Martin, 687
Stone, Matthew, 904, 905
Stowell, Timothy Angus, 427
Streeck, Jürgen, 895
Stroop, John Ridley, 781

Strunk, Jan, 428, 429
Stucky, Susan U., 429
Stump, Gregory, 83
Stump, Gregory T., 1042
Suppes, Patrick, 26, 817
Suñer, Margarita, 325
Svartvik, Jan, 851
Swingle, Kari, 503
Swinney, David A., 535
Sygal, Yael, 798
Szabolcsi, Anna, 415, 631, 1052

Tabor, Whitney, 426, 430, 436
Taghvaipour, Mehran A, 374, 401, 985
Takami, Ken-Ichi, 434
Takami, Ken-ichi, 438, 440
Tam, Wai Lok, 174, 1039, 1041, 1045–1047
Tanenhaus, Michael K., 779, 780, 987
Tang, C.-C. Jane, 627
Taraldsen, Knut T., 423
Tegey, Habibullah, 54
ter Meulen, Alice, 651
Thiersch, Craig L., 268, 1082
Thomas, Alun R., 975
Thuilier, Juliette, 175
Tian, Ye, 869
Toivonen, Ida, 1086, 1105
Tomasello, Michael, 48, 843, 1101
Tooley, Kristen M., 781
Torr, John, 798, 942–945
Torrego, Esther, 106
Torreira, Francisco, 843
Toutanova, Kristina, 26, 803, 809, 810, 821
transitive, 600
Traum, David, 848, 870
Trawiński, Beata, 518
Traxler, Matthew J., 429, 781, 788
Triesch, Jochen, 917
Troseth, Erika, 210
Trueswell, John C., 780, 781
Truswell, Robert, 441
Tseng, Jesse, 670
Tsiamtsiouris, Jim, 437
Tsujii, Jun'ichi, 777, 811, 1091
Tsujii, Jun-ichi, 25
Tsuruoka, Yoshimasa, 811
Tuite, Kevin, 891
Tyler, Lorraine, 779

Uhmann, Susanne, 721
Uszkoreit, Hans, 16, 22, 25, 270, 271, 275, 277, 333, 429, 971, 1083, 1102

Vaillette, Nathan, 374, 985
Valentín, Oriol, 1049
Vallduví, Enric, 710, 712, 713, 715, 717, 718, 725, 730, 731, 737, 848, 906, 908, 968
Vallduví, Enric, 712
Van Benthem, Johan, 651, 652, 1017
Van der Beek, Leonoor, 815
van der Sluis, Ielka, 903
Van Eynde, Frank, 29, 187, 195, 198, 200, 207, 208, 213, 215, 218, 359, 427, 428, 847, 940, 945
Van Koppen, Marjo, 976
Van Noord, Gertjan, 25, 26, 274, 275, 281, 283, 317, 330, 754, 759, 810, 811, 815, 971, 1082
van Noord, Gertjan, 427
Van Riemsdijk, Henk, 167, 956
Van Riemskijk, Henk, 402
Van Schijndel, Marten, 426
Van Trijp, Remi, 48, 1091, 1106
Van Valin, Robert, 237, 238, 251

Van Valin, Robert D. Jr., 438, 440
Varlokosta, Spyridoula, 618
Vasishth, Shravan, 567, 789
Veenstra, Mettina Jolanda Arnoldina, 1082
Velldal, Erik, 810, 817
Verbeke, Saartje, 105
Vergnaud, J.-R., 372
Vermaat, Willemijn, 1033
Versante, Laura, 895
Verspoor, Cornelia Maria, 1104
Vicente, Luis, 421
Vieira, Renata, 641
Vijay-shanker, K., 11
Vikner, Sten, 552
Villata, Sandra, 426
Villavicencio, Aline, 490–492, 530
Vincent, Nigel, 750
Vivès, Robert, 423
Von der Malsburg, Christoph, 917
Von Stechow, Arnim, 719, 721

Wada, Hajime, 615
Wagner, Petra, 887
Wahlster, Wolfgang, 22, 808, 916
Wald, Benji, 138
Waldron, Benjamin, 810
Wali, Kashi, 358
Walkden, George, 750, 768
Walker, Heike, 369, 393, 395
Walther, Markus, 1098
Wang, Rui, 26
Wanner, Eric, 788, 790
Warner, Anthony, 553, 558, 559
Wasow, Thomas, 15, 17, 24, 431, 437, 451, 520, 522, 523, 525, 775, 781, 951, 967, 986, 987, 1083, 1090, 1091, 1095, 1101
Wasow, Tom, 451, 778, 847, 869, 937, 940, 951, 1114
Wax, David, 814
Webelhuth, Gert, 65, 69, 70, 326, 333, 373, 569, 710, 719, 733, 736
Wechsler, Stephen, 61, 97, 100, 105, 109, 133, 134, 137, 138, 141–147, 149, 190, 192, 233–235, 237, 238, 241, 251, 278, 279, 318, 341, 373, 489, 492, 615, 813, 846, 855, 1101
Wechsler, Stephen Mark, 18, 226, 254, 258, 1039, 1090, 1102, 1104, 1106
Weinberg, Amy S., 422
Weinreich, Uriel, 519, 549
Weir, David J., 11
Weissmann, John, 917
Wells, Rulon S., 287
Wesche, Birgit, 283
Westergaard, Marit, 1084
Wetta, Andrew C., 286, 287, 297
Wexler, Kenneth, 501, 618
Whitman, Neal, 1041
Wichmann, Søren, 104
Wilder, Chris, 1051
Wilkins, David, 895
Williams, Edwin, 47, 504, 946
Willis, David, 984
Wiltshire, Anne, 888
Winkler, Susanne, 299, 711
Winter, Yoad, 419, 909
Wintner, Shuly, 25, 798
Witten, Edward, 428
Wittenburg, Peter, 915
Worth, Christopher, 1036
Wright, Abby, 401
Wu, Ying Choon, 892

