

Humboldt-Universität zu Berlin  
Philosophische Fakultät II  
Institut für deutsche Sprache und Linguistik  
M.A. Linguistik

Erstgutachterin: Prof. Dr. Anke Lüdeling  
Zweitgutachter: Prof. Dr. Hans Boas, University of Texas at Austin, USA

**Mehr Chunks!**  
**Entwicklungsperspektiven für die Konstruktionsgrammatik unter  
Einbeziehung von Phraseologie, Psycholinguistik und L2-  
Erwerbsforschung**

Arbeit zur Erlangung des akademischen Grades „Master of Arts“ im Fach Linguistik

vorgelegt von Anna Shadrova  
Matrikelnr. 542283  
shadrova@hu-berlin.de

eingereicht am 03.06.2013

# Gliederung

1. Einleitung.....	2
2. Konstruktionsgrammatik .....	4
2.1 Theoretische Unschärfen .....	7
3. Was sind Chunks? - Begriffsklärung und Forschungsstand.....	9
3.1 Phraseologie.....	10
3.2 Psycholinguistik und L2-Forschung.....	14
3.3 Pragmatik und Diskursanalyse .....	18
3.4 Korpuslinguistik.....	19
3.5 Ein übergreifender Ansatz.....	21
3.6 Chunks in der Grammatik .....	25
3.7 Überblick: Begriffe für Chunks und ihre Flexibilitätseigenschaften.....	31
3.8 Funktionen von Chunks.....	33
4. Zwischenfazit: Chunks als Basic Level Category.....	38
5. Einbindung in die Konstruktionsgrammatik.....	44
5.1 Schema.....	46
5.2 Attribut-Wert-Matrizen.....	52
5.3 Frames.....	54
5.4 Chunks und Exemplare.....	55
6. Zusammenfassung und Fazit.....	56
7. Literatur.....	58

## 1. Einleitung

Erman und Warren (2000) schätzen, dass etwa 55% der von ihnen erhobenen gesprochenen Daten aus Einheiten bestehen, die nicht neu aus ihren Einzelteilen zusammengesetzt sind, sondern „am Stück“ produziert wurden. Aus der phraseologischen Forschung geht eine Vielzahl verschiedener Begriffe für diese Art von Einheiten hervor, die unterschiedliche Merkmale haben – zum Beispiel können sie kompositionell ableitbar sein oder nicht, ritualisierte pragmatische oder andere soziale Funktionen besitzen oder diskursstrukturierend wirken.

Ein Ziel dieser Arbeit ist, diese Begriffe zu finden, disziplinübergreifend darzustellen und Gemeinsamkeiten und Unterschiede in Definition und Gebrauch der Begriffe herauszuarbeiten – „*Was sind Chunks?*“ ist eine zentrale Fragestellung der Abschnitte 3 und 4.

In der Zweitspracherwerbsforschung ist man darauf aufmerksam geworden, dass gerade weit fortgeschrittene Lernerinnen<sup>1</sup> trotz hoher Expertise auf dem Gebiet der grammatischen Regelmäßigkeiten typischerweise durch eine geringe Natürlichkeit ihrer Sprache auffallen, weil sie nicht die konventionalisierte Form verwenden, um etwas zu ausdrücken, sondern eine andere: So wäre es ein der deutschen Grammatik entsprechender Satz, zu sagen „*Wie ist der Besuch deiner schulischen Einrichtung an diesem Vormittag verlaufen?*“, typischerweise würde eine Sprecherin aber wahrscheinlich „*Wie war's heute in der Schule?*“ fragen, und die Frage „*Willst du meine Frau werden?*“ oder „*Willst du mich heiraten?*“ ist wahrscheinlicher in der entsprechenden Situation anzutreffen, als der Satz „*Entspricht es deinen Zukunftsvorstellungen, eine Ehe zwischen dir und mir schließen zu lassen?*“. Es lassen sich viele Beispiele dieser Art konstruieren. Aber besser als konstruierte Beispiele zeigen Korpus- und experimentelle Daten, dass Wörter und grammatische Konstruktionen eine starke Tendenz haben, ausgewählte lexikalische Partner anzuziehen, und nicht sehr frei kombiniert zu werden. Das heißt nicht, dass nicht auch alles sprachlich Denkbare weitgehend frei kombinierbar wäre, es ist lediglich eine Feststellung, dass es nicht frei kombiniert wird.

Ein weiteres Ziel dieser Arbeit ist daher, die Arbeiten einer Reihe von Autorinnen zusammenzutragen und gemeinsame Tendenzen herauszuarbeiten, um damit die obige Behauptung anhand von psycho- und korpuslinguistischen Daten zu untermauern. „*Wo und wie treten Chunks auf?*“ ist die zweite zentrale Fragestellung, die vor Allem in Abschnitt 3 bearbeitet wird.

---

1 Ich werde der Einbeziehung weiblicher Subjekte und der Einheitlichkeit wegen im Folgenden ausschließlich die weiblichen Formen generisch verwenden. Männliche Subjekte wie Sprecher, Autoren, Linguisten und Lerner sind ausdrücklich mitgemeint.

Drittens sollen die zusammengetragenen Erkenntnisse auf die Konstruktionsgrammatik angewandt werden. Diese erhebt an verschiedener Stelle den Anspruch, eine Gesamtheorie der Sprache werden zu wollen – im Rahmen der konstruktionsgrammatischen Annahmen arbeiten Linguistinnen an diachronen, psycholinguistischen, lexikographischen, semiotischen, spracherwerbsbezogenen und vielen anderen Fragestellungen. Als vergleichsweise junge Theorie besitzt die Konstruktionsgrammatik allerdings bislang eher das Potential, zu einer Gesamtheorie heranzuwachsen. Der Entwicklungsspielraum hinsichtlich der Formalisierung, der Einarbeitung verschiedener Ergebnisse aus anderen Subdisziplinen, der kognitiven, semantischen und pragmatischen Modellierung ist noch groß. Abschnitt 2 stellt die Theorie und ihren Entwicklungsstand kurz vor.

Die in dieser Arbeit zitierten Studien deuten stark darauf, dass Chunks eine zentrale Rolle in vielen sprachlichen Teilbereichen zukommt. Daraus leite ich in Abschnitt 4 ab, dass Chunks als *Basic Level Category* nicht nur im Spracherwerb, sondern in der gesamten Sprache verstanden werden können. Die Konstruktionsgrammatik bezieht diese Studienlage bislang wenig, und nur hinsichtlich des Spracherwerbs, in ihre Theoriebildung ein. In Abschnitt 5 soll daher skizziert werden, wie eine Modellierung unter Einbeziehung von Chunks aussehen kann. Neben der Betrachtung von Chunks als zentrale Einheit schlage ich zwei weitere Modifikationen vor: Erstens, die Annahme eines verschobenen Schwerpunkts hinsichtlich des Gebrauchs von flexiblen und statischen Einheiten im mündlichen und schriftlichen Gebrauch; und zweitens, die Annahme eines Kontinuums von Abstrakt zu Konkret in der frame-semantischen Betrachtung. Die Erläuterungen hierzu finden sich in Abschnitt 5.

Der Rahmen dieser Arbeit erlaubt nur den Beginn einer Diskussion um die Formalisierung der Überlegungen. Dafür werde ich eine Notation unter Verwendung von Attribut-Wert-Matrizen ähnlich zur Sign-Based Construction Grammar (SBCG, vgl. Sag, Boas und Kay 2012) verwenden um beispielhaft das entwickelte Schema zu implementieren, ohne einen Anspruch auf Vollständigkeit erheben zu können. Die tiefergehende Ausarbeitung muss Teil anderer Arbeiten sein, diese Arbeit soll vielmehr den Anstoß für Überlegungen in die beschriebene Richtung geben.

## 2. Konstruktionsgrammatik

Die Konstruktionsgrammatik, oft abgekürzt als CxG (in deutschen Texten zuweilen auch KxG, vgl. z.B. Deppermann 2011, 205), ist eine Sammlung von Ansätzen zu verschiedenen linguistischen Fragestellungen (u.a. grammatische, typologische, diachrone und spracherwerbsbezogene), die sich parallel und in enger Verbindung zur Kognitiven Grammatik (Langacker 1987; 2008) und anderen, neueren Theorien der Kognitiven Linguistik<sup>2</sup> in Abgrenzung zur Generativen Grammatik Chomsky'scher Prägung (GG) seit den späten 80er Jahren herausgebildet hat. Erste umfassende Beschreibungen finden sich bei Fillmore (1988), Fillmore, Kay und O'Connor (1988) und Goldberg (1995). Seit Beginn der 2000er Jahre erhält die Konstruktionsgrammatik zunehmend Aufmerksamkeit und wird von vielen Autorinnen als Ausgangspunkt ihrer Arbeiten betrachtet. So erschienen Crofts (2001) *Radical Construction Grammar*, für die er aus typologischer und methodischer Sicht argumentiert (vgl. auch Croft 2004), Goldbergs (2006) Diskussion um Argumentstrukturen: *Constructions at Work*, die Sammelbände *Konstruktionsgrammatik I, II und III*, hgg. von Fischer und Stefanowitsch (2006a (I)), Stefanowitsch und Fischer (2008 (II)) und Lasch und Ziem (2011 (III)) zu einer Bandbreite an Fragestellungen, sowie auf Grundlage der *Head Driven Phrase Structure Grammar (HPSG)* (vgl. dazu u.a. Müller 2008) der stärker auf die grammatische Formalisierung orientierte Ansatz der *Sign Based Construction Grammar* (vgl. Sag, Boas und Kay 2012). Darüber hinaus werden eine Vielzahl von Fragestellungen um spezifische Konstruktionen (vgl. den Sammelband von Fried und Boas 2005; Boas (2003; 2005; 2011); Staffeldt 2011; Höche und Shahrokny-Prehn 2011 u.A.), kognitive Implikationen (vgl. z.B. den Sammelband von Östman und Fried 2005; Bybee 2006; Gries und Wulff 2005; Ellis und Larsen-Freeman 2009) und methodische Aspekte wie die Beurteilung quantitativer Daten und statistische Berechnungsmöglichkeiten (u.A. Stefanowitsch und Gries, 2003; Gries und Stefanowitsch 2004; Tomasello und Stahl 2004; Stefanowitsch 2006) diskutiert.

Wie Stefanowitsch (2011, 11) ausführt, ist

„[d]ie Konstruktionsgrammatik (...) gleichzeitig weniger und mehr als eine Grammatiktheorie. Sie ist weniger, weil sie bis heute keinen vollständigen oder gar allgemein akzeptierten Formalismus bereitstellt, auf dessen Grundlage sich ein erschöpfendes formales Modell der Grammatik einer natürlichen Sprache erstellen ließe. Die Konstruktionsgrammatik ist mehr als eine Grammatiktheorie, weil ihr Erkenntnisinteresse sich grundlegend von dem anderer

---

2 Auch die Generative Grammatik bezeichnet sich als kognitiv-linguistische Theorie. Gemeint ist hier jedoch ein Bündel von neueren Ansätzen, wie sie u.a. in Evans und Green (2006) beschrieben werden.

Theorien unterscheidet: Die Konstruktionsgrammatik will eine allgemeine Theorie vom Erwerb, der Repräsentation und der Verarbeitung sprachlichen Wissens und der Entstehung und dem Wandel grammatischer Systeme sein“.

In diesem Sinn kann die Konstruktionsgrammatik als Think-Tank verstanden werden, der sich aus unterschiedlicher Perspektive und mit unterschiedlichen Betonungen, Annahmen und Auslassungen von Annahmen diesem Ziel nähert. In dieser Arbeit kann kein umfassender Vergleich aller konstruktionsgrammatischen Ansätze erfolgen (für einen Überblick vgl. Fischer und Stefanowitsch 2006b). Die Grundlage aller Ansätze sind aber folgende Prämissen:

1. Die Grundeinheiten natürlicher Sprache sind Form-Bedeutungs-Einheiten auf verschiedenen Ebenen. Diese Einheiten heißen Konstruktionen und können Morphem-, Wort-, Wortgruppen-, Phrasen-, Satz- oder Textgröße annehmen. Unterschiede in der Form gehen mit semantischen, informationsstrukturellen oder pragmatischen Bedeutungsunterschieden einher. Die Konstruktionsgrammatik nimmt somit einen semiotischen Blickwinkel ein.

Während in frühen Arbeiten (z.B. Goldberg 1995) Nicht-Kompositionalität von grammatischen Konstruktionen als kritisches Merkmal gesehen wurde (eine Konstruktion war nur dann eine Konstruktion, wenn sie pragmatische oder semantische Idiosynkrasien aufwies und nicht vollständig kompositionell ableitbar war), gibt es mittlerweile eine breite Akzeptanz für die Existenz von Frequenzeffekten. Demnach ist eine Konstruktion weiterhin eine Form-Bedeutungs-Einheit, kann allerdings auch ganzheitlich gespeichert<sup>3</sup> sein, wenn sie keine über die Bestandteile hinausgehende Bedeutung aufweist, sofern sie häufig genug vorkommt.

2. Konstruktionen sind auf verschiedenen Abstraktionsebenen angesiedelt. Eine scharfe Trennung zwischen Syntax und Lexikon wird nicht angenommen. Die verschiedenen Konstruktionsebenen sind als strukturiertes Inventar zu verstehen, Form- und Bedeutungsbestandteile werden hierarchisch vererbt. Die Konstruktionsgrammatik bezeichnet sich, trotz Abgrenzung zur GG, als generative Theorie, weil sie von der Generierbarkeit neuer Äußerungen durch Interaktion der Konstruktionen ausgeht und zumindest einige Autorinnen erklären wollen, welchen Gesetzmäßigkeiten diese Interaktion folgt.

---

3 Zum Problem der „Speicherung“ s. Abschnitt 3.5 dieser Arbeit.

3. Sprache lässt sich erschöpfend in Form von Zeichen beschreiben. Transformationen, Bewegungen, Prinzipien und leere Kategorien werden nicht angenommen. Über den Grad der Abstraktionen und die Rolle von Komposition gibt es unterschiedliche Auffassungen (vgl. Fischer und Stefanowitsch 2006b).
4. Abstrakte und konkrete Einheiten werden für den Gebrauch vereint. Die Bedeutung entsteht dabei kompositionell aus den semantischen Bestandteilen der Konstruktion, der Füllwörter und des relevanten Kontexts.
5. Sprache sollte gebrauchsbasiert betrachtet werden: Sie ist, wie sie gesprochen und geschrieben wird, Performanz und Kompetenz sind nicht zu trennen. Konstruktionen treten stets in einem reichhaltigen (pragmatischen, syntaktischen, semantischen, kotextuellen etc.) Kontext auf, der die Form der Äußerung mitbestimmt und als relevant anerkannt wird. Es gibt keine peripheren Bereiche der Sprache (vgl. z.B. Fried und Östman 2005, 1753). Versuche einer Kontextmodellierung haben vor Allem im Rahmen der Kognitiven Semantik (vgl. Evans and Green 2006, 695ff.) und der Frame-Semantik (vgl. Fillmore 1982/2011; Martin 2007; Boas 2010; Antonopolou und Nikiforodou 2011 u.a.) stattgefunden.
6. Natürliche Sprache ist aus der sprachlichen Umgebung (gebrauchsbasiert) lernbar. Angeborene grammatische Kategorien oder grammatisch-generative Mechanismen, wie in der Generativen/Universalgrammatik angenommen, sind nicht notwendig (vgl. insb. Tomasello 1995). Der Lernprozess ist ein Konstruktions- und Abstraktionsprozess, bei dem mit gestaltwahrnehmungs- und anderen allgemein-psychologischen Lernmechanismen Formen und Bedeutungen zusammengeführt werden. Sprache und Sprachenlernen ist neurophysiologisch nicht modularisiert.

Im frühen Spracherwerb werden zunächst feste Einheiten gelernt, die der Kommunikation dienlich sind. Im weiteren Erwerb findet eine Zerlegung der Einheiten statt, indem durch Erfahrung aus dem Input gelernt wird, dass die Gesamteinheit aus austauschbaren Teilen besteht. Am Ende steht der Abstraktionsprozess über verschiedene vergleichbare Einheiten, der in der Entstehung abstrakter Konstruktionen mit nur wenigen oder ganz ohne feste Bestandteile mündet (etwa in spezifischen Verargumentstrukturen: [NP<sub>AKK</sub> zu NP<sub>DAT</sub> *einladen*], oder in verbunabhängigen Konstruktionen, wie der Ditransitivkonstruktion: [AGENS VERB BENEFITEE PATIENS]).

7. Die meisten Ansätze sind sich darüber hinaus einig darüber, dass die Betrachtung gebrauchsbasierter Daten für die Betrachtung von unabhängig erhobenen Daten spricht, vor Allem für Korpusdaten. Deren Betrachtung wird im Allgemeinen der Betrachtung introspektiv erdachter Daten vorgezogen.<sup>4</sup>

Arbeiten, die auf der Konstruktionsgrammatik fußen, aber oft nicht explizit grammatiktheoretisch, orientiert sind, variieren im Grad der Einbeziehung weiterer Annahmen. Dazu zählen Embodiment (das Wurzeln sprachlicher Konzepte in körperlichen Erfahrungen) und Simulation (das neurophysiologisch-reale Nachempfinden bei der/zwecks Bedeutungsdekonstruktion), vgl. z.B. Bergen und Chang (2005), die Rolle aufmerksamkeitssteuernder Elemente (Relevanztheorie, Salienz, vgl. Enikő und Károly 2010; Ellis und Collins 2009), sowie Konnektionismus (neurophysiologische Netzwerkmodelle, in denen Frequenz, Salienz und weitere Faktoren eine Rolle beim Erwerb spielen, vgl. z.B. Ellis 1996; 1998; 2003; 2011; Hoey 2005).

## **2.1 Theoretische Unschärfen**

Als vergleichsweise junge Theorie muss die Konstruktionsgrammatik ihre Wissenschaftlichkeit seit ihrer Entstehung gegenüber einer starken GG-Tradition verteidigen. Das gilt insbesondere aufgrund der grundsätzlichen theoretischen und methodischen Unterschiede, die ihr weiterhin häufig als Unwissenschaftlichkeit ausgelegt werden. In der Darstellung wurden von den bekanntesten Theorieschreiberinnen Aspekte der Lernbarkeit von Sprache und der Kontinuität von Syntax und Lexikon, Pragmatik und Kontextwissen in den Vordergrund gestellt (vgl. z.B. Goldberg 1995; 2006; Croft 2001; 2004; Tomasello 1995; 2003).

Die gebrauchsbasierte Ausrichtung der Konstruktionsgrammatik bezieht sich derzeit hauptsächlich auf den (Erstsprach-)Erwerb aus dem Gebrauch einerseits und die Betrachtung von in Korpusstudien gefundenen, und dadurch als im Gebrauch nachgewiesenen, Strukturen andererseits. Zu weiteren Aspekten des Gebrauchs, zum Beispiel der Interaktion und Integration von Konstruktionen verschiedener Abstraktionsgrade, der konkreten Konstruktionsauswahl im

---

4 Zum Teil aus praktischen Gründen – der weitaus überwiegende Teil von Sprache wird nicht in Korpora erfasst, vielmehr erfassen Korpora nur sehr spezifische Ausschnitte – und zum Teil wegen mangelnder Stringenz trifft das allerdings nicht auf alle Arbeiten zu. So arbeitet z.B. Goldberg (1995; 2006) nicht ausschließlich mit Korpusdaten. Das ist problematisch, da, wie u.A. Sinclair (1991, 4) und Wray (2002, 21) herausarbeiten, Intuitionen über Sprache äußerst unzuverlässig und instabil sind, und die exemplarische Darstellung zu fehlgeleiteten Schlussfolgerungen über die Zentralität der gewählten Beispiele führen kann.



Gebrauch, kommunikativen, sowie Kontext- und Diskursaspekten, wird die Theorie erst in jüngster Zeit weiterentwickelt. Das ist zum Teil der Methodik geschuldet, die aus der bestehenden umfangreichen Datenlage viele Hinweise und kleinere Erkenntnisse gewinnt, die eine umfassende Extrapolation über die Daten hinweg aber aus Praktikabilitätsgründen erschwert. Ein zweiter Grund liegt in der Unübersichtlichkeit der zahlreichen sinnverwandten Ansätze, die zu einer gewissen Begriffsinflation führen, ohne dass Implikationen, Gemeinsamkeiten und Unterschiede ausdiskutiert und integriert würden. So stehen allein in der Grammatiktheorie Pattern Grammar (Hunston und Francis 2000), New Word Grammar (Hudson 2007) und Linear-Unit Grammar (Sinclair und Mauranen 2006) nebeneinander, ohne die theoretische Vereinheitlichung zur Konstruktionsgrammatik – die größer und ausführlicher modelliert ist – zu suchen, wo sie sich anbietet. Dazu kommen zahlreiche Beobachtungen aus der Erst- und Zweitsprachforschung, der Korpus- und Psycholinguistik, der Pragmatik und der Fremdsprachendidaktik, die nur in wenigen Fällen aufeinander bezogen und in einen gemeinsamen theoretischen Rahmen eingebunden werden. Gerade weil die Konstruktionsgrammatik, wie im Zitat von Stefanowitsch (2011, s.o.) ausgeführt, mehr sein will als eine Grammatiktheorie – eine Gesamtheorie der Sprache, im Idealfall – und bisher weniger ist als eine Grammatiktheorie, wäre das umso wünschenswerter.

