

4. Theoretische Grundlage und Messmethode

Dieses Kapitel stellt die theoretischen Grundlagen vor und wie diese zur Komplexitätsmessung übernommen und/oder angepasst werden müssen (Kap. 4.1.). Davon wird abgeleitet, was ein System mehr oder weniger komplex macht (Kap. 4.2.) und wie in dieser Arbeit Komplexität gemessen wird (Kap. 4.3.).

Zu einem einfacheren Verständnis wird hier sehr kurz zusammengefasst, welche Komponenten die nominale Flexionsmorphologie mehr oder weniger komplex machen. Diese Zusammenfassung basiert auf Baechler & Seiler (2012: 28) und Baechler (2016: 23).

Ein System wird durch Folgendes **komplexer**:

- Anzahl der morphosyntaktischen Eigenschaften, die in der Flexion markiert werden, und zwar unabhängig davon, wie sie ausgedrückt werden (z.B. Affigierung, Subtraktion, Modifikation der Wurzel etc.).
- Je mehr Allomorphie, desto komplexer. Z.B.: Pluralallomorphe im Substantiv der deutschen Standardsprache (*–ə*, *–ər*, *–n* etc.).
- Mehrfachausdruck derselben morphosyntaktischen Eigenschaft. Z.B.: In *Wälder* wird der Plural durch den Umlaut und durch das Suffix *–ər* ausgedrückt.
- Ein bestimmter Synkretismus: Wenn die Werte von mindestens zwei morphosyntaktischen Eigenschaften variieren. Z.B.: In der starken Adjektivflexion der deutschen Standardsprache markiert *–ər* u.a. den Genitiv Singular Feminin und den Genitiv Plural. Folglich variieren die Werte von zwei morphosyntaktischen Eigenschaften, nämlich Genus (Feminin vs. nicht spezifiziert) und Numerus (Singular vs. Plural).

Ein System wird durch Folgendes **simpler**:

- Ein bestimmter Synkretismus: Wenn die Werte von maximal einer morphosyntaktischen Eigenschaften variieren.
- Wenn keine overte Markierung vorhanden ist (Unterspezifikation). Z.B.: Der Nominativ Singular der Substantive in der deutschen Standardsprache.
- Unterscheidungen, die in einer Wortart gemacht werden, aber nicht in einer anderen. Z.B.: Im Kaiserstuhl Alemannischen wird Kasus am Substantiv nicht markiert, jedoch an den Determinierern und am Adjektiv.
- Phonologisch erklärbare Allomorphie. Z.B.: Die Allomorphe des Genitiv Singular Suffixes *–əs/–s* in der deutschen Standardsprache hängen von der phonologischen Struktur des Wortes ab (Auslaut, Anzahl Silben, Akzent).

4.1. Theoretische Grundlage

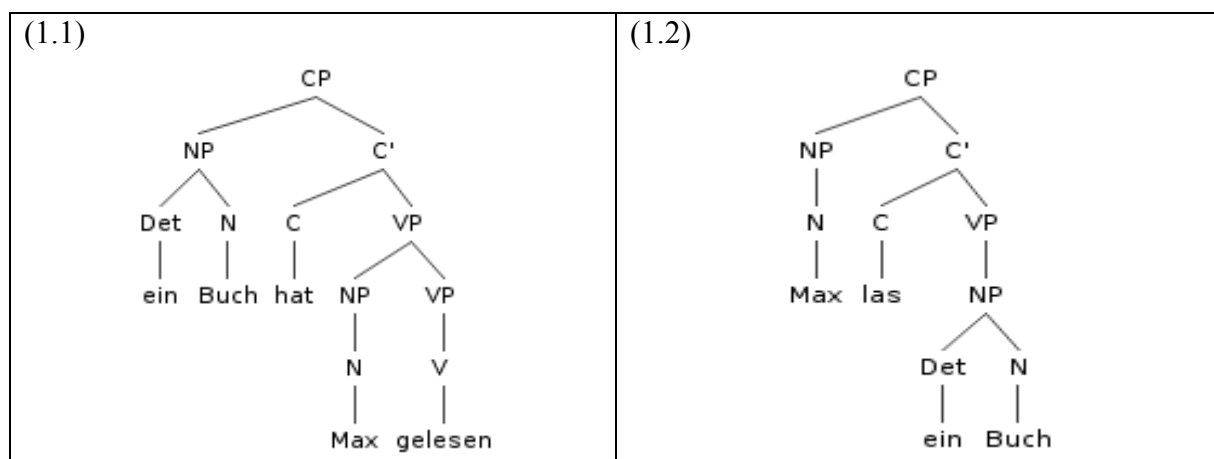
4.1.1. LFG und Morphologie

Grundsätzliches zu LFG: In der Lexical-Functional Grammar (LFG) werden neben einem Lexikon verschiedene Subsysteme angenommen, die jeweils eigene Einheiten und Regeln aufweisen. Zu diesen Subsystemen, die Strukturen genannt werden, gehören beispielsweise die

semantische Struktur, Informationsstruktur, phonologische Struktur, Argumentstruktur, funktionale Struktur und Konstituentenstruktur (Falk 2001: 22–25). Diese Subsysteme bilden keine hierarchische, sondern eine parallele Architektur, d.h., sie existieren gleichzeitig parallel, keine Struktur geht einer anderen voraus (Falk 2001: 23). Die Entsprechungen und Beziehungen zwischen den verschiedenen Strukturen werden durch ein System von Funktionsgleichungen abgebildet (Bresnan 2001: 51). Wie in anderen Frameworks wird auch in LFG diskutiert, welche Subsysteme in einer Sprache existieren und folglich modelliert werden müssen. Jedoch gehen alle mir bekannten Publikationen davon aus, dass zumindest eine funktionale Struktur (f-Struktur) und eine Konstituentenstruktur (c-Struktur) angenommen werden müssen. Da diese im weiteren Verlauf dieser Arbeit wieder vorkommen, sollen sie kurz skizziert werden.

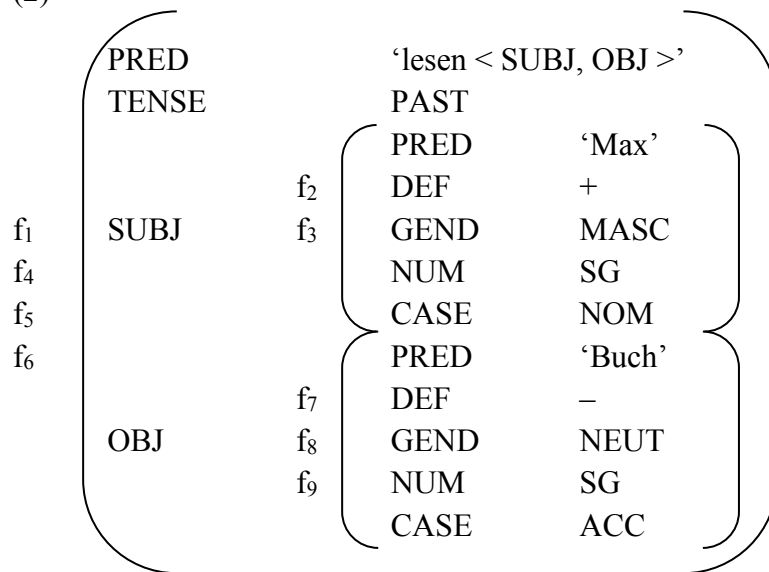
Die Unterscheidung zwischen c-Struktur und f-Struktur beruht auf der Beobachtung, dass die Hierarchie von Phrasen und deren Abfolge nichts mit den Funktionen in einem Satz zu tun haben müssen. Man betrachte die beiden Sätze, deren c-Strukturen in (1) und deren f-Struktur in (2) abgebildet ist: (1.1) *Ein Buch hat Max gelesen*, (1.2) *Max las ein Buch*.¹⁰ Gehen wir davon aus, dass Perfekt und Präteritum in der deutschen Standardsprache die Zeitstufe Vergangenheit ausdrücken, weisen beide Sätze dieselben Funktionen auf: Prädikat *lesen* in der Vergangenheit, Subjekt *Max* (mit morphosyntaktischen Eigenschaften), Objekt *ein Buch* (mit morphosyntaktischen Eigenschaften). Die Unterschiede betreffen ausschließlich den Aufbau und die Abfolge der Konstituenten. Dieser Tatsache kann Rechnung getragen werden, indem für beide Sätze eine f-Struktur angenommen wird (abgebildet in (2)) und zwei c-Strukturen (abgebildet in (1)).

(1)



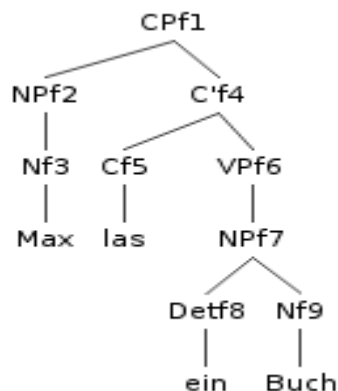
¹⁰ Die c-Struktur basiert auf Berman (2003: 23-44).

(2)



Um die c-Struktur und die f-Struktur miteinander zu verbinden, wird ein System von „correspondence functions, or [...] projection [...] functions“ (Falk 2001: 24) benötigt. Dazu wird jedem Knoten in der c-Struktur eine Variable f_x zugewiesen, wie in der c-Struktur in (3). Die einem bestimmten Knoten entsprechende f-Struktur wird mit denselben Variablen versehen (vgl. (2)).

(3)



Damit kann nun das Mapping zwischen der c-Struktur und der f-Struktur definiert werden, und zwar anhand einer Reihe von Funktionalgleichungen, was als funktionale Deskription (f-Deskription) bezeichnet wird (Falk 2001: 68–69). Die f-Deskription für die f-Struktur in (2) und die c-Struktur in (3) steht in (4). Aus der f-Deskription in (4) wird ersichtlich, dass die Gleichungen die Beziehung zwischen Mutter- und Tochterknoten ausdrücken. Die Variablen f_x können also durch Metavariablen ($\uparrow\downarrow$) ersetzt werden, wobei die Variable \uparrow auf den Mutterknoten verweist, die Variable \downarrow auf den Tochterknoten (Falk 2001: 71). Beispielsweise kann die Gleichung (f_1 SUBJ) = f_2 durch die Gleichung (\uparrow SUBJ) = \downarrow ersetzt werden.

(4)

(f₁ SUBJ) = f₂

f₂ = f₃

(f₃ PRED) = 'Max'

(f₃ DEF) = +

(f₃ GEND) = MASC

(f₃ NUM) = SG

(f₃ CASE) = NOM

f₁ = f₄

f₄ = f₅

(f₅ PRED) = 'lesen < (f₅ SUBJ) (f₅ OBJ) >'

f₄ = f₆

(f₆ OBJ) = f₇

f₇ = f₈

f₇ = f₉

(f₈ DEF) = –

(f₈ NUM) = SG

(f₈ CASE) = ACC

(f₈ GEND) = NEUT

(f₉ PRED) = 'Buch'

(f₉ NUM) = SG

(f₉ CASE) = ACC

(f₉ GEND) = NEUT

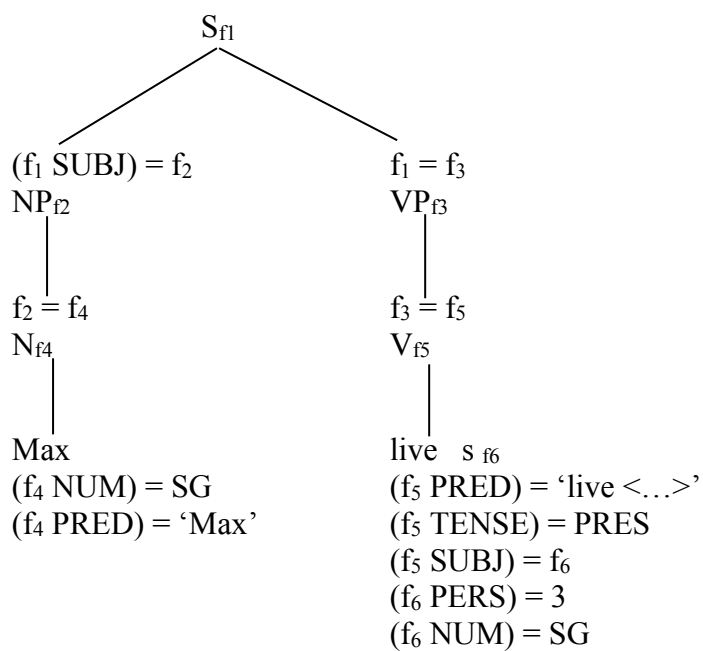
Gehen wir von einem Modell mit paralleler Architektur aus, kann auch die Morphologie ein Subsystem bilden. Für die Komplexitätsmessung ist eine parallele Architektur vor allem deswegen von Vorteil, weil nur so in Zukunft eventuelle Ausgleichstendenzen zwischen den Subsystemen (z.B. zwischen Syntax und Morphologie) gemessen werden können. Natürlich stellt sich ganz grundsätzlich die Frage, ob Morphologie als autonome Repräsentationsebene existiert oder sie Teil der Syntax und Phonologie ist. Die unterschiedlichen Positionen dazu können hier nicht umfassend dargestellt werden. Vielmehr soll hier dafür argumentiert werden, dass die Morphologie vielleicht keine universelle Komponente in der Grammatik bildet, aber dass sie, wenn sie in einer Sprache existiert, nach unabhängigen Prinzipien und Regeln strukturiert ist (Börjars, Vincent & Chapman 1997: 156). Neben vielen anderen konnte auch Anderson (1992, vor allem Kapitel 2) zeigen, dass es ein System an Regeln gibt, das Wörter konstruiert, und ein anderes System an Regeln, dass die syntaktische Struktur von Phrasen und Sätzen organisiert (Anderson 1992: 22). Nehmen wir also die Morphologie als Subsystem an, stellt sich die Frage, wie die Einheiten und Regeln dieses Subsystems innerhalb von LFG ausschauen. Traditionell wird in LFG eine morphembasierte Auffassung von Morphologie vertreten; genauer eine lexikalisch-inkrementelle Auffassung (Ackerman & Stump 2004: 112). Lexikalisch heißt, dass Stämme und Affixe im Lexikon gelistet sind. Unter inkrementell versteht man den Vorgang, dass ein flektiertes Wort eine bestimmte grammatische Eigenschaft erhält, indem es jenen Marker zu sich nimmt, der genau diese grammatische Eigenschaft trägt.

Folgendes leicht abgeändertes Beispiel aus Bresnan (2001: 55) soll dies kurz illustrieren. Die Verbform *lives* besteht aus einem Stamm und einem Suffix, die mit funktionalen Schemata im Lexikon gelistet sind:

live (↑ PRED) = 'live <...>'

–s (↑ TENSE) = PRES
 (↑ SUBJ) = ↓
 (↓ PERS) = 3
 (↓ NUM) = SG

Werden beide Lexikoneinträge miteinander kombiniert, entsteht die flektierte Wortform *lives* mit den morphologischen Eigenschaften, die vom Suffix –s stammen. Dieses flektierte Wort kann in die c-Struktur eingefügt werden, wobei eine funktionale Deskription generiert wird (zum besseren Verständnis wird hier auch ein Subjekt angefügt). Leicht abgeändertes Beispiel aus Bresnan (2001: 58):



Diese funktionale Deskription definiert das Mapping zwischen c- und f-Struktur. Zur Veranschaulichung wird in der Folge noch die f-Struktur dargestellt (Bresnan 2001: 59; leicht abgeändert):

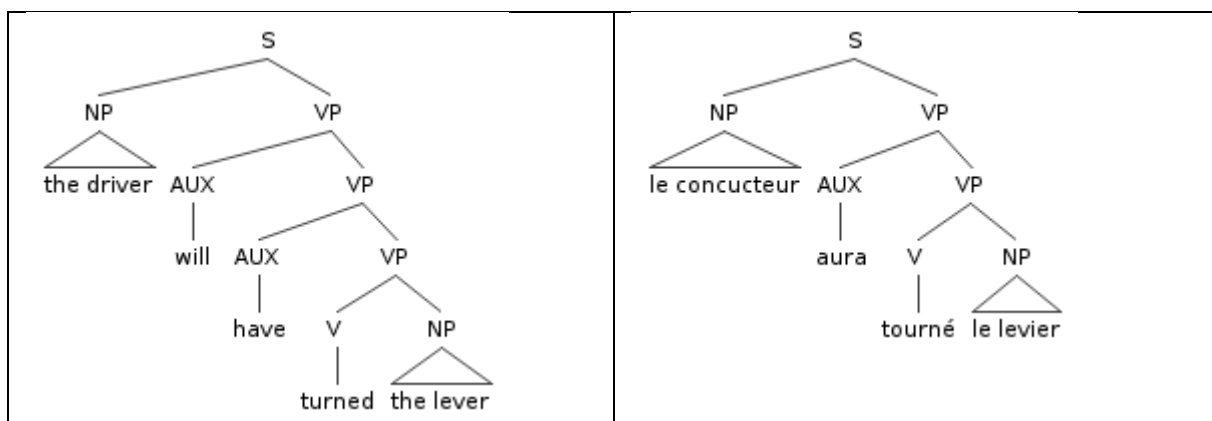
$$f_1, f_3, f_5: \left(\begin{array}{c} \text{SUBJ} \\ \\ \text{TENSE} \\ \text{PRED} \end{array} \right) \quad f_2, f_4, f_6: \left(\begin{array}{c} \text{PERS } 3 \\ \text{NUM } \text{SG} \\ \text{PRED 'Max'} \end{array} \right)$$

Zwei Dinge können hier festgehalten werden. Erstens können nur vollständige Wortformen in die c-Struktur eingefügt werden, was einen zentralen Grundsatz von LFG darstellt. Somit bilden Syntax und Morphologie zwei verschiedene Module mit eigenen Regeln. Zweitens wird hier aber noch von keinem autonomen morphologischen Subsystem ausgegangen. Dass es ein solches aber braucht, geht auf die Diskussion zur Schnittstelle zwischen Syntax und Morphologie sowie auf die konsequente Trennung von Form und Bedeutung zurück. Diese Diskussion soll hier kurz an einigen Beispielen skizziert werden.

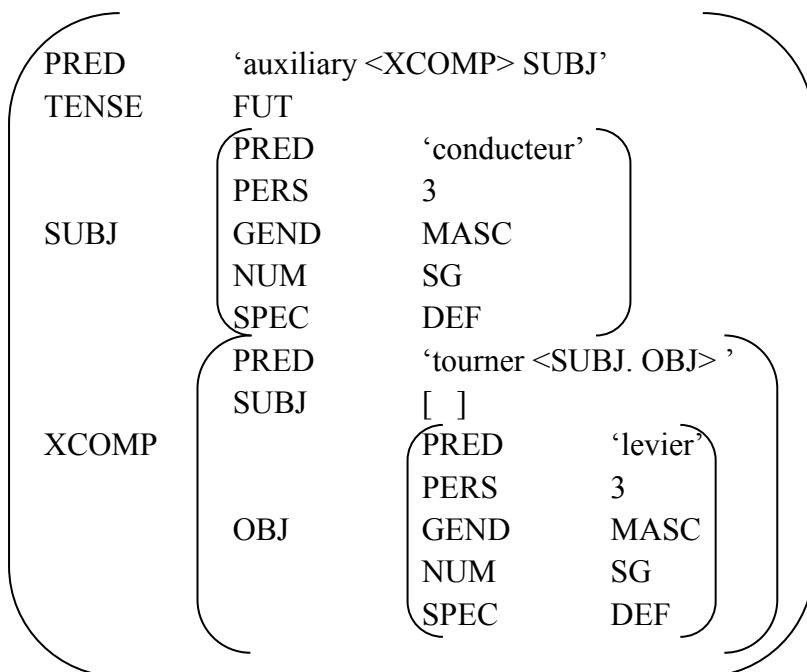
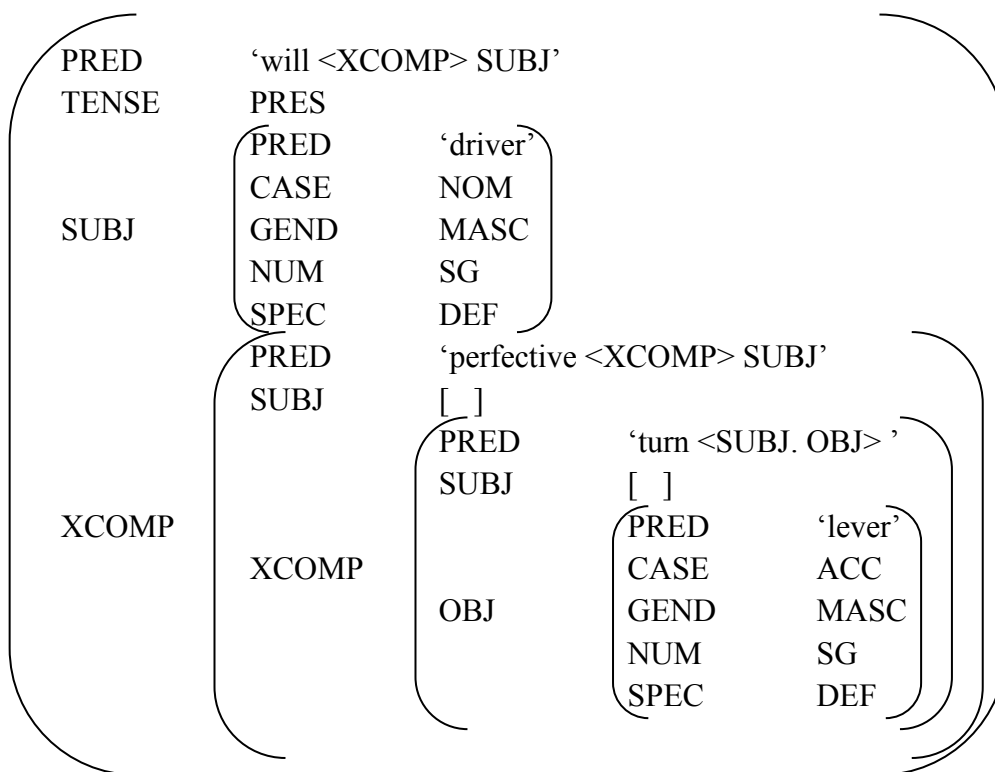
M-Struktur: Butt, Niño & Segond (2004) schlagen eine Modellierung von Hilfsverben vor, die die spezifische Kombination von Hilfs- und Vollverben zur Bildung von Periphrasen in der m(orphologischen)-Struktur verortet. Traditionell werden in LFG Hilfs- und Modalverben wie Vollverben behandelt, wobei jedes verbale Element ein Komplement zu sich nimmt (Butt, Niño & Segond 2004: 13). Nach dieser Analyse haben die Sätze in (5) die c-Struktur in (6) und die f-Struktur in (7) (Butt, Niño & Segond 2004: 14–15.)