Wurmbrand, Susanne, 946

Xia, Fei, 814
Xue, Ping, 615, 627

Yang, Chunlei, 93
Yang, Jaehyung, 25, 383
Yankama, Beracah, 942, 1085
Yatabe, Shûichi, 173, 174, 474, 497–499, 502, 1039, 1041, 1045–1047
Yatsushiro, Kazuko, 983
Yip, Moira, 165
Yngve, Victor, 474
Yoo, Eun Jung, 162, 402, 681, 682
Yoo, Eun-Jung, 317, 330, 337, 338, 401, 402
Yoon, Jae-Hak, 398
Yoshida, Masaya, 1050
Ytrestøl, Gisle, 816
Ytrestøl, Gisle, 810
Yu, Jiye, 817

Zaenen, Annie, 157, 165, 503, 1044, 1083
Zalila, Ines, 377
Zamaraeva, Olga, 814, 818
Zanuttini, Raffaella, 552, 562, 572
Zeevat, Henk, 696
Zeijlstra, Hedde, 552, 554
Zhang, Niina Ning, 421
Zhang, Yi, 809, 810
Zimmermann, Thomas Ede, 857
Zlatic, Larisa, 489, 492
Zlatić, Larisa, 133, 137, 138, 141, 142, 145
Zribi-Hertz, Anne, 624, 625
Zwart, Jan-Wouter, 471
Zwarts, Frans, 415
Zwarts, Joost, 909
Zwicky, Arnold, 210
Zwicky, Arnold M., 167, 470, 471, 812

Özyürek, Aslı, 887, 888, 896, 919
Ørsnes, Bjarne, 26, 93, 279, 283, 807

Żywiczyński, Przemysław, 894

Language index

Algonquian, 174
Alpino Treebank, 26
American Sign Language, 901
Arabic, 110[24], 373, 375–377, 377[28], 388[41], 401
Archi, 105[18], 975[29]
Austronesian, 174

Balinese, 61
Bantu, 1084
Basque, 94, 94[3], 99, 100, 1084
Bulgarian, 25, 373[23]
Burushaski, 94

Catalan, 730
Chinese, 789
Classical Armenian, 164
Coast Tsimshian, 94, 114

Danish, 26, 283, 401, 712, 807
Dogon, 397[53]
Dutch, 11, 25, 283, 810, 815, 980[34]

English, 22–24, 162, 167, 168, 274, 356–358, 359[1], 361, 361[7], 362, 362[8], 367, 367[15], 373, 374, 376, 377, 378[29], 382–384, 386, 386[39], 387, 389, 390, 396, 398, 398[55], 399, 401, 613, 619, 625, 626, 632, 635, 777, 788, 790, 807, 808, 810, 811, 815–818, 821, 940, 975
African American Vernacular, 1083

Finnish, 165, 166
French, 162, 274, 283[7], 287, 288, 374, 379, 380[32], 381, 382[33], 386[38], 389[44], 395, 396, 396[51], 807

Georgian, 26
German, 22, 25, 26, 69, 156, 159–161, 163, 167, 168, 267, 268, 270, 271, 274, 275, 277, 278, 278[5], 279–281, 283, 285, 287, 288, 290–292, 294–299, 361[7], 362[8], 364[11], 373[23], 376[25], 389, 390, 395, 399, 399[56], 400, 401, 401[59], 623, 732, 779, 807, 808, 811, 891, 903, 948, 973, 979, 981, 983, 1083, 1102, 1113
Germanic languages, 812
Gothic, 164
Greek, 25, 618, 731
Gujarati, 105
Gungbe, 1084

Hausa, 25, 373[23]
Hebrew, 25
Hindi, 358[1], 394[48], 807
Hindi-Urdu, 105
Hungarian, 175, 949

Icelandic, 157, 160, 163, 165[5], 618
Indonesian, 25, 110[24]

Italian, 162

Japanese, 22, 25, 382, 383, 383[34], 387[40], 396, 397[52], 397[54], 789, 808, 815[30], 973, 983
Javanese, 110[24]

Kimaragang, 94, 113, 115
Korean, 25, 162, 175, 382, 383, 383[34], 384–386, 396, 397, 397[53]

Lakhota, 397[53]
Latin, 4
Lezgian, 94, 106
Lummi, 776

Macushi, 94, 96, 98
Malagasy, 115
Maltese, 26, 94, 94[3], 807
Mandarin, 807
Mandarin Chinese, 25, 26, 948
Marathi, 105, 358[1], 394[48]

Ngan'gityemerri, 140
Nias, 174
Norwegian, 23, 25, 627

Oneida, 94, 94[3], 993

Pashto, 54
Persian, 26, 401, 807
Polish, 25, 162, 166, 169, 170, 373[23]
Portuguese, 25, 616, 623, 624, 644, 649

Quechua, 397[53]

Romanian, 389[44]
Russian, 169

Serbo-Croatian, 133, 134
Spanish, 25, 26, 737, 738, 807, 815[30]
Sundanese, 110[24]
Swiss German, 11
Syrian Arabic, 112[26]

Tagalog, 110, 174
Thai, 25
Toba Batak, 626
Tongan, 94, 94[3], 111
Turkish, 373[23]
Turkish , 53

Udi, 54

Wambaya, 25, 299
Warlpiri, 94, 94[3], 269, 290–292, 299–300
Welsh, 94, 111, 112[26], 400[58], 401, 975

Yiddish, 26, 807
Yucatec Maya, 94, 94[3]