Die größte theoretische Unschärfe besteht darin, dass es keinen gemeinsamen Formalismus gibt, der Gemeinsamkeiten und Unterschieden in den Annahmen verdeutlichen würde. Auch deswegen werden Unklarheiten in der Begriffsschreibung leicht übersehen, wie etwa die Betrachtung von Lexikon und Grammatik als Kontinuum, die, wie sich im Folgenden zeigen wird, auch für die Betrachtung von Chunks interessant ist. So beschreiben alle Konstruktionsgrammatiken dieses Kontinuum, argumentieren bei der Analyse allerdings sehr weitgehend mit einer Interaktion von abstraktesten und konkretesten Einheiten – Satzmustern und Wörtern – was, wie Boas (2010) herausstellt, de facto eine Unterscheidung zwischen Lexikon und Syntax ausmacht.

Weiterhin gibt es bislang jenseits der grundsätzlichen Feststellung, dass das Konstruktionsinventar strukturiert sein muss und hierarchisch vererbt, keine umfassende Diskussion um die Organisation dieses Inventars (oft als Konstruktikon bezeichnet) und die Verknüpfung zu Lexikon-, Kategorisierungs-, Psycholinguistik- und Pragmatiktheorien ist bislang nicht breit vorgenommen worden. Ein zusätzlicher Diskussionsaspekt betrifft die Frage des semiotischen Charakters von abstrakten Konstruktionen – die zentrale Annahme der Konstruktionsgrammatik ist, dass auch abstrakte Konstruktionen „echte“ Zeichen mit abgrenzbaren Form- und Bedeutungsseiten sind. Während bei Wörtern die Formseite durch die phonetische oder graphemische Form repräsentiert

wird, ist die genaue Formseite von abstrakten Konstruktionen, wie Verhagen (2009) herausarbeitet, unklar. Es gibt keine einheitliche phonetische oder graphemische Repräsentation, sondern nur Symbole, die auf Symbole verweisen. Diese Schwierigkeit wird selten angesprochen oder diskutiert, hat aber Implikationen für die Organisation des Konstruktions.

Die derzeitige Praxis der Konstruktionsgrammatik ist daher von starker Heterogenität hinsichtlich der Forschungsinteressen und des theoretischen und methodischen Anspruchs geprägt. Es existieren wenige theoretische Überarbeitungen und Modellierungen unter Einbeziehung der geführten Diskussionen, obwohl reichhaltiges empirisches Material zu einer Reihe von Fragestellungen vorliegt.

### 3. Was sind Chunks? - Begriffsklärung und Forschungsstand

Während die CxG Schwierigkeiten mit der Erklärung ihres Lexikon-Grammatik-Kontinuums hat, aber doch zumindest relativ sicher sehr statische Konstruktionen und Wörter von flexiblen abgrenzt, zeichnen sich bei näherer Betrachtung Probleme selbst damit ab. Aus denjenigen Subdisziplinen der linguistischen Forschung, die sich genauer mit solchen Einheiten befassen, geht keine einheitliche Definition der „konkreten Konstruktion“ hervor:

„One of the reasons it is difficult to define formulaic sequences lies in their diversity. For example, formulaic sequences can be long (*You can lead a horse to water, but you can't make him drink*) or short (*Oh no!*), or anything in between. They are commonly used for different purposes. They can be used to express a message or idea (*The early bird gets the worm* = do not procrastinate), functions (*[I'm] just looking [thanks]* = declining an offer of assistance from a shopkeeper), social solidarity (*I know what you mean (...)*), and to transact specific information in a precise and understandable way (*Wind 28 at 7* = in aviation language this formula is used to state that the wind is 7 knots per hour from 280 degrees). (...) These sequences can be totally fixed or have a number of 'slots' which can be filled with appropriate words or strings of words (...)" (Schmitt und Carter 2004, 3).

Es werden zahlreiche unterschiedliche Eigenschaften und Klassifikationssysteme beschrieben – Wray (2002, 44f.) nennt über 20 Klassifikationen – und auch die Zuschreibung von Funktionen, die kognitive Modellierung und die Relevanz von holistischen Einheiten im Sprachgebrauch wird sehr unterschiedlich bewertet (vgl. Gries 2007; Granger und Paquot 2007). Die Problematik beginnt bereits bei der Begriffsfindung: Ausgehend von verschiedenen Forschungstraditionen und Erklärungsinteressen werden rekurrierende, statische oder teilweise statische Einheiten

unterschiedlich definiert und genannt. Wray (2002, 9) zählt die folgenden 56 Begriffe auf:

„amalgams - automatic - chunks - clichés - co-ordinate constructions - collocations - complex lexemes - composites - conventionalized forms - F[ixed] E[xpressions including] I[dioms] – fixed expressions - formulaic language - formulaic speech - formulas/formulae - fossilized forms - frozen metaphors - frozen phrases - gambits - gestalt - holistic - holophrases - idiomatic - idioms - irregular - lexical simplex - lexical(ized) phrases - lexicalized sentence stems - listemes - multiword items/units - multiword lexical phenomena - noncompositional - noncomputational - nonproductive - nonpropositional - petrifications - phrasemes - praxons - preassembled speech - precoded conventionalized routines - prefabricated routines and patterns - ready-made expressions - ready-made utterances - recurring utterances - rote - routine formulae - schemata - semipreconstructed phrases that constitute single choices - sentence builders - set phrases - stable and familiar expressions with specialized subsenses - stereotyped phrases - stereotypes - stock utterances - synthetic - unanalyzed chunks of speech - unanalyzed multiword chunks - units“.

Diesen wären mindestens noch die Begriffe *formulaic sequence* (nach der Aufzählung von Wray 2002 verwendet), *collostruction/Kollostruktion* (Stefanowitsch und Gries 2003), *colligation/Kolligation* (Sinclair 1991; Hoey 2005; Römer 2005), *phraseologism* (Gries 2007), *conventionalized phrases* (Erman und Warren 2000) und *syntagmatic pattern* (Sinclair 2007) hinzuzufügen. Obwohl es zahlreiche Überschneidungsflächen gibt, sind die Begriffe theoretisch unterschiedlich gefärbt und nicht immer synonym zu verstehen. Erschwerend kommt hinzu, dass viele Begriffe undefiniert und wechselweise verwendet werden, und dass theoretische Implikationen undiskutiert bleiben.

Der Definition und Diskussion vorausgreifend werde ich im Folgenden die beschriebenen Einheiten zum Zweck der Einheitlichkeit als Chunks bezeichnen. Zum Ende des Abschnitts soll diese Wahl hergeleitet und erklärt werden.

### **3.1 Phraseologie**

Die Phraseologie ist, auch wenn sie keine eigenständige Forschungstradition ist, die ureigene Disziplin zur Betrachtung von Chunks. Sie ist hauptsächlich aus der Lexikographie hervorgegangen, heute werden aber oft auch Erkenntnisse zu Chunks aus psycho- und korpuslinguistischer sowie l2-didaktischer Forschung unter dem Begriff der Phraseologie subsumiert (vgl. Schmitt 2004; Granger und Meunier 2007b; Gries 2007). Die L2-Didaktik versucht, häufig beschriebenen Schwierigkeiten von weit fortgeschrittenen L2-Lernerinnen,

Natürlichkeit bzw. Muttersprachenähnlichkeit in der Zielsprache zu erreichen, beizukommen (vgl. Mauranen 2009; Handwerker und Madlener 2009; Li und Schmitt 2009; Crossley und Salsbury 2011). Interessanterweise dringt die Erkenntnis, dass L1-Expertise also wahrscheinlich mit Chunks zu tun hat, nicht weit bis in die L1-Beschreibung vor, sondern bleibt auch in modernen Theorien wie der Konstruktionsgrammatik hinter der Annahme zurück, Sprache sei in erster Linie kreativ oder generativ.

Aus lexikographischer Perspektive besteht ein besonderes Interesse, lexikalische Einheiten und Wendungen und Ausdrücke, in denen sie vorkommen, sinnzusammenhängend darzustellen. Das trifft z.B. dann auf Schwierigkeiten, wenn unklar ist, unter welchem seiner Bestandteile ein Chunk, etwa ein Idiom wie *ins Gras beißen* oder eine Routine wie *Guten Tag!*, geführt werden soll. Die Schwierigkeit allein wäre durch schlichte Konvention lösbar, aber sie macht deutlich, dass Chunks Fragen zur Kompositionalität und Konventionalität von Sprache im Gebrauch ansprechen. So hat sich die Phraseologie in Abgrenzung zur GG-Tradition entwickelt hat, indem sie sich auf den der „sprachlichen Peripherie“ zugewiesenen Bereich der idiomatischen Wendungen und Kollokationen konzentriert hat – gerade weil die Generative Grammatik von einer weitgehenden Analysierbarkeit der Sprache in ihre Bestandteile ausgeht, sich die Beobachtungen in der Lexikographie und der L2-Didaktik aber häuften, dass sich Sprache im Gebrauch anders verhält. Der Peripherie zugewiesen wurden in erster Linie Einheiten, die semantisch oder syntaktisch nicht kompositionell ableitbar sind. Entsprechend folgt für die Definitionen aus dem Bereich der Phraseologie, insbesondere aus der Lexikographie, eine Konzentration auf Nicht-Kompositionalität (Idiome) und Konventionalität des gemeinsamen Auftretens (Kollokation, konversationelle Routine, Floskel u.a.).

Der Kollokationsbegriff beschreibt lexikalische Präferenzen von lexikalischen Einheiten, wie etwa *starker Kaffee* statt *?kräftiger Kaffee* oder *eine große Frau* statt *??eine hohe Frau*. Er wird von verschiedenen Autorinnen unterschiedlich verwendet, zum Teil in Abgrenzung von Idiomen, zum Teil auch als Oberbegriff für verschiedene Arten von Chunks, und mit oder ohne Bezug auf statistische Eigenschaften wie Nähe der Einheiten im Text oder Kollokationsstärke (ein Maß, das die Stärke der Bindung zwischen zwei lexikalischen Einheiten gemessen an ihrem jeweiligen Einzelvorkommen ausdrückt, vgl. dazu Stefanowitsch und Gries (2003), Gries und Stefanowitsch (2004)). Sinclairs (1991) Definition von Kollokation wird besonders oft herangezogen:

„Collocation is the occurrence of two or more words within a short space of each other in a text. The usual measure of proximity is a maximum of four words intervening. Collocations can be dramatic and interesting because unexpected, or they can be important in the lexical structure of

the language because of being frequently repeated. This second kind of collocation, often related to measures of statistical significance, is the one that is usually meant in linguistic discussions (...). **Collocation in its purest sense (...) recognizes only the lexical co-occurrence of words.** This kind of patterning is often associated with grammatical choices as well, leading to the wealth of idioms and fixed phrases that are found in everyday English. Some writers on collocation (...) include matters of grammatical relations in their specification of collocation“ (S. 170, meine Herv. - AS).

In der Regel werden Kollokationen als kompositionell und Idiome als nicht-kompositionell betrachtet, wobei die Abgrenzung unscharf ist:

„An idiom is a group of two or more words which are chosen together in order to produce a specific meaning or effect in speech or writing. (...) The individual words which constitute idioms are not reliably meaningful in themselves, because the whole idiom is required to produce the meaning. Idioms overlap with collocations, because they both involve the selection of two or more words“ (ebd., 172).

Sinclair formuliert vergleichsweise vorsichtig die von vielen Autorinnen weiterhin als kritische Eigenschaft betrachtete Nicht-Kompositionalität. Obwohl seit einiger Zeit zunehmend darauf hingewiesen wird, dass Kompositionalität und Idiomatizität gradierbar sind (vgl. z.B. Svensson 2007; Wulff 2008), ist der Grundtenor vieler Arbeiten weiterhin eine klare Abgrenzung kompositioneller von idiomatischen Einheiten (vgl. z.B. Eрман und Warren 2000; Dobrovol'skij 2011; Stathi 2011).

Während Kollokationen und Idiome relativ fix erscheinen (obwohl auch sie flexibel gehandhabt werden und in Korpora vielfach nur in abgewandelter Form zu finden sind, vgl. Moon 1999), gibt es die Kategorie der lexikalisch nur teilweise oder gar nicht spezifizierten Konstruktionen, die ebenfalls ins Interessengebiet der Phraseologie fallen: Dobrovol'skij (2011, 114f.) führt z.B. [*es/das* IST zum N<sub>inf</sub>] – *Es ist zum Verrücktwerden; Es war zum Kotzen*; [PP HAT gut V<sub>inf</sub>] – *du hast gut lachen*; oder [N<sub>1</sub> hin, N<sub>1</sub> her] – *Freund hin, Freund her; Krise hin, Krise her* für die teilweise spezifizierten Konstruktionen, und [X KOPULA X] – *futsch ist futsch, Krieg ist Krieg, gestohlen bleibt gestohlen*, oder [von N<sub>1</sub> zu N<sub>1</sub>] – *von Tag zu Tag* für die nicht-spezifizierten an. Im Widerspruch zu diesen Beobachtungen kommentiert er zuvor:

„Die Produktivität dieser Konstruktionen [Resultativa, Ditransitiva, *caused motion constructions* etc., AS] ist zwar semantisch und zum Teil auch lexikalisch begrenzt, **sie haben aber keine lexikalische Spezifizierung, d.h., sie weisen keine Elemente auf, die lexikalisch**

**fixiert sind, was eines der wichtigsten Kriterien für Phraseme ist. Ein Phrasem darf nicht aus Slots einer syntaktischen Struktur bestehen, die nach bestimmten Kriterien gefüllt sein müssen. Mit anderen Worten, eine Konstruktion ohne lexikalische Anker ist kein Phrasem.** Da aber neben den phraseologischen und quasi-phraseologischen Entitäten im Rahmen der KxG auch solche sprachlichen Phänomene untersucht werden, die nicht zur Phraseologie gerechnet werden können, sondern traditionell in den Gegenstandsbereich der Syntax gehören, darf die KxG nicht mit der Phraseologie identifiziert werden“ (113, meine Herv. - AS).

Er zählt als Bezeichnungen für diese Art Konstruktionen die Begriffe *modellierte Bildungen*, *syntaktische Phraseme*, *syntaktische Idiome*, *Phraseoschemata*, *lexikalisch offene Idiome*, *Phraseoschablonen* und *Konstruktionsidiome* auf und entscheidet sich selbst für den Begriff der *Phrasem-Konstruktion* (ebd.). Er unterscheidet hinsichtlich der anderen Kategorien zwischen Idiomen („bildliche[n] Phraseme[n] mit stark ausgeprägten Irregularitäten in Bezug auf das Form-Bedeutungs-Verhältnis“), Kollokationen („die mit ihrer „freien“ lexikalischen Bedeutung (Basis) und einer gebundenen, entweder semantisch abgeschwächten oder metaphorisch umgedeuteten Konstituyente (Kollokator) bestehen“), situativen Klischees bzw. Routineformeln (*Guten Tag!*), Sprichwörtern und grammatischen Phrasemen (*wie dem auch sei*, *geschweige denn*) (111f.). Dobrovolskijs Arbeit ist ein gutes Beispiel für den inflationären Umgang mit Begriffen in Arbeiten, die sich der Phraseologie zuordnen, die methodisch, theoretisch und begrifflich nicht klar definiert oder von anderen Bereichen abgegrenzt arbeitet.

Oft in der Phraseologie zitiert, aber deutlich anderen Charakters sind die englischen *phrasal verbs*, die deutliche Überschneidungen mit dem Interessensgebiet der Grammatiktheorie haben. Ebenfalls in diesem Bereich liegt die *Kolligation* (z.B. Sinclair 1991; Hunston 2001; Hoey 2004; Römer 2005), die das Auftreten von lexikalischen Elementen in spezifischen grammatischen Zusammenhängen oder Positionen, wie etwa Subjekt- oder Objektstatus, aber auch im Kontext von spezifischen Tempora, Modi oder Aspekten etc. beschreibt. Man könnte sagen, Kolligationen sind grammatische Präferenzen lexikalischer Einheiten. Hier bieten sich Überschneidungsflächen zur *Kollostruktion*, ein Begriff, der in der Konstruktionsgrammatik hauptsächlich von Stefanowitsch und Gries (2003; vgl. auch Gries und Stefanowitsch 2004) geprägt wurde und das gemeinsame Auftreten von (abstrakter) Konstruktion und (spezifischem) lexikalischem Element beschreibt. Der Unterschied besteht in der theoretisch größeren Offenheit des Kolligationsbegriffs, der mit verschiedenen Grammatiktheorien vereinbar ist und keine theoretischen Annahmen über den semiotischen Status (zum Beispiel den Unabhängigkeit und den Abstraktionsgrad) der syntaktischen

Umgebung macht.

Der Kollostruktionsbegriff ist im Vergleich dazu recht eindeutig in der Konstruktionsgrammatik verwurzelt und übernimmt alle Annahmen über die umgebende Konstruktion. Zusätzlich ist deren Blickwinkel etwas verschoben: Es geht eher um lexikalische Präferenzen grammatischer Einheiten als, wie beim Kolligationsbegriff, umgekehrt. Außerdem wird im Zusammenhang mit Kollostrukturen im Regelfall die Kollostruktionsstärke von Lexemen und Konstruktionen in Korpora gemessen. Die Kollostruktionsanalyse ist daher stärker in der Korpuslinguistik verortet, indem sie eine Methodik für die quantitative Auswertung darstellt. Beide Begriffe gehen von der Abstraktion und damit einer gewissen Eigenständigkeit der Umgebungsgrammatik und der *Einfügung* von ebenfalls vergleichsweise unabhängigen lexikalischen Elementen aus, einem *Merge*-Vorgang also. Modelle, in denen etwas zusammengefügt wird, implizieren ein vorheriges Einzelbestehen der Bestandteile. Ausgangspunkt der Kollostruktion ist damit nicht das gesamte Chunk, sondern die Betrachtung der einzelnen Teile und deren Interaktion, wobei die Methodik natürlich geeignet sein kann, um das stetig gemeinsame Vorkommen von lexikalischen Einheiten zu beweisen bzw. zu widerlegen.

### **3.2 Psycholinguistik und L2-Forschung**

Während die lexikographische und didaktische Phraseologie ihre Beschreibungen hauptsächlich in der deskriptiven und präskriptiven Beschreibung des *Sprachsystems*, nach Saussure der *langue* (Saussure 1916/1967),<sup>5</sup> verortet, erweitert der Blick auf Verarbeitung, Produktion und Funktionalität von Sprache die Inhalte von Chunks. Die Psycholinguistik betrachtet vor Allem die *parole* oder *I-Language* (nach Chomsky 1965). Bei der Betrachtung von psycholinguistischen Ansätzen ist zunächst augenfällig, dass hier andere Einheiten betrachtet werden. Die lexikographisch-phraseologischen Ansätze befassen sich ausschließlich mit Einheiten, die größer als ein Wort sind, und sind im Regelfall nicht an Äußerungsgrenzen interessiert. Die L2-Didaktik kann sich im Rahmen der pragmatischen Ausbildung zwar für Interjektionen, die keine strikte Wortgröße haben (wie *oh!* oder *aha!*), interessieren, im Regelfall werden aber größere Einheiten betrachtet. Die

---

5 Die *langue/parole*-Dichotomie wird unterschiedlich interpretiert und manchmal mit der *Kompetenz/Performanz*-Dichotomie nach Chomsky gleichgesetzt *parole* mit *Performanz* (vgl. Bybee 2006). Eine andere Lesart, und die hier gemeinte, bezeichnet mit *langue* die Sprache als gemeinsames System oder, weniger strukturalistisch, als Schnittfläche der Idiolekte der Einzelsprecherinnen, und verortet die *parole* relativ nah am Idiolekt-Begriff, wobei Aspekte der Sprachverarbeitung einbezogen werden. Insofern ist *parole* eher vergleichbar mit Chomskys *I-Language*, die Kompetenz und Performanz umfasst, als mit Performanz allein.

Bezeichnungen variieren dabei erheblich (zum Beispiel sind im Sammelband von Meunier und Granger 2008 u.a. die Begriffe *memorized language, phrases, formulaic sequences* und *multiword expression* in synonymer Verwendung zu finden). Strikter aus der psychologischen Tradition kommende Ansätze betrachten dahingegen auch deutlich kleinere und, vor Allem, nicht nur sinntragend-abgeschlossene Einheiten. So zählt Ellis (1996; 2003) auch Lautabfolgen wie typische Konsonantencluster oder Silbenbestandteile (z.B. V-C-V oder C-V-C) als Chunk. Der Chunkbegriff bei Ellis nicht auf sprachliche Phänomene beschränkt, er zählt auch Zahlenreihen oder komplexe musterhafte Buchstabenreihen, die keinen erkennbaren Sinn ergeben, aber trotzdem von Probandinnen als Reihen erkannt werden, auf. Allerdings betont Ellis (1996; 1998; 2011) die Bedeutung von gleichbleibenden Einheiten für den Spracherwerb und die Sprachproduktion. So stellt Ellis (1998) fest, dass die Sprachproduktionsfähigkeit in L1 und L2 stark mit der Segmentierungsfähigkeit, der Fähigkeit zur Entnahme von gleichbleibenden und abweichenden Einheiten aus dem Redefluss, korreliert. Er betont, wie Sinclair (2007), besonders die Fähigkeit des Menschen, in zwei Richtungen verstehen und abstrahieren zu können: Vom Kleinen zum Großen (Chunking) und vom Großen zum Kleinen (Sequenzierung) (vgl. auch Ellis 2003). Entsprechend ist die erste Aufgabe im Spracherwerb, die jeweils größten verarbeitbaren Bestandteile zu isolieren – je nach Kapazität des Arbeitsgedächtnisses sind das Phoneme, Ein-Wort-Chunks<sup>6</sup> oder größere Chunks – um sie kommunikativ gebrauchen zu können. Erst im späteren Erwerb folgt die Erkenntnis, dass die verwendeten Einheiten weiter zerlegbar sind. Dieser Prozess ist, wie Wray (2002, Teil III) beschreibt, nicht linear und kann zu Über- und Untergeneralisierungen und dadurch zur nicht-zielsprachlichen Zerlegung oder Beibehaltung von Chunks führen. Ellis wechselt zwischen den Begriffen *chunk* (1996) und *formula, formulaic chunk, memorized chunk* (2008). Schmitt (Hg., 2004) übernimmt in seinem Sammelband zu hauptsächlich psycholinguistischen Fragestellungen den Begriff *formulaic sequences* von Wray (2002) (Schmitt und Carter 2004; Underwood et al. 2004 u.a.).