- (5) The driver will have turned the lever.
 Le conducteur aura tourné le levier
 Der Fahrer wird den Hebel gedreht haben.

(6)

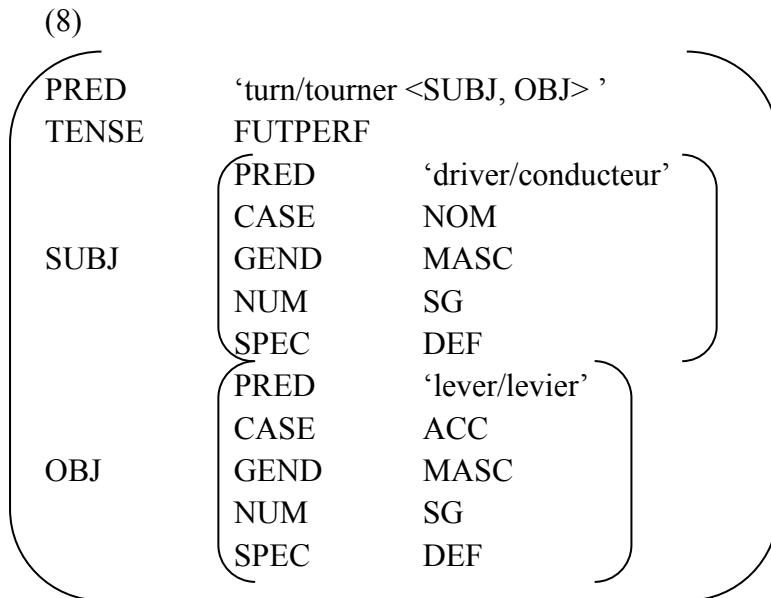


(7)

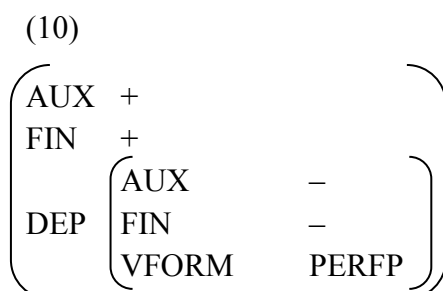
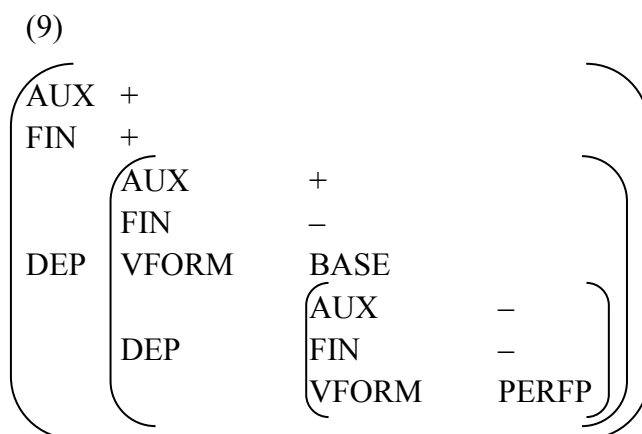


Eine wichtige Grundidee in LFG ist, grammatische Funktionen (f-Struktur), wie z.B. Subjekt und Tempus, vom Aufbau und der Abfolge der Konstituenten (c-Struktur) zu trennen. Die Analyse in (7) suggeriert nun, dass sich die beiden Sätze in ihrer funktionalen bzw. hier in ihrer prädikationellen Struktur unterscheiden. Diese Analyse ist jedoch falsch, da in beiden Sätzen ein Futur 2 ausgedrückt wird. Der einzige Unterschied in diesen Sätzen besteht darin, dass das Futur 2 verschieden kodiert wird. Es gilt also auf der einen Seite die Gleichheit auf der

funktionalen Ebene (Bedeutung) zu modellieren und auf der anderen Seite die Unterschiede in der Kodierung (Form) dieser Funktion/Bedeutung. Dazu schlagen Butt, Niño & Segond (2004) eine f-Struktur (8) für die Sätze in (5) vor (Butt, Niño & Segond 2004: 16):



Die sprachspezifischen Unterschiede in der Form, bei denen es sich hier um morphologische Unterschiede handelt, werden in der m-Struktur kodiert. (9) ist die m-Struktur für den englischen Satz, (10) die m-Struktur für den französischen Satz (Butt, Niño & Segond 2004: 18; aus Gründen der Verständlichkeit fallen die m-Strukturen hier ausführlicher aus).



Für die Sätze in (5) können wir also festhalten, dass sie die gleichen grammatischen Funktionen aufweisen, was mit einer identischen f-Struktur modelliert werden kann (8). Der Unterschied besteht ausschließlich in der verschiedenen Kodierung der Funktion Futur 2, was durch unterschiedliche m-Strukturen repräsentiert wird (9 und 10).

Durch diese konsequente Trennung von syntaktischer, funktionaler Bedeutung und morphologischer Form können nicht nur Periphrasen adäquat modelliert und sprachübergreifend verglichen werden. Auch Wortformen und Periphrasen, die eine bestimmte Form haben, welche jedoch ihrer Bedeutung widerspricht, stellen kein Problem mehr da. Der englische Satz in (5) beispielsweise weist kein Verb auf, das im Futur steht: *will*=Präsens, *have*=Infinitiv, *turned*=Partizip Perfekt. Es stellt sich also die Frage, wie die Periphrase *will have turned* zu ihrer Bedeutung Futur 2 kommt. Nimmt man jedoch eine f-Struktur mit der Funktion Futur 2 an und eine m-Struktur mit der spezifischen Kombination dieser drei Verben, können diese beiden Strukturen z.B. mit Funktionalgleichungen miteinander verbunden werden (Projektionen). Ein weiteres Beispiel sind die lateinischen Deponentia. Es handelt sich dabei um eine Gruppe von Verben, die eine passive Form, aber eine aktive Bedeutung haben, z.B. *loquor* (Tabelle 4.2):

Tabelle 4.1: Flexion des Verbs *laudo* (Sadler & Spencer 2001: 74)

Imperfektiv	Aktiv	Passiv
Präsens	laudat	laudatur
Vergangenheit	laudabat	laudabatur
Futur	laudabit	laudabitur
Perfektiv	Aktiv	Passiv
Präsens	laudavit	laudatus/a/um est
Vergangenheit	laudaverat	laudatus/a/um erat
Futur	laudaverit	laudatus/a/um erit

Tabelle 4.2: Flexion des Deponens *loquor* (Sadler & Spencer 2001: 75)

	Imperfektiv	Perfektiv
Präsens	loquitur	locutus/a/um est
Vergangenheit	loquebatur	locutus/a/um erat
Futur	loquar	locutus/a/um erit

Das Verb *laudat* (Tabelle 4.1) und das Deponens-Verb *loquitur* (Tabelle 4.2) haben für das gleiche f-Feature [VOICE ACTIVE], aber zwei unterschiedliche m-Features: für *laudat* [voice active], für *loquitur* [voice passiv].

Es konnte hier kurz aufgezeigt werden, dass durch die klare Trennung der Komponenten (Funktion, Aufbau und Abfolge von Konstituenten, morphologische Form) Phänomene innerhalb einer Sprache wie auch sprachübergreifende Phänomene adäquat modelliert werden können. Es stellt sich aber nach wie vor die Frage, wie die Verbindung zwischen den

morphosyntaktischen Eigenschaften eines flektierten Wortes (=Bedeutung) und seiner Struktur (=Form) aussehen bzw. wie ein flektiertes Wort zu seiner Bedeutung und Form kommt. In allen hier vorgestellten Studien werden die Wortformen innerhalb des Lexikons gebildet, was impliziert, dass die Stämme und Affixe im Lexikon gelistet sind. Es wird also von einer lexikalisch-inkrementellen Auffassung ausgegangen. Dass diese Auffassung gerade für flektierende Sprache problematisch ist und eine inferentielle-realisierende (engl. inferential-realizational) Auffassung mehr leistet, soll im anschließenden Kapitel referiert werden.

4.1.2. Lexikalische-inkrementelle vs. inferentielle-realisierende Morphologie

Definition: Die grundsätzlichen Unterschiede in der morphologischen Theoriebildung kann nach Stumps (2001: 1–2) Taxonomie in vier Kategorien eingeteilt werden (lexikalisch vs. inferentiell, inkrementell vs. realisierend), die in der Folge kurz vorgestellt wird. **Lexikalische Theorien** gehen davon aus, dass Stämme und Affixe mit ihrer jeweiligen Bedeutung im Lexikon gelistet sind. Im Lexikon stehen also Stämme mit ihren grammatischen und semantischen Bedeutungen, Affixe mit ihren morphosyntaktischen Bedeutungen. In **inferentiellen Theorien** werden Form und Bedeutung eines Affixes voneinander getrennt. Von Wurzeln¹¹ werden Wortformen durch Regeln oder Formeln abgeleitet. Diese Regeln definieren nicht Form und Bedeutung eines Affixes, sondern sie verknüpfen das Auftreten eines Affixes bzw. einer phonetische Modifikation des Wortes mit bestimmten morphosyntaktischen Eigenschaften. In **inkrementellen Theorien** erhalten Wörter ihre morphosyntaktische Bedeutung, indem sie Affixe mit denselben morphosyntaktischen Eigenschaften zu sich nehmen. Im Gegensatz dazu sind in **realisierenden Theorien** Wörter mit morphosyntaktischen Eigenschaften verknüpft und diese Verknüpfung lizenziert das Einfügen eines Markers. Alle vier logisch möglichen Kombinationen sind in Theorien zur Flexionsmorphologie zu finden, viele lassen sich aber grundsätzlich entweder der lexikalischen-inkrementellen oder der inferentiellen-realisierenden Richtung zuordnen. Diese Unterscheidung wird auch morphem- oder wortbasiert genannt und kann wie folgt zusammengefasst werden:

[...] what Stump (2001) calls a *lexical-incremental* conception of morphology: such treatments are LEXICAL by virtue of the assumption that affixes, like stems, possess their own separate representations in the lexicon, and they are INCREMENTAL in that the grammatical properties of a fully inflected word are associated with it only as an effect of its acquiring the morphological markers bearing those properties. (Ackerman & Stump 2004: 112)

[...] morphology of the type designated as inferential-realizational in Stump's taxonomy. Recently there has been a resurgence of interest in the so-called Word & Paradigm approach to morphology [...]; what distinguishes this approach from traditional morpheme-based approaches is its premise that language's inflectional system is INFERENTIAL rather than lexical (in the sense that it represents inflectional exponents not as lexically listed elements, but as markings licensed by rules by which complex word forms are deduced from simpler roots and stems) and is REALIZATIONAL rather than incremental (in the sense that it treats a word's association with a particular set of morphosyntactic properties as a precondition for – not a consequence

¹¹ Genauerer zu Wurzeln und Stämmen vgl. Kap. 4.1.3.3.

of – the application of the rules licensing the inflectional exponents of those properties). (Ackerman & Stump 2004: 116)

Vorteile der inferentiellen-realisierenden Theorie: Eine inferentielle-realisierende Theorie kann Phänomene in der Flexionsmorphologie genauer und adäquater beschreiben als eine lexikalische-inkrementelle. Die wichtigsten Argumente dafür sollen hier Stump (2001) folgend kurz aufgezeigt werden. Für eine **realisierende Theorie** sprechen vor allem zwei Beobachtungen. Erstens kann eine morphosyntaktische Eigenschaft durch mehrere Affixe ausgedrückt werden: „The morphosyntactic properties associated with an inflected word may exhibit EXTENDED EXPONENCE in that word’s morphology. That is, a given property may be expressed by more than one morphological marking in the same word“ (Stump 2001: 4). Ein Beispiel dafür aus der Flexion der deutschen Standardsprache ist die Pluralmarkierung in Wörtern wie *Wälder*: Hier wird der Plural sowohl durch das Suffix *-er* als auch durch den Umlaut kodiert. In inkrementellen Theorien darf eine morphosyntaktische Bedeutung durch maximal einen Affix ausgedrückt werden und Mehrfachausdruck ist zu verhindern: „[b]ecause operations are informationally additive, multiple additions of identical information are precluded“ (Steele 1995: 280, zitiert in Stump 2001: 4). In realisierenden Theorien existiert diese Einschränkung nicht, denn eine morphosyntaktische Eigenschaft kann das Einfügen eines oder mehrerer Marker verursachen (Stump 2001: 4).

Zweitens kommt auch der umgekehrte Fall vor, d.h., eine morphosyntaktische Eigenschaft wird durch kein Affix ausgedrückt: „The morphosyntactic properties associated with an inflected word’s individual inflectional markings may underdetermine the properties associated with the word as a whole“ (Stump 2001: 7). In der deutschen Standardsprache kommt dies häufig vor. Beispielsweise wird in der 1. und 3. Person Singular Präteritum die Person nicht markiert: *ging* (1./3. Person Singular Präteritum) vs. *ging-st* (2. Person Singular Präteritum). Stark flektierte Substantive tragen keine Akkusativmarkierung (*Tag* = Nominativ und Akkusativ), während Nominativ und Akkusativ der schwach flektierten Substantive unterschieden werden (*Mensch* = Nominativ, *Mensch-en* = Akkusativ). Da in inkrementellen Theorien die morphosyntaktischen Eigenschaften einer Wortform von den morphosyntaktischen Eigenschaften der Marker abgeleitet werden, müssen Nullmorpheme angenommen werden (bzw. in inferentiellen Theorien Regeln, die keine Formveränderungen verursachen) (Stump 2001: 7–9). Für die starke Substantivflexion des Deutschen bedeutet das, dass nicht nur ein Nullmorph für den Akkusativ angenommen werden muss, sondern auch ein anderes Nullmorph für den Dativ, also zwei verschiedene Nullmorpheme. Realisierende Theorien brauchen „nothing so exotic to account for these facts“ (Stump 2001: 9), denn die Marker müssen nicht alle Eigenschaften realisieren. Die Wurzel *Tag* ist verknüpft mit der morphosyntaktischen Eigenschaft Genitiv und erst diese Verknüpfung ermöglicht es Regeln, ein Flexionsaffix zu lizenzieren, nämlich *-es*. Gibt es aber für eine bestimmte morphosyntaktische Eigenschaft (z.B. Akkusativ) keine Regel, die ein Flexionsaffix lizenziert, passiert mit der Wurzel nichts.

Es wurde hier gezeigt, dass die Eins-Zu-Eins-Beziehung zwischen Form und Bedeutung eines Morphems nicht zutrifft und dass folglich Wörter ihre Bedeutung nicht durch das Aneignen von Morphemen erhalten können. In den zwei besprochenen Beispielen gab es im ersten eine

Bedeutung, aber mehrere Formen, im zweiten ebenfalls eine Bedeutung, aber keine Form. Auch die beiden umgekehrten Fälle treten auf, nämlich eine Form/keine Bedeutung und eine Form/mehrere Bedeutungen (Spencer 2004: 80–82). Für den ersten Fall sind die Fugenelemente in den deutschen Varietäten ein gutes Beispiel. Bei gewissen Komposita muss obligatorisch ein Element eingefügt werden, das aber keine Bedeutung trägt, z.B. *Universität-s-zeitung*. Der zweite Fall (eine Form/mehrere Bedeutungen) tritt typischerweise in flektierenden Sprachen auf. In der standarddeutschen Verbflexion trägt das nicht weiter segmentierbare Suffix *-st* die Bedeutung 2. Person Singular, das Suffix *-t* 2. Person Plural, die Information zu Person und Numerus sind also in einem Suffix kodiert.

Eine **inferentielle Theorie** im Gegensatz zu einer lexikalischen hat vor allem in zweierlei Hinsicht Vorteile. Erstens wird in inferentiellen Theorien die Unterscheidung zwischen konkatenativer und nicht-konkatenativer Flexion verworfen. In lexikalischen Theorien werden diese zwei Arten von Flexion unterschieden, da nur Affixe mit ihren morphosyntaktischen Eigenschaften aus dem Lexikon stammen. Erst, nachdem Wurzel und Affix sich verbunden haben, werden an der Wurzel Modifikationen vorgenommen. Als Beispiel nennt Stump (2001) die Analyse des englischen Simple Past von Halle & Marantz (1993), wobei es um die komplementäre Verteilung des Ablauts (z.B. *sing – sang*) und des Default-Suffixes *-ed* geht. Halle & Marantz (1993) nehmen an, dass *sang* ein Nullsuffix für die Vergangenheit trägt, das mit dem Default-Suffix *-ed* in Konkurrenz steht. Des Weiteren löst dieses Nullsuffix eine Modifikation am Wurzelsuffix aus. Da das Nullsuffix nun eine kleinere Klasse an Verben subkategorisiert, setzt sich dieses Suffix nach Pāṇinis Prinzip durch, wenn es mit dem Suffix *-ed* konkurriert. In einer inferentiellen Theorie werden nur zwei Regeln benötigt, die miteinander in Konkurrenz stehen: Eine Regel für den Ablaut (in diesem Bsp. die Ersetzung von *i* durch *a*) und eine für die Suffigierung von *-ed*. Da die Ablautregele beschränkter Anwendung findet als die *-ed*-Regel, setzt sie sich nach Pāṇinis Prinzip durch (Stump 2001:10). Dass eine Unterscheidung zwischen konkatenativer und nicht-konkatenativer Flexion nicht nötig ist, bringt Stump (2001) wie folgt auf den Punkt:

[...] but although concatenative and nonconcatenative inflection differ in their phonological expression, there is no convincing basis for assuming that they perform different functions or occupy different positions in the architecture of a language's morphology [...] Thus, in inferential theories, the morphological rule associated with a given set of morphosyntactic properties may be either affixational or nonconcatenative; the difference between affixational rules and nonconcatenative rules has no theoretical importance. (Stump 2001: 9)

Dies lässt sich auch auf das Deutsche übertragen, z.B. auf die Pluralmarkierung. Wird der Plural markiert, gibt es dafür drei Möglichkeiten: a) Suffix (*Tisch – Tisch-e*), b) Umlaut (*Apfel – Äpfel*), c) eine Kombination von beidem (*Wald – Wäld-er*). Wie gezeigt wurde, stellen weder unterschiedliche Mechanismen in der Flexion (konkatenativ, nicht-konkatenativ) noch die Kombination dieser Mechanismen und der Mehrfachausdruck derselben Funktion (hier Plural) für eine inferentielle-realisierende Theorie ein Problem dar.

Eine zweite strittige Unterscheidung, die in lexikalischen Theorien gemacht wird, ist die Unterscheidung zwischen Bedeutungs- und Kontexteigenschaften eines Affixes. Ein Affix

kann mit zwei Arten von morphosyntaktischen Eigenschaften im Lexikon gelistet sein: a) mit seiner Bedeutung, b) mit Restriktionen seine Subkategorisierung betreffend, d.h., in welchen Kontexten das Affix eingesetzt werden darf (Stump 2001: 10). Die deutsche Standardsprache weist im Dativ Plural ein *-n*-Suffix auf (Variation von *-en/-n* ist phonotaktisch erklärbar). In lexikalischen Theorien muss nun entschieden werden, ob das Suffix *-n* die Bedeutung Dativ trägt und eine Beschränkung, nur an Pluralstämme suffigiert zu werden oder ob das Suffix *-n* die Bedeutung Dativ Plural hat. Da in inferentiellen Theorien Regeln eine Relation zwischen den morphosyntaktischen Eigenschaften eines Wortes und seiner Morphologie herstellen, ist die Unterscheidung zwischen Bedeutungs- und Kontexteigenschaften nicht nötig. Wie diese Regeln aussehen, wird im nachfolgenden Kapitel 4.1.3. gezeigt.

Es konnte hier also gezeigt werden, dass ein inferentieller-realisierender Ansatz die Phänomene in der Flexion adäquater beschreiben kann. Traditionell wird in LFG ein lexikalischer-inkrementeller Ansatz vertreten (nicht aber z.B. Butt, Niño & Segond 2004, vgl. Kap. 4.1.1.). Im folgenden Kapitel soll dargestellt werden, dass ein inferentieller-realisierender Ansatz durchaus in LFG implementierbar ist.