Subject index

◯, 288
⊕, 272[1], 1093
bigcirc, *see* shuffle
*SEM 2012 Shared Task, 818
+, 853, 1112
/H, 806[16]
/QEQ, 806[16]
DELPH-IN Consortium, 801, 801[9], 805, 808–810, 813
DELPH-IN MRS Dependencies (DM), 805, 806, 818, 819
[incr tsdb()], 808, 809
SLASH, 789, 790
Focus Inheritance Principle, 711
pro-drop, 134
ARG-ST, 59, 80

A-accent, 848
abstract gesture, 897
accents, 737
accept, 869
acceptability, 783
acceptance, 852, 853
accessibility hierarchy, 788[11]
ACE, 809
acknowledgment, 852
ACQUILEX, 22
acquisition, 844, 988–990, 1084
Across the Board Extraction, 943
adaptor gesture, 894
addressee, 847
adjacency pair, 868
adjectives, 821
adverb, 561
adverbs, 817
affiliate, 898, 900
affiliation, 890–892
affix, 964
affixation, 53, 77
AGGREGATION, 813–814
Agree, 809, 953–954, 990
agreement, 21, 287
  object, 1084
alignment, 104
Alpino, 804[14], 810–811, 815
alternations, 625
ambiguity, 780, 785, 802, 804, 809, 815, 816, 821
  lexical, 821
  modifier attachment, 802, 804
  part-of-speech, 802
  spurious, 816
  syntactic, 821
American structuralism, 820
AMI Meeting Corpus, 916
analysis order
  effect of, 822
anapher, 972
anaphora resolution, 640
anaphoric binding, 615
annotation, 804, 815, 816
  automatic, 816
anthology search, 26

Anvil, 915
aphasia, 918
Arc Pair Grammar, 960[20]
argument indexing, 96
argument label, 806
argument realization, *see* linking, *see* linking
argument structure, 812, 1102
art, 809
artificial datasets, 820
attitude reports, 859
attributive use, 862
augmented/adaptive communication, 23
Austinian proposition, 863
auxiliary verbs, 811

B-accent, 848
Babel, 811
background condition, 850
background gesture, 894
background information, 845–847
basic integration scheme, 899
basic type, 860
baton, 895
beat gesture, 895
behavioral properties, 105[17]
belief objects, 859
bi-directionality, 799
bimodal integration, 900
binary-branching, 811
binding, *see* anaphora, *see* anaphora, 615
Binding Domains Principle, 636
binding exemption, 622
binding local domain, 617
binding principles, 615
Binding Theory, 136, 277[4], 972
binding theory, 21
biomedical NLP, 817, 818
BioNLP 2009 Shared Task, 817
bound anaphora, 641, 645
bound variable, 806
branching, 964
    binary, 959
bridging, 843
bridging anaphora, 641
British lexicology, 820
brnaching
    unary, 972[27]
broad-coverage grammar, 810, 821
BV, 806

c-command, 620, 953
canonical SYNSEM, 80
Cartography, 968
case, 155–176
case frames, 814
Categorial Grammar, 62, 170, 269, 943, 1101
Categorial Grammar (CG), 282, 1081, 1105, 1110, 1113
categorisation game, 903
category
    functional
        T, 963
Celex, 810
CFG, 815
character viewpoint, 913
Chomsky hierarchy, 992
chronemics, 888
chunking, 815
circumstantial features, 845–847, 862
CKY algorithm, 11
clarification potential, 854–856, 863
clarification potentional, 858
clarification request, 854–856, 859
classifications, 860

clause type, 968
CLIMB, 809, 822
clitic left dislocation, 731
cliticization, 11
clitics
  endoclitics, 53–56
  in Pashto, 54
  in Romance languages, 54, 57, 79, 81
  in Udi, 54
co-verbal gesture, 888–894
coding properties, 105[17]
Cognitive Grammar, 28, 1090
coindexation, 617
Combinatory Categorial Grammar (CCG), 803, 812[25]
command, 619
comment, 848
commitment-stores, 853
common ground, 850–851, 853
common knowledge, 853
communicative interactions, 862
community membership, 850, 852
comparative correlative construction, 941
competence, 775, 843
competence hypothesis, 775, 777, 778, 782
complements, 801
Complenent Extraction Lexical Rule, 79
complex antecedents, 807
complexity
  average-case, 799, 801
  computational, 799–801
  of grammars, 798
  time, 799
  worst-case, 799, 800, 801[9], 812
compositionality, 820
comprehension, 778, 780–782, 784, 788[11], 951
computational complexity, 631
computational linguistics, 940
computational tractability, 631, 778
concatenation, 11
conceptual literacy, 842
conceptual orality, 842
conceptual vector meaning, 909
conduit metaphor, 897
conjunct constraint, 787
constituent
  discontinuous, 1102
constituent boundaries, 815
Constituent Order Principle, 11
constituent ordering, 21
constructicon, 1086
Construction, 1080
construction, 786
Construction Grammar, 28, 62, 85, 798, 985, 1077–1115
  Berkeley, 1089–1091
  Cognitive, 1090
  Embodied, 1090
  Fluid, 1090, 1106
  Radical, 1090
  Sign-Based, 1090–1091
Construction Grammar (CxG)
  Sign-Based, 284, 939, 1090, 1091
Construction Morphology, 1107
Construction-based HPSG, 21
construction-like object, 1090
constructions, 21
context, 781, 784, 851
context dependence, 843
contextual anchors, 847
contingent gesture, 893

control, 21, 801, 811, 821
conventional interpretations, 853
conversational analysis, 868
conversational rule, 864–868
coordinate constructions, 9
Coordinate Structure Constraint, 787
coordination, 24, 166, 169–174, 287, 811
Core, 1081
core, 798, 821
core grammar, 807, 808, 813, 941
coreference, 641
CoreGram, 51, 807–808, 812–814, 940
corpus search
  structure-based, 815
cosupposition, 914
coverage, 821
cross-index, 99
cross-linguistic comparison, 807
cross-serial dependencies, 810
cycle
  in feature description, 289
cyclic feature structures, 807