Hinsichtlich der Abstraktionsfähigkeit von Chunks zu Konstruktionen im späten L2-Erwerb schreiben Handwerker und Madlener (2009, 12):

---

6 Wie Sinclair (1991, 176) feststellt, ist es nicht immer treffend, vom Gebrauch von „Wörtern“ im frühen L1-Erwerb zu sprechen, da ein „Wort“ eine analysierte Einheit in Abgrenzung zu anderen Einheiten darstellt. Es ist eher davon auszugehen, dass Sinneinheiten aus dem Redefluss extrahiert und übernommen werden, deren Analyse in „Wörter“, „Sätze“ und andere Einheiten erst später erfolgt.



„Im Unterschied zum frühkindlichen Erwerb nutzen Schüler und Erwachsene ihre Chunks zwar teilweise bewusst zu vielfältigen Zwecken, **sie nutzen die Chunks aber nicht systematisch und sie brechen die Chunks nicht lernbegierig auf, um eine Basis für Abstraktionen und Generalisierungen zu schaffen**. Sie brauchen – so inzwischen die wohl einhellige Meinung in Forschung und Lehre – den steuernden Eingriff mit expliziter Unterweisung im Umgang mit dem Vorgefertigten (...)“ (meine Herv., AS).

Insbesondere der hervorgehobene Teil dieser Aussage wird nicht unbedingt von Datenmaterial gestützt. Neben der zweifelhaften Aussage, dass Lernbegierigkeit und Systematizität direkt an der flexiblen Verwendung von Chunks gemessen werden können, sprechen einige Untersuchungen dafür, dass auch L2-Lernerinnen in der Lage sind, abstrakte Strukturen in Chunks zu erkennen bzw. sie daraus abzuleiten. Gries und Wulff (2005) diskutieren anhand von L2-Daten von fortgeschrittenen Lernerinnen Konstruktionsbegriffs in der Zweitsprache und kommen zum Ergebnis, dass ihre Daten am ehesten durch vorliegende muttersprachenähnliche Abstraktionen des Englischen erklärbar sind:

The experimental study has yielded several interesting results: First, we have seen that our foreign language learner subjects perceive sentences as much more similar to each other when they instantiate identical constructions than when they feature the same verbs. This tendency was surprisingly strong and clearly emerged both from the comparison of reclassification frequencies (across all experimental sentences) and, more graphically, from the dendrogram (for each experimental sentence). Apparently, then, the subjects relied more on the complement configurations and the corresponding semantics than on the superficial lexical similarity for sorting and that, even for language learners, argument structure constructions “are psychologically real linguistic categories that speakers use in comprehension” (Bencini & Goldberg, 2000, pp. 649f.)“ (194)

Wray und Fitzpatrick (2008) stellen bei der Untersuchung von Chunks bei L2-Lernerinnen auf nicht-zielsprachliche Abweichungen eine absteigende Reihenfolge von phrasalen, lexikalischen und morphologischen Veränderungen fest. Diese Veränderungen durch die Lernerinnen deuten darauf, dass Chunks durchaus „aufgebrochen“ und – wenngleich nicht immer zielsprachlich akzeptiert – angepasst werden. Wible (2008) weist auf die modalen Unterschiede beim Erwerb als mögliche Erklärung für die unterschiedlichen Beobachtungen im L1- und L2-Erwerb hin:

„(...) the „whole to parts“ view of lexical chunks is that it rests in an assumption which, while convincing for child L1 acquisition, is unwarranted in the case of a substantial portion of the

second language learners in the world. Specifically, it assumes that the primary target language input to the learner (L1 or L2) is spoken input. In the case of spoken input, word boundaries are not reliably flagged in the speech stream. It is for precisely this reason that learners who rely on spoken input must discover these boundaries and the identity of words between them through such delicate, indirect means, for example by contrasting them with related variations (...) [T]ext-oriented learners [adult classroom learners, AS] require from researchers a different construal of the challenge that multiword expressions pose for the learner. Basically, for them the task is not one of unbundling what they first considered to be unanalyzed wholes without parts and discovering the word boundaries and the words inside, but the reverse: discovering that some sequences of the discrete units occurring between white spaces in text are in some respects best considered as bundled wholes despite the lack of typographical evidence that this is so“ (167).

Zur Einordnung der Erkenntnisse, auf die sich Handwerker und Madlener (s.o.) beziehen, sollte zusätzlich die normative Kraft des L2-Unterrichts berücksichtigt werden, der in nicht selten vom Experimentieren mit und Ausprobieren von Strukturen abrät, und zusätzlich in der Regel keinen direkt kommunikativen Zweck verfolgt. Es ist denkbar, dass dieser letzte, funktionale, Aspekt eine Rolle bei der Auswahl der Erwerbsstrategien spielt.

Aus der L2-Forschung wird auch deutlich, dass Chunks offenbar eine Rolle bei der Herstellung von Sprachflüssigkeit spielen, was mit den Grenzen des Arbeitsgedächtnisses zu tun haben kann. Stengers et al. (2011) haben in ihrer Analyse L2-Lernerinnen, denen sie gezielt Chunks für spezifische Konversationssituationen beibrachten, festgestellt, dass die Bewertung der Sprachflüssigkeit und Natürlichkeit durch Muttersprachlerinnen dann besser ausfiel, wenn mehr Chunks verwendet wurden. Es geht dabei anscheinend nicht darum, wie viele Sprechpausen gemacht wurden, sondern um die Verteilung derselben: Sprechpausen an Chunkgrenzen wurden nicht als störend empfunden, wohl aber Sprechpausen innerhalb von Chunks. Dieser Aspekt der Ergebnisse, der eine Hörerwartung ausdrückt, spricht dafür, dass Chunks kognitive Realität besitzen.

Eine ähnliche Herausforderung, die Rezipientenerwartungen der Zielsprache zu erfüllen, betrifft die Fähigkeit stark normierte Texte, insbesondere wissenschaftliche Texte in der L2 zu verfassen. Auch hier finden Chunks besondere Beachtung, die anscheinend bei der Bewältigung dieser Herausforderung stark von Nöten sind. Das findet sich in zahlreichen Artikeln zur Thematik von *English for Academic Purposes* wieder (vgl. z.B. Kerz und Haas 2009; Ellis et al. 2008; Simpson-Vlach und Ellis 2010; Coxhead 2012).

### 3.3 Pragmatik und Diskursanalyse

Ein wichtiger Begriff der Phraseologie, der eigentlich pragmatischen Inhalt hat, ist der der *Formelhaftigkeit (formulaicity)*. Obwohl auch dieser Begriff selten diskutiert wird, ist er nicht theorieneutral. Formelhaftigkeit impliziert eine Anwendungs- bzw. Interaktionsperspektive, und darüber hinaus eine gewisse Abgeschlossenheit einer Einheit. Das wird verstärkt durch die Verwendung in *routine formula*, und in der Verknüpfung *conversational routine* (vgl. z.B. Bladas 2012). Als formelhaft wird verstanden, was als zusammengehörig und funktional einsetzbar erkannt wird – allerdings nicht notwendigerweise, gemessen an Maßstäben, die objektiver und zuverlässiger sind als die Intuition der Sprecherinnen, was zusammengehörig *ist*. Die Gefahr ist entsprechend, von weitaus geringeren Anteilen von „formelhaften Ausdrücken“ auszugehen, als tatsächlich in der Sprache zu finden sind.

Von besonderer Bedeutung für die Diskursanalyse sind solche Chunks, die informationsstrukturierend wirken: Biber, Connor und Upton (2007, 4) zählen Diskursmarker, Konnektoren, Diskurspartikeln, *lexical phrases* und *cue phrases* auf. Antonopoulou und Nikiforidou (2011) zählen als weitere Begriffe *discourse construction*, *discourse pattern* und *conventional discourse* auf, wobei diese Begriffe nicht auf lexikalisch identische Chunks beschränkt sind, sondern eine Breite von Phänomenen bezeichnen: So wird anfangs (S. 2594) auf die Verwendung von *discourse construction* für die Beschreibung von Assoziationen zwischen Textsorte (argumentativ, narrativ, deskriptiv etc.) und Genre (Rezept, Märchen, Nachruf etc.) verwiesen. Hier ist das Bezeichnete relativ abstrakt und umfassend. Später (S. 2596f.) wird derselbe Begriff für die Beschreibung einer Auflistung von Eröffnungsfloskeln in Telefongesprächen verwendet. Diese Verwendung wiederum ist sehr konkret und relativ spezifisch.

Die Relevanz von Chunks für die Erstellung von Texten gemäß einer Textsorte, wie zum Beispiel Abschiedsfloskeln für Briefe oder narrative Erzählmuster wie *and they lived happily ever after* für Märchen, ist intuitiv zugänglich. Die sprachliche Intuition wird in diesem Fall vom Bildungssystem gestützt, das das Erlernen von geeigneten Textbausteinen für Textsorten wie Zusammenfassungen, Interpretationen oder wissenschaftliche Arbeiten vorsieht.

Argumentationsmuster, unabhängig von ihrer konkreten lexikalischen oder konstruktionellen Ausformung, werden von Kanoksilapatham (2007a und 2007b) als *moves* bezeichnet, konkrete lexikalisch basierte Elemente von Biber und Jones (2007) und Csomay (2007) als *vocabulary based discourse units (VBDU)*. *Moves* sind dabei als kommunikationsfunktionale Kategorie zu verstehen, während VBDUs eine stärker sprachsystematische Kategorie darstellen.

Chunks, die für schriftliche Diskurse gezielt gelernt werden, in schriftlichen Diskursen zu suchen, ist ein eher zirkuläres Vorgehen (obwohl die statistische Auswertung interessant sein kann). Interessanter ist die Betrachtung von informationsstrukturierenden Einheiten und anderen Chunks im Rahmen der Pragmatik, insbesondere der Interaktionalen Linguistik (IL, vgl. Deppermann 2011) und der Konversationsanalyse (vgl. z.B. Fried und Östman 2005), die sich mit der Sprache im Gespräch befassen. Diese Arbeit bietet leider nicht den Rahmen, näher auf diesen Bereich einzugehen, es sei aber darauf verwiesen, dass gerade im Gespräch Aspekte wie Alignment (die Anpassung der Sprecherinnen aneinander, was zum Beispiel räumliche, zeitliche, syntaktische oder semantische Bezugsrahmen angeht; vgl. Pickering und Garrod 2006; Watson et al. 2004), Ko-Konstruktion (syntaktische Anpassung bei der Fortführung von begonnenen Sätzen durch die zweite Sprecherin, vgl. Helasvuoto 2004) und Priming auf syntaktischer, konstruktioneller und lexikalischer Ebene (vgl. Melinger und Döbel 2005; Messenger et al. 2012; Santesteban et al. 2010) anders als im geplanten, monologisch und solitär geschriebenen Text eine besondere Rolle spielen. Diese vier Effekte beschreiben eine Tendenz zur Gemeinsamkeit und dadurch Anpassung der Sprecherinnen bis zur Gleichheit der Sprache – eine Tendenz, die auch das Auftreten von Chunks erklären kann (für weitere Ausführungen vgl. die Abschnitte 3.5 und 4 dieser Arbeit).

### **3.4 Korpuslinguistik**

Die Korpuslinguistik, die eher eine Methode ist als dass sie einen eigenen Forschungsgegenstand betrachtet, macht weniger theoretische Annahmen zur Frage von Kompositionalität, kognitiver Realität/Salienz, Koventionalität etc. von Chunks. Stattdessen werden (*fully fixed*) *multi-word expressions/units* (Stefanowitsch und Gries 2003), *multi-word chunks*, *lexical bundles* (vgl. z.B. Crossley und Salsbury 2011, Csomay 2012), *n-grams* oder *clusters* (vgl. Cheng et al. 2009, 240) meist automatisch extrahiert und statistisch ausgewertet. Weil hier die statistische Auswertung möglich und erwünscht ist, verschieben Definitionen für Chunks sich weg vom Sprachsystem und weg von der Sprache des Individuums hin zum beobachteten Sprachgebrauch. Das hat im Rahmen der Betrachtung von Chunks den Vorteil, dass intuitiv nicht erkennbare Chunks gefunden werden. Nicht alle Wortgruppen, die in dieser Weise gefunden werden, werden allerdings als kognitiv real zusammengehörige Einheiten betrachtet (vgl. Handwerker und Madlener 2009, 13; Schmitt et al. 2004; Wray 2002, 31). Ein typisches Beispiel wäre die Folge *is the*, die nicht als Chunk sondern als Schnittstelle zweier Phrasen gesehen wird, und der somit kein „unabhängiger“ Status zugeschrieben wird<sup>7</sup>. Allerdings ist diese Zuschreibung problematisch, weil sie erstens post hoc geschieht (die

<sup>7</sup> Csomay (2012) beschreibt *lexical bundles* umgekehrt zu diesen Ausführungen als text- und diskursstrukturierende

Daten werden zuerst anhand einer speziellen Theorie analysiert und anschließend die kognitive Realität der Teilstrukturen definiert), und zweitens intuitiv. Indes übersteigt die Menge der gefundenen Chunks im Korpus intuitive Vorstellungen:

„In this study, we use the term lexical bundles to refer to groups of two or more words that repeatedly appear in language as fixed items more frequently than expected by chance and much more frequently than phrasal verbs and idioms (...). It is the frequency of these multi-word lexical units that provides meaning in word combinations (...). **Lexical bundles are so frequent in fact, that they occur in over 80 % of the words produced in spoken English (based on the London Lund Corpus). Two word lexical bundles (i.e., bigrams) are the most common lexical bundles (...) and may occur more frequently than high frequency words (...), especially in spoken language (...). Word combinations are therefore important lexically because of their repetition.** Conversely, from a second language perspective, a higher frequency of significantly infrequent word combinations in second language production might signal abnormal word patterns production“ (Crossley und Salsbury 2011, 3, meine Herv. - AS).

Akzeptiert man Frequenzeffekte als kognitiv real, erschwert das die Argumentation gegen die kognitive Realität von solchen Chunks, die scheinbar nicht in die Kategorisierung des Sprachsystems passen.

Mit Hilfe grammatisch annotierter Korpora ist es neben konkreten, lexikalisch fixierten Einheiten auch möglich, auf abstrakter Ebene zu suchen, und somit Chunks in verschiedenen Varianten wiederzufinden. Tatsächlich deuten einige Funde darauf, dass „klassische“ Chunks, wie Sprichwörter oder Idiome, in Korpora oft häufiger lexikalisch verändert zu finden sind (vgl. z.B. Moon 1999 für *look for a needle in a haystack*). Das lässt sich möglicherweise auf den Versuch zurückführen, Klischees zumindest in der öffentlichen Schriftsprache (viele Korpora basieren auf Zeitungs- oder literarischen Texten) zu vermeiden. Obwohl die Betrachtung von Korpora das beste Verhältnis von Praktikabilität zu Natürlichkeit der Daten bietet, ist durchaus möglich, dass auch diese Daten zu speziell sind um „Alltäglichkeiten“ wiederzufinden. Cheng et al. (2009) beschreiben eine Analysemethode, bei der sogenannte *concgrams* extrahiert werden: Chunks, die auch erkannt werden, wenn sie nicht formidentisch auftreten, sondern in veränderter Reihenfolge oder mit zusätzlichen, zwischengeschobenen Elementen (etwa *work hard – hard work – work very hard*, ebd.).

Die Korpuslinguistik dokumentiert nicht nur auftretende Chunks, sondern berechnet mit Hilfe der Kollostruktionsanalyse die Wahrscheinlichkeit ihrer Elemente, zusammen aufzutreten, im Verhältnis

Bausteine, auch wenn sie nicht notwendigerweise im Sinne der Analyse abgeschlossen sind.

zum Einzelauftreten der Elemente jeweils. Ähnlich zur Kollostruktionsanalyse ist die Kollexemanalyse. Als Kollexeme bezeichnen Gries und Stefanowitsch (2004) und Stefanowitsch und Gries (2005) gemeinsam auftretende Lexeme in verschiedenen Slots ein- und derselben Konstruktion. So wären in der Äußerung „There's a bone in my nose that's slightly bent and it's progressively caused slight breathing problems“ (Stefanowitsch und Gries 2003, 221, Korpusbeispiel) *cause* und *problem* Kollexeme, und die Kollostruktionsstärke wäre im Rahmen der transitiven Kausalkonstruktion bemessen (für *cause* + *problem* + [TRANSITIVE KAUSALKONSTRUKTION] höher als für *cause* + *damage* + [TRANSITIVE KAUSALKONSTRUKTION], vgl ebd.). In einer anderen, verwandten Konstruktion treten die Kollexeme nicht in gleicher Häufigkeit auf.

Eine Konstruktion, für die eine bestimmte Gruppe von Kollexemen eine besonders hohe Kollostruktionsstärke aufweist, müsste demnach als besonders statisch verstanden werden. Chunks sind korpuslinguistisch betrachtet Einheiten, deren Bestandteile als Kollexeme auftretend hohe Kollostruktionsstärken haben.

### **3.5 Ein übergreifender Ansatz**

Wray (2002) definiert in ihrer umfassenden Analyse unter Einbeziehung von Daten aus Aphasieforschung, L1- und L2-Erwerb sowie Pragmatik Chunks (die sie *formulaic sequences* nennt) wie folgt:

„a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar“ (9).

Wray bezieht grammatische (Nicht-Analysierbarkeit, Diskontinuität), psycholinguistische (mentales Lexikon) und in der Diskussion auch korpuslinguistische Aspekte ein, und diskutiert theoretische Implikationen ausführlicher als viele andere Autorinnen (eine Ausnahme ist Gries (2007)). Dennoch enthält die Definition einige theoretische Implikationen, die auch bei Wray nicht oder nur wenig diskutiert werden:

Die ganzheitliche Speicherung einer Einheit zum kritischen Merkmal für die Definition zu machen ist nicht unproblematisch. Es besteht keine direkte Möglichkeit, die Speicherung nachzuweisen, da nur der Gebrauch empirisch zu beobachten ist. Im Gebrauch wiederum ist nicht ersichtlich, ob eine Einheit „am Stück“ oder „in Stücken“ produziert wurde. Das gilt besonders im Schriftlichen – lediglich der wiederkehrende Gebrauch weist darauf hin, dass sprachliche Einheiten

„zusammengehören“. Im Mündlichen weisen außerdem Intonationskurven, Sprechpausen und Interjektionen/Verzögerungslaute darauf hin, dass bestimmte Einheiten voneinander weniger abgetrennt sind als von anderen. Es ist aber nicht auszuschließen, dass diese Tatsache auf eine beschleunigte Produktion zurückzuführen ist, induziert durch Langzeitpriming/präferierte Sequenzierung (vgl. Hoey 2005, Ellis 1996; 2011) oder andere Faktoren. Möglicherweise gibt es überhaupt keinen Unterschied zwischen Produktion und Entnahme aus dem Lexikon bei Chunks, weil der neurologische Weg so vorgepflastert ist, dass beinahe reflexartig der lexikalische Partner aufgerufen wird. Wray diskutiert diese Frage am Rande (Wray 2002, 7-9), kommt aber in ihrer Definition nicht darauf zurück. Obwohl sie sich nicht ausdrücklich dazu äußert, scheint der Definition ein Listenmodell des mentalen Lexikons zu Grunde zu liegen. Das ist insofern interessant, als sie sich vielfach auf Forschungsergebnisse der Psycholinguistik bezieht, die wiederum tendenziell mit Netzwerkmodellen arbeitet (Hoey 2005, Ellis 1996; zu verschiedenen Modellen des mentalen Lexikons vgl. z.B. Aitchison 2012).<sup>8</sup>

Erschwerend kommt hinzu, dass es, wie Bybee (erscheint) ausführt, logisch inkonsistent ist, *keine* Repräsentation *aller* gehörten Sequenzen anzunehmen: Es gäbe keine Möglichkeit, etwas bei ausreichender Häufigkeit abzuspeichern, wenn es nicht bereits beim ersten Mal abgespeichert würde, da es schlicht unmöglich wäre, das zweite Mal vom ersten zu unterscheiden. Das lässt nicht darauf schließen, dass alles Gehörte gleichwertig bzw. gleich stark repräsentiert oder bewusst zugänglich ist (vgl. Ellis 1996, insb. 110). Eine Form von Verankerung, möglicherweise durch Erstverknüpfung oder Aktivierung von Synapsen (vgl. Hoey 2005) ist jedoch notwendig. Dieses Problem lässt sich nur teilweise mit einer Theorie der kritischen Inputmasse umgehen, da es offensichtlich Strukturen und Wörter gibt, die ab der ersten Aufnahme bewusst verfügbar sind, und andere, die immer wieder „vergessen“ werden. Zur Erklärung ließe sich das Konzept der Salienz heranziehen (vgl. z.B. Yorio 1980). Allerdings erschwert das die Analyse eher als sie zu vereinfachen, da Salienzfaktoren auf psychosozialer, kultureller, ko- und kontextueller, situationeller, pragmatischer, semantischer, syntaktischer oder anderer Ebene auftreten können. Eine Ausdifferenzierung der benötigten Salienzfaktoren bis zur „Speicherung“ wäre notwendig. Meine Vermutung wäre allerdings, dass Salienzfaktoren zu vielfältig und zu stark individuell und situativ verortet sind, als das man sie sinnvoll in überprüfbare Hypothesen einbetten könnte. Dafür spricht auch Carrolls (2012) Artikel, in dem sie in einer Analyse von Satzpositionen und Wortlänge

---

<sup>8</sup> Am Ende von Wray (2002) erarbeitet sie ein Modell der heteromorphen Verteilung des Lexikons, das gespeicherte Einheiten mit Balkendiagrammen darstellt. Das deutet stark auf ein Listenmodell, es ist aber nicht unmöglich, Netzwerkmodelle in dieser Weise quantitativ darzustellen (ab S. 263).

nach perzeptueller Salienz sucht und keine eindeutigen Faktoren erkennen kann.