4.1.3. LFG und inferentielle-realisierende Morphologie

4.1.3.1. Content-Paradigm und Form-Paradigm und ihre Relationen

Es wurde gezeigt, dass nicht nur auf der Satzebene Form und Bedeutung auseinandergehalten werden sollen (c-Struktur und f-Struktur), sondern auch auf der Wortebene. Folglich darf auch nur die Bedeutungsseite (und nicht die Formseite) eines Lexems mit der f-Struktur verbunden werden. Am Beispiel periphrastischer und analytischer Formen konstatieren Ackerman & Stump (2004) Folgendes:

Our claim here, however, is that the syntactic atoms constituting a periphrase may be nothing more than form-theoretic exponents of a unitary content-theoretic element, and that it is this latter element – not its exponents – that determines the periphrase’s f-structure. In particular, we claim that rules of morphology define the (potentially periphrastic) realization of a lexeme’s pairing with a particular set of morphosyntactic properties, and that the association of such a pairing with an f-structure is insensitive to the manner of its realization. (Ackerman & Stump 2004: 117)

Dazu schlagen Ackerman & Stump (2004) ein inferentielles-realisierendes Modell für die Morphologie vor, welches in LFG implementiert werden kann. Dieses Modell stellt in dieser Arbeit die Grundlage zur Messung morphologischer Komplexität dar und wird nun vorgestellt. Die folgenden Ausführungen beziehen sich auf Ackerman & Stump (2004: 117–124), wobei die englischen Ausdrücke nicht ins Deutsche übersetzt werden.

Um Bedeutung von Form zu unterscheiden, ist das Lexikon zweigeteilt und besteht aus einem Lexemikon (Bedeutung), das Lexeme mit einer lexikalischen Bedeutung beinhaltet, und einem Radikon (Form), das Wurzeln, also ausschließlich Formen beinhaltet (schematische

Darstellung in Tabelle 4.3, Beispiele zu Teilparadigmen in Tabelle 4.4). Jedes Lexem L des Lexemikons ist mit einem *Content-Paradigm* C-P (L) verbunden. Eine Zelle des *Content-Paradigm* besteht aus einer Kombination von L und einem kompletten¹² Set an morphosyntaktischen Eigenschaften $\langle L, \sigma \rangle$. Diese Kombination wird *Content-Cell* genannt. Nur die *Content-Cell* beinhaltet syntaktisch und semantisch interpretierbare Informationen. Das *Content-Paradigm* bildet also das Interface mit der Syntax (f-Struktur) und Semantik: „A lexeme L’s content paradigm lists the morphosyntactic property sets with which L may be associated in syntax and which determine L’s semantic interpretation in a particular sentential context“ (Stump 2016: 104).

Parallel dazu hat jedes r des Radikons ein *Form-Paradigm* F-P (r), das aus einer Kombination aus r und einem Set an morphosyntaktischen Eigenschaften zusammengesetzt $\langle r, \sigma \rangle$ ist. Diese Kombination heißt *Form-Cell*. Die *Form-Cells* beinhalten also nicht nur die Wurzeln (r), sondern auch die Informationen (σ), die nötig sind, um die morphologische Realisierung (d.h. die Wortform) herzuleiten. Sie besteht folglich aus jenen morphosyntaktischen Eigenschaften, die morphologisch, also durch die Flexion des Stammes realisiert werden (Stump 2016: 104). Kurz gesagt, hat das *Form-Paradigm* Informationen zur Form, das *Content-Paradigm* Informationen zur Bedeutung.

Tabelle 4.3: Lexikalische Repräsentation (Ackerman & Stump 2004: 124)

	Lexical representations:		Realizations (word forms)
	content-cell ($\langle L, \sigma \rangle$ pairings)	form-cell ($\langle r, \sigma \rangle$ pairings)	
Associations:	content-cell associated with form-cells by <i>rules of paradigm linkage</i>		
		form-cell associated with realizations by <i>realization rules</i>	
Information represented:	Contentive (functional-semantic content, grammatical functions, morphosyntactic properties	Formal (morphosyntactic property labels); Diacritic/indexical (inflection class membership, root phonology)	Purely phonological

Wichtig ist hier jedoch, dass keines dieser Paradigmen Wortformen beinhaltet. Die Wortformen werden von sogenannten Realisierungsregeln generiert, die ausschließlich die phonologische Substanz der Wortform definieren. Die Realisierungsregeln werden genauer in Kapitel 4.1.3.2.

¹² Ein Set an morphosyntaktischen Eigenschaften ist komplett, wenn es alle für die Kategorie/Wortart relevanten morphosyntaktischen Eigenschaften enthält und diese morphosyntaktischen Eigenschaften sich nicht widersprechen (wohlgeformt). Für ein deutsches Substantiv ist das Set {NOM, SG, MASC} komplett, aber nicht {NOM, SG}. Ein komplettes Set muss auch wohlgeformt sein: Das Set {NOM, SG, MASC} ist wohlgeformt, das Set {NOM, SG, PL, MASC} ist es nicht. Genauer beschrieben wird dies im Kapitel 4.1.3.2.

eingeführt. Wie die Zellen des *Content-Paradigm* und *Form-Paradigm* sowie die Wortformen realisiert und miteinander verknüpft werden, wird gleich anschließend erklärt. Zur Illustration sollen hier zuerst einige Beispiele vorgestellt werden (Tabelle 4.4), die zugleich die Vorteile zeigen, wenn ein *Content-Paradigm*, ein *Form-Paradigm* und die Wortformen voneinander getrennt werden.

Für das Verb *machen* ist grundsätzlich festzuhalten, dass jede *Content-Cell* aus einem Lexem L (MACHEN) und morphosyntaktischen Eigenschaften besteht (z.B. 2 sg ind pres act). Der *Content-Cell* entspricht eine *Form-Cell*, die aus der Wurzel (*mach*) und ebenfalls morphosyntaktischen Eigenschaften besteht (2 sg ind pres act). Die Wortform, also die Realisierung, lautet *machst*. Die Teilung in *Content-Paradigm*, *Form-Paradigm* und Realisierung scheint hier erst mal überflüssig. Die Vorteile zeigen sich in den folgenden Beispielen.

In der Verbflexion der deutschen Standardsprache werden die Formen sowohl synthetisch als auch analytisch gebildet, z.B. das Präsens synthetisch (*machst*), das Futur analytisch (*wirst machen*). Im Italienischen dagegen wird das Futur synthetisch gebildet (*farai*). Für die Bedeutung eines Lexems ist es jedoch unerheblich, wie eine Wortform realisiert wird (vgl. dazu auch die Diskussion in Kap. 4.1.1.). Im Deutschen und Italienischen gibt es also ein Lexem mit der Bedeutung ‘machen’ und mit der Bedeutung ‘2. Singular Indikativ Futur Aktiv’ (*Content-Paradigm*). Im Deutschen hat dieses Verb die Wurzel *mach*, im Italienischen *far* (*Form-Paradigm*). Realisiert wird das Verb im Deutschen analytisch (*wirst machen*), im Italienischen synthetisch (*farai*). Des Weiteren ist hier wichtig, dass sich die Bedeutung ‘Futur’ nicht aus der Form *wirst machen* ableiten lässt, da weder *wirst* noch *machen* futurische Bedeutung trägt. Vielmehr wird in diesem Modell eine bestimmte Form (synthetisch oder analytisch, *Form-Cell* und Realisierung) mit einer bestimmten Bedeutung (*Content-Cell*) verbunden.

Tabelle 4.4: Beispiele für Content-Cell, Form-Cell und Realisierung

Content-Cell <L,σ>	Form-Cell <r,σ>	Realisierung (Wortformen, rein phonologisch)
<MACHEN, {2 sg ind pres act}>	<mach, {2 sg ind pres act}>	machst
<MACHEN, {2 sg ind fut act}>	<mach, {2 sg ind fut act}>	<u>wirst machen</u>
<FARE, {2 sg ind fut act}>	<far, {2 sg ind fut act}>	<u>farai</u>
<SEIN, {1 sg ind pres act}>	< <u>b</u> , {1 sg ind pres act}>	bin
<SEIN, {1 pl ind pres act}>	< <u>s</u> , {1 pl ind pres act}>	sind
<LAUDARE, {3 sg ind pres act}>	<laud, {3 sg ind pres act}>	laudat
<LAUDARE, {3 sg ind pres pass}>	<laud, {3 sg ind pres pass}>	laudatur
<LOQUI, {3 sg ind pres act}>	<loq, {3 sg ind pres pass}>	loquitur

Im Deutschen haben fast alle Verben nur eine Wurzel (z.B. *machen*). Es gibt jedoch auch Verben mit Suppletivstämmen, die also mehr als eine Wurzel haben, z.B. *sein*, *ich bin*, *wir sind*. Diese Wurzeln sind im *Form-Paradigm* gelistet.

Schließlich kommen wir hier auf die lateinischen Deponentia zurück, z.B. *loqui* (vgl. Tabelle 4.2, Kap. 4.1.1.). Dabei handelt es sich um Verben, die von der Form her wie ein Passiv aussehen, jedoch eine aktive Bedeutung haben. Die Verben *laudat* und *loquitur* haben dieselbe Bedeutung (3. Singular Indikativ Präsens Aktiv, vgl. *Content-Cell*), die Form *loquitur* ist jedoch gleich gebildet wie *laudatur*, das eine passive Bedeutung hat (3. Singular Indikativ Präsens Passiv, vgl. *Form-Cell*). Für das Deponens *loqui* lautet σ in der *Content-Cell* also u.a. Aktiv (weil *loquitur* aktive Bedeutung hat), in der *Form-Cell* jedoch Passiv (weil *loquitur* wie ein Passiv gebildet wird).

Es stellt sich nun die Frage, wie diese Zellen realisiert werden und miteinander verknüpft sind. Zur Veranschaulichung ist dies ebenfalls in Tabelle 4.3 dargestellt. Die Relation zwischen den Zellen des *Form-Paradigm* und der Realisierung dieser Zellen wird durch sogenannte Realisierungsregeln (engl. Realization Rules, RR) ausgedrückt, wie diese in inferentiellen-realisierenden Theorien üblich sind (vgl. z.B. Zwicky 1985, Anderson 1992). Wie Realisierungsregeln genau aussehen, wird im Kapitel 4.1.3.2. erklärt. Zur Verdeutlichung soll hier nur ein Beispiel angefügt werden. Der Genitiv Singular *Tag-es* hat folgende Realisierungsregel (vereinfachte Darstellung): $RR_{\{CASE:GEN, NUM:SG, GEND:MASC\}, N(\langle X, \sigma \rangle)} =_{\text{def}} \langle Xes', \sigma \rangle$.¹³ Die Realisierung einer *Form-Cell* $\langle r, \sigma \rangle$ ist also exakt jene Form, die dadurch definiert ist, dass alle Realisierungsregeln auf r angewendet werden, die das Set morphosyntaktischer Einheiten σ realisieren.

Auch die Relationen zwischen den Zellen des *Content-Paradigm* und jenen des *Form-Paradigm* sind durch Regeln definiert, aber durch sogenannte *Rules of Paradigm Linkage*. Jeder Zelle des *Content-Paradigm* entspricht eine Zelle des *Form-Paradigm*. Diese Zelle des *Form-Paradigm* wird *Form-Correspondent* (FC) von $\langle L, \sigma \rangle$ genannt. Die Realisierung einer *Content-Cell* $\langle L, \sigma \rangle$ besteht also in der Realisierung des FC von $\langle L, \sigma \rangle$, wobei FC von $\langle L, \sigma \rangle$ durch die *Rules of Paradigm Linkage* ermittelt wird.

Es wird eine *Universal Default Rule of Paradigm Linkage* angenommen: „If root r is stipulated as the primary root of a given lexeme L , then the FC of the content-cell $\langle L, \sigma \rangle$ is the form-cell $\langle r, \sigma \rangle$ “ (Ackerman & Stump 2004: 120). Davon weichen insbesondere zwei Fälle ab, für die spezifischere *Rules of Paradigm Linkage* definiert werden müssen. Erstens gibt es Lexeme, die zwei Wurzeln aufweisen (vgl. Tabelle 4.5 *Form-Correspondent*). Das *Content-Paradigm* besteht aus zwei Zellen: $\langle L, \sigma \rangle$ und $\langle L, \sigma' \rangle$. Die Zelle $\langle L, \sigma \rangle$ hat den FC $\langle r, \sigma \rangle$ und die Zelle $\langle L, \sigma' \rangle$ den FC $\langle r', \sigma' \rangle$, wobei $r' \neq r$ (d.h. zwei verschiedene Wurzeln). Z.B. hat das Nomen HRD im Sanskrit zwei heteroklitische Stämme: *hrdaya* und *hrd*, wobei *hrdaya* in den direkten und *hrd* in den indirekten Kasus verwendet wird. Sanskrit braucht also eine *Rule of Paradigm Linkage*, die den Zellen des *Content-Paradigm* mit direktem Kasus einen FC aus dem *Form-Paradigm* mit der Wurzel *hrdaya* zuweist und den Zellen des *Content-Paradigm* mit indirektem Kasus einen FC aus dem *Form-Paradigm* mit der Wurzel *hrd*.

¹³ Vereinfacht ausgedrückt, besagt diese Realisierungsregel, dass einer Wurzel *-es* suffigiert wird.

Tabelle 4.5: Flexion des Nomens HRD in Sanskrit (Ackerman & Stump 2004: 121)

Cells in HRD's content-paradigm	Form-correspondents	Realizations
<HRD, {neut <u>nom</u> sg}>	<hrdaya, {neut <u>nom</u> sg}>	hrdaya-m
<HRD, {neut <u>loc</u> sg}>	<hrd, {neut <u>loc</u> sg}>	hrd-i

Dies kann auf Suppletivstämme übertragen werden. Die Stämme des Verbs *sein* gehen auf unterschiedliche Wurzeln zurück, die im *Form-Paradigm* verortet werden können, wie oben dargestellt wurde (vgl. Tabelle 4.4).

Die zweite Abweichung von der *Universal Default Rule of Paradigm Linkage* betrifft jene Fälle, in denen der FC von <L,σ> <r,σ'> lautet, wobei $\sigma \neq \sigma'$. Ein Beispiel dafür sind die lateinischen Deponentia, deren Aktiv (Bedeutung) mit einem passiven Stamm (Form) gebildet wird (vgl. Tabelle 4.2 und 4.4). Diese Verben brauchen also eine *Rule of Paradigm Linkage*, die jeder Aktivzelle des *Content-Paradigm* einen FC aus dem passiven *Form-Paradigm* zuweist.

Durch dieses Modell von Ackerman & Stump (2004) wird gewährleistet, dass Form und Bedeutung auf der Wortebene voneinander getrennt werden. Die Verbindungen zwischen Form und Bedeutung sind nicht 1:1 im Lexikon gelistet, d.h., die Form eines Wortes wird nicht direkt mit seiner Bedeutung assoziiert, die Bedeutung eines Wortes ist nicht von seiner Form ableitbar. Vielmehr wird die Relation zwischen Form und Bedeutung durch Regeln abgebildet.

Nun ist noch die Frage nach den Verbindungen dieser morphologischen Komponente des Systems mit der f-Struktur und der c-Struktur zu klären. Die f-Struktur eines Lexems wird ausschließlich von den Informationen im *Content-Paradigm* dieses Lexems projiziert. Die Realisierungen hingegen, die mit einer Zelle des *Content-Paradigm* verknüpft sind, bestimmen nicht die f-Struktur, sondern bilden nur den c-strukturellen Ausdruck dieser Zelle. Bezogen auf die Beispiel in Tabelle 4.4 sind es also die *Content-Paradigms* von *machen*, *fare* (2. Singular Indikativ Futur Aktiv) und *loqui* (3. Singular Indikativ Präsens Aktiv), die die f-Struktur projizieren. Die c-Struktur wird jedoch von den Realisierungen gebildet: *wirst machen*, *farai* (2. Singular Indikativ Futur Aktiv), *loquitur* (3. Singular Indikativ Präsens Passiv). Damit ist die Darstellung von Ackerman & Stump (2004: 117–124) abgeschlossen.

4.1.3.2. Realisierungsregeln und Komplexität

Form der Realisierungsregeln: Es soll nun noch beschrieben werden, wie die Realisierungsregeln aussehen, und zwar basierend auf Stump (2001: 40–46 und 50–53), was durch eigene Beispiele aus der deutschen Standardsprache ergänzt wird. In den vorherigen Abschnitten wurde gezeigt, dass die Realisierungsregeln flektierte Wörter ableiten. Formal sind Realisierungsregeln Funktionen, wie in (11) dargestellt:

$$(11) \quad \text{RR}_{n,\tau,C} (\langle X, \sigma \rangle) =_{\text{def}} \langle Y', \sigma \rangle$$

RR steht für Realisierungsregel. Gefolgt wird eine RR von drei Indizes: Der Blockindex n gibt den Block an, zu dem die RR gehört; der Klassenindex C (Class Index) bezeichnet die Klasse an Lexemen, deren Paradigma die RR definiert; der Eigenschaften-Set-Index τ (Property-Set Index) bestimmt das Set an morphosyntaktischen Eigenschaften, die diese Regel realisiert. Weiter besteht die RR aus drei Variablen. Die Variable σ steht für die realisierten morphosyntaktischen Eigenschaften, X für die Wurzel des Lexems und Y' ist eine phonologische Kette und das Resultat der RR, wobei Y' durch weitere RR erweitert werden kann. Z.B. werden für die Wortform *Wäldər* zwei RR benötigt, um den Umlaut (12, *Wäld*) und das Suffix *-ər* (13, *Wäldər*) zu produzieren. Zur Veranschaulichung werden die beiden RR für *Wälder* aufgeführt:

$$(12) \quad \text{RR}_{A, \{\text{NUM:PL}\}, \text{N[IC: } 1 \vee 3 \vee 7 \vee 8]} (\langle X, \sigma \rangle) =_{\text{def}} \langle \ddot{X}', \sigma \rangle$$

$$(13) \quad \text{RR}_{B, \{\text{NUM:PL}\}, \text{N[IC: } 3]} (\langle X, \sigma \rangle) =_{\text{def}} \langle Xər', \sigma \rangle$$

Bedingungen: Bevor die RR weiter referiert wird, muss noch ausgeführt werden, unter welchen Bedingungen eine RR definiert ist. Diese sind in der Bedingung **Rule-Argument Coherence** (14) zusammengefasst:

(14)

„Rule-argument coherence:

$\text{RR}_{n,\tau,C}(\langle X, \sigma \rangle)$ is defined iff (a) σ is an extension of τ ; (b) $\text{L-index}(X) \in C$; and (c) σ is a well-formed set of morphosyntactic properties for $\text{L-index}(X)$ “ (Stump 2001: 45).

Diese Bedingung besteht also aus drei Bedingungen, nämlich Extension und Wohlgeformtheit, die die morphosyntaktischen Eigenschaften betreffen, und **L-index**, womit die Verbindung zwischen einer Wurzel und dem Lexem beschrieben wird. Wenn phonologisch identische Wurzeln unterschiedlichen Lexemen entsprechen, müssen Wurzel und Lexem mit einem Index miteinander verbunden werden. Z.B. haben die Verben MALEN_1 und MAHLEN_2 identische Wurzeln, ihre Paradigmen weisen jedoch einen wesentlichen Unterschied auf: Das Lexem MALEN_1 bildet sein Partizip Perfekt schwach (*gemalt*), das Lexem MAHLEN_2 jedoch stark (*gemahlen*). Findet nun eine RR Anwendung auf eine Wurzel, so erbt Y' den Index von X (15):

(15)

„Persistence of L-indexing:

For any realization rule $RR_{n,\tau,C}$, if $RR_{n,\tau,C}(\langle X, \sigma \rangle) =_{\text{def}} \langle Y', \sigma \rangle$, then $L\text{-index}(Y') = L\text{-index}(X)$ “ (Stump 2001: 45).

Die Bedingungen der Wohlgeformtheit und der Extension betreffen die morphosyntaktischen Eigenschaften. Da die Unifikation eng mit der Wohlgeformtheit und der Extension zusammenhängt, wir auch diese hier beschrieben. **Wohlgeformtheit** ist wie folgt definiert (16), wobei Feature mit F und Value mit v abgekürzt wird:

(16)

„A set τ of morphosyntactic properties for a Lexeme of category C is WELL-FORMED in some language $[l]$ only if τ satisfies the following conditions in $[l]$:

- a. For each property $F:v \in \tau$, $F:v$ is available to lexemes of category C and v is a permissible value for F .
- b. For any morphosyntactic feature F having v_1, v_2 as permissible values, if $v_1 \neq v_2$ and $F:v_1 \in \tau$, then $F:v_2 \notin \tau$ “ (Stump 2001: 41).

Die Bedingung a. verhindert, dass z.B. ein Verb die Eigenschaft $\{\text{CASE:NOM}\}$ und dass das Feature $\{\text{CASE}\}$ den Wert $\{\text{SG}\}$ bekommt. Die Bedingung b. stellt sicher, dass ein Set an morphosyntaktischen Eigenschaften keine sich widersprechende Werte enthält. Z.B. kann ein flektiertes Wort nicht gleichzeitig im Singular und Plural stehen. Ein Set $\{\text{NOM, SG, PL, MASC}\}$ ist also nicht wohlgeformt.