D-structure, 1082
database query, 26
deep processing, 808
default, 939
deferred ostension, 905
deferred reference, 843, 905
deictic gesture, 902–907
deictic gestures, 895
deictic word, 903
delivery gesture, 914
demonstration act, 895
demonstratum, 843
dependencies
  semantic, 804[14], 804–806, 811
  syntactic, 804[14], 811
dependency, 117, 1102, 1104
Dependency Grammar, 1101
Dependency Grammar (DG), 1110, 1113
dependency graph, 804, 815, 818
Dependency Minimal Recursion Semantics (DMRS), 804, 806, 819
dependency parsing, 811
  semantic, 819
dependent type, 860
dependent-marking, 104[15]
derivation, 1107
derivational morphology, 77
Derivational Theory of Complexity, 775[1]
derived word, 77
dialog history, 862
dialogue agent, 853
dialogue gameboard, 853, 862–868
dialogue gameboard parameters, 862
dialogue progress, 853
diathesis alternation
  locative alternation, 68
  voice alternations
    Balinese, 61
direct reference, 847
directed acyclic graph, 811
directionality, 782
disambiguation, 26, 815
discontinuous constituency, 11
discontinuous constituents, 811
Discourse Representation Theory, 847, 848
discriminants, 804, 815, 816
disfluencies, 778, 781, 843
disjoint union, 731
Distributed Morphology, 50, 52
distributional hypothesis, 820

distributional semantics, 820
documentation
  of grammars, 799, 809
domain, 857
domain objects, 849
downdate, 853
downdating, 869
dual nominals, 652
dual point, 905
DUEL Corpus, 916
dynamic semantics, 848

e-commerce customer email response, 26
e-type anaphora, 641, 645, 652, 656
Education Program for Gifted Youth, 817
educational applications, 816–817
effects, conversational rule, 864
ELAN, 915
element constraint, 787
Elementary Dependency Structures (EDS), 804, 819
ellipsis, 287, 391
  word internal, *see* affixation, suspended
Eloquent Software, 14
emblem, 895
emblematic gesture, 895
Embodied Construction Grammar, 1091
emotions, 888
empty element, 939, 955, 960, 964, 972[27], 980, 991
empty elements, 807
empty operator, 965, 984
endangered languages, 813
English Resource Grammar, 25
English Resource Grammar (ERG), 801[9], 802, 805, 808, 817, 818, 821
Enju, 811, 817, 819
EQ, 806
equality modulo quantifiers, 805
ergative, 104, 108
evaluation
  automatic, 817
exemplification, 908
EXMARaLDA, 915
experiencer, 850
experiment, 945
experimental semiotics, 914
expletive pronoun, 21
extension condition, 954
extra-grammatical mark-up, 810
extraction, 79, 811
Eye-tracking, 779
eye-tracking, 780, 790

face to face interaction, 888
facial expressions, 888
facts, 852
feature
  accent, 737
  addressee, 361, 363
  aff, 908
  anchors, 394
  arg-st, 365, 367[23], 383
  arg1, 806
  arg2, 806
  background, 361[23], 370[23], 845
  boh, 901
  c-inds, 845
  carrier, 909
  case, 377, 379
  cat, 357, 360, 365, 376, 380, 381
  comps, 359[23], 365, 374[23], 376[23], 376, 380, 381, 387, 391
  content, 360, 361, 365, 368, 370–372, 375, 376[23], 380, 381, 388,

392[23], 846
contextual-indices, 845
context, 717, 845, 849
correlative, 394[23]
cvm, 909
deps, 801[23]
dgb-params, 862
dgb, 853
dir, 901
domain, 288, 391–393
experiencer, 850
extent, 901
extra, 292[23], 393, 394
facts, 852
first, 801[23]
focus, 848
frames, 1100
g-dtr, 906
gender, 377
ground, 848
handshape, 901
hcons, 713
head-daughter, 369, 375, 384, 386
head, 357, 360, 365, 368–371, 375, 376, 380, 384, 386–388
icons, 713, 805
independent-clause, 966
index, 359[23], 359–361, 363, 365, 368, 371, 376, 381, 388
instance, 363
inv, 367, 368, 966
latest-move, 852
lid, 1087
link, 848
local, 357, 360, 365[23], 365, 369, 370, 376, 379, 380, 393
marking, 377
mc, 357, 367, 368, 384
mode, 908
mod, 357, 360, 361, 365, 366, 368–371, 376[23], 376, 378[23], 380, 383, 384[23], 384, 386–390
ms, 853
non-head-daughters, 369, 375, 384
non-local, 359[23], 360, 376, 380
nucleus, 361, 363
number, 377
orient, 901
palm, 901
path, 901, 902, 910
phon, 737
position, 901
pos, 384[23]
prop, 363[23], 363, 365, 371
q-params, 862
qud, 852
que, 360[23], 362, 370[23]
rels, 1100
relzr, 391
rel, 358, 359[23], 359, 360[23], 360, 361[23], 361, 362, 365, 366, 368, 369[23], 369, 370[23], 370, 372, 375, 377, 378[23], 378, 382, 388, 390[23], 394[23], 396, 397[23], 403
restr, 363, 365, 371, 376, 388
rest, 801[23]
rstr, 806
s-dtr, 906, 910
select, 357
slash, 358, 359[23], 359, 360, 365, 366, 369, 370[23], 370, 372, 375, 376, 377[23], 377, 378, 380, 381, 383–386, 387[23], 387, 388, 391, 393, 403, 801[23]