Dennoch ist davon auszugehen, dass der Verlauf des L1- und L2-Erwerbs neben frequentuellen auch von Salienzfaktoren geprägt ist (Tomasello 2003; Ellis und Collins 2009). Ergänzend oder alternativ kann der Aspekt der Aufmerksamkeitssteuerung zur Erklärung herangezogen werden (vgl. Tomasello et al. 2005/2011, 543). Aus Wrays Definition und Arbeit wird nicht ersichtlich, wie sie diese Faktoren bewertet. Meiner Einschätzung nach meint sie weniger die gemeinsame Speicherung als die gemeinsame Aufrufbarkeit: „(...) stored and **retrieved whole from memory** (...)“ (s.o., meine Herv. - AS). Zusätzlich ungeklärt ist die Frage, in welcher Form Einheiten gespeichert seien: 1) Als „große Wörter“ (bestehend aus analytisch reduzierbaren Einheiten, mit graphemischer, phonetischer, semantischer und, theorieabhängig, syntaktischer Information) bzw. 2) phonetische Sequenzen (die erst in einem weiteren Schritt die genannten Informationen abrufen kann) oder 3) als ein Aktivationsmuster von semantischen Inhalten (die nicht unbedingt unmittelbar an ein phonetisches Muster gebunden sind). Die erste Annahme wird von lexikonbasierten Chunkmodellen favorisiert, so auch von Wray (2002) in ihrem heterogen verteilten Lexikonmodell. Für die zweite spricht die häufige Beobachtung von unanalysierten, „fehlerhaften“ Chunks insbesondere im L1- und L2-Erwerb (vgl. z.B. Wray 2002, Teile III und IV; Tomasello 2003). Für die dritte These, dass es sich um Aktivationsmuster für komplexe Zeichen handelt, sprechen eine Reihe von Beobachtungen: Umschreibungen von Idiomen und Sprichwörtern, von denen in einer gegebenen Situation nur ein Teil abrufbar ist oder die falsch aufgerufen werden („*Ich hab vergessen, wie das Sprichwort geht, aber es war etwa so: "Was ich nicht weiss, macht dich nicht heiss", "Ich bin heiss, was du nicht weisst" oder auch "Was du nicht weisst, macht mich heiss". Kann jemand helfen?*“, gefunden im Internet<sup>9</sup>), das „Abschleifen“, die phonetische Reduktion von hochfrequenten Einheiten – also der phonetische Wandel bei syntaktischen und semantischen Erhalt (vgl. Bybee 2006), und die Übertragbarkeit (und Übertragung) von Sprichwörtern und Redewendungen aus einer Sprache in eine andere (wie die Werbeanzeige einer Hotelkette, gesehen am 11.2.13 am Berliner Hauptbahnhof, mit dem Spruch „...*macht deinen Tag!*“, offenbar übertragen vom englischen *makes my day/made my day* → *makes your day*).

Die von Wray formulierte Definition enthält zwei weitere Merkmale: Die Elemente sollen *prefabricated* sein, also vorgefertigt, und nicht der grammatischen Generation oder Analyse unterliegen. Der Begriff der Vorfertigung findet sich in zahlreichen phraseologischen Arbeiten

---

9 [http://kamelopedia.mormo.org/index.php/Frage:Ich\\_hab\\_vergessen,\\_wie\\_das\\_Spruchwort\\_geht,\\_aber\\_es\\_war\\_etwa\\_so;\\_22Was\\_ich\\_nicht\\_weiss,\\_macht\\_dich\\_nicht\\_heiss%22,\\_22Ich\\_bin\\_heiss,\\_was\\_du\\_nicht\\_weisst%22\\_oder\\_au%20%80%A6%3F](http://kamelopedia.mormo.org/index.php/Frage:Ich_hab_vergessen,_wie_das_Spruchwort_geht,_aber_es_war_etwa_so;_22Was_ich_nicht_weiss,_macht_dich_nicht_heiss%22,_22Ich_bin_heiss,_was_du_nicht_weisst%22_oder_au%20%80%A6%3F), zuletzt aufgerufen am 13.3.13



wieder und wird häufig für die Beschreibung des Phänomens insgesamt gebraucht (*prefabs* (vgl. z.B. Bybee und Cacoullos 2009), *prefabricated units/sequences*). In aller Regel wird er verwendet, ohne seine theoretische Implikation auszusprechen: Dass Chunks vorgefertigt, also aus kleineren Einheiten zusammengesetzt und für den späteren Gebrauch abgelegt sind. Das ist, anders als es erscheinen mag, keine triviale Aussage. Zunächst sollte geklärt werden, wer produziert und was produziert wird. Produziert die Sprecherin eine für sie häufig benötigte Sequenz vor? Produziert die Sprachgemeinschaft Sequenzen für alle Sprecherinnen? Mit anderen Worten: Bewegt sich der Begriff auf Ebene der *langue* oder der *parole*? Werden die Einheiten unter Einbeziehung von syntaktischen und semantischen Ebenen vorproduziert und dann abgelegt, oder werden ganze Sequenzen als gemeinsames Zeichen gespeichert und unterliegen auch im Produktionsprozess keiner Analyse? Der Begriff der Vorfertigung impliziert einen bewussten Synthesevorgang (wenn er auf der Ebene der *parole* verwendet werden soll), der in Antizipation bevorstehender kommunikativer Situationen ausgeführt wird. Gerade diese Vorstellung vertritt Wray nicht. Sie geht zwar davon aus, dass Einheiten im mentalen Lexikon vollständig abgelegt werden, wenn sie sich als nützlich in der Kommunikation erweisen (S. 102). Sie beschreibt den Chunkerwerb aber nicht als bewussten, regelbasierten Vorgang.<sup>10</sup>

Auf Ebene der *langue* ist der Begriff kaum weniger problematisch. Zwar ist es intuitiv einleuchtend, dass später als Chunks verwendete Einheiten, wie zum Beispiel diskursstrukturierende Einheiten oder konversationelle Routinen, zunächst von einer Sprecherin regelbasiert gebildet wurden. Allerdings erschließt sich daraus noch keine *Prefabrication*, sondern zunächst nur eine einfache Produktion. Zwar kann man den Begriff durchaus als Abkürzung von *previously fabricated*, also „zuvor (bereits einmal) produziert“ verstehen. Die Bezeichnung wäre dann allerdings erstens redundant, weil jede vergangene Äußerung *prefabricated* wäre, und zweitens wenig hilfreich, weil die kritischen Merkmale weiterhin undefiniert blieben.

Meine Intention an dieser Stelle ist weniger, Mängel an Wrays Definition aufzuzeigen. Vielmehr will ich auf die unausgesprochenen Inhalte hinweisen, da diese durchaus diskutierenswert sind, und eine präzise Begriffswahl vorteilhaft für die Modellierung des sehr komplexen Phänomens ist. Ich gehe nicht davon aus, dass Wray und andere Autorinnen der Auffassung sind, dass der Ursprung eines Chunks in erster Linie die regelbasierte, bewusste Vorfertigung ist.

Was den letzten Aspekt der Definition betrifft, den der grammatischen Nicht-Analyse bzw. der

---

10 Etwas anders verhält es sich im L2-Erwerb, in dem besonders anfangs und in vielen Fällen auch im weiteren Verlauf vorgefertigte, regelbasierte Strukturen gezielt gelernt und in Antizipation späteren Gebrauchs abgelegt werden (vgl. z.B. Taguchi 2008; Handwerker und Madlener 2009).

Annahme, Chunks würden nicht regelbasiert geäußert werden, so wirft dieser Bereich vielleicht die meisten Fragen auf. Zwar wird deutlich, was wahrscheinlich gemeint ist – dass ein Chunk keiner grammatischen Analyse bedarf, weil es (hypothetisch) als Ganzes bekannt ist und nicht segmentiert werden muss. Das schließt eine zusätzliche grammatische Analyse allerdings nicht aus, insbesondere in der Rezeption. Zwar gibt es Chunks, die syntaktisch und/oder semantisch nicht einfach kompositionell ableitbar sind (z.B. die englische Wendungen „*Long time no see*“, vgl. Zeldes (erscheint), 34). Allerdings zeichnen sich viele Chunks, wie Pawley und Syder (1983) beschreiben, dadurch aus, dass sie gerade die übliche von vielen grammatischen Varianten darstellen – sie fallen also nicht unbedingt durch geringere Grammatikalität auf, weshalb die Annahme einer Nicht-Analyse auf grammatischer Ebene sich zumindest nicht aufdrängt. Natürlich ist nicht ausgeschlossen, dass die Analyse trotzdem ausfällt. Jedoch müsste man eine Vorstellung dazuliefern, welche Rolle die Analyse in anderen Äußerungen spielt und wann sie einsetzt – denn um etwas nicht-analysieren zu können, muss der Unterschied zu anderen, zu analysierenden Einheiten erkannt werden und auf andere (?) semantische (?) Extraktionsmethoden umgestellt werden.

### **3.6 Chunks in der Grammatik**

Chunks spielen in den meisten Grammatiktheorien eine untergeordnete Rolle, weil sie als unsystematisch (in der GG auch als peripher) und daher nicht für die Generierung neuer Phrasen interessant betrachtet werden.

Einige grammatiktheoretische Ansätze, der Gegenpol zur GG sozusagen, stellen konkrete Wörter und Wortgruppen in den Mittelpunkt ihrer Beschreibung, so z.B. Hudsons (2007) *New Word Grammar* und Hunston und Francis' (2000) *Pattern Grammar*. Diese sind parallel zur Konstruktionsgrammatik und Kognitiven Grammatik entstanden und teilen viele grundsätzliche Annahmen beider. Die Details können im Rahmen dieser Arbeit nicht diskutiert werden, es sei aber darauf hingewiesen, dass diese Ansätze es leichter haben, mit idiosynkratischen Einheiten und Konstruktionen, die von ihren lexikalischen Partner abhängig sind, umzugehen, indem sie von vorneherein weniger Struktur und Systematik und mehr Redundanzen annehmen :

„In short, I believe that the syntax of a language does in fact consist of a very large number of constructions, each with its own peculiar interactions with other constructions and with lexical items. This means that the syntax is basically a rather messy collection of inductive generalizations on which we impose some order by generalization, rather than a small set of

very simple, very abstract, and very elegantly interacting patterns on which the imperfections of language have imposed some mess. Since we obviously cope with a great deal of mess in the rest of cognition, I see no reason not to assume the same for language“ (Hudson 2007, 153).

Der Konstruktionsgrammatik hingegen fällt es weiterhin schwer, offenbar lexikalisch – und nicht abstrakt oder schematisch semantisch-kategorisch – motivierte Sprache zu beschreiben. Entsprechend werden Chunks als ausgefüllte Konstruktionen (*filled constructions* (Barðdal, erscheint)), *fixed constructions*, Antonopoulou und Nikiforidou (2011)), holistische Einheiten (*holophrases*, Tomasello 2003), Chunks (Bybee, erscheint), Kollostruktionen (s.o.), *phraseologisms* (Gries 2007) oder *multi-word units* (Gries 2011) bezeichnet, vor Allem aber auf der Beschreibungsebene an den Rand gedrängt. Zwar ist man sich einig, dass es keine sprachliche Peripherie geben sollte, implizit wird aber ein Unterschied zwischen Syntax und Lexikon angenommen (s. o.) und viele Autorinnen befassen sich fast ausschließlich mit relativ abstrakten Konstruktionen (z.B. Argumentstrukturen bei Goldberg (1995; 2006); Resultativkonstruktionen bei Boas (2003; 2005; 2011), Alternationen bei Gries und Stefanowitsch (2004)). Chunks fanden vor Allem Platz in der Beschreibung des Erstspracherwerbs, wo sie in ihrer Steigbügelfunktion als Exemplare für den Erwerb abstrakter/schematischer Konstruktionen betrachtet werden. Interessant ist, dass für den weit fortgeschrittenen Erwerb keine Annahmen über diese „alten“ Chunks gemacht werden – es wird nicht diskutiert, ob sie weiterhin vorhanden sind, sich auflösen oder ob neue Chunks hinzukommen. Gerade in jüngster Zeit zeigen Untersuchungen von Korpusdaten jedoch, dass die Interaktion von abstrakteren und konkreteren/lexikalischen Konstruktionen nicht zufällig verteilt ist, sondern stärkere und schwächere Assoziationen bestehen. So stellt Römer (2005, Zahlen S. 118-125) anhand von Korpusdaten gesprochener Sprache fest, dass der Gebrauch von Progressivformen der englischen Verben stark lexikalisch abhängig ist. Einige Beispiele:

	Present	Past	Present Perfect	Past Perfect
accepting	90,91%	7,58%	1,52%	0
betting	71,43%	14,29%	0	14,29%
hearing	52,24%	23,88%	23,88%	0

Tabelle 1: Verteilung auf Tempusformen

	I	you	he	she	it	we	they
accepting	7,58%	10,61%	1,52%	1,52%	0	18,18%	16,67%
hearing	35,82%	31,34%	1,49%	0	0	23,88%	1,49%

Tabelle 2: Verteilung auf Subjekte

Römer schlussfolgert:

„The analysis (...), which took the progressive forms of 100 individual high-frequency verbs as the starting point, clearly demonstrated that average values and statements about „the progressive“ as such, as can be found in most existing accounts of this construction, do not tell us a lot about the actual use and behaviour of a large number of progressive forms. What we obviously need is an account of progressives, and perhaps of language in general, that pays more attention to lexical preferences and co-selection patterns; in other words, we need an account which individualises more and generalises less“ (S. 297).

Faulhaber (2011) analysiert 87 englische Verben, die sie 22 semantischen Gruppen zuordnet, daraufhin, in welchen syntaktischen Konstruktionen sie auftreten. Sie stellt fest, dass selbst semantisch sehr ähnliche Wörter – wie beispielsweise ihre Gruppe 1, *answer, reply, respond* – sich syntaktisch nicht identisch verhalten, was am Beispiel *When no one answered/\*responded/\*replied her, she came stomping into the room with a scowl on her face* (S. 93) deutlich wird. Faulhaber stellt fest, dass es zwar eine Überschneidungsfläche von ca. 70% pro Verbgruppe gibt, die restlichen 30% allerdings starke idiosynkratische Tendenzen zeigen (S. 114). Sie schlussfolgert:

„It was not possible to link the restrictions which were identified for the verbs examined here to the meaning of these verbs in any systematic way. What the data suggests is that conventionalization, and thus storage of entities larger than single words, plays a very important role. Further evidence for this comes from the lexical properties of some of the patterns examined here, which have a clear effect on complementation. If a verb is strongly associated with a certain context or co-text, i.e. if it is part of an idiom or a collocation, this can also influence how strongly it is associated with a specific pattern; for instance, of the 241 occurrences of the lemma *keep* + [NP/REFLPRON] + [to\_X] the element functioning as complement in the *to*-phrase is *a minimum* in 91 cases (e.g. *keep something to a minimum*) and a reflexive pronoun in 140 cases (e.g. *keep it to yourself* or *she keeps herself to herself*). Thus with *keep* the pattern is severely restricted to certain lexical fillings, which is not the case with semantically similar verbs, for example *sustain*“ (S. 313),

und weiter:

„While verb meaning (participant inventory, aspectual quality, selection restrictions), complement type meaning, as well as the meaning associated with certain patterns (i.e. constructions) can account for a considerable number of a verb's complementation patterns, the number of restrictions found is, by far, too high for these to be regarded as a peripheral phenomenon“ (S. 325).

Obwohl Faulhaber sich zunächst explizit von der Konstruktionsgrammatik abgrenzt und sich eher der *Pattern Grammar* von Hunston und Francis (2000) zuordnet, sind die Daten nicht unverträglich mit der Konstruktionsgrammatik. Ihre Daten und Überlegungen berühren eine wichtige Frage für die CxG, denn sie deuten darauf, dass nicht die Bedeutung der abstrakten Einheiten – Kategorien, Konstruktionen, Verbgruppen – sondern die Bedeutung der konkreten Einzelverben, und ihrer prototypischen lexikalischen Partner, die syntaktische Umgebung bestimmen:<sup>11</sup>

„(...) it seems that speakers, when planning an utterance, are influenced towards choosing one pattern over another parallel pattern by the lexical choices they make, i.e. semantic features in the lexemes occurring as complements are responsible for the choice of complement type“ (S. 77)..

Faulhaber diskutiert Chunks nicht explizit, bezeichnet sie aber dort, wo sie sie nennt, als *speech chunk* (z.B. S. 17).

Cacoullos und Walker (2009) untersuchen in einer multifaktoriellen Analyse gesprochene Korpusdaten des kanadischen und amerikanischen Englischen die Verteilung von Futurformen (*present, present progressive, will-* und *going-to-future*) auf lexikalische und grammatische Partner. Dabei stellen sie keine eindeutige lexikalische Präferenz fest, jedoch wird eine Verteilung der Formen auf jeweils eigene „Nischen“ deutlich:

„The niche occupied by the futurate Present (including both the more frequent Progressive and the quite rare Simple Present) is small but well-defined: going, coming and other motion verbs and scheduled events, as shown in the magnitude of the verb class and temporal adverbial effects, which are independent of each other. The lexical verb *go*, which makes up 70% of the motion verb class, and temporal adverbials indicating a definite time in the future, such as *this* phrases (*this afternoon/week/summer/year*), both highly favor the choice of futurate Present. Unlike the Romance languages, in which the go-future is taking over, English *will* and *going to* are equally viable. Importantly, neither form is strongly favored in the most frequent contexts: declaratives, main clauses and absence of adverbial specification. Thus, putative invariant meanings are not supported by these facts, since most of the time neither form is ‘expected’ to appear. Nevertheless, vestiges of earlier source constructions persist, as revealed by distribution patterns “ (S. 30)<sup>12</sup>

---

11 Faulhaber nennt das *triggern* (S. 17), allerdings erläutert sie nicht, welche psychologischen Automatismen, oder allgemeiner, Vorgänge sie damit beschreiben will. Ich übernehme den Begriff daher nicht.

12 Zitiert aus der zum Download bereitstehenden Version von Jan. 2009, [test.scripts.psu.edu/users/r/c/rct11/.../FUTURE.DV.Jan%202009.pdf](http://test.scripts.psu.edu/users/r/c/rct11/.../FUTURE.DV.Jan%202009.pdf), 37 S., zuletzt aufgerufen am 13.02.13, später

Zwar tritt hier keine reine Klumpung von lexikalischen Elementen zu Tage. Allerdings zeigt sich, dass Gebrauchsmuster nicht allein konstruktionsbasiert oder lexikalisch-semantic erklärbar sind – es scheint eine starke Anziehungskraft von konkreten Elementen zueinander zu geben: *come* und *go* zeigen besondere Nähe zu *present* und *present progressive*-Varianten; *I + 'll + never* treten besonders häufig gemeinsam auf, Fragekonstruktionen sind stark mit *going to* assoziiert.

Zu ähnliche Ergebnissen kommt die Analyse von gesprochenen Daten des neuseeländischen Englisch von Calude (2009), die verschiedene *Cleft*-Konstruktionen betrachtet. Sie stellt fest, dass, obwohl sich keine relevante Anzahl an exakt gleichbleibenden Chunks finden lässt, sich doch 90% der Demonstrativ-Clefts (DEM. PRO. + BE + [WH-WORD + REL CL], in ihren Korpusbeispielen etwa '*That's what I don't like about her*'; '*That's what he thought*' oder '*This is what it's all about*') in ein relativ enges Schema einordnen lassen: *That + 's + {what/why/where} + {PERS. PRO./NAME} + {VERB OF COGNITION/COMMUNICATION/MOVEMENT}*. Das könnte als Hinweis auf eine flexible oder abstrakte Konstruktion gewertet werden, aus meiner Sicht liegt jedoch angesichts der weitaus größeren kombinatorischen Möglichkeiten der Schluss nahe, dass es sich um eine Reihe überlappender Chunks handelt: '*That's why he*' und '*he left*', zum Beispiel.