Um die Bedingung der **Extension** zu verstehen, sind noch zwei Arten morphosyntaktischer Features zu unterscheiden: *atom-valued* und *set-valued*. Ein morphosyntaktisches Feature ist *atom-valued*, wenn seine Werte nicht weiter analysiert werden können, z.B. $\{\text{CASE:NOM}\}$. Von einem *set-valued* morphosyntaktischen Feature spricht man, wenn der Wert des Features selbst ein Set an morphosyntaktischen Einheiten ist, z.B. $\{\text{AGR:}\{\text{PERS:1, NUM:SG}\}\}$. Die Extension wird in (17) definiert:

(17)

„Where σ and τ are well-formed sets of morphosyntactic properties, σ is an EXTENSION of τ iff (i) for any atom-valued feature F and any permissible value v for F , if $F:v \in \tau$, then $F:v \in \sigma$; and (ii) for any set-valued feature F and any permissible value ρ for F , if $F:\rho \in \tau$, then $F:\rho' \in \sigma$, where ρ' is an extension of ρ “ (Stump 2001: 41).

Ein Beispiel für ein *set-valued* Feature ist in Stump (2001: 41) zu finden. Da die Nominalphrase des Deutschen nur *atom-valued* Features aufweist, wird das Konzept der Extension anhand

einer möglichen Nominalphrase veranschaulicht: {CASE:NOM, NUM:SG, GEND:F} ist eine Extension der folgenden Sets an morphosyntaktischen Eigenschaften:

(18)

- a. {CASE:NOM, NUM:SG, GEND:FEM}
- b. {CASE:NOM, NUM:SG}
- c. {NUM:SG, GEND:FEM}
- d. {CASE:NOM, GEND:FEM}
- e. {CASE:NOM}
- f. {NUM:SG}
- g. {GEND:FEM}

Unifikation ist ein zentrales Konzept in LFG, aber auch in anderen Feature basierten Modellen (Falk 2001: 17). Bezogen auf die RR wird Unifikation wie folgt definiert (19):

(19)

„Where σ and τ are well-formed sets of morphosyntactic properties, the UNIFICATION ρ of σ and τ is the smallest well-formed set of morphosyntactic properties such that ρ is an extension of both σ and τ “ (Stump 2001: 41).

Beispielsweise ist die Unifikation von {CASE:NOM, NUM:SG} und {GEND:FEM} definiert als {CASE:NOM, NUM:SG, GEND:FEM}, wohingegen die Unifikation von {CASE:NOM, NUM:SG} und {NUM:PL} nicht definiert ist. Die Idee der Unifikation ist von großer Bedeutung, wenn mehrere RR angewendet werden. Z.B. das Wort *Wäldärn* (Dativ Plural) wird von drei RR gebildet, nämlich von den Regeln in (12) und (13), die hier wiederholt werden, sowie von der Regel in (22).¹⁴ Die Unifikation der Sets an morphosyntaktischen Eigenschaften der Regeln (12), (13) und (22) lautet {CASE:DAT, NUM:PL}.

$$(12) \quad \text{RR}_A, \{ \text{NUM:PL} \}, \text{N[IC: } 1 \vee 3 \vee 7 \vee 8] \} (<X, \sigma>) =_{\text{def}} <\ddot{X}', \sigma>$$

$$(13) \quad \text{RR}_B, \{ \text{NUM:PL} \}, \text{N[IC: } 3] \} (<X, \sigma>) =_{\text{def}} <X\partial r', \sigma>$$

$$(20) \quad \text{RR}_C, \{ \text{CASE:DAT, NUM:PL} \}, \text{N[IC: } 1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6 \vee 7 \vee 8] \} (<X, \sigma>) =_{\text{def}} <X\partial n', \sigma>$$

¹⁴ $-\partial n/-n$ -Variation ist phonotaktisch bedingt und muss deswegen in der Morphologie nicht berücksichtigt werden.

Blöcke: Nach diesen Ausführungen zu den Bedingungen der RR, sollen nun die drei Indizes (n, τ , C) der RR näher erörtert werden. Zur Erinnerung ist die Form einer RR in (11) wiederholt.

$$(11) \quad RR_{n,\tau,C} (<X,\sigma>) =_{\text{def}} <Y',\sigma>$$

Basierend auf Anderson (1992) geht auch Stump (2001) davon aus, dass die RR in Blöcken organisiert sind. Die Blöcke bestimmen, in welcher Reihenfolge die RR angewendet werden, wozu jeder Block einen Index A–Z bekommt: Z.B. werden zuerst die Regeln des Blocks A angewendet, dann jene des Blocks B etc. Innerhalb eines Blocks konkurrieren die RRs folglich um die gleiche Position in der Abfolge der Regeln (RRs unterschiedlicher Blöcke konkurrieren nicht miteinander). Für die Substantivflexion der deutschen Standardsprache sind folgende Blöcke denkbar: Block A beinhaltet die Regeln zur Modifikation der Wurzel, Block B die Regeln zum Plural, Block C die Regeln zum Kasus (vgl. Regeln (12), (13) und (20)). Da die deutsche Standardsprache unterschiedliche Plural-Suffixe aufweist, befinden sich alle RRs dieser Pluralsuffixe in Block C und konkurrieren miteinander. Die Regeln für die Pluralsuffixe konkurrieren jedoch nicht mit den Regeln für den Umlaut, da diese nicht im selben Block sind. Dadurch wird der Tatsache Rechnung getragen, dass der Plural nur durch den Umlaut oder nur durch ein Suffix oder durch eine Kombination von beiden ausgedrückt werden kann, aber nie durch zwei Umlaute oder zwei Suffixe.

Zur Konkurrenz der RR stellen sich nun noch zwei Fragen: a) Was passiert, wenn mehrere Regeln auf die gleichen morphosyntaktischen Eigenschaften zutreffen? b) Was passiert, wenn keine Regel aus einem bestimmten Block angewendet wird?

Frage a) kann mit **Pāṇinis Prinzip** beantwortet werden, das besagt, dass immer zuerst die engste oder spezifischste Regel für ein bestimmtes Set an morphosyntaktischen Einheiten angewendet wird. Z.B. weist der Dativ Singular der starken Adjektivflexion in der deutschen Standardsprache zwei Suffixe auf: *-am* im Maskulin und im Neutrum, *-ar* im Feminin. Dazu können zwei RR ((21) und (22)) gebildet werden:

$$(21) \quad RR_{A, \{CASE:DAT, NUM:SG\}, ADJ[STRONG]} (<X,\sigma>) =_{\text{def}} <Xam',\sigma>$$

$$(22) \quad RR_{A, \{CASE:DAT, NUM:SG, GEND:FEM\}, ADJ[STRONG]} (<X,\sigma>) =_{\text{def}} <Xar',\sigma>$$

$$(23) \quad RR_{A, \{CASE:ACC, NUM:SG, GEND:MASC\}, ADJ[]} (<X,\sigma>) =_{\text{def}} <Xan',\sigma>$$

Da die RR (22) spezifischer ist als die RR (21), wird sie zuerst angewendet und blockiert die Anwendung einer weiteren RR (z.B. der RR (21)) aus demselben Block. Somit wird für den Dativ Singular Feminin die RR (22) verwendet und danach die RR (21) für die übrigen gebliebenen Positionen des Dativ Singular, und zwar für das Maskulin und Neutrum. Dieses Beispiel zeigt auch, dass die RRs im Gegensatz zum *Content-Paradigm* unterspezifiziert

bleiben können. Im *Content-Paradigm* muss das Set an morphosyntaktischen Eigenschaften komplett sein, d.h. z.B. für ein Adjektiv, dass die Features Kasus, Numerus und Genus spezifiziert sein müssen. In den RRs ist dies nicht nötig, hier können Features auch unterspezifiziert bleiben. In den RRs trifft dies nicht nur auf die morphosyntaktischen Eigenschaften zu, sondern auch auf die Wortklasse. Z.B. wird im Akkusativ Singular Maskulin der starken und schwachen Adjektivflexion *-ən* suffigiert (23). Im Gegensatz zu den RRs in (21) und (22) muss folglich die Art der Flexion nicht angegeben werden, die eckige Klammer nach ADJ bleibt also leer. Stump (2001) definiert die spezifischste RR, wie in (24) wiedergegeben wird. Dabei bezieht sich a. auf die morphosyntaktischen Eigenschaften und b. auf die Wortklasse. Die Anwendung einer RR ist in (25) definiert:

(24)

- „a. $RR_{n,\sigma,C}$ is NARROWER than $RR_{n,\tau,C}$ iff σ is an extension of τ and $\sigma \neq \tau$.
- b. Where $C \neq C'$, $RR_{n,\sigma,C}$ is NARROWER than $RR_{n,\tau,C}$ iff $C \subseteq C'$ “ (Stump 2001: 52)

(25)

„ $RR_{n,\tau,C}$ is APPLICABLE TO $\langle X, \sigma \rangle$ iff $RR_{n,\tau,C}(\langle X, \sigma \rangle)$ is defined (according to [(24)])“ (Stump 2001: 52).

Die Frage b), was geschieht, wenn keine Regel aus einem bestimmten Block angewendet wird, kann dadurch beantwortet werden, dass die Wortform nicht verändert wird. Beispielsweise zeichnet sich die Flexionsklasse 2¹⁵ der standarddeutschen Substantive dadurch aus, dass sie keinen Umlaut aufweist (z.B. *Tag – Tagə*). Ist nun die Wurzel *tag* mit den morphosyntaktischen Eigenschaften {CASE:NOM, NUM:PL} verbunden, findet sich in Block B keine RR für die Flexionsklasse 2 (vgl. RR in (12)), folglich passiert mit der Wurzel nichts. Stump (2001) nimmt für diese Fälle die RR ***Identity Function Default*** an:

(26)

„Identity Function Default [...]
 $RR_{n,\{\},U}(\langle X, \sigma \rangle) =_{\text{def}} \langle X, \sigma \rangle$ “ (Stump 2001: 53).

Sie betrifft alle Blocks und alle Lexeme einer Sprache und die morphosyntaktischen Eigenschaften sind maximal unterspezifiziert, d.h. \emptyset . Diese RR definiert, dass die Wurzel nicht verändert wird. Es handelt sich dabei also um die allgemeinste RR, die nach Pāṇinis Prinzip

¹⁵ Zur Definition und Einteilung der Flexionsklassen siehe Kapitel 5.1.1.

immer dann angewendet wird, wenn sonst in einem Block keine RR für eine bestimmte morphosyntaktische Eigenschaft vorhanden ist.

Realisierungsregeln und Komplexität: Damit möchte ich die Darstellung der RRs basierend auf Stump (2001: 40–46 und 50–53) schließen. Bis hier wurde gezeigt, dass ein Lexem aus einem *Content-Paradigm*, einem *Form-Paradigm* und einer realisierten Wortform besteht. Die Relationen zwischen *Content-Paradigm* und *Form-Paradigm* werden durch *Rules of Paradigm Linkage* ausgedrückt, die Relationen zwischen *Form-Paradigm* und der Realisierung durch RRs (vgl. Tabelle 4.3 in Kap. 4.1.3.1). Da für die Modifikation einer Wurzel zu einem flektierten Wort die RRs zuständig sind, kann folglich die Flexionsmorphologie im System der RRs verortet werden. Daraus folgt, dass, wenn man die Komplexität der Flexionsmorphologie berechnen möchte, die RRs den Ort der Messung bilden. Im Kapitel 3.1.2. wurde erörtert, dass die absolute Komplexität eines linguistischen Phänomens durch die Länge der Beschreibung dieses Phänomens gemessen werden kann (Miestamo 2008: 24). Je kürzer diese Beschreibung ist, desto weniger komplex ist das linguistische Phänomen. Übertragen auf RRs heißt das Folgendes: Da die RRs die Beschreibung der Flexionsmorphologie darstellen, ist jenes Flexionssystem weniger komplex, das weniger RRs aufweist. Es wird hier also mit einem formalen Instrument die Komplexität der Flexionsmorphologie gemessen. Dies und besonders die Vorteile dieser Methode werden im Kapitel 4.3.2. ausgeführt.

4.1.3.3. Variation, Synkretismus und Wurzelalternation

Drei Fragen müssen noch beantwortet werden: 1. Wie wird (freie) Variation modelliert?; 2. Wie werden Synkretismen durch die RRs erfasst?; 3. Wie werden Stämme von Wurzeln abgeleitet und wie wird die Verteilung der Stämme erfasst?

Variation: Bis jetzt wurden nur Beispiele gezeigt, in denen keine Variation vorkommt, d.h., dass eine Zelle eines Paradigmas nur eine Wortform enthält. Jedoch gerade die Nicht-Standardvarietäten des in dieser Arbeit untersuchten Samples weisen immer wieder Variation auf. Tabelle 4.6 zeigt die Flexion des bestimmten Artikels in Jaun (Höchstalemannisch). Hier sind drei Arten von Variation zu unterscheiden: 1. phonologisch bedingte Variation, 2. syntaktisch bedingte Variation, 3. freie Variation. Die phonologisch bedingte Variation betrifft die Formen *əm* und *m* im Dativ Singular Maskulin, welche nur nach einer Präposition vorkommen. Endet die Präposition auf einen Vokal, folgt die Form *m*, endet sie auf einen Konsonanten, folgt *əm*. Die Variation ist also rein phonologisch, d.h., das Subsystem Phonologie steuert diese Variation und diese braucht deswegen hier nicht weiter berücksichtigt zu werden. Das Paradigma (vgl. Tabelle 4.6) enthält also nur die Formen *dəm* und *əm*. Syntaktisch bedingt sind alle Varianten dieses Paradigmas. Die Artikel *ə/əm* (Akkusativ/Dativ) treten nach einer Präposition auf, die Artikel *dər/dəm* (Akkusativ/Dativ), wenn ihnen keine Präposition vorangeht. Der Artikel *di* wird vor einem Adjektiv verwendet, der Artikel *t* vor einem Nomen. Auch wenn die Verteilung syntaktisch gesteuert ist, müssen die unterschiedlichen Formen, die die Syntax benutzt, von irgendwo herkommen, und zwar von der Morphologie. Freie Variation findet sich im Dativ Singular Maskulin: Geht dem Artikel keine Präposition voraus, kann sowohl *dəm* als auch *əm* verwendet werden. Sowohl die

syntaktisch bedingte als auch die freie Variation sind innerhalb der Morphologie als freie Variation anzusehen, d.h., sie müssen durch RRs produziert werden.

Tabelle 4.6: Flexion des bestimmten Artikels in Jaun (Stucki 1917: 282)

			Nominativ	Akkusativ	Dativ	Genitiv
Singular	mask	Artikel	dər	dər	dəm, əm	ts
		P + Artikel		ə	əm, (m)	
Singular	neut	Artikel	ts	ts	dəm, əm	ts
		P + Artikel			əm, (m)	
Singular	fem	Artikel + Adj	di	di	dər	dər
		Artikel + N	t	t		
Plural		Artikel + Adj	di	di	də	dər
		Artikel + N	t	t		

Diese Tatsache muss das System an RRs abbilden. Dies bedeutet, dass es in einem Block für dieselbe Wortklasse zwei RRs mit denselben morphosyntaktischen Eigenschaften gibt:

$$(27) \quad \text{RR}_{A, \{ \text{CASE:DAT, NUM:SG, GEND:MASC} \vee \text{NEUT} \}, \text{ART[DEF]}} (<X, \sigma>) =_{\text{def}} <Xdəm', \sigma>$$

$$(28) \quad \text{RR}_{A, \{ \text{CASE:DAT, NUM:SG, GEND:MASC} \vee \text{NEUT} \}, \text{ART[DEF]}} (<X, \sigma>) =_{\text{def}} <Xəm', \sigma>$$

Genau dies ist aber laut Stump (2001) nicht möglich, denn in jedem Block gibt es eine RR, die am spezifischsten ist. Die spezifischste RR ist nicht definiert, „if block *n* contained two or more applicable rules no one of which was narrower than all the others“ (Stump 2001: 53). Des Weiteren war keine Stelle in Stump (2001) zu finden, in der Variation modelliert wird. Die Frage, die sich hier auftut, ist eine viel grundsätzlichere: Wenn zwei Varianten vorhanden sind, bilden diese den Output einer oder zwei Grammatiken? Traditionell wird in der generativen Grammatik davon ausgegangen, dass Variation und Optionalität auf die Präsenz unterschiedlicher Grammatiken zurückgeführt werden kann (Seiler 2004: 385), d.h. also, interne Variation und Diglossie werden gleich behandelt (Seiler 2004: 384). Dass dies aus prinzipiellen und empirischen Gründen nur wenig Sinn macht, dafür argumentiert Seiler (2004), wobei die für diese Arbeit wichtigsten Gründe zitiert werden:

First, diglossia and grammar-internal variation are sociolinguistically different facts, and we need a formal instrument in order to account for this difference. Second, we need a ‘meta-grammar’ telling us the right arrangement of the two co-present grammars. Third, if two co-present grammars are needed in order to account for one variable feature (i.e., one feature with two options), the number of parallel grammars exponentially increases with every variable feature – which is a highly undesired result. (Seiler 2004: 384)

Geht man davon aus, dass die Artikel *dəm* und *əm* die Outputs von zwei verschiedenen Grammatiken darstellen, bräuchte es für jede dieser Varianten wie auch für jede weitere Variante ein zusätzliches Paradigma. Dies würde zu einer sehr hohen Zahl an Paradigmen führen. Die Morphologie kann ökonomischer, d.h., mit weniger Mitteln modelliert werden, wenn in einer Grammatik Varianten erlaubt sind.

Dazu wird hier eine Modifikation von Stumps (2001) Regeln vorgeschlagen, die die Anwendung der spezifischsten RR betrifft. Die Definition der spezifischsten RR (24) und ihrer Anwendung (25) nach Stump (2001) sind hier wiederholt:

(24)

- „a. $RR_{n,\sigma,C}$ is NARROWER than $RR_{n,\tau,C}$ iff σ is an extension of τ and $\sigma \neq \tau$.
- b. Where $C \neq C'$, $RR_{n,\sigma,C}$ is NARROWER than $RR_{n,\tau,C}$ iff $C \subseteq C'$ “ (Stump 2001: 52)

(25)

„ $RR_{n,\tau,C}$ is APPLICABLE to $\langle X, \sigma \rangle$ iff $RR_{n,\tau,C}(\langle X, \sigma \rangle)$ is defined (according to [(24)])“ Stump 2001: 52).

(25) impliziert, dass immer nur eine RR appliziert werden kann. Diese Beschränkung muss also dahingehend erweitert werden, dass alle RRs, die am spezifischsten sind (also der Definition in (24) entsprechen), angewendet werden, egal ob es sich dabei um eine oder mehrere RRs handelt. In (29) wird also eine Erweiterung von (25) vorgeschlagen:

(29)

Every $RR_{n,\tau,C}$ is APPLICABLE to $\langle X, \sigma \rangle$ iff $RR_{n,\tau,C}(\langle X, \sigma \rangle)$ is defined (according to (24)).

Durch diese Erweiterung können sowohl die RR (27, *dəm*) als auch die RR (28, *əm*) angewendet werden und füllen bzw. definieren die Zelle Dativ Singular Maskulin und Neutrum des bestimmten Artikels.

Synkretismus: Eine weitere Frage ist, wie mit Synkretismen umgegangen wird. Stump (2001) unterscheidet vier Arten von Synkretismen, wobei diese mit verschiedenen Typen von Regeln definiert werden, was kurz vorgestellt wird. Anschließend wird gezeigt, weshalb dies zur Komplexitätsmessung dieses Samples nicht übernommen werden kann. Dafür wird ein Vorschlag gemacht, wie die Synkretismen in den Varietäten dieses Samples einheitlich erfasst und gemessen werden können.

Synkretismus bei Stump (2001): Stump (2001) unterscheidet vier Arten von Synkretismen, die teils u.a. an Zwickys (1985) Unterscheidung zwischen systematischem und zufälligem Synkretismus erinnern. Stumps vier Arten von Synkretismen und ihre Typen von Regeln sind in Tabelle 4.7 zusammengefasst.

Tabelle 4.7: Typen von Synkretismen basierend auf Stump (2001: 212–217)

Synkretismus				
	direktional		nicht direktional	
	unidirektional	bidirektional	symmetrisch	nicht stipuliert
Abhängigkeit	A=Determ, B=Dep	A=Determ, B=Dep B=Determ, A=Dep	–	
teilen natürl. Klasse	unwichtig		–	+
Regeltyp	rules of referral		metarule	RR

Ein direktonaler Synkretismus ist dann gegeben, wenn eine Form nach dem Vorbild einer anderen gebildet wird, wobei ersterer das abhängige (*dependent*) Mitglied und letzterer das determinierende (*determinant*) Mitglied des Synkretismuspaares bilden (Stump 2001: 213). Stump (2001) illustriert dies am folgenden Beispiel. Im Bulgarischen weisen die Verben der 3. Person Singular unabhängig vom Tempus ein *–e* als Suffix auf. In der 2. Person Singular wird ebenfalls ein *–e* suffigiert, jedoch nur in den präteritalen Tempora, d.h. im Imperfekt und Aorist. Die 3. Person Singular bildet also das determinierende Mitglied des Synkretismuspaares, die 2. Person Singular das abhängige Mitglied (Stump 2001: 212–213). Direktionale Synkretismen werden durch eine *Rule of Referral* definiert. In (30) ist die *Rule of Referral* für das bulgarische Beispiel aufgeführt:

(30)

Rule of referral:

„Where *n* is any of rule Blocks **A** to **D**,

$RR_n \leftarrow \{PRET:yes, AGR:\{PERS:2, NUM:sg\}\} \rightarrow, v(\langle X, \sigma \rangle) =_{def} \langle Y, \sigma \rangle$, where

$Nar_n(\langle X, \sigma / \{AGR:\{PERS:3\}\} \rangle) = \langle Y, \sigma / \{AGR:\{PERS:3\}\} \rangle$ “ (Stump 2001: 218).