soa, 363[23], 363, 365, 367[23]
speaker, 361[23], 361, 363
specified, 1098[23]
specifier, 359[23]
spec, 359[23]
standard, 850
subj, 284[23], 365, 374[23], 374, 375, 376[23], 388, 396
sync, 901
synsem, 357, 360, 365, 371, 376, 379
tail, 848
to-bind, 366, 377[23], 380, 381
traj(ectory), 909
vec, 909
vform, 375
wh, 370[23]
wrist.mov, 909
wrist, 901
feature cooccurrence restrictions, 16
feature descriptions, 20
feature geometry, 57, 66
feature structures, 20
feedback, 864
feedback signal, 864
FIGURE Corpus, 917
File Change Semantics, 848
filler, 15
filler-gap constructions, 21
finality, 46, 314, 352, 353, 674, 774, 886, 1072
fingerspelling alphabet, 901
finite closure, 16
first person pronoun, 846
first-order logic, 817
Fluid Construction Grammar, 1091
focus, 847–849, 1084
foreground gesture, 894
formalism, 809
differentiating from theory, 798
formalization, 28, 940
fractal architecture, 848
fractal design, 854
fractality, 848, 854
fragmentary utterance, 391
Frame Semantics, 1100–1101
frame semantics, 910
free relatives, 166
function type, 861
functional application, 1081

garden paths, 781
gaze, 888
gaze-pointing, 895
GB, 1081
gender, 133
generalised quantifier, 859
Generalized Phrase Structure Grammar, 9, 937
Generalized Phrase Structure Grammar (GPSG), 269, 1081, 1102, 1107
generation, 798, 799, 809, 811, 817, 820
Generative Grammar, 711
Generative Semantics, 48, 49, 962
genetics, 1084
gerunds, 73
gesture, 986[36]
gesture affiliation, 891, 892, 898
gesture classes, 894
gesture classification, 896
gesture meaning, 898
gesture representation, 900–902
gesture space, 901
gesture stroke, 897
gesture typology, 907
gesture-as-byproduct, 894

gesture-as-product, 894
gesture-speech relations, 892
gestures, 843
GF Resource Grammar Library, 814
GG, 808
given information, 848–849, 852
gold parse, 803
gold standard, 803
good continuation, 889
Government and Binding, 156
GPSG, 270, 296, 1083[4]
gramm-index, 101
grammar correction, 26
grammar engineering, 812[25]
multilingual, 807–810
grammar exploration, 798
Grammar Matrix, 25, 174, 808, 813–814
customization system, 813, 822
libraries, 808, 813
grammar tutoring, 26
grammar-dialogue interface, 862
Grammatical Framework (GF), 812[25]
grammaticalisation process, 914
grammaticality, 809
grammaticality judgments, 808
Grammix, 26
ground, 848
grounded reasoning, 820
grounding, 853, 862, 863, 869

handshape, 901
head, 959
functional, 955, 964
Head Feature Principle, 21, 66
head information, 811
head movement, 942
Head-All-Valents Schema, 112
Head-Driven Phrase Structure Grammar (HPSG), 747, 750, 754, 758, 769
Head-Feature Principle, 57
head-marking, 104[15]
heads, 801
hedges, 817
heterogeneous universe, 857
hierarchical lexicon, 63–71
Hinoki Treebank, 815[30]
homogeneous universe, 857
HP labs, 798
HP-NL, 24
HPSG-TTR, 870–873

i-within-i effect, 637
iconic gesture, 894, 907–914
idiom, 1081
illustrator gesture, 894
imagistic gesture, 888
immediate dominance, 269
implicational constraints, 66, 67
incomplete demonstratives, 890
incremental, 869
incremental processing, 869
incrementality, 778
index, 843
index finger pointing, 895
index palm down, 895
index palm vertical, 895
indexical expressions, 846
indirect reference, 843, 905
infinitival VP, 821
inflectional morphology, 82
information extraction, 817
biomedical, 818
information packaging, 847–849
information retrieval, 817
information state, 852, 853

information state model, 853
information structure, 813, 847–968
inheritance, 63–69
    default inheritance, *see* YADU, 71–76
    multiple inheritance, 69–71
insertion, 11
intensional entities, 859
interactive gesture, 893
interactive gestures, 896, 914–915
interface, 968
Interface model, 919
interjections, 863–864
interlinear glossed text (IGT), 814
internal structure of words, *see* morphology, lexical integrity
interrogative, 954
interrogatives, 22
intrinsic argument, 804, 805[15]
introspection, 945
irregular morphology, 72
island, 992
island constraints, 21, 786, 787

Jacy, 808
judgement, 860

Kahina, 808
kairemics, 888
KoS, 862–868

label, 806
Labeling, 990
labeling, 955
language acquisition, 1083–1085, 1101–1102
language documentation, 812–816
language system, 843–844
language teaching, 816
language understanding, 820
Lassy treebanks, 815
latest move, 852
laughter, 843
Lexical Access model, 918
lexical boost, 786
lexical classes, 814
lexical entries, 56–59, 798[3], 803, 815
    induction of, 811
Lexical Functional Grammar, 13, 94, 101, 165–166, 937
Lexical Functional Grammar (LFG), 750, 754, 769, 812[25], 821[34], 1081
lexical integrity, 50–52, 55
lexical item, 805
lexical representation, 784
Lexical Resource Semantics, 369[17], 393–394, 858
lexical rule, 972[27]
lexical rules, 21, 50, 76–86, 798[3], 814
lexical sign, 59
lexical types, 785
lexicalism, 49–56, 803
LFG, *see* aso Lexical Functional Grammar48, *see* Lexical Functional Grammar, *see* Lexical Functional Grammar
linear precedence, 269
Linear Precedence Rule, 269
linearisation, 391
linearization grammar, 11
linguistic competence, 844, 852, 864
linguistic evidence, 851
linguistic generalizations, 809
linguistic hypothesis testing, 807, 812–816
linguistic knowledge, 843–844, 864
linguistic universals