Gries (2011) stellt fest, dass englische Chunks – er nennt sie in diesem Artikel *multi-word units* (S. 491), *prefabs* (S. 505) und *conventionalized utterances/structures* (S. 505), und das Phänomen *chunking* (S. 505) – eine starke Tendenz zur Alliteration haben. Er nennt eine mnemotechnische Funktion als mögliche Erklärung. Da Gries hier neben vollständig statischen Chunks auch die *way*-Konstruktion: „a.) SUBJ<sub>theme</sub> V<sub>move</sub> POSS way [PP P NP/S]<sub>path</sub>;; b.) ...*as the British Task Force made its way across the Atlantic*. (BNC: FNX); c.) ...*some expanse of water found its way into the picture*. (BNC: C9W)“ (S. 499f.) untersucht, scheint auch die flexible Neukonstruktion und nicht nur die gemeinsame statische Aufrufbarkeit davon betroffen zu sein. Ein möglicher Faktor ist ein phonetisches Priming, das die Aufrufgeschwindigkeit für gleiche oder ähnliche Phoneme erhöht und somit die Wahrscheinlichkeit dafür, dass aus einer Menge semantisch ähnlicher Möglichkeiten eine bestimmte Variante als lexikalischer Partner gewählt wird. Sollte dieser Effekt auch in Konstruktionen mit mehreren offenen Slots standhalten, würde das interessante Fragen zur Verwendung von ausgewählten Wörtern, auch Funktionswörtern, in lexikalisch spezifizierten Umgebungen aufwerfen.

Sinclair (1991) formuliert zwei gegenüberliegende Pole: Das *idiom principle* und das *open-choice principle*:

---

erschienen in *Language* 85(2), S. 321-354.

„[The open choice principle] is a way of seeing language text as the result of a very large number of complex choices. At each point where a unit is completed (a word, phrase, or clause), a large range of choice opens up and the only restraint is grammaticalness. (...) It is often called a 'slot-and-filler' model, envisaging texts as a series of slots which have to be filled from a lexicon which satisfies local restraints. At each slot, virtually any word can occur“ (S. 109).

und weiter:

„The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments. (...) At its simplest, the principle of idiom can be seen in the apparently simultaneous choice of two words, for example, *of course*. (...) However, the principle of idiom is far more pervasive and elusive (...):

- a. Many phrases have an indeterminate extent. As an example, consider *set eyes on*. This seems to attract a pronoun subject, and either *never* or a temporal conjunction like *the moment*, *the first time*, and the word *has* as an auxiliary to *set*. How much of this is integral to the phrase, and how much is in the nature of collocational attraction?
- b. Many phrases allow internal lexical syntactic variation. For example, there seems to be little to choose between *in some cases* and *in some instances*; or between *set x on fire* and *set fire to x*.
- c. Many phrases allow internal lexical syntactic variation (...).
- d. Many phrases allow some variation in word order (...).
- e. Many uses of words and phrases attract other words in strong collocation. (...)
- f. Many uses of words and phrases show a tendency to co-occur with certain grammatical choices. (...)
- g. Many uses of words and phrases show a tendency to occur in a certain semantic environment. For example, the verb *happen* is associated with unpleasant things – accidents and the like“ (S. 110ff.)

Zwar stellt Sinclair an dieser Stelle keine eigene Grammatiktheorie auf, er berührt aber wichtige Fragen für andere Grammatiktheorien. Die Konstruktionsgrammatik diskutiert die Aspekte b (eingeschränkt), c, d und g (vor Allem über Frames). Sie bezieht die Überlegungen a, b, e und f allerdings nur begrenzt in ihre theoretische Diskussion ein und nimmt dafür eine Art unausgesprochenes *open-choice principle* an.

Die zitierte Datenlage deutet dem entgegenstehend stark darauf, dass auch in Bereichen der Sprache, die bislang überwiegend als grammatisch und frei kombiniert betrachtet wurden, starke Anziehungskräfte zwischen konkreten Einheiten wirken.

### 3.7 Überblick: Begriffe für Chunks und ihre Flexibilitätseigenschaften

	KL/CL	GT	PL	L2	L1	PH	PR	Größe	Typische Merkmale	fix/ flex	abg.
(lexical) Chunk	x		x	x	x			<W bis >S		fix/ flex	
conventionalized pattern	x	x				x		?		fix	
Formel/ formula/ formulae/ formulas	x	x	x	x		x	x	≥2W bis ~S	oft: Überbegriff	fix/ flex	x
formulaic sequence	x		(x)	x	x	x		≥2W bis ~S	Überbegriff	fix/ flex	x
gefüllte Konstruktion	x	x						>W	Post-hoc-Zuschreibung (Konstruktion wird übergeordnet betrachtet); muss nicht lexikalisch spezifiziert sein	flex	x
holophrase				x	x			<W bis >S	v.A. im L1-Erwerb	fix	(x)
Idiom	x	x		x		x		≥2W bis >S	Semantik/Syntax nicht kompositionell	fix/ (flex)	x
Idiosynkrasie	x	x						?		fix	
Kolligation	x	x						~P	grammatische Präferenz von Lexemen	flex	x
Kollokation	x	x				x		≥2W bis ~P	lexikalische Präferenz von Lexemen; oft: Überbegriff	Fix/ (flex)	x
Kollostruktion	x	x				(x)		~2W	Kollostruktionsstärke und Frequenzgrenze entscheidend	flex	x
lexical bundles	x							xW	Größe wird festgelegt; Frequenzgrenze entscheidend	Fix/ (flex)	
lexikalisch basierte Konstruktion	x	x						>W	Post-hoc-Zuschreibung (Konstruktion wird übergeordnet betrachtet)	flex	x
move							x	?	diskursfunktional	fix/ (flex)	x
multi word unit/chunk	x	x						≥2W	Frequenzgrenze entscheidend	fix/ flex	
N-gram	x	x						(N)W	Größe wird festgelegt; Frequenzgrenze entscheidend	flex	
Phrasem/ Phraseologismus	x	(x)	(x)	(x)	x	x		≥2W		fix/ flex	
Phraseoschablone								≥2W bis ~S		flex	x
(konversationelle) Routine				x	x	x	x	≥W bis >S		fix	x
syntagmatic pattern	x							≥2W bis ~S		flex	x
unanalyzed Chunk			x	x	x			<W bis >S	v.A. im L1- und L2-Erwerb	fix	(x)

Tabelle 3: Die gängigsten Bezeichnungen für Chunks in Computer- und Korpuslinguistik, Grammatiktheorie, Psycholinguistik, L1- und L2-Erwerbsforschung, Phraseologie und Pragmatik mit Angaben über typische Zuschreibung von Flexibilisierbarkeit und Abgeschlossenheit.



Wegen der zahlreichen Verwendungsweisen und Definitionen der Begriffe kann die Tabelle nur einen ungefähren Überblick verschaffen. Es wird deutlich, dass fast alle Begriffe in verschiedenen Unterdisziplinen verwendet werden, was allerdings nicht auf Definitionsgleichheit deutet. Es wird auch deutlich, dass die meisten Verwendungsweisen vage, unterschiedliche oder widersprüchliche Angaben zu Flexibilisierbarkeit und Größe machen. Die Kategorie „Abgeschlossenheit“ betrifft den oben diskutierten kommunikativen und/oder semantischen Aspekt der Vollständigkeit oder relativen Geschlossenheit gegenüber anderen Einheiten. Dieser Aspekt wird in der Literatur selten diskutiert, daher basiert dieser Teil der Tabelle auf meiner Einschätzung aus den Literaturbeispielen und der zugehörigen Argumentation.

Es ist denkbar, die Chunkbegriffe auf einem Flexibilitätskontinuum zu verorten:

fix	eher fix	eher flexibel	flexibel
Holophrase	lexical chunk	Phrasem	Phraseoschablone
conventionalized pattern	formula(ic sequence)	Kollostruktion	gefüllte Konstruktion
unanalyzed chunk	Idiom	Kolligation	n-gram
koversationelle Routine	Kollokation	lexical bundle	
	Kollexem		
	move		
	multi-word unit/chunk		

Tabelle 4: Mögliche Einteilung von Chunkbegriffen in ein Flexibilitätskontinuum

In der Einteilung finden sich phonetisch identische Einheiten, wie sie hauptsächlich im Spracherwerb und in sehr stark konventionalisierten bzw. ritualisierten Bereichen der Sprache auftreten am fixen Ende und stärker als grammatisch variabel verstandene Einheiten am flexiblen Ende. Die Begriffe *lexical bundle* und *n-gram* sind hier am flexiblen Ende des Kontinuums angesiedelt, weil sie für die Suche nach verschiedenen wiederkehrenden Einheiten und ihren jeweiligen Varianten gebraucht werden. Das einzelne gefundene *lexical bundle* ist sicherlich fix. Zugleich besteht noch eine Unterscheidung hinsichtlich der benötigten Gleichheit zwischen diesen Begriffen, da *lexical bundle* zwei oder mehr gemeinsam auftretende Lemmata erfordert, während *n-gram* auch alle n-großen Zusammenhänge um ein Lemma bezeichnen kann.

Die Übersicht kann allerdings nur als Annäherung verstanden werden, da sie intuitiv und somit äußerst Streitbar ist, zumal der Gebrauch, wie oben diskutiert, von ebenso viel Intuition und wenig theoretischer Abgrenzung geprägt ist. Es wird dennoch deutlich, dass der Begriff der „konkreten

Konstruktion“ besser als Abstraktion über eine Reihe von Einheiten mit unterschiedlichen Eigenschaften zu verstehen ist, deren Betrachtung „im Paket“ nicht unbedingt zielführend ist.

### **3.8 Funktionen von Chunks**

Erkenntnisse aus der Forschung um den L1-, L2- und Schriftspracherwerb, der Psycho- und Korpuslinguistik sowie der Diskursanalyse deuten auf eine Vielzahl von Funktionen von Chunks.

Im L1-Erwerb werden Chunks als erster Einstieg in die Sprache und als Grundlage für darauffolgende Abstraktionen verwendet. Casenhiser und Goldberg (2005) stellten fest, dass eine optimale Lernbarkeit von Konstruktionen dann gewährleistet ist, wenn ca. 80% des Inputs auf ein Exemplar entfallen und die restlichen ca. 20% auf möglichst viele verschiedene (Zipfsche Verteilung). Das ist ein starker Hinweis auf die Orientierung an konkreten Einheiten im Spracherwerb. Chunks reichen von Unter-Wortgröße bis zu Mehrwortäußerungen im frühen Erwerb und bis hin zu weit darüber hinaus reichenden Einheiten wie Liedern, Sprichwörtern, Gedichten etc. im späten Erwerb. Ein bestimmender Aspekt dabei ist die Sequenzierungsfähigkeit, also die Fähigkeit, gehörte Abschnitte in kleinere „Bröckchen“ (=Chunks) zu zerlegen und nachzuahmen. Viele Chunks haben eine direkt kommunikative Funktion: Koversationelle Routinen wie Begrüßungen, Verabschiedungen oder Wünsche (sowohl eigene als an andere gerichtete). Außerdem sind Chunks zunächst die einzige Kommunikationsoption, so dass im Prinzip gesagt werden kann, dass alle kindliche Kommunikation im Prinzip aus Chunks besteht, wenngleich erste Slots entdeckt und „neu“ gefüllt werden. Lieven et al. (2003) stellten bei der Beobachtung eines 2jährigen Kindes über sechs Wochen fest, dass zum 79% der Äußerungen der letzten Aufnahme zuvor identisch vorgekommen waren. Bei den Mehrwortäußerungen war der Anteil der nicht identisch vorgekommenen Äußerungen höher, nämlich bei 37%. Allerdings enthielten von diesen neuen Mehrwortäußerungen 74% einen Teil, der bereits in anderen Äußerungen vorgekommen war, z.B. „*Where is X*“ mit X=neu (Lieven et al. 2003; Tomasello 2003, 308). Selbst dieser relativ geringe Anteil dürfte überschätzt sein, folgt man den Ausführungen von Tomasello und Stahl (2004) dass gerade seltene Äußerungen kaum mit gängigen Aufnahmeintervallen zu erfassen sind. Somit ist nicht nur nicht ausgeschlossen, sondern sogar wahrscheinlich, dass nur sehr wenige Äußerungen von Kleinkindern eigenständig ausgefüllte Anteile enthalten. Dazu passt die Beobachtung von Wray (2002), dass Erwachsene nicht einfach grammatische, sondern konventionelle Äußerungen erwarten:

„Parental behavior [such as correcting a child when she says „*Gimme more milk*“ instead of „*Please, may I have more milk*“, AS] signals that there is a significance not just to *what* you say, but *how* you say it. This, in turn, alerts the child that, as with the rhymes and songs, some strings are not to be broken down, but simply remembered whole and produced at the right time, in order to achieve a speech act. Saying a routine incorrectly can be counted as a failure, even when the intention is clear, for only the permitted expressions are considered appropriate within the speaker culture“ (110)

Da Kinder diese Beobachtung regelmäßig machen dürften, liegt es nahe, dass sie nicht den *Sinn* des Gesagten zu entnehmen versuchen, sondern dem auszudrückenden Sinn eine *bestimmte phonetische Form* zuweisen. Das passt zu Wrays nächster Beschreibung von vier Phasen mit unterschiedlicher Betonung von Chunks im L1-Erwerb. Die erste, vom Beginn der Sprachproduktion bis zum Alter von zwei Jahren ist eine Zeit des „reinen“ Chunkgebrauchs mit sehr wenig Flexibilität. Es folgt eine Phase, bis ca. zum achten Lebensjahr, in der der Chunkanteil abnimmt und eine vergleichsweise „wörtliche“ Sprachproduktion und -rezeption einsetzt. Diese Phase ist zum Beispiel zu erkennen an der Unfähigkeit, Ironie, Witze, Wortspiele oder Homonyme zu erkennen bzw. zu verstehen. Im mittleren Grundschulalter beginnt eine Phase der erneuten Konzentration auf Chunks, des Erwerbs der erwachsenen und der Schriftsprache. Die Schriftlichkeit spielt dabei eine besondere und interessante Rolle: Sie hat eine normalisierende, an die Sprachgemeinschaft anpassende Funktion, die *langue* und *parole* einander näher bringen soll. So entsteht ein gemeinsames Inventar an Konstruktionen und Chunks. Insbesondere diskursive und somit informationsstrukturierende Muster werden gelehrt und normiert. Gleichzeitig greift die Schriftlichkeit sehr direkt in die Repräsentation sprachlicher Einheiten ein, indem sie sie in Wörter und zusammen- bzw. getrenntgeschriebene Einheiten unterteilt. Wörter werden so zu einer Kategorie, in die sich zuvor bekannte Chunks fügen müssen. Das Lesenlernen dürfte hier eine besondere Rolle spielen: Das Zusammenfügen von Wörtern aus Buchstaben und Sätzen aus Wörtern ist im Prinzip der umgekehrte Vorgang zum L1-Erwerb, bei dem Wörter aus größeren Einheiten heraussegmentiert werden, und die Phon-Phonem-Unterscheidung nicht bewusst zugänglich ist – weil auch sie durch die Schriftlichkeit nachträglich geprägt wird. Viele Aspekte der systematischen und im strengeren Sinne analytischen (also nicht synthetischen) linguistischen Betrachtung erhalten ihre Grundlage im Bewusstsein erst mit Einsetzen der Schriftlichkeit. Es wäre in dieser Hinsicht interessant, Bücher für Leseanfängerinnen und für Jugendliche auf ihren Chunk-Gehalt hin zu untersuchen.

Wrays Phasenmodell wird allerdings gestört vom Erwerb von Höflichkeits- und Gesprächsfloskeln und dem Beginn des Verständnisses für Humor, der von Guckelsberger (2008) auf das mittlere

Vorschulalter datiert wird, und nicht erst auf die späte Grundschulzeit. Möglicherweise besteht hier keine so deutliche U-Kurve im Chunkerwerb, wie sie von Wray gezeichnet wird, sondern ein differenzierter Erwerb in Abhängigkeit von Kommunikationsbereichen.

Neben der explizit kommunikativen Funktion haben Chunks auch individualisierende und gruppenidentitätsstiftende Funktionen.<sup>13</sup> So beschreibt Wray (2002, 72-75; 87-89) die gezielte Verwendung von Chunks, die für eine Person oder eine Gruppe charakteristisch sind. Sinclair (1991, 111) zählt fachsprachliche und Jargonausdrücke als Chunks auf, dem lassen sich weitere Chunks, die dialektale, Schicht-, oder andere Gruppenzugehörigkeit signalisieren (Szene, Organisation, Partei, Religion, oder Fangemeinde etc.), hinzufügen. Anders als die explizit kommunikative oder die grammatikstiftende ist diese Eigenschaft keine spezifische Chunk-eigene: Auch die Verwendung bestimmter Wörter, (flexibler) Konstruktionen oder phonetische Anpassung (Jargon, Akzent, Dialekt) können identitätsstiftend gebraucht werden, wie auch diverse andere semiotische Ausdrucksmittel. Dennoch können auch Chunks diese Funktion einnehmen, und ihre diskursive Funktion ermöglicht das Aufrufen ganzer semantischer Frames, die situationelle, identitäre und zahlreiche weitere Inhalte aufrufen können. Zum Beispiel kann eine Sprecherin durch den Gebrauch eines Zitats wie „Die Geister die ich rief, werde ich nicht mehr los!“ aus Goethes Zauberlehrling den metaphorischen Inhalt der Aussage aufrufen, aber auch deutlich machen, dass sie das Zitat kennt und sich so einer bestimmten Gruppe von Menschen zuordnen (diejenigen, die Goethe kennen),<sup>14</sup> und sie kann außerdem (angenommen, sie sagt das zu einem Kind, das stört, und von dem klar ist, dass es die Äußerung nicht versteht, aber in Anwesenheit eines Erwachsenen), einen ironischen Gehalt und gleichzeitig eine soziale Verortung vornehmen (weg vom Kind, hin zum Erwachsenen). Das ist nicht in gleicher Weise mit kleineren oder abstrakteren Einheiten möglich. Chunks haben daher die besondere Eigenschaft, dass sie zugleich perzeptuell, kulturell oder (sozial-)psychologisch stark verankert sein können und Situationsrahmen für ganze Gemeinschaften von Sprecherinnen aufrufen können.

Obwohl eins der meistgenannten Argumente für das Aufrufen von Chunks und gegen die generative Neuproduktion von höherfrequenten Äußerungen ist, dass es prozessökonomischer sei (vgl. Schmitt und Carter 2004, 5; Stengers et al. 2011; Wray 2002, 14-18), ist die Datenlage hierzu widersprüchlich. Underwood et al. (2004) fanden in einer Eye-Tracking-Studie an Chunks weniger

---

13 Natürlich ist die Stiftung von Identitäten jedweder Art auch eine Form der semiotischen Kommunikation.

14 Das trifft sicherlich nur begrenzt zu, es ist auch möglich, eine Äußerung zu verwenden, ohne den Kontext zu kennen, so dass der realisierte semiotische Gehalt von Intention, Vorwissen und Kontextwissen von Sprecherin und Hörerin abhängt.

Sakkaden und weniger Rücksprünge, also Hinweise auf ganzheitliche und schnelle Verarbeitung, in einer weiteren Studie aber keinen Lesezeit-, und somit keinen Verarbeitungsvorteil. Taguchi (2008) hat Lernerinnen des Japanischen in einem Sprachkurs gezielt Chunks beigebracht und zum Semesterende Dialoge abgefragt, in denen Chunks verwendet werden konnten (ohne gezielte Abfrage). Sie fand sowohl für Dialoge, die persönliche Informationen der Lernerinnen betreffen, als auch für narrative Erzählmuster Vorteile hinsichtlich der Komplexität, Flüssigkeit und lexikalischen Variabilität (S. 450).

Wray (2002) weist an verschiedenen Stellen darauf hin, dass Chunks oft zur Verlängerung und nicht zur Verkürzung von Äußerungen genutzt werden (*draw a conclusion = conclude; take/make a decision = decide; up the hill and down the dale = everywhere*, S. 74). Obwohl das zunächst widersprüchlich erscheint – wenn nicht für die Sprechplanung, so zumindest für die phonetische Umsetzung müsste eine Verlängerung den prozeduralen Aufwand steigern – kann es gute Gründe geben, eine längere einer kürzeren Beschreibung vorzuziehen. Stengers et al. (2011) argumentieren, dass es auch von prozeduralem Vorteil sein kann, ein längeres Chunk zwischen flexiblere Konstruktionen einzubauen, um sich Zeit für den Aufbau letzterer zu verschaffen. Wray (2002, 75) zählt die Sicherung eines gleichmäßigen Sprechtempos und Sprechrhythmus sowie die Herstellung der Gruppenidentität durch eine Art „ihr wisst schon was ich meine“-Effekt auf. Viele weitere Gründe kommen zusätzlich in Frage: Etwa die Bestätigung oder der Wechsel des Registers, die diskursive Fortführung, das Aufrufen anderer/neuer Frames oder die Bestätigung des vorherigen, die Vermeidung von Wiederholung (Konversationsmaximen), der Wunsch, auf andere Weise zu wirken (eloquent, kreativ, angepasst etc.) und so fort.

Gleichzeitig ist bekannt, dass in schnellen Sprechsituationen (Sportkommentare oder Auktionen z.B.) besonders chunklastig gesprochen wird, und Chunks zur Überbrückung von Sprechpausen anstelle von Interjektionen/Verzögerungslauten wie *äh* verwendet werden (vgl. z.B. Kuiper 2004).