Bei dieser Regel handelt es sich um einen unidirektionalen Synkretismus, d.h. für dieses Beispiel, dass immer die 2. Person Singular nach der 3. Person Singular gebildet wird und nicht umgekehrt. Bei einem bidirektionalen Synkretismus kommen beide Richtungen vor. Z.B. im Rumänischen (Tabelle 4.8) weisen die Verben aller Konjugationsklasse mit Ausnahme der ersten Konjugationsklasse einen Synkretismus zwischen der 1. Person Singular und der 3. Person Plural auf. Dabei handelt es sich um einen bidirektionalen Synkretismus. Im Paradigma von *umplea* bildet die 3. Person Plural die abhängige und die 1. Person Singular die determinierende Form, da das Suffix *–u* in der ersten Konjugationsklasse nur in der 1. Person Singular vorkommt. Im Paradigma des Verbs *fi* gilt der umgekehrte Fall: Die 1. Person Singular

ist die abhängige Form, da sie auf der Basis des Pluralstamms gebildet wird (Stump 2001: 213). Der bidirektionale Synkretismus wird ebenfalls mit einer *Rule of Referral* ausgedrückt, die komplexer ausfällt als die Regel in (30) (vgl. Stump 2001: 219). Für die Belange dieser Arbeit braucht sie aber nicht weiter ausgeführt zu werden.

Tabelle 4.8: Präsens Indikativ rumänischer Verben aus (Stump 2001: 214, hier gekürzt)

	<i>a invita</i> einladen	<i>a sufla</i> atmen	<i>a umplea</i> füllen	<i>a fi</i> sein
Konjugationsklasse	1	1	2	4
1.Sg.	invít	súfl-u	úmpl-u	sínt
2.Sg.	invít-i	súfl-i	úmpl-i	ést-i
3.Sg.	invít-ă	súfl-ă	úmpl-e	ést-e
1.Pl.	invită-m	suflă-m	úple-m	sínte-m
2.Pl.	invită-ți	suflă-ți	úple-ți	sínte-ți
3.Pl.	invít-ă	súfl-ă	úmpl-u	sínt

Bei den nicht direktionalen Synkretismen können keine determinierenden und abhängigen Formen ausgemacht werden, es gibt keine Default-Form. Sie werden weiter in nicht-stipulierte und symmetrische Synkretismen eingeteilt. Als Beispiel für den nicht-stipulierten Synkretismus nennt Stump (2001) die 3. Person Singular und Plural der ersten Konjugationsklasse im Präsens Indikativ des Rumänischen (vgl. Tabelle 4.8). Erstens bildet das Set an morphosyntaktischen Einheiten eine natürlich Klasse, nämlich 3. Person Präsens Indikativ (Stump 2001: 213). Zweitens entstehen die beiden Formen durch eine Regel, die –ă in der 3. Person suffigiert, die aber für Numerus unterspezifiziert ist (Stump 2001: 213–215). Es handelt sich also um eine Arte „poverty in the system of realization rules [...] the system happens not to have any rule that is sensitive to number in the inflection of third-person present indicative forms of first-conjugation verbs“ (Stump 2001: 215). Diese Art von Synkretismus wird also durch eine für Numerus unterspezifizierte RR wie in (11) definiert.

Der symmetrische Synkretismus zeichnet sich schließlich dadurch aus, dass weder eine Richtung ausgemacht werden kann, noch bilden die Sets an morphosyntaktischen Eigenschaften eine natürliche Klasse (Stump 2001: 216). Stump (2001) veranschaulicht dies am Beispiel der 2. Person Singular und der 1. Person Plural der Verben in Hua (Sprache von Neu Guinea). Die beiden Formen fallen unabhängig vom Modus immer zusammen, es ist aber weder eine Richtung zu identifizieren, noch formen die Sets an morphosyntaktischen Eigenschaften eine natürliche Klasse (Stump 2001: 216–217). Der symmetrische Synkretismus kann also nicht durch eine *Rule of Referral* und auch nicht durch eine RR nach (11) ausgedrückt werden. Um den symmetrischen Synkretismus zu definieren, schlägt Stump die *Symmetrical Syncretism Metarule* vor (Stump 2001: 222). Auf das diskutierte Beispiel bezogen sieht diese Regel wie folgt aus:

(31) Symmetrical Syncretism Metarule

„Where τ is an extension of $\{AGR(su): \{PER:2, NUM:sg\}\}$,

$$RR_{II,\tau}(<X,\sigma>) =_{def} <Y,\sigma>$$

\updownarrow

$$RR_{II,\tau/\{AGR(su): \{PER:1, NUM:pl\}\}}(<X,\sigma>) =_{def} <Y,\sigma>“ \text{ (Stump 2001: 223).}$$

Wie die *Rule of Referral* muss auch die *Symmetrical Syncretism Metarule* nicht näher besprochen werden. Für beide Regeln ist aber festzustellen, dass sie von der RR wie in (11) abweichen, folglich einen anderen Typ von Regeln darstellen und dass sie komplexer sind als die RRs. Möchte man also mit diesen Regeln die Komplexität der Flexion messen, stellt sich sehr schnell das Problem der Vergleichbarkeit. Können verschiedene Typen von Regeln miteinander verrechnet werden, und wenn ja, wie? Wenn, wie bereits konstatiert wurde, die *Rules of Referral* und die *Symmetrical Syncretism Metarules* in ihrer Form deutlich komplexer sind als die RRs, wie viel komplexer als die RRs sind sie? Sind die *Rules of Referral* und die *Symmetrical Syncretism Metarules* gleich komplex? In den hier vorgestellten Beispielen fallen immer zwei Zellen des Paradigmas zusammen. Je mehr Zellen jedoch zusammenfallen, desto umfangreicher werden die *Rules of Referral* und die *Symmetrical Syncretism Metarules*. Bildet eine komplexere Regel auch einen komplexeren Sachverhalt ab? Es ist kontraintuitiv zu sagen, dass, wenn z.B. drei von vier Zellen eines Paradigmas zusammenfallen, dies komplexer ist, als wenn zwei von vier Zellen zusammenfallen. Es zeigt sich hier, dass, obwohl Stumps Analyse sehr reizvoll erscheint, sie zur Komplexitätsmessung nicht verwendet werden kann. Zur Komplexitätsmessung ist es also fundamental wichtig, ein einheitliches 'Messinstrument' zu benutzen, damit die Objektivität, Reliabilität und besonders die Validität gewährleistet werden können.

Im vorangehenden Kapitel 4.1.3.2. wurde argumentiert, dass die Komplexität durch die Anzahl RRs gemessen werden kann. Würde man dies erweitern, folglich durch unterschiedliche Regeln (RR, *Rules of Referral* etc.), die Komplexität messen und würde man annehmen, dass alle Regeln unabhängig von ihrem Typ gleich komplex sind, wäre ein Synkretismus genauso komplex wie wenn zwei unterschiedliche Suffixe vorkommen würden. Man vergleiche dazu die starke Adjektivflexion in der deutschen Standardsprache (Tabelle 4.12 unten). Das Maskulin unterscheiden zwei Formen für den Nominativ ($-\bar{a}r$) und für den Akkusativ ($-\bar{a}n$). Das Neutrum weist für die beiden Kasus ein Suffix auf ($-\bar{a}s$). Für das Maskulin bräuchte man also zwei RRs, für das Neutrum eine RR für den Nominativ und eine Synkretismusregel für den Akkusativ. Damit würden Maskulin und Neutrum die gleiche Komplexität aufweisen, obwohl das Maskulin zwei unterschiedliche Suffixe hat und das Neutrum nur eines. Es muss also eine Messmethode gefunden werden, die genau diese Beobachtung abbildet.

Synkretismus in dieser Arbeit: Da zur Komplexitätsmessung, wie bereits argumentiert wurde, eine einheitliche Methode verwendet werden muss, soll die RR so erweitert werden, dass sie auch Synkretismen abbilden kann. Dazu wird hier das Einführen eines Junktors vorgeschlagen, nämlich der ausschließenden Disjunktion (auch Kontravalenz genannt). Es handelt sich dabei um die Entweder-Oder-Verbindung, d.h., die ausschließende Disjunktion ist wahr, wenn entweder die eine Aussage oder die andere wahr ist, aber nicht, wenn beide Aussagen wahr

oder falsch sind (Zoglauer 2008: 38). Als Operator wird in dieser Arbeit das Symbol \vee verwendet. Dadurch können der Nominativ und Akkusativ Maskulin und Neutrum der starken Adjektivflexion adäquat abgebildet werden, d.h. zwei RRs für das Maskulin ($-\bar{a}r$, $-\bar{a}n$) ((32)–(33)) und eine RR für das Neutrum ($-\bar{a}s$) ((34)):

$$(32) \quad \text{RR}_{A, \{ \text{CASE:NOM, NUM:SG} \}, \text{ADJ[STRONG]}} (<X, \sigma>) =_{\text{def}} <X\bar{a}r', \sigma>$$

$$(33) \quad \text{RR}_{A, \{ \text{CASE:ACC, NUM:SG} \}, \text{ADJ[STRONG]}} (<X, \sigma>) =_{\text{def}} <X\bar{a}n', \sigma>$$

$$(34) \quad \text{RR}_{A, \{ \text{CASE:NOM} \vee \text{ACC, NUM:SG} \}, \text{ADJ[STRONG]}} (<X, \sigma>) =_{\text{def}} <X\bar{a}s', \sigma>$$

Es stellt sich nun die Frage, ob das vorgestellte Modell von Ackerman & Stump (2004) mit der Einführung einer ausschließenden Disjunktion vereinbart werden kann. Ackerman & Stump (2004) plädieren dafür, Form und Bedeutung klar voneinander zu trennen. Wie bereits erörtert wurde, stellen dann beispielsweise lateinische Deponentia kein Problem mehr dar (Tabelle 4.9). Auch wenn sie eine passivische Form haben (FC und Realisierung), weisen sie eine aktivische Bedeutung auf (C-P). Genauso problemlos fügen sich Synkretismen in dieses Modell (Tabelle 4.10): Auf der Bedeutungsebene macht es einen Unterschied, ob ein Wort im Nominativ oder Akkusativ steht (C-P), auf der Ebene der Form hingegen können eben diese Formen auch zusammenfallen (FC und Realisierung). Es soll hier kurz daran erinnert werden, dass nur von den Informationen des C-P aus die f-Struktur projiziert wird, während die Realisierungen lediglich den c-strukturellen Ausdruck bilden.

Tabelle 4.9: Flexion der lateinischen Verben *fatēri* und *monēre*, aus Ackerman & Stump (2004: 122)

Cells in content-paradigm	Form-correspondents	Realizations
<FATĒRĪ, {1 sg pres <u>act</u> indic}>	<fat, {1 sg pres <u>pass</u> indic}>	fateor
MONĒRE, {1 sg pres <u>pass</u> indic}>	<mon, {1 sg pres <u>pass</u> indic}>	moneor

Tabelle 4.10: C-P, FC und Realisierung des Nominativs und Akkusativs Maskulin und Neutrum

Cells in content-paradigm	Form-correspondents	Realizations
<SCHÖN, {nom sg m}>	<schön, {nom sg m}>	schöner
<SCHÖN, {acc sg m}>	<schön, {acc sg m}>	schönen
<SCHÖN, {nom sg n}>	<schön, {nom \vee acc sg n}>	schönes
<SCHÖN, {acc sg n}>		

Die Herausforderung liegt vielmehr in der Vereinbarkeit mit der Bedingung der Wohlgeformtheit. Es wurde gezeigt, dass die RRs nur unter bestimmten Bedingungen definiert

sind: *Persistence of L-indexing* (15), Wohlgeformtheit (16), Extension (17) und Unifikation (19). Die Bedingung der Wohlgeformtheit ist hier in gekürzter Version wiederholt:

(16)

„A set τ of morphosyntactic properties for a Lexeme of category C is WELL-FORMED in some language [l] only if τ satisfies the following conditions in [l]:

[...]

- b. For any morphosyntactic feature F having v_1, v_2 as permissible values, if $v_1 \neq v_2$ and $F:v_1 \in \tau$, then $F:v_2 \notin \tau$ “ (Stump 2001: 41).

Auf den ersten Blick scheint also eine RR wie in (34) nicht wohlgeformt zu sein. Nimmt man aber die Trennung von Bedeutung und Form ernst, ist die vorgeschlagene Erweiterung der RRs durch eine ausschließende Disjunktion mit der Bedingung der Wohlgeformtheit kompatibel. Erstens bezieht sich die Wohlgeformtheit primär auf Lexeme, d.h. auf das C-P, von dem aus die f-Struktur projiziert wird. Die Bedingung der Wohlgeformtheit ist folglich besonders wichtig, denn sie verhindert, dass im Satz eine Konstituente z.B. gleichzeitig im Nominativ und im Akkusativ steht, was einen Satz uninterpretierbar macht. Die RR dagegen stellt nur eine Form zur Verfügung. Zweitens lässt die ausschließende Disjunktion offen, ob es sich z.B. um einen Nominativ oder Akkusativ handelt. Sie beschränkt nur, dass eine bestimmte Form entweder im Nominativ oder im Akkusativ, aber in keinem anderen Kasus (z.B. Dativ) stehen kann. Ob eine bestimmte Form im Satz schließlich im Nominativ oder Akkusativ steht, hängt von der Valenz des Verbs, von der f-Struktur und dem C-P ab. Einfach ausgedrückt auf die RR (34) übertragen heißt das, dass die Form *schönas* verwendet werden kann, wenn in einem Satz ein Nominativ oder ein Akkusativ, aber nicht wenn ein Dativ gebraucht wird.

In der Folge werden zwei weitere wichtige Vorteile der Implementierung der ausschließenden Disjunktion vorgestellt. Erstens wurde bezüglich der *Rules of Referral* und der *Symmetrical Syncretism Metarules* darauf hingewiesen, dass diese Typen von Regeln umfangreicher werden, je mehr Zellen eines Paradigmas zusammenfallen. Dies impliziert, dass der Zusammenfall von drei Zellen ein System komplexer macht als der Zusammenfall von zwei Zellen. Dies ist kontraintuitiv, denn je mehr Zellen zusammenfallen, desto weniger unterschiedliche Formen weist ein Paradigma auf und desto kürzer sollte also die Beschreibung des Systems ausfallen. Dieses Problem kann vermieden werden, indem Synkretismen mit RRs abgebildet werden, welche eine ausschließende Disjunktion zulassen. Als Beispiel dient die Flexionsklasse 4 der standarddeutschen Substantivflexion (Tabelle 4.11, s.u.). Im Singular fallen Akkusativ, Dativ und Genitiv zusammen, was mit einer RR definiert werden kann:

(35) $RR_C, \{CASE:ACC \vee DAT \vee GEN, NUM:SG\}, N[IC:4] \langle X, \sigma \rangle =_{def} \langle X_{\partial n'}, \sigma \rangle$

Es ist also egal, ob von vier Formen zwei oder drei zusammenfallen. In beiden Fällen kann derselbe Typ von RR verwendet werden, womit die einzelnen RRs miteinander vergleichbar sind und zur Komplexitätsmessung miteinander verrechnet werden können. Auf die gleiche Art und Weise können auch Genusynkretismen abgebildet. Fallen Singular und Plural zusammen, bleibt der Numerus unterspezifiziert.

Zweitens ist zu beobachten, dass dasselbe Suffix über die Flexionsklassen hinweg wiederverwendet wird. Beispielsweise wird in der Substantivflexion der deutschen Standardsprache *-as* als Genitiv Singular und *-ən* als Dativ Plural in mehreren Flexionsklassen benutzt (Tabelle 4.11). Angenommen jede Flexionsklasse weist im Dativ Plural ein anderes Suffix auf, würde die Beschreibung jenes Systems länger ausfallen als die Beschreibung des Systems der deutschen Standardsprache. Auch dies kann durch die folgende RR adäquat abgebildet werden, welche bereits weiter oben eingeführt wurde:

$$(20) \quad \text{RR}_{C, \{ \text{CASE:DAT, NUM:PL} \}, \text{N[IC: } 1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6 \vee 7 \vee 8 \text{]}} (<X, \sigma>) =_{\text{def}} <X_{\text{ən}}, \sigma>$$

Tabelle 4.11: Flexion der Substantive in der deutschen Standardsprache basierend auf Eisenberg (2006: 158–167)

	Singular				Plural			
Flexionsklasse	Nom	Akk	Dat	Gen	Nom	Akk	Dat	Gen
1	gast	gast	gast	gast-əs	gäst-ə	gäst-ə	gäst-ən	gäst-ə
2	tag	tag	tag	tag-əs	tag-ə	tag-ə	tag-ən	tag-ə
3	wald	wald	wald	wald-əs	wäld-ər	wäld-ər	wäld-ər-n	wäld-ər
4	matrosə	matrosə-n	matrosə-n	matrosə-n	matrosə-n	matrosə-n	matrosə-n	matrosə-n
5	staat	staat	staat	staat-s	staat-ən	staat-ən	staat-ən	staat-ən
6	blumə	blumə	blumə	blumə	blumə-n	blumə-n	blumə-n	blumə-n
7	stadt	stadt	stadt	stadt	städt-ə	städt-ə	städt-ən	städt-ə
8	muttər	muttər	muttər	muttər	müttər	müttər	müttər-n	müttər
9	zoo	zoo	zoo	zoo-s	zoo-s	zoo-s	zoo-s	zoo-s
10	pizza	pizza	pizza	pizza	pizza-s	pizza-s	pizza-s	pizza-s

Tabelle 4.12: Flexion der starken Adjektive in der deutschen Standardsprache basierend auf Eisenberg (2006: 178)

		Nominativ	Akkusativ	Dativ	Genitiv
Singular	mask	–ər	–ən	–əm	–ən
	neut	–əs	–əs	–əm	–ən
	fem	–ə	–ə	–ər	–ər
Plural		–ə	–ə	–ən	–ər

Die Synkretismen betreffend bleibt noch ein letzter Fall zu besprechen. In der starken Adjektivflexion der deutschen Standardsprache (Tabelle 4.12) weisen der Dativ und der Genitiv Feminin Singular sowie der Genitiv Plural dieselbe Form auf, nämlich $-\bar{a}r$. Man könnte nun auf die Idee kommen, das Suffix $-\bar{a}r$ der drei Zellen durch die RR (36) zu definieren. Dabei kommen jedoch falsche Formen heraus, denn die RR (36) stellt die Formen für die folgenden Bündel an syntaktischen Eigenschaften zur Verfügung: Dativ Singular Feminin, Genitiv Singular Feminin, Dativ Plural Feminin, Genitiv Plural Feminin. Die RR (36) produziert also fälschlicherweise ein $-\bar{a}r$ für den Dativ Plural und ausschließlich feminine Formen für den Dativ und Genitiv Plural, obwohl im Plural Genus nicht unterschieden wird.

$$(36) \quad *RR_{A, \{CASE:DAT \vee GEN, NUM:SG \vee PL^{16}, GEND:FEM\}, ADJ[STRONG]} (<X, \sigma>) =_{def} <X\bar{a}r', \sigma>$$

Es werden hier also zwei RRs benötigt, nämlich eine für den Dativ und den Genitiv Singular Feminin (37) und eine für den Genitiv Plural (38), wobei Genus unterspezifiziert ist.

$$(37) \quad RR_{A, \{CASE:DAT \vee GEN, NUM:SG, GEND:FEM\}, ADJ[STRONG]} (<X, \sigma>) =_{def} <X\bar{a}r', \sigma>$$

$$(38) \quad RR_{A, \{CASE:GEN, NUM:PL\}, ADJ[STRONG]} (<X, \sigma>) =_{def} <X\bar{a}r', \sigma>$$

$$(39) \quad RR_{A, \{CASE:DAT, NUM:SG, GEND: MASC \vee NEUT\}, ADJ[STRONG]} (<X, \sigma>) =_{def} <X\bar{a}m', \sigma>$$

Auf den ersten Blick erscheint diese Lösung etwas ungünstig. Mit Stumps (2001) Analyse bräuchte es jedoch für diesen Fall auch zwei Regeln, und zwar eine RR für das Suffix $-\bar{a}r$ und eine *Symmetrical Syncretism Metarule* für die synkretisierten Formen. Die RRs (37) und (38) machen das System also nicht komplexer als Stumps Analyse. Dafür stellen sich die erörterten Probleme der Vergleichbarkeit bei verschiedenen Regeltypen nicht, wie sie Stump (2001) vorschlägt. Außerdem handelt es sich hier eventuell doch um einen komplexeren Synkretismus im Vergleich z.B. zum Dativ Singular Maskulin und Neutrum der stark flektierten Adjektive (vgl. RR (39)). Erstens wird im Plural im Gegensatz zum Singular kein Genus unterschieden. Zweitens fallen der Dativ und Genitiv Singular Feminin nur mit den Genitiv Plural und nicht mit dem Genitiv und Dativ Plural zusammen. Diese höhere Komplexität des Dativ/Genitiv Singular Feminin und des Genitiv Plural verglichen mit dem Dativ Singular Maskulin und Neutrum kann folglich durch die RRs abgebildet werden.