bottom-up exploration of, 807
link, 848
linking, 66, 80, 812
lip-pointing, 895
lists, 801
LKB, 809
local domain, 617
local domain reshuffling, 622
local management system, 868
locality, 29, 954
locutionary proposition, 863
logic instruction, 26
logical form, 817, 857
logical variable, 804
London-Lund corpus, 851
long-distance reflexives, 615, 652

machine learning, 819–820
machine translation, 22, 23, 26, 818
machine-learning, 803[12]
macros, 807
Mainstream Generative Grammar (MGG), 1081
maintability
of grammars, 798
mal-rules, 817
manifest field, 862
McGraw-Hill Education, 26
McGurk effect, 781
medium, 842
mental states, 853
Merge, 953, 1081
External, 952
Internal, 954
meta-grammar, 809, 813, 822
meta-grammar engineering, 822
Metagraph Grammar, 960[20]
metaphorical gesture, 897
metarules, 14, 16
metonymy, 144
metrical tree, 849
Micro-Cues, 1084
mildly context sensitive, 11
Minimal Recursion Semantics, 369[17], 393–394, 858, 907
Minimal Recursion Semantics (MRS), 804, 806, 1100
Minimalism, 277[4], 937–993, 1081, 1101, 1110, 1113
Minimalist, 132
Minimalist Grammars, 1083[4]
Minimalist Program, 27, 798
mixed syntax, 888
modals, 817
model theory, 856–859
model-theoretic semantics, 856–859
modularity, 779[5], 784, 968
morphological analyzers, 810
morphology, 807, 961, 1107
internal structure of words, 53
irregular morphology, 71
suppletion, 83
morphotactics, 814
Move, 954
movement
A, 954
A′, 954
movement trajectory, 901
multi-word expressions, 803
multichart, 899
multilogue, 869
multimodal chart parser, 898–899
multimodal communication, 888
multimodal ensemble, 914
multimodal grammar, 903
multimodal integration, 898
multimodal integration scheme, 899

multimodal therapy, 919
multimodal utterance, 892
multiple fronting, 735
multiple inheritance, 21
multiword expressions, 23
mutual beliefs, 850, 853

named entity recognizers, 810
naming game, 903
Natural Language Logic, 24
Natural Language Processing (NLP), 817–820
negation, 817, 818
NEQ, 806
new information, 848–849, 852
non-configurationality, 102
non-pronouns, 615, 652
non-sentential utterances, 843
non-verbal actions, 843
non-verbal behaviour, 887–888
non-verbal communication, 869
nouns, 821
NPN Construction, 978, 1110
NPN construction, 959
null subject, 623
numeration, 987[38], 990

o-binding, 616
o-bottom, 622
o=command, 619
obligatory gesture, 893
obliqueness, 619
obliqueness hierarchy, 60
obliqueness order, 21
obliquenss, 973
observer viewpoint, 913
online type construction, *see* type underspecified hierarchical lexicon, TUHL

ontology acquisition, 26
open hand pointing, 895
open-source, 808
opinion analysis, 818
order domain, 287–299
order domains, 55
organization of the lexicon, 21

pantomime, 897
paralinguistics, 887
parameter, 276
parasitic gaps, 21, 166
ParGram, 814
Parole, 810
parse forest, 803, 815
parse ranking, 802–804, 809, 810
parse selection, 26
parsing, 798, 799, 809, 811, 818
    neural, 815
    stochastic, 815
partial verb phrase fronting, 1105
passive, 17, 625
    impersonal, 980
    remote, 287
passives, 69
pending, 864
Penn Treebank, 811
performance, 775, 843, 951, 992
Performance-Grammar Correspondence Hypothesis, 776
peripheral, 798
Periphery, 1081
periphery, 821, 941
PET, 809
Phase, 990
phase, 1084
phase quantification, 652
phonetic speech-gesture constraint, 891

phonology, 964
phrase structure rules, 798[3]
pied-piping, *see* relative inheritance, 956
plit antecedent, 641
pointing, 895
pointing cone, 904
pointing gesture, 902–907
polar question, 868
polysemy, 820
polysynthetic, 102
position classes, 814
possible worlds, 859
post-stroke hold, 897
Poverty of the Stimulus, 1085
practical applications, 807–811
pre-stroke hold, 897
precision, 818, 821
preconditions, conversational rule, 864
predicate, 806
  abstract, 805
  lexical, 805
predicate-argument structure, 804, 805
predicate-argument structures, 819
preparation phase, 897
principle
  Head Feature, 1096–1097
  Sign, 1095[14]
Principle A, 615, 652
Principle B, 615, 652
Principle C, 615, 652
Principle Z, 615, 652
Principles & Parameters, 1084
private information, 853
PRO, 374, 375, 388
pro-drop, 99[10]
pro-index, 97
pro-speech gesture, 888
probe, 953
processing, 843
processing speed, 842
processing time, 810
production, 778, 781, 782, 784, 951
Project DeepThought, 808, 813
projection, 913
prolegomena, 4
pronoun, 845
  expletive, 980
pronouns, 615, 652
proper name, 847
proper names, 850
proposition, 863
prosodic constituency, 849
prosodic information, 849
prosody, 737, 780, 892
proxemics, 888
pruning, 803
PTT, 870
public information, 853
PyDelphin, 806, 809