Hinsichtlich der Prozessökonomie ergibt sich daher bislang ein gemischtes Bild. Auch, was die Vermeidung von Kreativität, zum Beispiel im L2-Erwerb, angeht, ist die Forschungslage widersprüchlich. So ist bekannt, dass gerade zu Beginn des L2-Erwerbs viele Chunks verwendet und durch die Verwendung vieler Chunks eine über den tatsächlichen Sprachstand hinausgehende grammatische Kompetenz suggeriert werden kann (vgl. z.B. Wray und Fitzpatrick 2008). Ähnliches ist von Aphasikerinnen und Alzheimer-Patientinnen bekannt: Bei beiden Gruppen bleiben oft Chunks erhalten, wenn die grammatischen Fähigkeiten zur Generierung derselben Äußerungen nicht mehr ausreichen (vgl. Wray 2002, Teil V für Aphasien; MacLagan et al. 2008 für Alzheimer-

Erkrankungen). Das deutet darauf hin, dass Chunks wenigstens in bestimmten Bereichen weniger grammatisch-generative Kompetenz oder einen geringeren Prozessaufwand erfordern. Andererseits beschreiben Myles et al (1998; 1999) in ihrer Untersuchung von Französisch-Lernerinnen mit L1 Englisch, dass gerade diejenigen, die die Chunks zu Beginn (Begrüßungsfloskeln, Beschreibung von Vorlieben mit formelhaften Ausdrücken) gut lernen konnten, später größere Flexibilität zeigten. Während „kleinere“ Chunks eine Reihe kognitiver, kommunikativer und grammatikstiftender Funktionen zu haben scheinen, sind „größere“ Chunks, wie Lieder (besonders National- und andere Hymnen), Gedichte, Sprichwörter, Schwüre/Eide oder ritualisierte Texte kulturell von besonderer Bedeutung<sup>15</sup>. Gerade bei solchen Texten geht es um das Aufrufen einer Situation oder einer Bedeutung in *exakt* gleicher Form. Wray (2002, 94) schlussfolgert, dass alle Funktionen von Chunks sich auf eine Sprecherbezogene reduzieren lassen: Effiziente Kommunikation sei am besten möglich, wenn der prozedurale Aufwand gering und der Inhalt möglichst eindeutig ist. Die Verwendung von möglichst vielen Chunks sichert diese Zielsetzung ab: Je weniger Aufwand die Hörerin aufbringen muss, um die Sprecherin zu verstehen, desto wahrscheinlicher kommt die Botschaft an. Je bekannter und „normaler“ die Bestandteile der Äußerung sind, desto wahrscheinlicher muss die Hörerin nur geringen Verständnisaufwand aufbringen. Somit konzentriert Wray die Kernfunktion von Chunks auf die Vermeidung von Missverständnissen in einem zur Missverständlichkeit neigenden kommunikativen System: Je neuer und kreativer eine Äußerung, desto mehr Spielraum für Interpretationen bleibt. Bekannte Äußerungen rufen bekannte Frames auf und erhöhen so die Wahrscheinlichkeit, dass das Gemeinte aufgerufen wird.

„The speaker, by virtue of deciding what is expressed and how, can exert control over the range of linguistic interpretations open to the hearer (...) and corner the hearer into the maximum likelihood of getting, and reacting to, the message. The more novel our output is for the hearer, the more likely it is to be misunderstood. Relevance Theory tells us that the hearer will engage in the least processing effort that results in a plausible (relevant) interpretation (...)“ (ebd.).

Dieser Effekt ist nicht nur für die gesprochene Sprache und die direkte Interaktion zu erwarten, sondern gerade in stark reglementierten Textsorten, wie wissenschaftlichen, juristischen oder medizinischen Texten, ebenfalls bedeutend. So schreiben Kerz und Haas (2009, 105f.):

„One might argue that whereas in spoken registers one of the major functions of prefabricated chunks is the reduction of online processing, in academic writing one of the main functions is to

---

15 Die menschliche Fähigkeit ganze Texte, wie zum Beispiel Dramenrollen oder hunderte Gedichte und Lieder, auswendig zu beherrschen spricht stark gegen eine Notwendigkeit zur Speicherökonomie.

contribute to the overall impression of a formal, impersonal style. Academic English is characterized by particular academic conventions such as the avoidance of personal language, judgemental words or emotive language. Academic texts are thus produced under highly controlled and edited circumstances, in which the author deliberately signals the impersonal, technical and formal style of this register. The need for a precise and dense packaging of information, as well as for planning and editing, is one of the distinguishing features of this text type. Awareness of these features triggers the use of prefabricated chunks. (...) **The author of a certain text deliberately uses such preconstructed patterns in order to signal the register. Furthermore, academic texts – regardless of which text type (e.g., article, monograph) – take a completed or ongoing research process with its key phases as their major topic, i.e., the pragmatic background of their production, namely a completed or ongoing research process and the academic discussion about it, constitutes an essential part of their content (...). This increases the probability of encountering formulaic sequences within the register**“ (meine Herv. - AS).

Die letztgenannten Überlegungen geben Anstoß für die Interpretation, dass gerade in Bereichen, in denen vergleichsweise viel *common ground* vorhanden ist – konkrete Sprechsituationen, Routinesituationen, professionelle (und daher von Kolleginnen nachvollziehbare) Situationen – Bedarf nach besonders hoher Kongruenz der sprachlichen Mittel besteht. In der Konsequenz würde das einen Konvergenzeffekt hinsichtlich der verwendeten abstrakten *und* konkreten Zeichen in ausgeprägten common-ground-Situationen und einen Adaptionseffekt vorwiegend zur Etablierung von common ground bedeuten. Es spricht nicht unbedingt dagegen, dass gerade Kennenlern-, Vorstellungs- oder Orientierungs- bzw. Fragesituationen (*Entschuldigung, wo ist denn hier die nächste Tankstelle?*) stark von Chunks geprägt sind, da, wenngleich die Inhalte eingefügt werden müssen, die Situationen dennoch stark ritualisiert sind. Die Idee kann an dieser Stelle nicht weiter diskutiert werden, es wäre aber im Rahmen anderer Arbeiten interessant, die Verknüpfung zur Pragmatik zu suchen.

#### 4. Zwischenfazit: Chunks als Basic Level Category

Es scheint so, als spielten Chunks auf verschiedenen Ebenen der Sprachlichkeit eine wichtige Rolle: auf der Verarbeitungsebene scheinen sie zum Sprachrhythmus, zur Flüssigkeit und (mit Einschränkungen) zur Entlastung des Arbeitsgedächtnisses beizutragen. Sie sind Steigbügel im Erwerb, aber sie haben anscheinend auch eine normative und damit regulierende Wirkung, insbesondere auf der diskursiven und pragmatisch-rituellen Ebene. Auf der grammatikalischen

Ebene haben sie damit eine stabilisierende Funktion: Nur, wenn ganz oder fast gleiche Einheiten für die Beschreibung ähnlicher Sachverhalte verwendet werden, können sie grammatikalisiert werden, und umgekehrt sind grammatikalisierte Einheiten neue Anker in der Grammatik. Vor Allem sind Chunks allgegenwärtig – Erman und Warren (2000) zählen in ihren Aufnahmen ca. 55% Chunks, andere Schätzungen gehen von 12-80% der gesprochenen Sprache aus (vgl. Wray 2002, 28ff.). Die extreme Varianz der Einschätzung basiert auf der unterschiedlichen Zählweise: Ob z.B. *phrasal verbs* oder phrasale Präpositionen (z.B: *of course*) gezählt werden, oder Konstruktionen mit lexikalischer Basis, oder nur im strikten Sinne Kollokationen und Idiome, und welche Untergrenze für die Frequenz im Korpus gesetzt wird, um eine Einheit als wiederkehrend zu identifizieren. Aus der Aphasie- und Demenzforschung ist bekannt, dass Chunks erhalten bleiben können, wenn stark regelbasierte Strukturen nicht mehr verfügbar sind.

Aus meiner Sicht ergeben sich aus dieser Breite der Bezeichnungen und Definitionen folgende Schlussfolgerungen:

1. Holistische Einheiten können eine große Bandbreite an Formen annehmen: Kollokationen, Idiome, *phrasal verbs*, *formulaic sequences*, Sprichwörter, Diskursmarker etc.. Alle diese Bezeichnungen haben theoretische Berechtigung, aber kein Begriff eignet sich wirklich gut als Oberbegriff für das Phänomen.
2. Im Bewusstsein der Tatsache, dass der Begriff insbesondere in der Psycholinguistik teilweise anders verwendet wird, entscheide ich mich dennoch für den Begriff *Chunk* als Oberbegriff. Er bietet den Vorteil, dass er die wenigsten Annahmen über Funktion und Speicherung/kognitive Modellierung und die wenigsten theoretischen Traditionen besitzt, und dass er daher das Potenzial hat, das Phänomen nur als „Brocken“ oder „Stück Sprache“ zu benennen und alle weiteren Annahmen gezielt zu prüfen und anzunehmen oder abzulehnen.
3. Chunks sind im hier gemeinten Sinn Einheiten, die eine beliebige Größe (im Regelfall über Wortgröße) annehmen können, die im Sprachgebrauch am Stück oder in enger zeitlicher Beziehung zueinander gleichbleibend auftreten, als Bausteine im Sprachgebrauch und als Exemplare für abstraktere und flexiblere Einheiten dienen. Sie können, Bybee und Cacoullos (2009, 189) folgend, auf den Achsen semantische Transparenz/Kompositionalität, grammatische Analysierbarkeit und Produktivität unterschiedlich verortet sein. Obwohl viele Chunks „groß“ sind und abgeschlossene oder zumindest konkrete Sinnabschnitte



markieren, würde ich auch sprachtypische Lautcluster als Chunks verstehen, zumal sie ebenfalls Inhalte transportieren (etwa Sprachzugehörigkeit). Allerdings können diese kleineren Chunks hier nicht weiter diskutiert werden.

4. Diese Definition enthält keine Annahmen über die gemeinsame oder ganzheitliche Speicherung von Chunks, sondern lediglich über ihre gemeinsame oder ganzheitliche Abrufbarkeit. Damit beziehe ich mich auf konnektionistische/Netzwerkmodelle in Anlehnung an Hoeys (2005) *Lexical Priming*-Ansatz. Er nimmt an, dass das gemeinsame Auftreten von verschiedenen Elementen die neurokognitive Verbindung zwischen ihnen stärkt und dadurch ein zukünftiges gemeinsames Auftreten wahrscheinlicher macht. Ein Auftreten eines Elements ohne seinen lexikalischen Partner schwächt die Verbindung. Ich würde in diesem Sinne auch nicht abgeschlossen sinntragende Einheiten wie *about the* oder *is the* als Chunks verstehen, obwohl sie üblicherweise aus der phraseologischen, psycholinguistischen, der Spracherwerbs- und selbst der korpuslinguistischen Forschung ausgeschlossen werden. Das Modell, das mir vorschwebt, ist entsprechend kein minimales, sondern lässt Redundanzen, Mehrfachverknüpfungen und Gegensätzlichkeiten zu.
5. Die jüngste Datenlage der Korpuslinguistik und Konstruktionsgrammatik deutet darauf hin, dass auch traditionell als grammatisch verstandene Phänomene wie Verbvalenzen und Progressivformen stark lexikalisch abhängig sind (vgl. Boas 2003, Römer 2005, Faulhaber 2011, Bybee (erscheint)). Ich deute das als Zeichen dafür, dass auch in der Erwachsenen-Grammatik Chunks – umdeutbar als gefüllte Konstruktionen – vorliegen. „Gefüllte Konstruktion“ ist eine post-hoc-Kategorie, die das Ergebnis eines Vereinigungsprozesses ausdrückt. Diese Beschreibung entspricht nicht dem Phänomen der Chunks, wie sie hier beschrieben wurden, da die Annahme ist, dass Chunks grund- und selbstständiger sind als die Vereinigung einer grammatischen Analyse (z. B. grammatischer Merkmale) mit lexikalischen Elementen. Daher übernehme ich den Begriff nicht.
6. Chunks spielen eine besondere und allgegenwärtige Rolle in der Sprache, sie sind kein Randphänomen. Viele Beobachtungen sprechen dafür, dass sie dem kreativen oder produktiven Gebrauch nicht neben-, sondern sogar übergeordnet sind. Das bedeutet aus meiner Sicht, dass sie nicht nur ein Ausschnitt aus dem mittleren Bereich eines Lexikon-Grammatik/Konkret-Abstrakt-Kontinuums sind, sondern eine besondere funktionale Kategorie darstellen (was keine scharfe Abgrenzung von Wörtern und/oder Konstruktionen

höheren Abstraktionsgrades bedeuten soll).

7. Ich schlage daher die Betrachtung von Chunks als *Basic Level Category (BLC)* im Sprachgebrauch vor. BLC sind Kategorien, die ontogenetisch und perzeptuell als „erste“, grundlegende Einheit verstanden werden: Etwa „Hund“ statt „domestiziertes Säugetier“ oder „43 cm großer Dackel“, oder „rot“ statt „warme Farbe“ oder „karmesinrot“ (vgl. Rosch 1978; Kawabe und Matsumoto 1998). In der Kognitiven Semantik spielen BLC eine wichtige Rolle bei der Beschreibung von Kategorisierungsmechanismen, sie sind eng verwandt mit Prototypen-/Exemplaransätzen, beschreiben aber eher die Mitglieder einer Kategorie, als ihre interne Organisation – so wäre eine Kategorie von BLC-Haustieren, wie „Hund“, „Katze“, „Vogel“ etc. dieser Sichtweise nach um die zentrale Gruppe der genannten häufigsten, prototypischen Haustiere organisiert, wobei dem deutschsprachigen Kulturraum gemäß „Waschbär“ oder „Tiger“ wahrscheinlich weiter außen angesiedelt wären. Die Prototypen-/Exemplarordnung beschreibt daher radiale Kategorien auf der horizontalen Ebene. BLC sind Ausschnitte aus der Vertikalen, obwohl sie dort auch Prototypen-Eigenschaften aufweisen.
8. Was bedeutet es, Chunks als BLC zu betrachten? Zunächst beinhaltet dieser Blick eine andere Aufteilung von sprachlichen Einheiten, als in der Konstruktionsgrammatik bislang vorgesehen: Sprache ist demnach nicht, grob, zweigeteilt – Wörter und Regeln/abstrakte Konstruktionen – sondern grob dreigeteilt: Wörter – Chunks – Regeln/abstrakte Konstruktionen. Die drei Bereiche sind nicht gleichberechtigt nebengeordnet, weil sie auch jenseits von ihrem Abstraktionsgrad unterschiedliche Eigenschaften aufweisen: Wörter und Chunks sind direkter zugänglich und situationell einsetzbar. Regeln/abstrakte Konstruktionen sind unbewusste Abstraktionen, neurologische Abdrücke ständigen Gebrauchs, deren Aufschlüsselung und Bewusstmachung weitergehendes kategorielles Wissen voraussetzt – was nicht bedeutet, dass sie nicht trotzdem gebraucht werden, so wie der Gleichgewichtssinn auch arbeitet, wenn man nichts davon weiß. Chunks und Wörter sind perzeptuell unmittelbar verfügbar (auditiv, visuell in der Schrift- und Gebärdensprache, taktil bei Braille-Schrift). In dieser Hinsicht sind sie ähnlich. Chunks haben jedoch, anders als abstrakte Konstruktionen, und anders als die meisten Wörter (oder Morpheme), die Fähigkeit, konkret-situative Frames zu evozieren, und ermöglichen somit ritualisierten, routinierten und stabilisierenden Sprachgebrauch. Als besondere Zwischenkategorie haben sie zugleich und direkt grammatische, semiotische, pragmatische

und kommunikativ-funktionale Eigenschaften. In dieser Betrachtungsweise sind sie die primären Bausteine auch der Erwachsenen- und Schriftsprache, obgleich sie in der Schriftsprache durch den verzögerten Sprachplanungsprozess und die spezifischen Herausforderungen wie Etablierung eines *common ground* und Bezugnahme auf Diskurse und Textsorten anders auftreten als in der mündlichen Sprache. Möglicherweise treten sie im schriftsprachlichen Gebrauch stärker in Erscheinung, was stark normierte Bereiche betrifft (Konnektoren, Diskursmarker, Textsortenspezifika), und zugleich schwächer, was stärker flexibilisierte betrifft (grammatische Konstruktionen/Auswahl lexikalischer Partner, Flexibilisierung von eindeutig identifizierbaren Chunks wie Redewendungen etc.). Diese Hypothese wäre gezielt zu überprüfen, die zitierte Datenlage gibt allerdings Hinweise darauf.

9. Chunks sind prinzipiell flexibel, so wie beinahe alles Sprachliche flexibel ist. Das bedeutet nicht, dass sie nicht dennoch in einer prototypischen Form abrufbar sind, und ist auch kein Hindernis für ihr regelmäßig exakt gleiches Auftreten. Nicht alles, was austauschbar ist, muss ausgetauscht werden. Tatsächlich erscheint die Annahme skurrill, ein Chunk würde, nachdem es – wegen seiner kommunikativen Funktionalität – ganzheitlich gelernt und flexibilisiert wurde, vergessen und durch die flexiblere Konstruktion ersetzt werden. Wahrscheinlicher ist, dass es in verschiedenen Varianten abrufbar bleibt und das Wissen um die neuen Möglichkeiten die alten nur dann gefährdet, wenn es frequentuell und salienzabhängig fast ausschließlich in neuen Varianten vorkommt, oder wenn es mit der Zeit als falsch segmentiert erkannt und zugunsten einer funktionelleren Einheit aufgebrochen wird. Es dürfte eine größere Menge prototypischer Äußerungen geben, die fortwährend wichtig und dadurch aktiviert bleiben. Diese müssen nicht immer in Korpora zu finden sein, weil gerade die üblichsten Äußerungen nicht schriftsprachlich und auch nicht für gesprochene Korpora aufgenommen werden dürften – Alltagsgespräche machen aber den größten Teil der verarbeiteten Sprache aus.

#### 10. Besitzen Chunks Slots?

Diese Frage stellt sich in doppelter Hinsicht: Erstens, können Chunks einzelne Positionen als Lücken lassen, die *prinzipiell* ergänzt werden müssen? Und zweitens, besitzen Chunks Positionen, die austauschbar sind, aber prototypisch besetzt? Meiner Ansicht nach treffen beide Aussagen auf unterschiedliche Fälle zu, es dürfte aber zum empirischen Schwierigkeiten kommen, wenn man versucht, sie zu trennen, weil nicht beweisbar ist, dass

ein Slot *keinen* prototypischen Filler besitzt.

Es gibt darüber hinaus zwei grundsätzliche Möglichkeiten, chunkbasiert zu argumentieren: Mit der Annahme von Slots, oder mit der Annahme von überlappenden Chunks mit zwischengeschobenen Wörtern: So ist die Äußerung '*that was where he was going to go apparently*' (Calude 2009, 66) beispielsweise aufteilbar in:

- a. [that was][where he was] [going to go] [apparently],
- b. [that was (<sub>1</sub>where [(<sub>2</sub>he)<sub>1</sub> was going to go] [apparently]<sub>2</sub>), (mit überlappenden Chunks),
- c. [that was][where he was going to X][apparently], (mit Slot),
- d. [that was][where X was][going to Y][apparently], (mit zwei Slots),

oder in eine Reihe weiterer Unterteilungen. Der entscheidende Unterschied zur bisherigen Tradition der Konstruktionsgrammatik besteht darin, dass ich nicht annehme, dass es sich *im Gebrauch* um eine lexikalisch basierte Konstruktion, etwa [[*that was*] [WH-word] [NP] [VP<sub>fin</sub>][Adv]], in Unifikation mit den zugehörigen lexikalisch(er)en Einheiten handelt. Zwar existiert diese und/oder eine Reihe ähnlicher Konstruktionen, sie bewegen sich meines Erachtens jedoch auf einer anderen, nicht unmittelbar und notwendig gebrauchten, Ebene.

Ein Chunk, das Slots besitzt, ist daher nicht gleichzusetzen mit einer mehrfach-lexikalisch basierten Konstruktion, insbesondere, weil Chunk-Slots mit Chunks gefüllt werden können, was Äußerungen wie '*actually that's what I worry about is um <, > eating too much no is I eat enormous quantities of chocolate every day I'd have chocolate*' (ebd., 70) erklärt:

Es scheint eine Überlappung von [*actually that's what I worry about*] und [*what I worry about is ... eating too much*] vorzuliegen, die durch die Aneinanderreihung von Chunks leicht nachzuvollziehen ist, aber bei reiner Betrachtung von Wörtern und Konstruktion nur als Fehlkonstruktion zu verstehen ist. Auch wenn die entstandene Äußerung für sich stehend muttersprachlichen Akzeptabilitätsurteilen<sup>16</sup> wahrscheinlich nicht standhalten würde, ist es

---

16 Die Rolle und Stabilität muttersprachlicher Akzeptabilitätsurteile kann an dieser Stelle nicht ausführlich diskutiert werden. Neben Aspekten wie Frequenz der abgefragten Strukturen und Priming – die also das Urteil abhängig machen von der Menge der abgefragten Strukturen zuvor –, ist Akzeptabilität ein unzuverlässiger Marker, weil sie in vielen Fällen einer Schriftbias unterliegt und weil, selbst andernfalls, jede Fragestellung nach der Richtigkeit/Schönheit/Wahrscheinlichkeit eines Satzes präskriptive und normierte Kategorien aufruft. Das deutet darauf, dass Akzeptabilitätsurteile nur als Hinweise in die eine oder andere Richtung gewertet werden dürfen.

m.E. unwahrscheinlich, dass sie im mündlichen Sprachgebrauch als fehlerhaft auffiele.