Wurzel-/Stammalternation: Schließlich ist die Frage zu klären, wie Wurzeln zu Stämmen werden und wie die Verteilung der Stämme erfasst werden kann. Stumps (2001) Modell bezüglich der Stämme kann aus unterschiedlichen Gründen, die erörtert werden, nicht

¹⁶ Der Numerus könnte unterspezifiziert bleiben. Zur besseren Verständlichkeit werden Singular und Plural explizit genannt.

übernommen werden. Dafür wird ein Vorschlag gemacht, der mit dem bisher vorgestellten Modell kompatibel ist und mit dem Komplexität gemessen werden kann.

Es soll an dieser Stelle außerdem vorausgeschickt werden, was unter Wurzel und Stamm verstanden wird. Wie im Kapitel 4.1.3.1. dargestellt wurde, stammt die Wurzel aus dem Radikon. Durch RRs werden von Wurzeln neue Formen abgeleitet, indem phonologisches Material affigiert oder die Wurzel modifiziert wird (z.B. Umlaut). Dabei können Stämme entstehen, wie z.B. die Pluralstämme *äpfəl-* und *kind-ər-*. Es muss aber nicht zwangsläufig zuerst ein Stamm gebildet werden, bevor ein Flexionssuffix angehängt wird. Beispielsweise macht es sehr wenig Sinn, davon auszugehen, dass von der Wurzel *kind* ein Stamm *kind* geformt wird,¹⁷ an den *-əs* suffigiert wird (z.B. *Kind-əs*). Durch RRs können Stämme von anderen Stämmen (Stammalternation) wie auch von Wurzeln abgeleitet werden. In der Folge wird also von Wurzel- und Stammalternation gesprochen, da zwei Stämme, aber ebenfalls eine Wurzel und ein Stamm in einem Paradigma systematisch miteinander alternieren können.

Wurzel-/Stammalternationen bei Stump (2001): Stump (2001) unterscheidet u.a. *Stem-Formation Rules* und *Stem-Selection Rules*. Eine *Stem-Formation Rule* definiert die Form eines Stammes. Folgendes Beispiel aus Stump (2001) beschreibt die starken und schwachen Stämme der Possessivadjektive (Maskulin und Neutrum) *bhágavant* und *bhágavat* im Sanskrit:

(40)

„Stem-formation rule:

A possessive adjective has a Strong stem in *-vant* iff it has a Weak stem in *-vat*“ (Stump 2001: 172).

Die Verteilung der Stämme wird über *Stem-Selection Rules* definiert. Bezogen auf die Stämme in (40) beschreibt die *Stem-Selection Rule* in (41), dass der starke Stamm in den direkten Kasus des Maskulinums verwendet wird:

(41)

„Stem-selection rules:

[...] $RR_0 \{GEN:masc, DIR:yes\}, [C-stem\ nominal] \langle X, \sigma \rangle = \text{def } \langle Y, \sigma \rangle$, where Y is X's Strong stem [...]" (Stump 2001: 179).

Stumps (2001) Modell ist in zweierlei Hinsicht problematisch. Erstens beschreibt zwar die *Stem-Formation Rule* die Form des Stammes, aber auf eine eher informelle Weise. Vielmehr ähneln die *Stem-Formation Rules* einer Listung von Stämmen, was grundsätzlich unbefriedigend ist. Es wurde gezeigt, dass laut Ackerman & Stump (2004: 117) lediglich Wurzeln im Radikon gelistet sind und alle anderen Formen von diesen Wurzeln durch RR

¹⁷ In der inferentiellen-realisierenden Morphologie gibt es keine Nullmorpheme (vgl. Kap. 4.1.2.).

abgeleitet werden. Daraus kann geschlossen werden, dass auch die Stämme von den Wurzeln deriviert werden sollen. Der zweite Grund, weshalb Stumps (2001) Vorschlag nicht übernommen werden kann, betrifft die Komplexitätsmessung. Da die Stämme von den Wurzeln abgeleitet werden müssen, erhöht die Anzahl der Stämme (wie auch der Suffixe) die Komplexität der Flexion. Es wurde bereits dafür argumentiert, dass zur Komplexitätsmessung ein einheitliches und vergleichbares Instrument zentral ist. Die *Stem-Formation Rules* und die *Stem-Selection Rules* können jedoch miteinander nicht verglichen bzw. verrechnet werden, da sie völlig unterschiedliche Formen aufweisen (vgl. Diskussion zu den Synkretismen). Des Weiteren sind diese beiden Regeltypen nicht vergleichbar mit dem bisher vorgeschlagenen Regeltyp, nämlich der RRs. Zwar hat (41) die Form einer RR, der Nebensatz ist aber zu informell und beliebig erweiterbar, was für die Vergleichbarkeit problematisch ist. Aus diesen Gründen wird hier vorgeschlagen, die Form und Verteilung der Stämme über RRs zu definieren.

Wurzel-/Stammalternationen in dieser Arbeit: In den Kapiteln 4.1.2, 4.1.3.1. und 4.1.3.2. wurde gezeigt (basiert auf Anderson (1992), Stump (2001), Ackerman & Stump (2004)), dass ein Lexem mit seinen morphosyntaktischen Eigenschaften verbunden wird und dass erst diese Verknüpfung es den RRs erlaubt, die Flexionsaffixe anzubringen. Flexion ist also ein rein phonologischer Prozess, der aus einer Wurzel ein flektiertes Wort macht. Diese Idee soll hier auf die Stammbildung übertragen werden. Denn auch die Stammbildung ist nichts anderes als das Ableiten einer neuen Form von einer Wurzel. Auf das Resultat dieser Ableitung, d.h. auf den Stamm, können weitere RRs angewendet werden (z.B. für Suffixe), wie dies auch bei Suffixen möglich ist, wo mehrere Suffixe nacheinander angebracht werden können. So entstehen Stämme nicht durch Auflistung, sondern werden analog zu den Suffixen behandelt.

Im Gegensatz zu Stump (2001) wird also in dieser Arbeit, streng genommen, das Konzept des Stammes obsolet. Ein Lexem hat eine oder mehrere Wurzeln (im Radikon), die durch RRs modifiziert werden, wobei diese RRs wiederum in Blöcke eingeteilt sind, um die richtige Abfolge der RRs zu definieren. Trotzdem wird in der Folge zum einfacheren Verständnis von Wurzel-/Stammalternationen gesprochen, da diese Konzepte in der Morphologie weitverbreitet sind.

Affigierung und Stammbildung als phonologische Prozesse zu betrachten, hat den weiteren Vorteil, dass Operationen der nicht-konkatenativen Flexion wie Ablaut, Umlaut, Subtraktion etc. problemlos abgebildet werden können. Ein zentraler Anspruch der inferentiellen-realisierenden Morphologie ist, durch die strikte Trennung von Form und Funktion/Bedeutung auch Phänomene der nicht-konkatenativen Flexion adäquat abzubilden. Es wurde bereits gezeigt, dass RRs Umlaute produzieren können (vgl. RR in (12)). Hier wird dargestellt werden, wie eine RR aussieht, die Subtraktion definiert.

Die Wurzel-/Stammalternationen der in dieser Arbeit untersuchten Varietäten können alle durch die phonologische Umgebung erklärt werden. Wenn von einer Wurzel Stämme abgeleitet werden, gibt es drei logisch mögliche Operationen: Entweder wird der Wurzel phonologisches Material hinzugefügt oder getilgt oder die Wurzel wird Modifiziert (z.B. Umlaut). Beides kommt in den hier untersuchten Varietäten vor. Das Possessivpronomen der 1. Person Singular im Alemannischen von Bern weist zwei Stämme auf: *mi-* und *min-* (Marti 1985: 99). In dieser

Arbeit wird davon ausgegangen, dass, wenn die Wurzel und ein Stamm homophon sind, dieser Stamm den Defaultstamm bildet und mindestens im Nominativ verwendet wird. Im Alemannischen von Bern steht der Defaultstamm *mi-* wenn kein oder ein konsonantisch anlautendes Suffix folgt, der Stamm *min-*, wenn ein vokalisch anlautendes Suffix folgt (vgl. Paradigma 111). Durch eine RR muss also von der Wurzel *mi* der Stamm *min-* abgeleitet werden:

$$(42) \quad \text{RR}_{B, \{\}, \text{PRON.POSS}[\text{PERS:1} \vee 2, \text{NUM:SG}]} (<X, \sigma>) =_{\text{def}} <Xn/_V', \sigma>$$

Diese RR besagt, dass für das Possessivpronomen der 1. und 2. Person Singular der Wurzel ein *n* suffigiert wird, wenn ein Vokal folgt. Die RRs müssen also nur durch Kontextbedingungen erweitert werden, die durch phonologische Regeln formalisiert werden können. Dies ist mit dem bisherigen Modell kompatibel, da auch die Suffigierungsregeln ohne Kontextbedingungen nichts anderes als phonologische Prozesse sind, die durch phonologische Regeln definiert werden.

Ein Beispiel für Subtraktion findet sich im Alemannischen von Zürich. Auch hier weist das Possessivpronomen der 1. Person Singular zwei Stämme auf: *mīn-* und *mī-* (Weber 1987: 135). *Mīn-* bildet den Defaultstamm (also Wurzel), *mī-* den abgeleiteten Stamm, der dann auftritt, wenn das folgende Suffix konsonantisch anlautet (vgl. Paradigma 110). Dies drückt folgende RR aus (auch für die 2. Person Singular, die gleich variiert), indem sie besagt, dass *n* getilgt wird, wenn ein Konsonant folgt.

$$(43) \quad \text{RR}_{B, \{\}, \text{PRON.POSS}[\text{PERS:1} \vee 2, \text{NUM:SG}]} (<X, \sigma>) =_{\text{def}} <X *n \rightarrow \emptyset / _K', \sigma>$$

Wurzel-/Stammalternationen können auch mit morphosyntaktischen Eigenschaften in Zusammenhang stehen, z.B. wenn ein bestimmter Stamm nur in einem bestimmten Kasus oder Numerus auftritt. Auch dies kann dargestellt werden. Einige der untersuchten Varietäten zeigen in der Substantivflexion einen Singular- und einen Pluralstamm. Dies ist der Fall z.B. in der deutschen Standardsprache, in der in einigen Flexionsklassen der Plural mit dem Umlaut markiert wird. Dies ist also gleichzeitig ein Beispiel für Stammmodifikation. Folgende RR, die bereits oben eingeführt wurde, definiert den Umlaut für das Substantivparadigma der deutschen Standardsprache:

$$(12) \quad \text{RR}_{A, \{\text{NUM:PL}\}, \text{N[IC: } 1 \vee 3 \vee 7 \vee 8]} (<X, \sigma>) =_{\text{def}} <\ddot{X}', \sigma>$$

Ein weiteres Beispiel für den Zusammenhang zwischen Wurzelmodifikation und morphosyntaktischen Eigenschaften findet sich im Alemannischen von Huzenbach. Hier lautet der Singular des Wortes *Häuschen heisle*, der Plural *heisl-ə* (Baur 1967: 98). Da Kasus nicht am Substantiv markiert wird, gibt es keinen Grund das *e* in *heisle* als Suffix zu analysieren. Vielmehr weisen auf *e* endende Wurzeln einen Singularstamm auf, der gleich lautet wie die Wurzel, und einen Pluralstamm, bei dem das *e* getilgt wird. Auch der Plural zeigt keine Kasusmarkierung, sondern nur den Pluralmarker *-ə*. Die RR in (44) definiert genau dies: Stehen die Substantive der Flexionsklassen 4 und 5 im Plural, wird *e* vor *ə* getilgt.

$$(44) \quad \text{RR}_{C, \{ \text{NUM:PL} \}, \text{N[IC: 4} \vee 5]} (<X, \sigma>) =_{\text{def}} <X *e \rightarrow \emptyset / _ \text{ə}, \sigma>$$

Als Letztes muss noch überlegt werden, in welchen Blöcken die RRs zur Wurzel-/Stammalternation zu verorten sind. Für die Substantivflexion der deutschen Standardsprache wurden drei Blöcke vorgeschlagen: Block A für die Modifikationen an der Wurzel (Umlaut), Block B für die Pluralsuffixe, Block C für die Kassuffixe. So können die flektierten Substantive adäquat aufgebaut werden. Auf den ersten Blick scheint auch jede andere Abfolge unplausibel. Dies ändert sich jedoch, betrachtet man die Nicht-Standardvarietäten, was am Alemannischen von Huzenbach illustriert wird. Dieser Dialekt weist drei Typen von Flexion auf: Der Plural kann durch Suffixe oder Umlaut markiert werden und die Flexionsklassen 4 und 5 weisen eine Wurzelalternation auf. Man könnte also drei Blöcke annehmen, wobei Block A die Wurzelalternationen beinhaltet, Block B den Umlaut und Block C die Suffixe, was die Daten aber nicht genau erfasst. Denn die Wurzelalternation wird durch die Präsenz des Pluralsuffixes verursacht und nicht umgekehrt, es wird zuerst suffigiert und dann variiert die Wurzel je nach phonologischer Umgebung. Folglich müssen diese drei Blöcke angenommen werden: Block A für den Umlaut, Block B für die Suffixe, Block C für die Wurzelalternation. Es kann somit dargestellt werden, dass die Präsenz eines Suffixes der Auslöser der Wurzelalternation ist. Übrigens passiert diese Variation in der Morphologie und nicht in der Phonologie. Bei der Pluralsuffigierung entsteht zwar *heisle-ə*. Würde man aber die Tilgung in der Phonologie verorten, würde sie im gesamten Dialekt angewendet. Zur Vermeidung eines Hiats ist jedoch in den alemannischen Dialekten das Einfügen eines *n* zu erwarten. Da die Tilgung nur in diesem Paradigma gilt, ist die Regel dazu in der Flexionsmorphologie anzusetzen.

4.2. Was macht ein System komplexer oder simpler?

Im vorangehenden Kapitel wurde gezeigt, dass die Flexionsmorphologie im System der RRs zu verorten ist. Des Weiteren wurde im Kapitel 3.1.2. basierend auf Miestamo (2008) die absolute Komplexität eines linguistischen Phänomens definiert als die Länge der Beschreibung dieses Phänomens. Beides zusammengenommen bedeutet dies, dass je mehr RRs benötigt werden, um das System der Flexion zu beschreiben, desto komplexer ist dieses System. Stump (2001) schlägt neben den RRs weitere Regeltypen für Synkretismen und Wurzel-

/Stammalternation vor. Dass diese aber zur Komplexitätsmessung nicht übernommen werden können, da sie miteinander nicht vergleichbar sind, wurde bereits im Kapitel 4.1.3.3. erörtert. Was also ein System mehr oder weniger komplex macht und was ein System um wie viel mehr oder weniger komplex macht, wird ausschließlich aus dem System der RRs deduziert. Dadurch, dass die Definition und Messung der Komplexität theoriegeleitet vonstattengeht und nicht durch Introspektion o.Ä., wird eine maximal mögliche Objektivität gewährleistet.

Was die Flexion **komplexer** macht, ist folglich die Menge der Affixe, Wurzel-/Stammalternationen (z.B. Umlaute und Subtraktion) und freien Varianten, die alle durch RRs abgebildet werden. So wie die RRs hier definiert sind, resultiert bezüglich der Erhöhung der Komplexität des Flexionssystems automatisch Folgendes. Erstens erhöht die Anzahl der grammatischen Eigenschaften, die in der Flexion unterschieden werden, die Komplexität der Flexion, wie z.B. die Anzahl Kasus, die durch unterschiedliche Flexionsaffixe kodiert werden. Beispielsweise wird in der deutschen Standardsprache neben Plural auch der Dativ Plural markiert, während in den meisten alemannischen Dialekten nur der Plural markiert wird. Für die deutsche Standardsprache braucht es also eine RR für den Dativ Plural, dagegen ist dies in den meisten alemannischen Dialekten nicht nötig. Zweitens wird die Flexion komplexer, je mehr Allomorphe gezählt werden, z.B. die Anzahl unterschiedlicher Pluralaffixe. Für die Varietäten, die hier untersucht werden, bedeutet dies die Anzahl RRs für Suffixe und Wurzel-/Stammalternationen (Umlaut und Subtraktion), die durch die Verknüpfung einer Wurzel mit der Bedeutung Plural lizenziert werden. Drittens ergibt sich daraus, dass auch der Mehrfachausdruck zu höherer Komplexität führt. Während der Plural *Tag-ə* mit einer RR kodiert wird (45), werden für den Plural *Wäld-ər* zwei RRs benötigt, nämlich eine für das Suffix (12) und eine für den Umlaut (13).

$$(45) \quad \text{RR}_{B, \{ \text{NUM:PL} \}, \text{N[IC: } 1 \vee 2 \vee 7]} (\langle X, \sigma \rangle) =_{\text{def}} \langle X\partial', \sigma \rangle$$

$$(12) \quad \text{RR}_{A, \{ \text{NUM:PL} \}, \text{N[IC: } 1 \vee 3 \vee 7 \vee 8]} (\langle X, \sigma \rangle) =_{\text{def}} \langle \ddot{X}', \sigma \rangle$$

$$(13) \quad \text{RR}_{B, \{ \text{NUM:PL} \}, \text{N[IC: } 3]} (\langle X, \sigma \rangle) =_{\text{def}} \langle X\partial r', \sigma \rangle$$

Viertens erhöht eine bestimmte Art von Synkretismus die Komplexität der Flexion, der im Kapitel 4.1.3.3. ausführlich erörtert wurde. Kurz gesagt handelt es sich darum, dass, wenn die Werte von mehr als zwei Features variieren, der Synkretismus durch die RRs nicht einheitlich erfasst werden kann. Dargestellt wurde dies am Beispiel des Dativs und Genitivs Feminin Singular und des Dativs und Genitivs Plural der starken Adjektivflexion (vgl. Tabelle 4.7). Der Dativ und Genitiv Feminin Singular weisen das Suffix *-ər* auf, der Genitiv Plural ebenfalls das Suffix *-ər*, aber der Dativ Plural das Suffix *-ən*. Die Suffixe *-ər* können nicht durch eine RR erfasst werden, da der Plural im Gegensatz zum Singular kein Genus unterscheidet und der Dativ und Genitiv Plural nicht zusammenfallen wie im Feminin Singular. Es braucht also nicht eine, sondern zwei RRs, die in (36, Dativ und Genitiv Singular Feminin) und (37, Dativ Plural) aufgezeigt und hier wiederholt werden.

$$(37) \quad \text{RR}_{A, \{ \text{CASE:DAT} \vee \text{GEN}, \text{NUM:SG}, \text{GEND:FEM} \}, \text{ADJ}[\text{STRONG}]} (<X, \sigma>) =_{\text{def}} <X\partial r', \sigma>$$

$$(38) \quad \text{RR}_{A, \{ \text{CASE:GEN}, \text{NUM:PL} \}, \text{ADJ}[\text{STRONG}]} (<X, \sigma>) =_{\text{def}} <X\partial r', \sigma>$$

Was das System **simpler** macht, sind erstens alle anderen Synkretismen, bei denen nur die Werte eines Features variierten. Solche Synkretismen finden sich bezüglich aller Features: Kasus, Numerus, Genus, Wortart, Flexionsklasse. Ein Kasussynkretismus ist in der RR (37) dargestellt. In der starken Adjektivflexion fallen Dativ und Genitiv im Singular Feminin zusammen. Auf dieselbe Weise werden auch Genussynkretismen erfasst. Fallen Singular und Plural zusammen, bleibt das Numerusfeature unterspezifiziert. Auch können Unterscheidungen innerhalb einer Wortart aufgehoben werden. Beispielsweise wird das Suffix $-\partial n$ (Akkusativ Singular Maskulin) sowohl in der starken als auch in der schwachen Adjektivflexion verwendet. Dies zeigt die RR (23), die im Gegensatz zu den RRs (37) und (38) die Wortart nicht weiter spezifiziert.

$$(23) \quad \text{RR}_{A, \{ \text{CASE:ACC}, \text{NUM:SG}, \text{GEND:MASC} \}, \text{ADJ}[\]} (<X, \sigma>) =_{\text{def}} <X\partial n', \sigma>$$

Auch Synkretismen zwischen den Flexionsklassen können erfasst werden. Das Suffix $-\partial n$ für den Dativ Plural der Substantive in der deutschen Standardsprache wird in verschiedenen Flexionsklassen wiederverwendet. Dazu wird nur eine RR (20) gebraucht:

$$(20) \quad \text{RR}_{C, \{ \text{CASE:DAT}, \text{NUM:PL} \}, \text{N}[\text{IC: } 1 \vee 2 \vee 3 \vee 4 \vee 5 \vee 6 \vee 7 \vee 8]} (<X, \sigma>) =_{\text{def}} <X\partial n', \sigma>$$

Zweitens vereinfachen Wörter, die keine overte Markierung (Wurzel-/Stammalternation oder Affix) haben, das System der Flexion. Dies kommt gerade in der Substantivflexion der modernen deutschen Varietäten besonders häufig vor. Z.B. weisen in der deutschen Standardsprache im Singular nur der Genitiv Maskulin und Neutrum sowie die schwache Flexion Kasussuffixe auf (vgl. Tabelle 4.11). Diese Zellen werden durch unterschiedliche RRs gefüllt. Für alle anderen Zellen, d.h. für jene ohne overte Markierung, wird nach Stump (2001) nur eine RR benötigt, nämlich die RR *Identity Function Default*:

(26)

„Identity Function Default [...]