QEQ, 805, 806
quantification, 651
quantificational nominals, 652
quantificational parameters, 862
quantified noun phrases, 859
quantifier, 651, 805, 805[15], 806
querying relational databases, 25
question answering, 26
question under discussion, 852, 853, 862, 895
question-answering pair, 852
QuickSet system, 899

R-Principle E, 646
R-Principle V, 646
raising, 21, 801, 821, 954

realization ranking, 810
recall, 818
record, 860
record type, 860
record types, 859
records, 859
recursion, 801
Redbird Advanced Learning, 26
reduplication, 11, 978
Redwoods, 815, 816, 821
reference, 849
reference marker, 629
reference processing, 640
reference resolution, 820
referential interpretation, 862
referential nominals, 652
referential NPs, 850
reflexives, 615, 652
regulator gesture, 894
REL inheritance, *see* relative inheritance
relation
  $\bigcirc$, 288
  append, 272[23]
  shuffle, 288
relational constraint, 272[1], 273, 275
relational constraints, 807
relative clause, 21, 355–403, 940, 965
  antecedent of, 356, 358–365, 368, 370–372, 373[22], 377, 377[28], 380, 389–391, 393, 395[50], 398
  appositive, *see* relative clause, non-restrictive
  as complement
    in cleft construction, 395–396
    in dependent noun construction, 396–397
    in pseudo-relative construction, 397–398
  of *diejenigen*, 395
  of superlative adjective, 395
bare, 382–388
  in English, 384–388
  in Japanese, 382–384
  in Korean, 382–384
construction, 367–375, 376, 377–379, 382–388, 391
dont-relative, 379–382
empty relativiser, 364–367, 390
extraposition of, 392–394
free, 398–402, 402[60]
fused relative, *see* relative clause, free
headed by complementizer, 375–382
  in Arabic, 376–377
  in English, 377–379
  in French, 379–382
headed by verb, 382–384
headless relative, *see* relative clause, free
hydra, 368[17]
infinitival, *see* relative clause, non-finite
internally headed, 383[34], 397, 397[53]
matching analysis of, 372
non-finite, 374–375
non-restrictive, 389–392
predicative, 396
R, RP, *see* relative clause, empty relativiser
raising analysis of, 372, 373[22]
reduced, 386[39], 387–388
relative-corelative, 394[48]
relativised constituent, 356, 372, 380, 381

subject relative, 367[13], 367[15]
supplemental, *see* relative clause, non-restrictive
supplementary, *see* relative clause, non-restrictive
that-less, *see* relative clause, bare
transparent free, 401–402
wh-phrase in, 356–375, 377, 385, 389, 398–401
wh-relative, 356–372, 374–375, 376[25], 377
relative clauses, 788
relative inheritance, 358–359, 359[2], 361[7], 360–362, 366, 369[19], 370, 370[20], 372, 388, 397[53], 399–401, 401[59]
relative pronoun, 356–358, 358[1], 359, 359[1], 361, 362, 362[8], 364, 365, 369, 370, 373, 375, 376[25], 377, 377[29], 378, 378[30], 382, 385[37], 388, 390, 390[45], 391, 394[48], 396, 398, 398[55], 399, 400, 403
remnant movement, 991
representational gesture, 894, 907–914
reprise content hypothesis, 855
reprise fragment, 854–856
resemblance, 894
resumptive pronoun, 373, 377, 380–382
retraction phase, 897
robot control, 26
Robust Minimal Recursion Semantics, 907
robust parsing, 818
robust processing, 809

SaGA, 889[1]
SaGA Corpus, 916
schema
Head-Complement, 973
scope, 804–806, 981–983
of negation, 818
scope of quantificational NPs, 21
scrambling, 114
secondary predicate, 396
segmentation, 804
Segmented Discourse Representation Theory, 870
semantic analyses, 820
semantic representations, 804, 809, 810, 817, 856–859
variable-free, 804
semantic roles, 806
semantic type, 856
semantics, 961
sembank, 814–816
semiotics, 888
sentence chunking, 810
sentential extraposition, 17
Shakespeare, 890
shallow processing, 808
ShapeWorld, 820
shared assumptions, 849
shared attention, 843
shared information gesture, 914
shared knowledge, 843
Shared Task on Extrinsic Parser Evaluation, 818
short-distance reflexives, 615, 652
shuffle, 79
shuffle operator, 798
sign, 862
sign formation, 914
sign language, 897
Sign-Based Construction Grammar, 175

Sign-based Construction Grammar, 356, 367[15]
signature, 64
similarity, 894, 908
Simpler Syntax, 28
singleton type, 861
situated communication, 843
Situated Prosodic Word Constraint, 906, 911
situatedness, 842
situation, 860, 863
situation semantic, 863
Situation Semantics, 847
situation theory, 857
situation type, 863
sketch model, 918
slingshot argument, 859
SmartKom Corpus, 916
source language, 819
SOV, 979
speaker, 846
specificational pseudo-cleft, 401–402
speculation, 817
speech, 986[36]
speech-gesture ensemble, 908, 914
speech-gesture integration, 898, 903
speech-gesture production models, 888, 918–919
split antecedents, 617[6]
split ergativity, 109
Spoken Dutch Corpus, 811
spoken language, 844
square of duality, 650, 651
square of opposition, 650
standard, 850
standard meaning, 850
standardisation, 842
Stanford University, 817
statistical methods, 23
statistical model, 26, 811
string, 853
string theory of events, 853
stroke phase, 897
Stroop effect, 781
structural priming, 782, 785
structure sharing, 900
structured content, 859
subcategorisation frames, 810
Subcategorization Principle, 21
Subjacency, 992
subject extraction, 17
Subject Extraction Lexical Rule, 367[13]
subject orientedness, 625
subject-oriented anaphor, 627
subject-orientedness, 626
subjects, 801
summarization, 818
    abstractive, 818
summariziation
    extractive, 818
supertagging, 803, 810
surface-oriented, 783
suspended affixation, 53
SVO, 979
synchrony, 888

taboo, 914
taboo gesture, 914
tacesics, 888
tail, 848
target language, 819
temporal relationship, 900
tense, 847
test suites, 808
the lexical boost, 786
thumb pointing, 895
thumbs-up gesture, 895