Eine potentielle theoretische Schwierigkeit liegt darin, dass zahlreiche Möglichkeiten bestehen, die Äußerung in Chunks aufzuspalten, und im Prinzip nicht herauszuarbeiten ist, welche die von der Sprecherin gewählte ist. Zwar stellt sich dieses Problem auch für die traditionelle Konstruktionsgrammatik: Es ist nie klar, welche Konstruktionen notwendig und/oder hinreichend sind, und es bietet sich für jede Äußerung eine Fülle an Konstruktionen an. Darüber hinaus ist bisher nicht herausgearbeitet worden, welche Rolle genau den Abstraktionen über Konstruktionen im Gebrauch zukommt. Letzteres kann ein chunkbasierter Ansatz zumindest teilweise beantworten (eine geringe Rolle), ersteres bleibt als Problem bestehen. Sollte die Annahme, Chunks seien die Grundeinheiten der Sprache, sich allerdings bestätigen, so wäre dieser Einwand nicht bedrängend für die Theorie, nur ein Problem für die Erfassung und Dokumentation der möglichen Einheiten – eine Reduktion auf ein möglichst kleines Set an Abstraktionen ist über 40 Jahre in der GG versucht und von der Konstruktionsgrammatik verworfen worden, eben zu Gunsten psychologischer Plausibilität und eines Abbilds der Sprache im Gebrauch. Wenn sich nun herauszustellen scheint, dass Sprache auch den geordneten Rahmen von konkreten + abstrakten Konstruktionen sprengt, ist der Verzicht auf Redundanzen aus dem Wunsch nach Minimalismus heraus nicht förderlich für die akkurate Beschreibung von Sprache.

## 5. Einbindung in die Konstruktionsgrammatik

Angenommen, Chunks sind Grundbausteine der Sprache, *Basic Level Category*, und spielen eine Rolle bei der Begrenzung eines kombinatorisch unbegrenzten Sprachsystems und der kommunikativen Regulierung – warum wäre es sinnvoll, ihnen eine stark gewichtete Funktion in einer Grammatiktheorie zuzusprechen, die eigentlich das Ziel hat, Systematiken über Restriktionen, Produktivität und Akzeptabilität aufzudecken?

Aus meiner Sicht sprechen vier zentrale Gründe dafür, diesen Schritt zu gehen: Erstens steht Konstruktionsgrammatik unter dem Druck sich häufender Daten, die belegen, dass konkret-lexikalische und abstrakt-kategorielle/grammatische Konstruktionen nicht frei interagieren, sondern unterschiedlich starke Anziehungskräfte zwischen Einheiten existieren, und dass die Assoziationsstärke vor Allem lexikalisch und phonetisch determiniert ist. Bisher gibt es dafür keine

hinreichende Erklärung oder Modellierung – es wirkt fast wie Zauberei: Es gibt ein strukturiertes, hierarchisch vererbendes Inventar von Konstruktionen verschiedener Abstraktionsgrade, und die abstrakteren auf allen Ebenen interagieren mit den konkreteren. Dabei entstehen immer wieder weit überzufällig dieselben Einheiten – obwohl das Modell jenseits der Kontextabhängigkeit keine Kraft vorsieht, die dazu führen könnte. Außerdem interagieren trotz der Annahme eines Kontinuums von Abstrakt zu Konkret in aller Regel nur relativ abstrakte Konstruktionen mit relativ konkreten, mit Ausnahme von Satz schemata (abstrakt + abstrakt → + dann erst konkret), und es gibt auch dafür keine ausführliche Erklärung – so dass in der Konsequenz doch ein Modell von „Wörtern und Regeln mit einigen Zwischenfällen“ entsteht, auch wenn die Regeln semantische Aspekte und teilweise lexikalische Basen haben.

Zweitens strebt die Konstruktionsgrammatik an, ein Erklärungsmodell mit möglichst großer psychologischer Plausibilität zu werden. Ein komplexes und ubiquitäres Phänomen wie das des Chunkgebrauchs dabei unter „große Wörter“ zu verbuchen, zumal die eigene Theorie zum Erwerb auf selbigem fußt, widerspricht diesem Anspruch.

Drittens bietet die Betrachtung von Chunks als BLC fruchtbaren Boden für die Diskussion um die Organisation des Konstruktionsinventars. Bisher gibt es außer der Annahme, dass Konstruktionsbedeutungen hierarchisch vererbt und Konstruktionen netzwerkartig angeordnet sind, wenig konkrete Modelle zur Organisation und zum daraus folgenden, psychologisch plausiblen, Gebrauch. Der Vorschlag, Chunks ins Zentrum zu setzen, und Generalisierungen und Spezifizierungen aus ihnen heraus zu erklären, bietet einen Ansatz für die Diskussion. Man könnte weiter gehen und folgern, dass Chunks der Definition nach die ureigentlichen Konstruktionen sind: Sie sind, mehr als Wörter, und mehr als abstraktere Konstruktionen, Form-Bedeutungspaare mit einer relativ abgegrenzten und relativ umfangreichen Bedeutung.

Viertens ist der bisherige Versuch, Regularitäten und Restriktionen aus semantischen Kategorien abzuleiten, nur begrenzt erfolgreich gewesen, wie u.A. Boas (2003) für Resultativkonstruktionen, Römer (2005) für englische Progressivformen und Faulhaber (2011) für Verbvalenzen feststellen. Verfolgt man den Gedanken, Chunks eine zentrale Position zuzuweisen, liegt die Schlussfolgerung nahe, dass sie auch für Restriktionen und Systematizität verantwortlich sein können. Das wiederum ist tatsächlich ein Schritt, um die Lexikon-Grammatik-Grenze aufzulösen, wie es der CxG bislang schwerfällt. Für diese Annahme sprechen Beobachtungen aus der Grammatikalisierungsforschung. So verweisen Bybee und Cacoulios (2009) auf eine Reihe Studien, die darauf deuten, dass

allgemeine Konstruktionsmuster auch auf Chunks wirken, so dass Chunks auch nach ihrer Konventionalisierung grammatische Eigenschaften behalten. Umgekehrt besitzen Einheiten, die im Begriff sind, grammatikalisiert zu werden, weiterhin lexikalische Eigenschaften (insb. 192f.).

Die Präzisierung dieser Gedanken im Rahmen der Konstruktionsgrammatik ist traditionsbedingt eine schwierige Angelegenheit, da es wie bereits angesprochen keinen einheitlichen Formalismus gibt und die Entwicklung eines solchen nicht einstimmig als erstrebenswert betrachtet wird.

Es gibt allerdings zwei stärkere Traditionen: Die Darstellung in Schemata, der Kognitiven Semantik (vgl. Evans und Green 2006; Boroditsky 2000) und Kognitiven Grammatik (vgl. Langacker 1987; 2008) folgend, und die Darstellung in Attribut-Wert-Matrizen, verschiedenen Grammatiktheorien, wie z.B. der *Head Driven Phrase Structure Grammar* (vgl. Müller 2008) folgend, und insbesondere von der *Sign Based Construction Grammar* (vgl. Sag, Boas und Kay 2012) aufgegriffen.

Als vorletzten Punkt dieser Arbeit werde ich darstellen, was eine chunkbasierte Analyse in diesen Formalismen bedeuten kann, und wichtige Differenzen zu bisherigen Annahmen herausarbeiten. Zusätzlich sollen zwei weitere Aspekte kurz angesprochen werden: Die Frame-Semantik, die einen guten Ansatzpunkt für die Diskussion um die Organisation von Chunks bietet, und ein Trend der jüngsten CxG-Beschreibung, in dem zunehmend von einer exemplarbasierten Ausrichtung die Rede ist.

## 5.1 Schema

In der traditionellen Konstruktionsgrammatik würde die Äußerung *Sie streicht die Wand rot* in etwa so modelliert: Im linken, oberen Teil des Schemas liegt die abstrakte Konstruktion vor, die dreigeteilt ist in zwei Formseiten und eine Bedeutungsseite: die grammatische und semantische Formseite, die Informationen zu den zu erfüllenden Kriterien der Füllwörter enthält, und die semantische Seite, die pragmatische, diskursive und schematische Aspekte enthält.

Die Wörter, links unten, haben eine phonetische bzw. graphemische Form-Seite (wobei im Regelfall nach der Alphabetisierung beide Repräsentationen vorliegen. Es sind aber auch Fälle denkbar, in denen nur eine Repräsentation existiert, in der L1 nur die mündliche, in der L2 ist auch nur die schriftliche möglich), eine grammatische Analyseseite (NP, V<sub>change of state</sub>, etc.), die hinsichtlich ihrer Zugehörigkeit zu Form oder Bedeutung diskutiert werden müsste, und eine semantische Seite, die die Semantik des Wortes inklusive perzeptueller Inhalte beinhaltet.

Im Unifikationsprozess werden die abstrakten und die konkreten Einheiten zusammengebracht und erhalten zusätzlich kontextuelle und pragmatische Bedeutung, wie Implikaturen oder

Beziehungsbotschaften. Ich verzichte an dieser Stelle auf die genaue Beschreibung des semantischen Gehalts aller Einheiten, da die Diskussion um die möglichen, notwendigen und psychologisch plausiblen Inhalte an dieser Stelle nicht geführt werden kann. In der obigen Analyse lässt außerdem sich die Einheit '*die Wand*' natürlich in Determinierer und Nomen zerlegen, es müsste aber (an anderer Stelle) diskutiert werden, welche Rolle Determiniererphrasen in der Konstruktionsgrammatik spielen und ob sie für den Gebrauch vereinigt werden müssen.

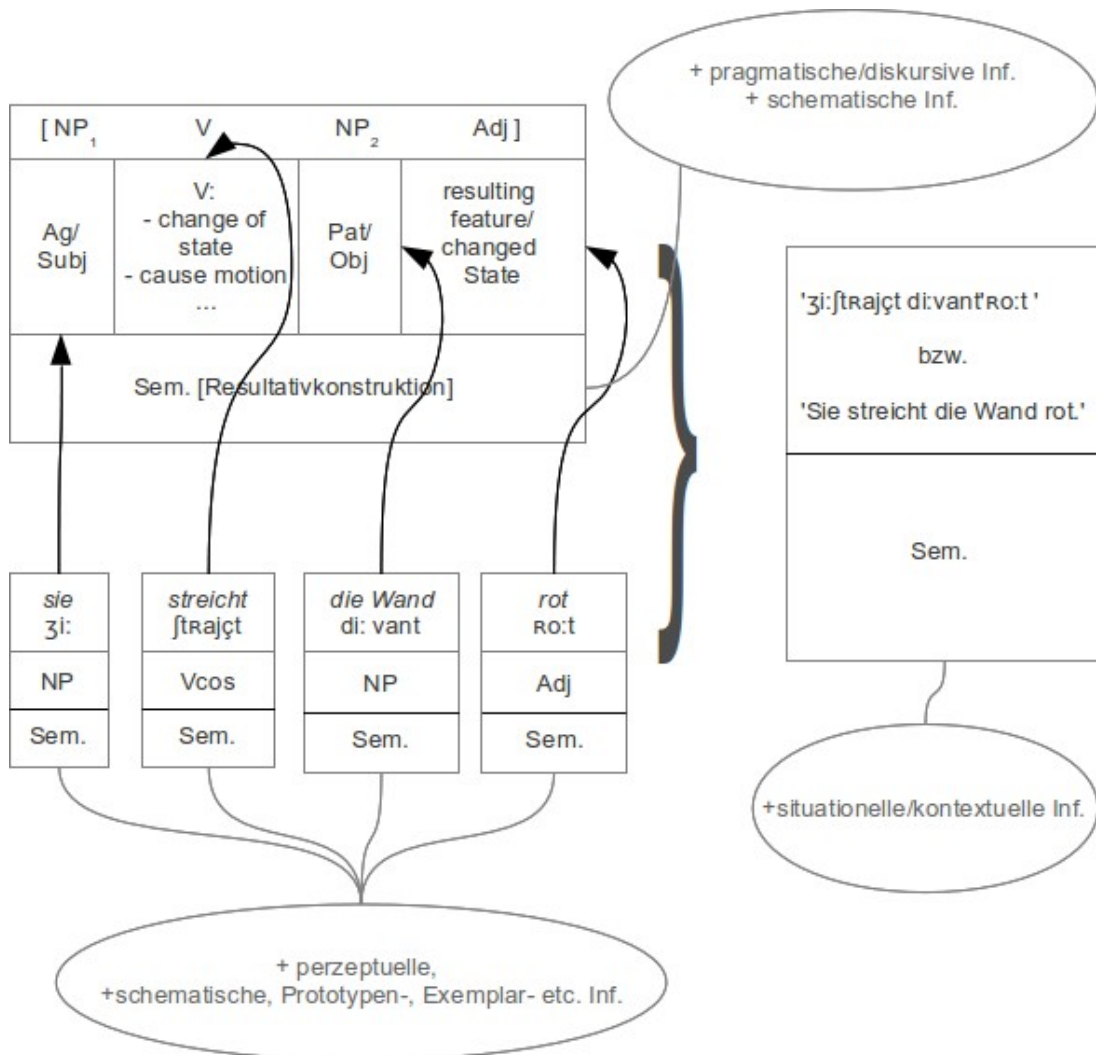


Abb. 1: Schematische Darstellung der CxG-Analyse von 'Sie streicht die Wand rot'

In der hier vorgeschlagenen alternativen, chunkbasierten Analyse (s. Abb. 2) entsteht diese Frage gar nicht erst, da 'X streicht die Wand Y' als Chunk mit zwei Slots betrachtet wird. Wie unten zu sehen, rückt hier eine mittlere Kategorie zwischen die abstrakten und die konkreten Konstruktionen. Dieses Chunk hat eine eigene Bedeutung, die perzeptuell von den Bedeutungen der Einzelwörter



abweichen kann. So kann die Erfahrung, dass eine Wand mit einem Pinsel, einer Malerrolle oder einem Quast bearbeitet wird und dadurch eine Farbe erhält, die sich von der vorherigen unterscheidet, diese Bearbeitung jedoch nicht durch die Sprecherin selbst erfolgt, ganzheitlich mit einem kommunikativen Mittel kombiniert werden, ohne in Einzelteile zerlegt werden zu müssen.

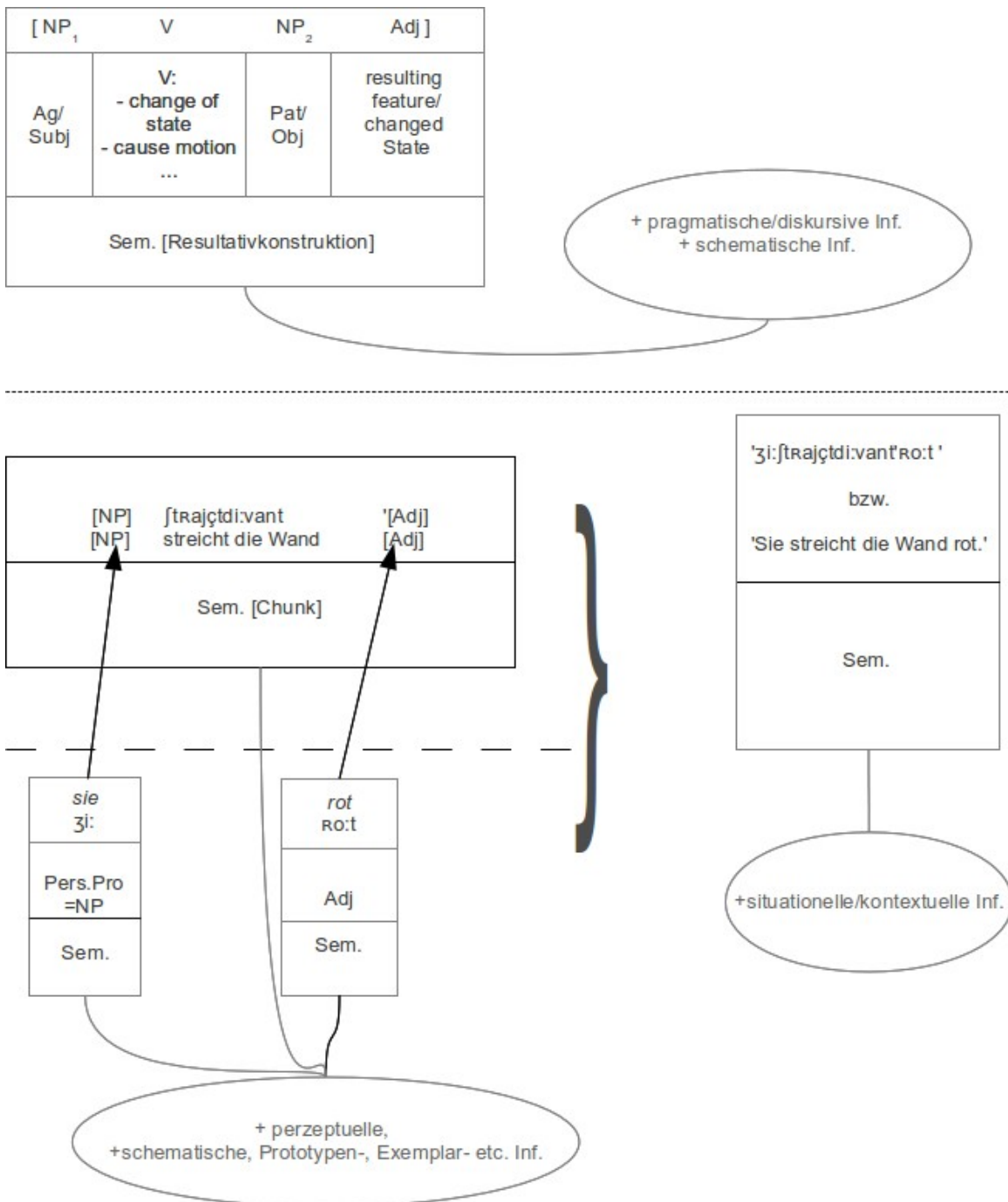


Abb. 2: Chunkbasierte CxG-Analyse von 'Sie streicht die Wand rot'

Ein zusätzlicher Vorteil liegt darin, dass Chunks direkt phonetisch repräsentiert sind und deshalb der Satzakkzent (in der betrachteten Resultativkonstruktion auf dem Resultat) unmittelbar und in der gleichen Form wie die restliche Formseite repräsentiert sein kann. In der klassischen Analyse ist das schwieriger, da der Satzakkzent in der abstrakten Konstruktion liegen muss, die aber keine eigentlich phonetische Repräsentation enthält, so dass er als einziger phonetischer Gehalt zusätzlich aufgerufen wird.

In der Abbildung ist außerdem dargestellt, dass die Trennlinie zwischen Chunks und Wörtern durchlässiger ist, als die von Chunks zu abstrakten Konstruktionen. Diese Darstellung basiert auf dem beschriebenen Unterschied in der perzeptuellen Zugänglichkeit und der Annahme, dass Chunks und Wörter Sprache in Verwendung, und abstrakte Konstruktionen Spuren dieser Verwendung durch die Erfahrung von Austauschbarkeit sind. Die Darstellung spiegelt auch die Annahme wider, dass abstrakte Konstruktionen dennoch analytisch zugänglich sind und keine scharfe Trennung zwischen den Kategorien anzunehmen ist.

Ist aber 'X *streicht die Wand* Y' überhaupt ein Chunk? Das Anstreichen von Wänden ist keine alltägliche Beschäftigung, über die mit hoher Frequenz im Präsens gesprochen werden dürfte („*Wo ist denn Paula gerade?*“ - ?„*Sie streicht die Wand rot*“), zumal bei dieser Fragestellung nicht der resultative, sondern der durative Aspekt im Vordergrund steht, der konventionell anders beantwortet würde (vgl. „*Wo ist denn Paula gerade?*“ - ?„*Sie streicht die Wand rot*“ mit „*Wo ist denn Paula gerade?* - „*Sie streicht*“). Es ist daher möglich, dass die abstraktere Resultativkonstruktion sich über das Chunk 'X *streicht die Wand*' mit nur einem Slot legt, und der zweite Slot durch die Verknüpfung entsteht. Dabei entsteht wieder die oben beschriebene Schwierigkeit um den Satzakkzent.

Ein wahrscheinlicheres Chunk wäre '*die Wand* 'X *gestrichen*', das sich in diesem Beispiel mit dem Chunk '*Ich habe*' (dessen kognitive Realität wegen des häufigen Auftretens höchstwahrscheinlich ist) verbindet (s. Abb.3). Die perfektive Verwendung erlaubt eine Form für alle Personen, was die Frequenz des identischen Auftretens potentielle erhöht (im Vergleich zum Präsens mit *streiche/streichst/ streicht/streiche*). Zusätzlich gibt es eine gewisse Wahrscheinlichkeit dafür, dass die resultative Verwendung eher nach Erreichen des Resultats berichtet wird. Beide Annahmen müssten allerdings an geeigneten Daten überprüft werden, das Beispiel soll vorwiegend der Illustration dienen, ohne dass eindeutig ist, dass das Chunk als solches vorliegt.