$\text{RR}_{n, \{ \}, U} (<X, \sigma>) =_{\text{def}} (<X, \sigma>)$ “ (Stump 2001: 53).

Diese RR besagt, dass mit der Wurzel nichts passiert, wenn keine RR gefunden wird. Da es diese RR in allen Varietäten braucht, sie also keine Varietät im Vergleich mehr oder weniger komplex macht, muss sie bei der Messung der Komplexität nicht berücksichtigt werden.

Drittens werden Unterscheidungen, die in einer Wortart nicht gemacht werden, jedoch in einer anderen Wortart ausgedrückt werden, nicht berücksichtigt. Zum Beispiel unterscheiden die meisten alemannischen Dialekte zwar Kasus an den Determinierern und Pronomen, aber nicht am Substantiv. Das Pronomen braucht also RRs für die Kasusunterscheidung, das Substantiv hingegen benötigt keine. Viertens wird alles, was nicht der Morphologie zugeordnet werden kann, nicht berücksichtigt. Dies betrifft u.a. Allomorphie, die phonologisch erklärt werden kann. Z.B. ist die *-as/-s* Variation der Substantive im Genitiv Singular Maskulin und Neutrum der deutschen Standardsprache phonologisch distribuiert. Die Verteilung der Varianten hängt vom Auslaut, der Betonung und der Anzahl Silben ab (Eisenberg et al. 1998: 224–225). Diese phonotaktischen Regeln gelten für das gesamte System und nicht nur für die Substantive, die Variation wird also nicht durch RRs, sondern durch phonotaktische Regeln ausgedrückt, weshalb sie die Komplexität der Phonologie und nicht der Morphologie erhöhen.

4.3. Messmethode

Dieses Kapitel widmet sich den Methoden zur Messung struktureller Komplexität. Im Kapitel 4.3.1. werden die bis jetzt entworfenen Messmethoden vorgestellt. Es wird gezeigt, dass keine dieser Methoden auf stark flektierende, eng verwandte Sprachen übertragen werden und die Komplexität in der Nominalflexion adäquat erfassen kann. Im Kapitel 4.3.2. wird die hier entwickelte und angewandte Messmethode präsentiert, welche auf der im Kapitel 4.1. dargestellten Theorie basiert.

4.3.1. Bisherige Messmethoden

Es wird nun eine Übersicht über die verschiedenen Typen an Methoden gegeben, die bis jetzt verwendet wurden, um strukturelle Komplexität zu messen. Jeder Typ an Methode wird kurz anhand eines Beispiels illustriert, wobei die meisten Beispiele aus Arbeiten stammen, die in den Kapiteln 2.1.2. und 2.2. vorgestellt wurden. Ziel ist es, das breite Spektrum an Methoden aufzuzeigen, und nicht, alle Methoden zu diskutieren. In der Folge werden zuerst die qualitativen und dann die quantitativen Methoden vorgestellt.

4.3.1.1. Qualitative Methoden

Dammel & Kürschner (2008) untersuchen die Komplexität der nominalen Pluralmorphologie in zehn germanischen Standardsprachen, und zwar an den folgenden fünf Parametern: Anzahl Allomorphe, formale Kodierung (Stammmodifikation, Redundanz, Null-Markierung, Subtraktion und Fusion Numerus/Genus), Zuweisung der Allomorphe zu Stämmen, Richtung der Bestimmung formaler Charakteristika zwischen Stamm und Suffix, Regularität (Dammel

& Kürschner 2008: 244, 257). Für all diese fünf Parameter wird bestimmt, was die Pluralmarkierung mehr oder weniger komplex macht, wobei qualitativ vorgegangen wird. Zum Beispiel werden für den Parameter 'Redundanz' zwei Gruppen an Sprachen gebildet: Sprachen, welche Mehrfachausdruck von Plural aufweisen (=komplex) oder nicht (=simpl) (Dammel & Kürschner 2008: 257). Wird mehr als einmal der Plural am Substantiv ausgedrückt, gilt dies als komplex, da es Ikonizität verletzt (Dammel & Kürschner 2008: 255–251). Der Parameter 'Stammmodifikation' wird anhand von vier weiteren Parametern untersucht: Suffigierung ohne Stammmodifikation, Modifikation des Stamm auslautenden Konsonanten, Modifikation des Wurzelsvokals und Suppletion. Für jede Sprache wird geprüft, wie sehr sie diese Kodierungstechniken verwendet (bezüglich Frequenz und Produktivität), wozu Werte von 1–4 vergeben werden (Dammel & Kürschner 2008: 250). Auf der Grundlage dieser Werte werden die Sprachen in fünf Gruppen auf einem Kontinuum eingeteilt, auf dem 'einfach' und 'komplex' die beiden Extreme bilden (Dammel & Kürschner 2008: 257). Auch der quantifizierbare Parameter 'Anzahl Allomorphe' wird qualitativ ausgewertet. Es werden vier Gruppen von einfach bis komplex gebildet: 1 Morph, 2–4 Allomorphe, 5–7 Allomorphe, Fusion von Kasus- und Numerusmarkierung (Dammel & Kürschner 2008: 247). Für jeden der fünf eingangs aufgezählten Parameter werden die Sprachen in zwei bis fünf Gruppen eingeteilt, welche auf einem Kontinuum von einfach bis komplex verortet werden (vgl. Tabelle 5 in Dammel & Kürschner 2008: 257). Mit dieser Methode ist es zwar möglich, die Sprachen innerhalb eines Parameters miteinander zu vergleichen. Da jedoch nicht quantifiziert wird, können die Parameter nicht miteinander verrechnet werden, d.h., dass keine genauen Aussagen über die Unterschiede in der Gesamtkomplexität der Pluralmarkierung zwischen den untersuchten Sprachen möglich sind. Des Weiteren lässt sich diese Methode nicht auf andere nominale Wortarten übertragen.

Auch **Camilleri (2012)** analysiert morphologische Komplexität aus einer qualitativen Perspektive. Die Grundlage bietet die kanonische Typologie und morphologische Komplexität wird definiert als das Ergebnis „of the divergence from the *canon*, where the further away from the canonical requirement a given example is, the more non-canonical, and the more morphologically complex it is“ (Camilleri 2012: 93). Zur Analyse eines Paradigmas werden zwei Perspektiven unterschieden: Erstens werden die verschiedenen Zellen im Paradigma eines Lexems verglichen, zweitens die Lexeme miteinander. Beide Perspektiven werden anhand von drei Parametern untersucht: 1) Zusammensetzung/Struktur, 2) lexikalisches Material (Form des Stammes), 3) Form der Affixe (Camilleri 2012: 94). Ein kanonisches Paradigma wird wie folgt definiert. Vergleicht man die Zellen eines Lexems, dann sind die Zusammensetzung/Struktur und der Stamm des Lexems immer gleich, die Form des Affixes ändert sich in Abhängigkeit der morphosyntaktischen Eigenschaften. Vergleicht man die Lexeme miteinander, weisen sie dieselbe Zusammensetzung/Struktur und dieselbe Form des Affixes für dieselbe morphosyntaktische Eigenschaft auf, während sich die Form der Stämme unterscheidet (Camilleri 2012: 94). Alles, was von einem solchen kanonischen Paradigma abweicht, gilt als komplex, wie z.B. Synkretismen und Flexionsklassen bezüglich des dritten Parameters. Einen Überblick, welche Phänomene als nicht-kanonisch gelten (nicht vollständige Liste), gibt folgende Tabelle:

Tabelle 4.13: Nicht-kanonische Phänomene in einem Paradigma (übernommen aus Camilleri 2012: 95)

	The content of the paradigmatic cell	Deviations	Comparisons across different lexical paradigms	Deviations
Composition/structure	different	fused exponence periphrasis	different	defectiveness overdifferentiation
Lexical material (stem-shape)	different	stem-alternation suppletion	same	heteroclisis
Affixal material (affix-shapes/forms)	same	syncretism uninflectability	different	deponency inflectional classes

Der Vorteil dieser Methode ist, dass ein Bezugspunkt bzw. ein Ideal aus rein formalen Überlegungen definiert wird, womit im Prinzip jedes Paradigma jeder Sprache verglichen werden kann. Damit können aber nur Aussagen der Art gemacht werden, inwiefern ein Phänomen von diesem Ideal abweicht. Camilleri (2012) zeigt nicht, wie diese Abweichungen quantifiziert werden könnten. Es ist folglich nicht möglich, den Grad der Komplexität von unterschiedlichen Sprachen zu vergleichen.

4.3.1.2. Quantitative Methoden

Anzahl Sprachen: Die Messmethode von Sinnemäki (2009) weicht ganz grundsätzlich von allen Methoden ab, die in der Folge noch vorgestellt werden. Untersucht wird die Markierung von Agens und Patiens, welche als komplex gilt, wenn sie das Eine-Bedeutung-Eine-Form-Prinzip verletzt. Dies ist der Fall, wenn eine Sprache mehr als eine Strategie zur Markierung verwendet (verletzt Ökonomie) oder wenn zu wenig markiert wird (verletzt Distinktheit) (Sinnemäki 2009: 130–133, vorgestellt im Kap. 2.2.2.). Es wird nicht die Komplexität an und für sich gemessen, sondern die Anzahl Sprachen, die dem Eine-Bedeutung-Eine-Form-Prinzip entsprechen und solche, die das nicht tun. Ein Resultat ist, dass die Anzahl Sprachen, die dem Eine-Bedeutung-Eine-Form-Prinzip folgen, parallel zur Größe der Sprachgemeinschaft zunimmt (Sinnemäki 2009: 135). Diese Methode misst also streng genommen nicht die Komplexität des Phänomens selbst, sondern beschreibt sie qualitativ (+/– Verletzung des Eine-Bedeutung-Eine-Form-Prinzips). Quantitativ ist nur die Anzahl Sprachen, die dem qualitativen Kriterium entsprechen bzw. nicht entsprechen.

Inventargröße – einzelne Phänomene und Kategorien: Eine weitere quantitative Methode, die ebenfalls die Komplexität nicht durch eine Zahl misst, ist sehr weit verbreitet und bezieht sich auf die Inventargröße. Es werden zumeist Aussagen gemacht, dass eine Kategorie mehr oder weniger grammatikalisiert ist und overt markiert wird. Die Komplexität der untersuchten Phänomene wird jedoch nicht durch Zahlen quantifiziert. Ein Beispiel dafür ist Schreier (2016),

dessen Arbeit im Kapitel 2.2.3. vorgestellt wurde. Er stellt u.a. fest, dass Tok Pisin im Gegensatz zu anderen englischen Varietäten im Personalpronomen der 1. Person Plural die Unterscheidung inklusiv/exklusiv grammatikalisiert hat, also mehr Formen hat als andere englische Varietäten in der 1. Person Plural, und in dieser Hinsicht folglich komplexer ist (Schreier 2016: 147). Oder das Verbalparadigma in Tristan da Cunha English hat durch Nivellierung Komplexität abgebaut, da *is* und *was* unabhängig von Person und Numerus verwendet werden (Schreier 2016: 149). Ähnlich verfahren beispielsweise Braunmüller (1984), McWhorter (2001), Kusters (2003) und Trudgill (u.a. 2011). Mit dieser Methode kann die Komplexität einzelner Phänomene und Kategorien beschrieben werden. Es ist jedoch nicht möglich, die Komplexität einer Kategorie oder eines Teilsystems zu messen und Aussagen darüber zu machen, wie viel komplexer oder weniger komplex eine Kategorie einer Varietät A im Vergleich zur selben Kategorie einer Varietät B ist. Auch kann die Komplexität verschiedener Komponenten eines Teilsystems nicht miteinander verrechnet werden (z.B. verschiedene Wortarten in der Nominalflexion).

Inventargröße – Präsenz/Absenz eines Phänomens oder einer Kategorie: In eine ähnliche Richtung geht ein Teil der Messmethode von Szmrecsanyi & Kortmann (2009) (vgl. Kap. 2.2.3.), welche jedoch klar quantifizierend ist. Als strukturell komplex gelten u.a. ornamentale Regeln und jene Regeln, die beim L2-Erwerb Schwierigkeiten verursachen (Szmrecsanyi & Kortmann 2009: 64–65). Als Datengrundlage dient der *World Atlas of Morphosyntactic Variation in English*, der 76 Charakteristika enthält. Von diesen 76 Charakteristika werden jene ausgesucht, welche dem linguistischen System Unterscheidungen und Asymmetrien hinzufügen, ohne dass diese kommunikative oder funktionale Vorteile haben (=ornamentale Regeln), wie z.B. Genus (Szmrecsanyi & Kortmann 2009: 68). Analog dazu wird bezüglich der L2-Schwierigkeit vorgegangen, nur dass hier von den 76 jene Eigenschaften extrahiert werden, welche in Intersprachen auftreten, z.B. fehlende Markierung der Vergangenheit (Szmrecsanyi & Kortmann 2009: 69). Anschließend wird für jede Varietät geprüft, wie viele der ausgesuchten Eigenschaften sie aufweist. Ein Beispiel: Hat eine Varietät zwei ornamentale Regeln, ist ihr Grad an Komplexität zwei, wobei eine Sprache komplexer ist, je mehr ornamentale Regeln sie aufweist. Das Gegenteil gilt für die L2-Schwierigkeit: Je mehr intersprachliche Eigenschaften eine Varietät hat, desto weniger komplex ist sie. Auch ein Teil der Methode von Nichols (2016) geht so vor, wobei ebenfalls den binären Eigenschaften die Zahl 1 vergeben wird, wenn eine bestimmte Eigenschaft in einer Sprache vorkommt, und die Zahl 0, wenn dies nicht der Fall ist (Nichols 2016: 137). Diese Methode basiert also auf der Präsenz bzw. Absenz gewisser Phänomene oder Kategorien, die als komplex gelten, wobei gilt, je mehr solche Phänomene/Kategorien vorkommen, desto komplexer ist eine Sprache. Der Vorteil dieser Methode ist, dass sie im Gegensatz zu den bis hierhin vorgestellten Methoden klar quantifiziert. Durch ihre Binarität (Absenz/Präsenz einer Kategorie oder eines Phänomens) ist sie jedoch nicht übertragbar auf die Messung der Komplexität in der Flexion. Zwar wäre es möglich, z.B. für die Pluralmarkierung 1 zu vergeben und für das Fehlen der Pluralmarkierung 0. Es macht jedoch einen Unterschied, ob der Plural immer mit derselben Form markiert wird oder nicht und ob der Plural mehrfach am selben Wort markiert wird oder nicht. Allomorphie und Mehrfachausdruck können mit dieser Methode folglich nicht erfasst werden, was jedoch bei stark flektierenden und eng verwandten Sprachen mit sehr feinen Unterschieden

wünschenswert ist. Schließlich kann diese Methode nicht die höhere bzw. geringere Komplexität innerhalb eines Teilsystems, einer Kategorie oder eines Phänomens messen.

Inventargröße – Anzahl Unterscheidungen innerhalb einer Kategorie: Vorschläge dazu kommen aus Arbeiten, in denen strukturelle Komplexität über die Inventargröße definiert wird: Je größer das Inventar ist, desto höher fällt die strukturelle Komplexität aus. Diese Grundidee wurde bereits auf verschiedene Teilsysteme der Grammatik angewandt, was hier exemplarisch veranschaulicht werden soll.¹⁸ Hay und Bauer (2007) berechnen die Größe des Phoneminventars durch die Anzahl Basismonophthonge, Extramonophthonge (Unterscheidung der Länge und Nasalierung), Diphthonge, Obstruenten und Konsonanten (Hay & Bauer 2007: 389). Die ersten Ideen in diese Richtung stammen bereits von Jakobson (1929). Nichols (2016) misst die Inventargröße von Komponenten unterschiedlicher linguistischen Beschreibungsebenen, wie z.B. die Anzahl der kontrastiven Obstruenten, die Anzahl der Wortabfolgen in den Basissatztypen sowie die Anzahl der unterschiedlichen Typen an Präfixen (Nichols 2016: 137). Shosted (2006) zählt die Anzahl der möglichen Silben und der Marker in der Verbflexion in einer Sprache (Shosted 2006: 9–17). Seine Methode kann nur auf die verbale Flexion angewendet und nicht auf die nominale Flexion übertragen werden (Shosted 2006: 16). Auch ein Teil der Messmethode von Garzonio (2016) kann hier dazugerechnet werden. Er untersucht die Komplexität von Entscheidungsfragesätzen und w-Fragesätzen in drei italienischen Dialekten. Dabei stellt er fest, dass in zwei Dialekten das w-Element verdoppelt werden kann und dass in einem das w-Element verschiedene Formen hat (Garzonio 2016: 105, 109–110). Diese freien Varianten werden gezählt und erhöhen folglich die Größe des Inventars. Trotz der großen Vielfalt an Messmethoden in diesem Bereich wurde noch keine für die Nominalflexion von stark flektierenden und eng verwandten Sprachen vorgeschlagen.

Stammformen: In eine völlig andere Richtung zielt die Typologie der Stammformen von Finkel & Stump (2007). Zwar geht es nicht darum, strukturelle Komplexität zu messen, ihre Typologie ließe sich aber einfach quantifizieren. Zuerst ist jedoch festzuhalten, dass sie drei verschiedene Schemen an Stammformsystemen unterscheiden: statische, adaptive und dynamische Stammformen. Dargestellt sind diese in Tabelle 4.14: I–VI stehen für Flexionsklassen, W–Z sind Sets an morphosyntaktischen Eigenschaften, a–o die Flexionsmarker. Werden in jeder Flexionsklasse die Stammformen durch dieselben Sets an morphosyntaktischen Eigenschaften ermittelt, so spricht man von statischen Stammformen (Finkel & Stump 2007: 2). In adaptiven Systemen teilen die Stammformen ein Set an morphosyntaktischen Eigenschaften, die anderen können unterschiedlich sein (Finkel & Stump 2007: 2). Lautet der Marker von W a, dann gehört dieses Lexem zur Flexionsklasse I. Lautet der Marker von W jedoch c, dann kann die Flexionsklasse nur durch den Marker von X bestimmt werden. Heißt dieser f gehört das Lexem zur Flexionsklasse III, heißt dieser g, gehört das Lexem zur Flexionsklasse IV. In einem dynamischen System hingegen müssen die Stammformen keine morphosyntaktischen Eigenschaften teilen (Finkel & Stump 2007: 4). Lautet der Marker von W b, ist das Lexem Teil der Flexionsklasse II, lautet der Marker von X f, ist das Lexem Teil der Flexionsklasse III. Es kann festgehalten werden, dass dynamische

¹⁸ Die folgenden Studien wurden bereits in den Kapiteln 2.2.2. und 2.2.3. ausführlicher vorgestellt, weshalb hier nur die Messmethoden erörtert werden.

Schemen zu bevorzugen sind, da sie mit kleineren Systemen an Stammformen auskommen als statische oder adaptive Schemen (Finkel & Stump 2007: 4).

Tabelle 4.14: Typen an Stammformsystemen (Finkel & Stump 2007: 2–4)

	statisch				adaptiv				dynamisch			
	W	X	Y	Z	W	X	Y	Z	W	X	Y	Z
I	a	e	i	m	a	e	i	m	a	e	i	m
II	b	e	i	m	b	e	i	m	b	e	i	m
III	c	f	j	n	c	f	j	n	c	f	j	n
IV	c	g	j	n	c	g	j	n	c	g	j	n
V	d	h	k	o	d	h	k	o	d	h	k	o
VI	d	h	l	o	d	h	l	o	d	h	l	o

Um die typologische Variation zwischen den Systemen von Stammformen verschiedener Sprachen darzustellen, werden fünf Kriterien vorgeschlagen (Finkel & Stump 2007: 9), wovon zwei zur Messung der Komplexität quantifiziert werden können: 1) Wie viele Stammformen sind nötig, um das Paradigma eines Lexems zu bestimmen; 2) Wie viele dynamische Stammformen werden gebraucht, um eine bestimmte Wortform im Paradigma eines bestimmten Lexems zu bestimmen (Finkel & Stump 2007: 9, 16). Für 1) wird die durchschnittliche Anzahl Stammformen für das gesamte Flexionssystem (also für alle Flexionsklassen zusammen) berechnet (Finkel & Stump 2007: 11). 2) geht von der Beobachtung aus, dass es Formen mit bestimmten morphosyntaktischen Eigenschaften gibt, die von mehr als einer dynamischen Stammform abgeleitet werden müssen. Berechnet wird also die durchschnittlich Anzahl an dynamischen Stammformen, die benötigt werden, um den Marker einer bestimmten Wortform ableiten zu können (Finkel & Stump 2007: 18).

Würde man mit den Kriterien 1) und 2) die Komplexität eines Paradigmas messen, zeichnet sich diese Methode durch ihre radikale Objektivität aus. Annahmen folgender Art müssen nicht gemacht werden: Was vom Eine-Form-Eine-Bedeutung-Prinzip abweicht oder als L2 schwierig gilt, ist komplex. Der Nachteil an dieser Methode ist, dass keine Aussagen über das Inventar der Formen und deren Verteilung (mit Ausnahme der Stammformen) auf die morphosyntaktischen Eigenschaften gemacht werden können. Auch jene Formen, die keine Stammformen sind, müssen von der Morphologie definiert werden. Diese Methode könnte also jene Fälle nicht adäquat erfassen, wenn z.B. zwei Flexionssysteme gleich viele Stammformen, aber unterschiedlich viele Synkretismen haben. Ein Beispiel (vgl. Tabelle 4.14, dynamisch): Von der Form a für W leite ich ab, dass dieses Lexem zur Flexionsklasse I gehört. Um das Paradigma zu vervollständigen, muss ich wissen, dass in der Flexionsklasse I die Form für X e lautet, für Y i und für Z m (also drei Regeln). Angenommen die Formen von X, Y und Z sind zu m zusammengefallen, ist der Aufwand jedoch kleiner, um das Paradigma zu vervollständigen, weil ich mir nur eine Regel merken muss. Dieser Tatsache würde diese Methode nicht Rechnung tragen. Es sei hier nochmal darauf hingewiesen, dass Finkel & Stump (2007) keine Messmethode, sondern eine Typologisierung vorschlagen. Mit der skizzierten,

hypothetischen Messmethode ließe sich jedoch die Komplexität der Stammformensysteme berechnen.