Tibidabo Treebank, 815[30]
token, 859
tokenization, 804
Top-down Phase-based Minimalist Grammar (TPMG), 1083[4]
topic, 848
topological field, 295
topological fields, 55
Trace-Principle, 367[13]
TRALE, 26, 807, 809
transfer grammar, 819
Tree Adjoining Grammar (TAG), 803, 1081, 1107
treebank, 26, 798[1], 803, 804, 809, 811, 814–816, 821
  searchable, 815
truth-conditional difference, 847
truth-conditional semantics, 820, 857
TUHL, *see* Type Underspecified Hierarchical Lexicon
Turing-complete, 799, 801
turn management, 893
turn-assigning gesture, 915
turn-taking, 843
type, 859, 965
  comp, 374[23]
  fact, 363[23], 371[23], 376[23]
  head-adjunct-phrase, 371[23]
  list, 801[23]
  local, 380[23]
  message, 363[23]
  naming, 361[23]
  nom-object, 363[23]
  non-empty-list, 801[23]
  nprl, 380[23]
  null, 801[23]
  outcome, 363[23]
  parameter, 360[23]
  plr, 380[23]
  proposition, 363[23], 367[23], 371[23]
  psoa, 363[23]
  question, 363[23]
  scope-object, 360[23], 363[23]
  simp-inf-rel-cl, 386[23]
  state-of-affairs, 363[23], 367[23]
  v-mod, 384[23]
  verbal, 374[23]
  verb, 374[23], 376[23]
  word, 370[23]
  arc, 902
  clause, 367, 368
  core-cl, 368
  decl-cl, 368
  diff-list, 713
  fact, 363, 364, 366, 371
  fin-wh-rel-cl, 368, 369, 390
  hd-fill-phrase, 369, 374
  hd-relative-mod-phrase, 384
  head-adjunct-phrase, 371, 403
  head-nexus-phrase, 369
  head-relative-mod, 384
  head-relative-phrase, 370, 371
  head-subject-phrase, 396
  head=comp=cx, 1095, 1096
  head=complement=phrase, 1093
  headed-phrase, 288
  imp-cl, 368
  inf-hd-fill-phrase, 375
  inf-hd-fill-rel-cl, 375
  inf-wh-rel-cl, 368, 375
  info-str, 713
  line, 902
  located_command, 900
  non-finite, 375
  non-wh-rel-cl, 368, 385
  noun, 357, 360, 365, 368, 371, 384,

388, 390
nplr, 380
nprl, 380
plr, 381
proposition, 367, 368
red-rel-cl, 368, 388
rel-cl, 367, 368, 371, 379, 385, 390
scope-object, 363, 364, 366, 372
simp-inf-rel-cl, 368
soa, 361, 363
state-of-affairs, 361, 364
that, 377
unmarked, 377
v-mod, 384
verb, 357, 365
verbal, 369
wh-rel-cl, 368, 369, 379
which, 391
type hierarchy, 24, 798
type shifting, 972[27]
type theory, 859–862
Type Theory with Records, 859–862
type underspecification, 81–86
type-theoretical semantics, 857–862
TypeGram, 814[28]
types, 858
typological properties, 814

ubertagging, 804, 810
unaccusative, 732
unbounded dependencies, 14, 24, 79
unbounded dependency, 356–403, 954
strong, 377
weak, 377
unbounded dependency constructions, 9
underspecification, 786
unification, 855, 1090
Uniform Theta Assignment Hypothesis (UTAH), 959, 961
Universal Dependencies (UD), 811, 815
Universal Grammar (UG), 940, 1084
universe, 857
update, 852, 853

V2 word order, 812
valence, 80
valence frames, 821
valence lists, 801
vector space model, 820
verb-penultimate word order, 812
Verbmobil, 808
Verb*mobil*, 22, 23, 25, 26
verbs, 821
of speculation, 817
visual question answering (VQA), 820
visual situation, 863
VOS, 730
VSO order, 110, 112

WeScience, 816
WeSearch, 815[29]
Wh-Inheritance Principle, 370[20]
wh-percolation, *see* relative inheritance
Wikipedia, 816
WikiWoods, 816, 820
witness, 860, 862
word
internal structure of, 52–56
word order, 814
word retrieval, 893
world knowledge, 779, 781
wrapping, 11
written language bias, 843

$\overline{\text{X}}$ theory, 991

YADU, 72

YY Software, 23

Did you like this book?

This book was brought to you for free

Head-Driven Phrase Structure Grammar

Head-Driven Phrase Structure Grammar (HPSG) is a linguistic framework that models linguistic knowledge on all descriptive levels (phonology, morphology, syntax, semantics, pragmatics) by using feature value pairs, structure sharing, and relational constraints. This volume summarizes work that has been done since the mid 80s. Various chapters discus formal foundations and basic assumptions, describe the evolution of the framework and go into the details of various syntactic phenomena. Separate chapters are devoted to non-syntactic levels of description. The book also handles related fields and research areas (gesture, sign languages, computational linguistics) and has a part in which HPSG is compared to other frameworks (Lexical Functional Grammar, Categorial Grammar, Construction Grammar, Dependency Grammar and Minimalim).