Anzahl Move- und Mergeoperationen: Eine rein theoriegeleitete Messmethode verwendet Garzonio (2016), dessen Arbeit bereits im Kapitel 2.2.4. vorgestellt wurde. Es sei hier nur die Messmethode wiederholt. Die Messmethode ist im Minimalismus verortet und gemessen wird die syntaktische Komplexität bzw. die Komplexität der Derivation: Je mehr Move- und Mergeoperationen benötigt werden, um eine bestimmte Struktur abzuleiten, desto komplexer ist die Derivation (Garzonio 2016: 99). Wie jede Theorie geht auch der Minimalismus von bestimmten Annahmen über die Architektur der Grammatik aus. Dies trifft jedoch nicht auf die Messmethode selbst zu, wodurch sie sich durch ein hohes Maß an Objektivität auszeichnet. Auf die Morphologie kann Garzonios (2016) Messmethode aber naturgemäß nicht übertragen werden.

Korpusbasiert – Textfrequenz von Morphemen: Schließlich sollen noch zwei korpusbasierte Messmethoden vorgestellt werden. Szmrecsanyi & Kortmann (2009) wie auch Maitz & Németh (2014) berechnen die Textfrequenz bestimmter Morpheme (vgl. Kap. 2.2.3.). Die Komplexität der Grammatizität wird durch die Summe des Synthese- und Analyseindex ermittelt (Szmrecsanyi & Kortmann 2009: 72): Der Syntheseindex ist die Anzahl (in Prozent) der gebundenen Morpheme pro 1.000 Tokens, der Analyseindex die Anzahl der freien Morpheme pro 1.000 Tokens (Szmrecsanyi & Kortmann 2009: 72). Des Weiteren wird ein Transparenzindex berechnet, und zwar durch den prozentualen Anteil der gebundenen regelmäßigen Allomorphe an allen gebundenen Allomorphen (Szmrecsanyi & Kortmann 2009: 74). Je höher die Grammatizität und je tiefer der Transparenzindex sind, desto komplexer ist eine Varietät. Diese Methode hat einen praktischen und prinzipiellen Nachteil. Erstens ist sie nur auf jene Sprachen und Varietäten anwendbar, für die große, annotierte Korpora existieren. Zweitens ist nur eine eher oberflächliche Analyse möglich, da wir nichts über freie Varianten, Synkretismen, Mehrfachausdruck, Flexionsklassen oder Stammalternationen erfahren. Präzise und detaillierte Aussagen über Zusammenhänge oder eventuelle Ausgleichsmechanismen innerhalb der Flexion sind nicht möglich. Jedoch zeichnet sich diese Methode durch eine hohe Objektivität aus und eignet sich durchaus, um einen ersten Überblick über die Flexionsmorphologie zu erhalten.

Korpusbasiert – Informationstheorie: Als letzte Messmethode sollen noch jene präsentiert werden, die auf einer informationstheoretischen Grundlage basieren, z.B. Kolmogorov-Komplexität, Ziv-Lempel-Komplexität etc. Arbeiten auf dieser Basis stammen u.a. von Juola (1998), Bane (2008) und Ehret & Szmrecsanyi (2016), um nur einige wenige zu nennen. Es soll hier zur Illustration nur eine Operationalisierung der Informationstheorien vorgestellt werden, nämlich jene von Ehret & Szmrecsanyi (2016). Allgemein geht es darum zu prüfen, wie stark Kompressionsprogramme (z.B. gzip) Texte komprimieren, wodurch die Kolmogorov-Komplexität approximiert werden kann. Dazu werden dieselben Texte in verzerrter und unverzerrter Form komprimiert und die Resultate dieser Komprimierung verglichen. Um morphologische Komplexität zu messen, werden 10% der Buchstaben gelöscht, wodurch neue Wortformen entstehen und so Regularitäten vermindert werden. Anschließend wird dieser verzerrte Text komprimiert, wobei eine schlechte Kompressionsrate des verzerrten Textes niedrige morphologische Komplexität bedeutet. Dies wird dadurch begründet, dass

morphologisch komplexe Sprachen auf jeden Fall viele verschiedene Wortformen aufweisen, eine Verzerrung richtet also weniger Schaden an als in morphologisch einfachen Sprachen, in denen die Verzerrung verhältnismäßig mehr statistisches Rauschen verursacht (Ehret & Szmrecsanyi 2016: 75–76). Der Grad der morphologischen Komplexität wird dadurch berechnet, dass die Größe der komprimierten Datei nach der Verzerrung durch die Größe der komprimierten Datei vor der Verzerrung geteilt wird (Ehret & Szmrecsanyi 2016: 76). Auf eine analoge Weise wird syntaktische Komplexität gemessen, wobei zur Verzerrung des Textes 10% der Wörter gelöscht werden (Ehret & Szmrecsanyi 2016: 76). Auch hier können dieselben praktischen und prinzipiellen Nachteile genannt werden, welche bereits oben bezüglich der Methode von Szmrecsanyi & Kortmann (2009) moniert wurden. Außerdem bilden Texte die Datengrundlage, welche den Output eines grammatischen Systems darstellen und somit nur einen indirekten Blick auf die Komplexität des Systems bieten (Sinnemäki 2011: 19). Es stellt sich also die Frage „[...] what the application of a mathematical algorithm on linguistic products (texts) can reveal about the complexities of the underlying systems (grammar, lexicon) that are needed to produce these texts“ (Miestamo 2008: 28).

Zusammenfassung: Wie dargestellt wurde, kann keine der bisherigen Messmethode in dieser Arbeit übernommen werden. Zusammenfassen lassen sich folgende Gründe:

- Die Messmethode ist nicht auf die Nominalflexion übertragbar (z.B. Shosted 2006, Garzonio 2016).
- Die Messmethode sammelt, beschreibt und vergleicht nur einzelne Phänomene (z.B. McWhorter 2001, Schreier 2016). Dieses Vorgehen kann zwar erste Hinweise geben, ermöglicht aber keine systematische Messung struktureller Komplexität.
- Mit der Quantifizierung der Präsenz/Absenz einer Kategorie kann strukturelle Komplexität nur grob gemessen werden und funktioniert bei feinen Unterschieden zwischen eng verwandten Sprachen und Varietäten nicht (z.B. Nichols 2009).
- Die korpusbasierten Methoden sind zu oberflächlich, d.h., es können keine präzisen und detaillierten Aussagen über Zusammenhänge oder eventuelle Ausgleichsmechanismen innerhalb des Systems (z.B. Flexion) gemacht werden (z.B. Ehret & Szmrecsanyi 2016).

In der vorliegenden Arbeit soll die Komplexität der Nominalflexion in deutschen Varietäten gemessen werden. Die Messmethode muss also folgende Kriterien erfüllen:

- Die Messmethode muss strukturelle Komplexität quantifizieren.
- Die Messmethode soll auf die Flexionsmorphologie anwendbar sein, und zwar unabhängig von der Wortart. Dies gilt innerhalb der Nominalflexion, z.B. für Substantive, Artikel etc. Sie soll aber grundsätzlich die Komplexität der gesamten Flexionsmorphologie messen können, also auch der Verbalflexion. Denn nur mit einer einheitlichen Messmethode kann zukünftig die Komplexität der ganzen Flexionsmorphologie ermittelt werden.
- Die Messmethode muss kleinste Unterschiede erfassen und quantifizieren. Nur so können feinste Unterschiede zwischen eng verwandten Sprachen und Varietäten gemessen werden.
- Die Messmethode muss auch in der Flexion häufig auftretende Phänomene erfassen, wie Synkretismen, Allomorphie, Mehrfachausdruck an einem Wort, Stammalternationen (z.B. Umlaut, Ablaut, Subtraktion), Flexionsklassen, freie Varianten etc. Inwiefern diese Phänomene die strukturelle Komplexität erhöhen oder senken, soll theoretisch hergeleitet

werden, um eine möglichst hohe Objektivität zu gewährleisten. Die Messmethode muss diese Herleitung quantifizieren.

4.3.2. Methode zur Komplexitätsmessung in der Flexion

Im vorangehenden Kapitel wurde gezeigt, dass keine der bisherigen Messmethoden für die hier untersuchten Varietäten verwendet werden kann. Entweder quantifizieren die Messmethoden nicht systematisch genug oder sie können kleinste Unterschiede stark flektierender Varietäten nicht erfassen (vgl. Kap. 4.3.1.2.). Im Kapitel 4.1.3.2. wurde dargestellt, dass die Flexion in den RRs zu verorten ist, folglich kann die Komplexität in der Flexion anhand der Anzahl RRs gemessen werden. Davon wurde im Kapitel 4.2. abgeleitet, was das System der Flexion mehr oder weniger komplex macht. In der Folge wird zuerst vorgestellt, wie hier Komplexität in der Flexion von nominalen Wortarten gemessen wird und dann, welches die Vorteile dieser Messmethode sind.

Messmethode: Gemessen werden folgende nominale Wortarten, die in sechs Kategorien eingeteilt sind: Substantive (Kategorie A), starke und schwache Adjektive (Kategorie B), Personalpronomen (Kategorie C), Interrogativpronomen *wer/was* (Kategorie D), bestimmter Artikel und einfaches Demonstrativpronomen *der* (Kategorie E), unbestimmter Artikel und Possessivpronomen (Kategorie F). Die Wahl für diese Wortarten fiel vorwiegend aus praktischen Gründen. Denn es handelt sich dabei um jene Wortarten, die in allen Grammatiken beschrieben werden. Es wäre sicher gerade für die Nicht-Standardvarietäten wünschenswert gewesen, z.B. auch die Numeralia zu berücksichtigen, die in vielen alemannischen Dialekten flektiert werden. Da aber nicht für jede untersuchte Varietät eine Beschreibung der Numeralia zur Verfügung steht, wäre die Gesamtkomplexität der Varietäten nicht mehr vergleichbar, hätte man auch die Numeralia berücksichtigt.

Die Grammatiken beschreiben nicht nur verschiedene Wortarten, sondern sie bauen die Paradigmen auch nach unterschiedlichen Kriterien auf. Würde man also die Paradigmen so übernehmen, wie sie in den Grammatiken stehen, könnte man die Varietäten nicht vergleichen. Deswegen werden zwar die Informationen aus den Grammatiken entnommen. Anhand dieser Informationen werden aber nach einheitlichen Kriterien Paradigmen erstellt, damit sie dann auch verglichen werden können. Die Details dazu werden im Kapitel 5 vorgestellt. Das allgemeinste Kriterium jedoch ist, dass die Paradigmen auf ein Minimum reduziert werden, d.h., es werden nur jene Unterscheidungen gemacht, die unbedingt nötig sind. Beispielsweise unterscheiden die alemannischen Dialekte Kasus an den Pronomen und Determinierern, aber die meisten nicht am Substantiv. In diesen Dialekten hat also im Substantivparadigma jede Flexionsklasse nur zwei Zellen, und zwar eine für den Singular und eine für den Plural. Dieses Vorgehen hat zwei Vorteile, die miteinander zusammenhängen. Erstens, wenn jedes Paradigma auf sein Minimum reduziert wird, sind diese Paradigmen auch miteinander vergleichbar. Zweitens soll hier absolute Komplexität gemessen werden. Nach Miestamo (2008: 24) kann die absolute Komplexität eines linguistischen Phänomens durch die Länge der Beschreibung dieses linguistischen Phänomens gemessen werden. Dabei gilt, dass je kürzer diese Beschreibung ausfällt, desto weniger komplex ist dieses Phänomen. Auf das Modell übertragen, das hier

verwendet wird, heißt das, dass je weniger RRs für ein Paradigma benötigt werden, desto weniger komplex ist dieses Paradigma. Um dies zu ermöglichen und die Anzahl RRs der verschiedenen Varietät vergleichen zu können, müssen aber auch die Informationen aus den Grammatiken zuerst auf ein Minimum reduziert werden. Nur so kann mit den RRs gemessen werden, wie stark ein Paradigma komprimiert werden kann.

Oben wurden die sechs Kategorien vorgestellt, in die die untersuchten Wortarten eingeteilt sind. Die Kategorien E und F beinhalten je zwei Wortarten: bestimmter Artikel und einfaches Demonstrativpronomen *der* (Kategorie E), unbestimmter Artikel und Possessivpronomen (Kategorie F). Dass gerade diese Wortarten in die zwei Kategorien eingeteilt werden, hat zwei Gründe. Erstens weisen der bestimmte Artikel und das einfache Demonstrativpronomen auf der einen Seite und der unbestimmte Artikel und das Possessivpronomen auf der anderen Seite jeweils die größten Ähnlichkeiten in ihrer Flexionsmorphologie auf. In der deutschen Standardsprache geht das sogar so weit, dass die Paradigmen vollständig zusammenfallen (mit der Ausnahme, dass das Possessivpronomen einen Plural hat, der unbestimmte Artikel nicht). Im vorangehenden Absatz wurde argumentiert, dass die Paradigmen auf ein Minimum komprimiert werden müssen. Dies ist am besten zu erreichen, indem ähnliche oder gar gleiche Paradigmen in derselben Kategorie stehen. Wichtig ist jedoch auch hier, dass für alle Varietäten die gleichen Kategorien gebildet werden, damit die Kategorien über die Varietäten hinweg miteinander verglichen werden können. Der zweite Grund für diese Einteilung findet sich in der diachronen Entwicklung. Erstens ist der bestimmte Artikel aus dem einfachen Demonstrativpronomen entstanden (Schrodt 2004: 23). Zweitens entwickelte sich die Flexion des unbestimmten Artikels und des Possessivpronomens parallel. Während im Althochdeutschen das Zahlwort und das Possessivpronomen stark flektiert wurden (Braune 2004: 234, 245), konnte im Mittelhochdeutschen der unbestimmte Artikel stark (Paul 2007: 217), das Possessivpronomen stark und schwach flektiert werden (Paul 2007: 216, 369). Außerdem sind im Mittelhochdeutschen der Nominativ Singular aller Genera, der Akkusativ Singular Neutrum und oft auch der Akkusativ Singular Feminin in beiden Wortarten endungslos (Paul 2007: 216-217). In der modernen deutschen Standardsprache flektieren beide Wortarten stark. Davon weichen der Nominativ Singular Maskulin und Neutrum und der Akkusativ Singular Neutrum ab, die endungslos sind (ähnlich wie im Mittelhochdeutschen), wie auch der Genitiv Singular Maskulin und Neutrum, der auf *-əs* endet (Eisenberg 2006: 176). Dies gilt ebenfalls für beide Wortarten.

Bei der Art der Flexion, die in den hier untersuchten Varietäten vorkommt, handelt es sich um Suffixe und Wurzel-/Stammalternationen (z.B. Umlaut). In den Kapiteln 4.1.3.2. und 4.1.3.3. wurde gezeigt, dass diese durch RRs erfasst werden können. Die Höhe der Komplexität der Nominalflexion von allen Kategorien kann also durch die Anzahl RRs berechnet werden. Dabei gilt, je mehr RRs gezählt werden, desto höher ist die Komplexität. Eine Ausnahme bilden die Substantive. Zwar werden die Suffixe und Wurzel-/Stammalternationen der Substantive auch anhand der RRs abgebildet. Bei den Substantiven kommen aber noch die Flexionsklassen dazu. In dieser Arbeit wird Flexionsklasse als eine Art Instruktion verstanden, die die verschiedenen RRs miteinander kombiniert. Zum Beispiel weisen das Alemannische von Zürich und des Sensebezirks acht RRs auf, aber das Alemannische von Zürich hat acht Flexionsklassen und das Alemannische des Sensebezirks zehn Flexionsklassen. Das Alemannische des Sensebezirks

enthält also zwei Kombinationen mehr als das Alemannische von Zürich bei gleich vielen RRs, dieselbe Größe des Inventars an RRs wird im Alemannischen des Sensebezirks öfter kombiniert als im Alemannischen von Zürich. Um die Komplexität der Substantive adäquat zu erfassen, müssen also neben der Anzahl RRs auch die Anzahl Flexionsklassen berücksichtigt werden. Es stellt sich nun die Frage, wie die RRs und die Flexionsklassen miteinander verrechnet werden können. Logisch möglich sind zwei Operationen, nämlich Addition und Multiplikation. Für die Addition sprechen zwei Beobachtungen. Erstens werden die RRs miteinander addiert, was auf die Flexionsklassen übertragen werden soll, damit mit einer einheitlichen Messmethode die Vergleichbarkeit gewahrt bleibt. Zweitens würde bei der Multiplikation von RRs und Flexionsklassen die Komplexität der Substantive im Vergleich zu den anderen Kategorien überproportional hoch.

Neben den RRs für die Suffixe werden auch die RRs für die Wurzel-/Stammalternationen gezählt. Eine Ausnahme bildet die RR *Identity Function Default* (RR (26), Kap. 4.1.3.2.). Sie definiert, dass, wenn in einem beliebigen Block keine RR gefunden wird, mit der Wurzel nichts passiert. Dies kommt gerade in der Substantivflexion moderner deutscher Varietäten besonders häufig vor, z.B. *Tag* (Nominativ, Akkusativ, Dativ Singular). Im Kapitel 4.2. wurde erörtert, dass diese RR in allen Varietäten anzunehmen ist, weshalb sie keine Varietät mehr oder weniger komplex macht. Ob sie also bei der Komplexitätsmessung mitgezählt wird oder nicht, macht keinen Unterschied. Deshalb wird sie der Einfachheit halber in der Folge bei der Komplexitätsmessung nicht berücksichtigt.

Praktisch wird nun so vorgegangen. Für jede Kategorie sind die RRs, für die Substantive zusätzlich die Flexionsklassen zu finden. Die Anzahl RRs für eine bestimmte Kategorie entspricht der Komplexität dieser Kategorie. Hat also z.B. eine Kategorie 10 RRs, beträgt die Höhe der Komplexität dieser Kategorie 10. Weist eine Kategorie einer Varietät A 10 RRs auf und dieselbe Kategorie einer Varietät B 20 RR, dann ist die Kategorie der Varietät B komplexer als jene der Varietät A. Nachdem die Komplexität jeder Kategorie berechnet wurde, kann die Gesamtkomplexität der Nominalflexion ermittelt werden. Die Gesamtkomplexität der Nominalflexion einer Varietät wird durch die Addition der Komplexität aller Kategorien definiert: Gesamtkomplexität der Nominalflexion = Komplexität der Substantive (Kategorie A) + Komplexität der starken und schwachen Adjektive (Kategorie B) + Komplexität des Personalpronomens (Kategorie C) + Komplexität des Interrogativpronomens (Kategorie D) + Komplexität des bestimmten Artikels und des einfachen Demonstrativpronomens (Kategorie E) + Komplexität des unbestimmten Artikels und des Possessivpronomens (Kategorie F).

Vorteile: Diese Messmethode zeichnet sich durch vier Vorteile aus: Ein einheitliches und sprachunabhängiges Messgerät, das quantifiziert und auch kleinste Unterschiede messen kann. Erstens ist die Einheitlichkeit, die durch die Messung anhand RRs gewährleistet wird, wohl der wichtigste Vorteil. Denn nur so können Wortarten oder Kategorien innerhalb einer Varietät und zwischen verschiedenen Varietäten verglichen werden, nur so ist ein Vergleich der Komplexität der Nominalflexion zwischen unterschiedlichen Varietäten möglich. Des Weiteren kann die Gesamtkomplexität der Nominalflexion nur ermittelt werden, indem die Komplexität jeder Wortart oder Kategorie mit demselben Messgerät berechnet wird. Einerseits geht es also um die Vergleichbarkeit der Komplexität von Wortarten und der gesamten Nominalflexion innerhalb einer Varietät und zwischen unterschiedlichen Varietäten, andererseits um die

Möglichkeit, Wortarten miteinander zu verrechnen, um die Gesamtkomplexität ermitteln zu können. Zweitens ist diese Messmethode auf die gesamte Flexionsmorphologie übertragbar. Damit ist gemeint, a) dass sie z.B. auch auf die Verbflexion übertragen werden kann, b) dass es möglich ist, damit auch andere Flexionstypen zu erfassen, z.B. Präfigierung, Infigierung etc., c) dass diese Methode nicht nur die Komplexität in der Flexionsmorphologie des Deutschen berechnen kann, sondern im Prinzip für jede flektierende Sprache verwendet werden kann. Drittens ermöglicht diese Messmethode auch kleinste Unterschiede in der Flexion zu messen. Sie ist also eine sehr genaue Methode, mit der auch eng verwandte Varietäten untersucht werden können. Viertens ist sie klar quantifizierend, d.h., es können problemlos Aussagen gemacht werden, wie z.B. eine Sprache mit 150 RRs ist komplexer als jene mit 100 RRs.