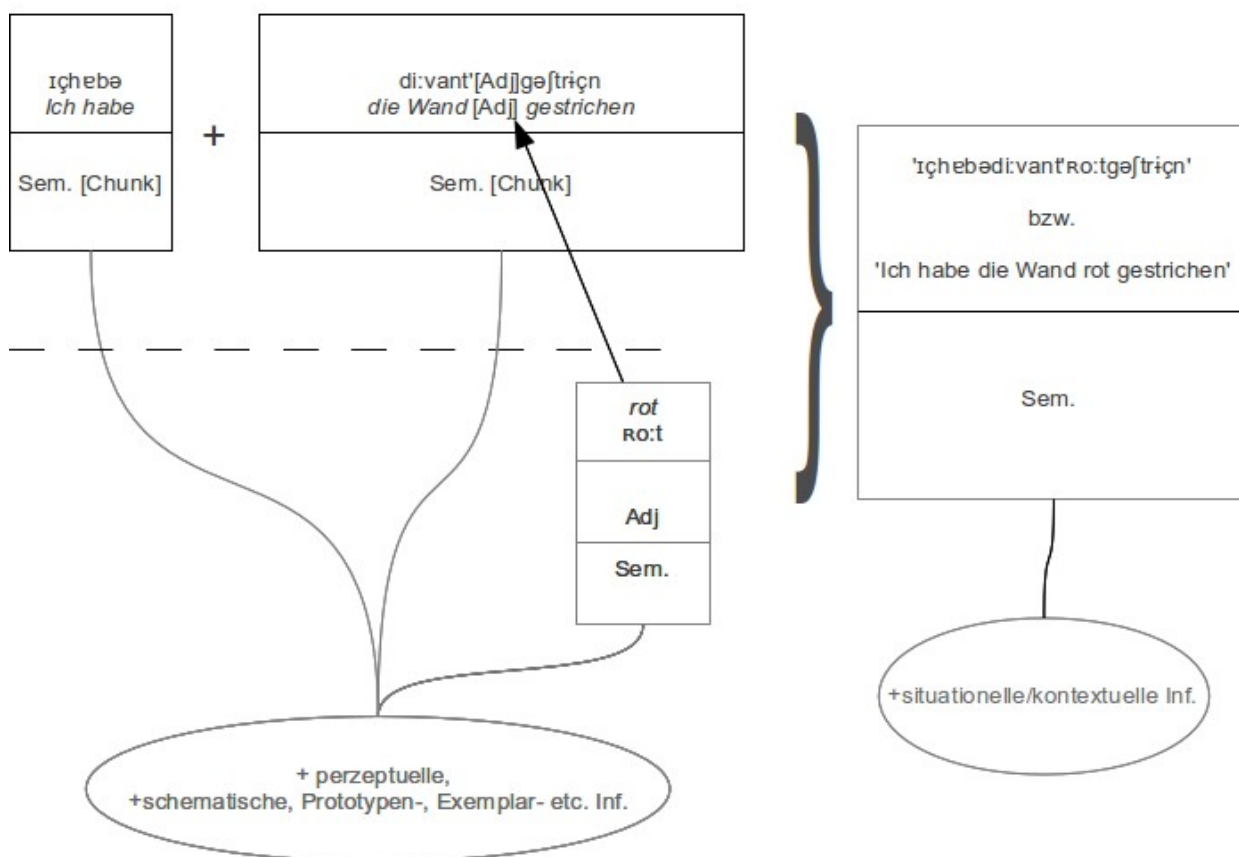
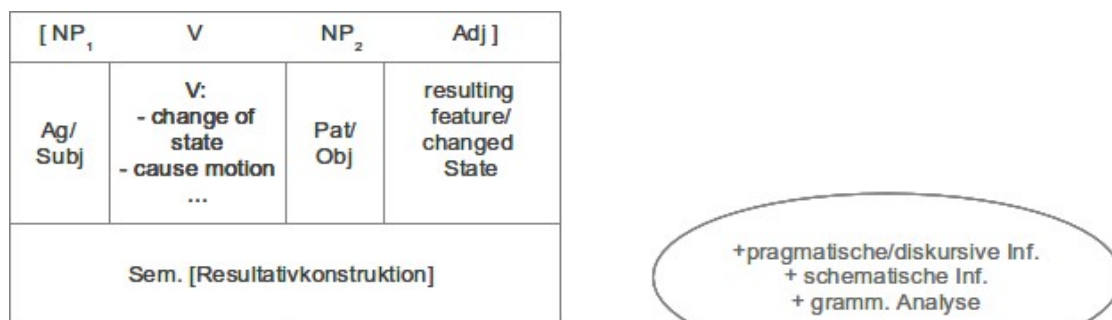


Abb. 3: Chunkbasierte Analyse von 'Ich habe die Wand rot gestrichen'

Die Annahme ist, dass Chunks für [*Ich habe*], [*du hast*], [*sie hat*] etc. vorliegen, und im Gebrauch nicht die Konstruktion [[NP] [HABEN]] verwendet wird. Dennoch ist die Auflösung von [*Ich habe*] in die Konstruktion [[NP] [HABEN]] möglich, so dass beispielsweise diachrone Veränderungen an [HABEN] früher oder später durch Regularisierungsprozesse auch auf das Chunk [*Ich habe*] angewandt werden. Ein solcher chunkbasierter Ansatz lässt Raum für die Integration von

Erkenntnissen aus der diachronen Sprachforschung, indem er Reduktionen und Regularisierungen über Chunks (und nicht nur über Abstraktionen) zulässt. Er lässt außerdem Raum für die Integration von unterschiedlichen Beobachtungen in mündlicher und schriftlicher Sprache. Eine Möglichkeit wäre die Annahme einer in Abb. 4 angedeuteten Abhängigkeit des Abstraktionsgrades von der Modalität und der Funktion. Die Abbildung zeigt einen deutlich höheren Anteil von sehr festen Chunks in der mündlichen Sprache, einen höheren Anteil für Äußerungen auf Grundlage von Konstruktionen mittlerer Abstraktionsgrade für die Schriftsprache, und deutet zusätzlich an, dass sehr flexibel gefüllte Strukturen vor Allem gezielt genutzt werden können.

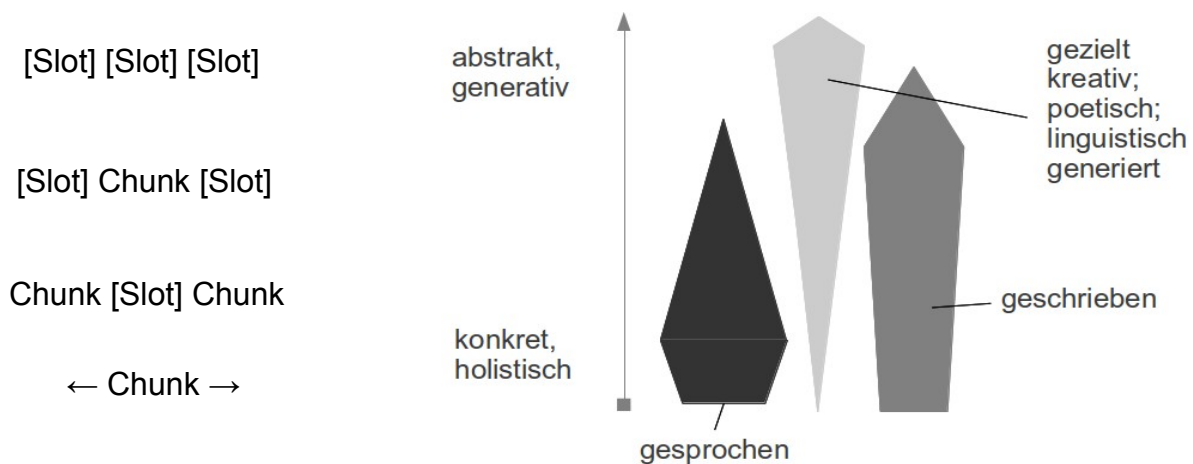


Abb. 4: Chunkanteil in Abhängigkeit von der Modalität und Funktion

Der Rahmen dieser Arbeit lässt keine Anwendung auf geeignete Korpusdaten zu. Den ausgeführten Überlegungen folgend bietet ein chunkbasierter CxG-Ansatz jedoch Ansatzpunkte für die Einbeziehung von mündlichen Daten in die Grammatikbeschreibung. Traditionell fällt es, wie Fried und Östman (2005) ausführen, den meisten Grammatiktheorien schwer, mit mündlichen Daten umzugehen. Grund dafür sind spezifische Anforderungen der Informationsstruktur und Begrenzungen des Arbeitsgedächtnisses, die zu Beschreibungsschwierigkeiten durch Satzabbrüche, Reformulierungen und Inkongruenzen (z.B. bei Sätzen, die in anderen Fällen, Personen, Tempora, Modalitäten etc. beendet werden als sie angefangen wurden) führen. Diese Veränderungen wurden lange als „nicht eigentlich“ sprachlich betrachtet. Ein chunkbasierter Ansatz könnte helfen, mündliche Sprache in die Grammatiktheorie zu integrieren, ohne auf grammatische Kategorien und Abstraktionen zu verzichten.

## 5.2 Attribut-Wert-Matrizen

Diese relativ informelle Darstellung ließe sich sehr gut auch in Attribut-Wert-Matrizen übertragen, wie sie von der *Sign Based Construction Grammar* (SBCG, vgl. Sag, Boas und Kay 2012) verwendet werden:

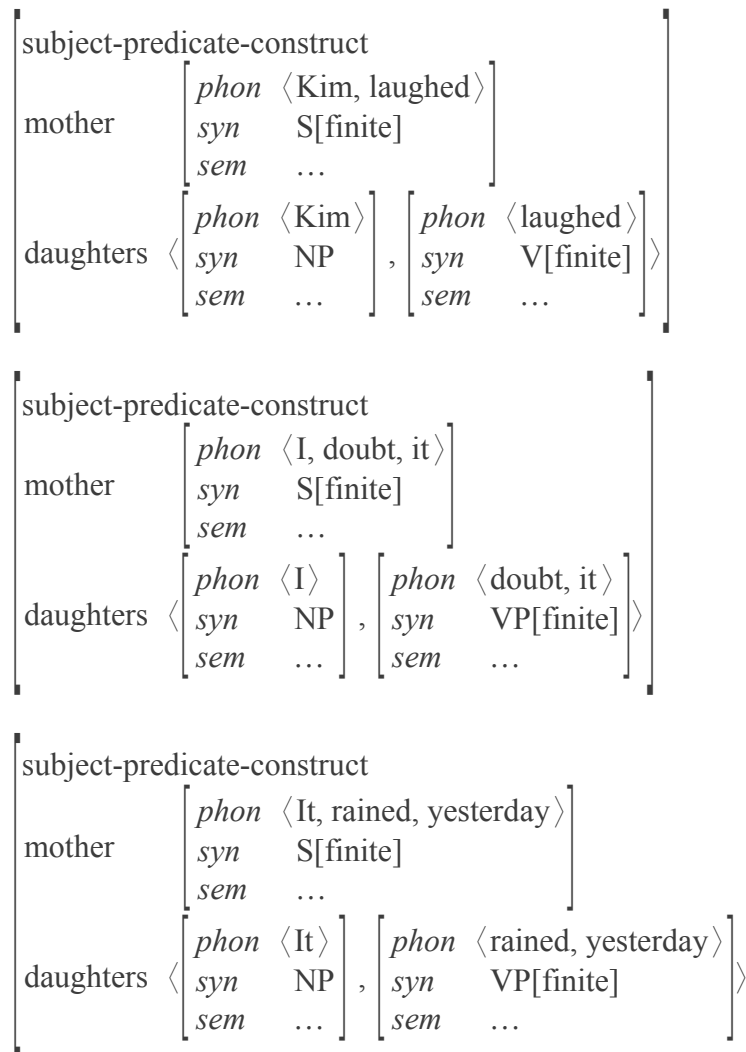


Abb. 5: Sag, Boas und Kay 2012, 17: Darstellung der Äußerungen als vererbende Konstruktionen

Diese Darstellung zeigt sehr schön den angenommenen Unifikationscharakter der SBCG, indem selbst die Phonologie der Gesamtäußerungen 'I doubt it' oder 'It rained yesterday' als zusammengefügt und noch in Einzelteilen zu betrachten verstanden wird. Interessant ist, dass Sag (2012) im selben Band an anderer Stelle (S. 75) eine phonologische (und nicht nur phonetische) Verschmelzung darstellt:

<i>Phonology</i>	/ εvri #liŋwɪst /
<i>Syntax</i>	NP
<i>Semantics</i>	'the set of properties all linguists share'

Abb. 6: Zeichencharakter der Äußerung 'every linguist'

Tatsächlich ist die von Sag (2012) beschriebene Formalisierung von abstrakten Konstruktionen 1:1 auf das obere chunkbasierte Schema übertragbar, während ein chunkbasierter Ansatz der SBCG helfen würde, elegant natürliche Sprache zu modellieren, ohne auf Abstraktionen verzichten zu müssen. Die Übertragbarkeit der Formalisierung bedeutet jedoch keine theoretische Gleichheit: Denn auch für diese Theorie besteht das Problem, dass die unzufällige Verteilung der Unifikationen unerklärt bleibt, und somit zwar eine sehr exakte Beschreibung von Konstruktionen vorliegt – was von großer Bedeutung für die Entwicklung der Konstruktionsgrammatik ist – aber weiterhin kein Modell, das die Selektion der Konstruktionen und ihrer lexikalischen Partner aus der bestehenden Menge erklären würde. Das steht im Widerspruch zum postulierten Anspruch an psychologische Plausibilität und exakte Beschreibung der Daten (Sag, Boas und Kay 2012, 14):

„SBCG embodies a strong commitment to psycholinguistic plausibility (...). This means that the mechanisms of grammar are motivated not by descriptive economy, lack of redundancy, or mathematical abstractions of questionable direct relevance to linguistic science (...). Rather, linguistic proposals are motivated and evaluated in terms of how well they comport with models of language use (e.g. production and comprehension), language learning, and language change (...).“

Die Integration von Chunks in eine SBCG-Analyse ist formal einfach, allerdings stört sie die Eleganz der Beschreibung erheblich, da Chunks zahlreich, redundant und mehrfachverknüpft sind.

Abb.7 (s.u.) ist eine Annäherung an eine klammerbasierte Formalisierung einer chunkbasierten Analyse, die sich von den in der SBCG vorgeschlagenen Analyse in folgenden Punkten unterscheidet: 1. Phonetische, also Formbestandteile, sind phonetisch dargestellt. Konsequenterweise sind graphemische Bestandteile extra aufgeführt. 2. Der Vereinigungsprozess findet zwischen Chunks statt, was zum Beispiel die Vereinigung zweier finiter Verbalphrasen erlaubt. Der Unifikationsprozess ist damit stärker semiotisch und weniger strukturalistisch gedacht. 3. Der Resultativanteil der Äußerung besteht aus einem Chunk, das einen Slot enthält, und dem Slotfiller ('rot').

Die Nachstellung des Slotfillers hat keinen theoretischen Gehalt, sondern dient der Übersichtlichkeit. Es wäre ebenso möglich, das Chunk als übergreifende Klammer mit einer Einfügung des Chunks von oben oder unten bzw. vorne oder hinten anzunehmen. Es sollen keine Annahmen über die Verarbeitungsreihenfolge oder andere kognitive Prozesse ausgedrückt werden.



Abb. 7: Formalisierung einer Resultativkonstruktion in einer Attribut-Wert-Matrix

Diese Darstellung der Äußerung, insbesondere was den phonetischen Anteil betrifft, ist nicht die einzig mögliche. So ist eine Verkürzung von 'ich habe' zu 'ich hab', 'ichab' oder 'chab' bzw. 'hab' zusätzlich aufzulisten, zumal es diverse Varianten der phonetischen Realisierung aller dieser Möglichkeiten gibt; und neben der semantischen Analyse ist auch die syntaktische Tiefe nicht ausgeschöpft. Allerdings ist die Annahme in dieser Arbeit, dass Chunks relevante und reale Einheiten im Sprachgebrauch sind, so dass sich die Analyse vorläufig auf den Chunkaspekt beschränken darf. Tiefere Analyseebenen betreffen die abstrakteren Teile der Konstruktion, deren Existenz und Analysebedarf hier nicht angezweifelt werden sollen.

### 5.3 Frames

Einen wichtigen Anknüpfungspunkt für die Organisation von Chunks bietet die Frame-Semantik. Sie nimmt an, dass semantische Einheiten nicht für sich stehen, sondern assoziative Verkettungen

mit sich bringen, und erklärt daraus zum Beispiel die Rollenvergabe von Verben. So ist die Annahme, dass das Verb *verkaufen* einen Verkaufsframe evoziert, der die typischen Bestandteile der Verkaufssituation beinhaltet: Eine Verkäuferin, eine Käuferin, das Gekaufte. Dieser Teil des Rahmens erklärt Verbvalenzen und Argumentstrukturen. Der Verkaufsframe kann auch situationelle Inhalte aufnehmen, wie etwa Ablaufroutinen, die als Skripte bezeichnet werden. Ein Beispiel wäre: Geschäft betreten – Ware auswählen – Bezahlen – Ware einpacken – Geschäft verlassen. Skripte und Argumentstrukturen sind von mittlerem Abstraktionsgrad. Es sind abstraktere Teile derselben Frames vorstellbar, etwa der Ditransitivframe (der sich teilweise mit der semantischen Seite der abstrakten Ditransitiv-Konstruktion deckt), oder der Interaktionsframe (der abstraktere Informationen zum Umgang mit Verkäuferinnen oder Käuferinnen enthält). Ebenso sind konkretere Frameausschnitte denkbar, die etwa spezifische Skripte zum Verhalten im Supermarkt, Sportgeschäft oder auf dem Markt beinhalten. Außerdem wären Chunks gut als Auflistungen von konkreten, konventionellen Äußerungen im Rahmen von Frames denkbar. Bisher ist mir eine solche Auflistung nur bei Antonopoulou und Nikiforidou (2011) zu Gesprächseröffnungen am Telefon begegnet. Dabei bieten Frames ein hervorragendes Mittel für die Modellierung der Organisation des Konstruktions und zur Erklärung der Selektion von spezifischen Chunks in gegebenen Situationen, zumal sie sich ideal mit konnektionistischen und exemplarbasierten Ansätzen verbinden lassen. So würde ein Skript, das Chunks als konkrete Zeichen enthält, also neurologisch aktiviert und somit verfügbar macht, die Auswahl derselben wahrscheinlich machen. Aus der Erfahrung der Austauschbarkeit von Untereinheiten könnte es aber flexibilisiert werden und Chunks als Exemplare für verwandte Äußerungen „freigeben“.

#### **5.4 Chunks und Exemplare**

Seit Kurzem zeichnet sich in der Konstruktionsgrammatik ein Trend zu einer exemplarbasierten Ausrichtung ab, der insbesondere auf der oben beschriebenen Häufung von Daten fußt, die auf eine starke lexikalische Restriktion von als abstrakt formulierten Konstruktionen verweisen. So beschreiben Hampe und Schönfeld (2006), Yoshikawa (2012) und Bybee (erscheint) die Notwendigkeit, Konstruktionen aus konkreten Exemplaren heraus zu erklären, und andere Autorinnen arbeiten an der Erstellung von exemplarbasierten Modellen, die zum Beispiel phonetische und syntaktische Aspekte verbinden (Walsh et al. 2010). Obwohl diese Ansätze in eine ähnliche Stoßrichtung gehen wie der hier vorgestellte, bin ich nicht überzeugt, dass Chunks und Exemplare identische Einheiten sind. Insbesondere im von Yoshikawa (2012) beschriebenen Modell



wird deutlich, dass es bei den Exemplaren vor Allem um prototypische Beispiele für Abstraktionen geht, deren Anteile austauschbar sind. Chunks, mit den funktionalen Eigenschaften, wie sie hier beschrieben sind, sind mehr als das, sie können aber auch weniger sein. Denn nicht jedes Chunk muss ein (prototypisches) Exemplar für die Abstraktion sein, was wahrscheinlich insbesondere für grammatisch oder semantisch intransparente Chunks gilt. Im bisher beschriebenen Sinn ist jedes Exemplar sicher ein Chunk, aber nicht jedes Chunk ein Exemplar. Das bedeutet nicht, dass die hier vorgestellte These mit einer exemplarbasierten Ausrichtung unvereinbar wäre. Im Gegenteil besteht ein großes Synergiepotential, dessen Realisierung jedoch die genaue Abgrenzung der Kategorien und der Abgleich mit empirischen Daten vorausgehen muss.

## 6. Zusammenfassung und Fazit

Im Rahmen dieser Arbeit wurde nach der Einleitung in Abschnitt 2 die Konstruktionsgrammatik in ihrem gegenwärtigen Stand vorgestellt und theoretische Unschärfen angesprochen, die vorwiegend die Organisation des Konstruktionskontinuums, den widersprüchlichen Umgang mit der Annahme eines Lexikon-Grammatik-Kontinuums und die Einbeziehung der mündlichen Sprache betreffen. In Abschnitt 3 wurde die Studienlage zu Chunks aufgearbeitet, indem vorwiegend phraseologische, korpuslinguistische, psycholinguistische und L2-Erwerbsstudien herangezogen wurden. Daraus wurde deutlich, dass Chunks sehr große Teile der gebrauchten Sprache, insbesondere der mündlichen Sprache, ausmachen, und dass sie zahlreiche Funktionen haben: mnemotechnische, stabilisierende, identitätstiftende, frame-evozierende und wahrscheinlich zumindest teilweise verarbeitungserleichternde. Es wurde außerdem gezeigt, dass auch Studien, die nicht gezielt Chunks untersuchen, deutliche Hinweise auf starke Anziehungskräfte zwischen lexikalischen Einheiten in wiederkehrenden grammatischen Formen finden – Hinweise auf bislang unerkannte Chunks. Diese Studien betrachten grammatisch sehr unterschiedliche Phänomene. Aus dieser Studienlage wurde in Abschnitt 4 eine Definition von Chunks als wiederkehrende, gleichbleibende, in enger Zeitfolge zusammenstehende Einheiten, die zentral für Sprache im Gebrauch sind, mündlich und schriftlich, und Exemplarfunktion haben, abgeleitet. Daraus erwächst der Vorschlag, Chunks als *Basic Level Category* für die Grammatik- und gebrauchsbasierte Sprachtheorie zu betrachten. Chunks wären damit die eigentlichen Grundbausteine der Sprache im Gebrauch, die zum konkreteren und abstrakteren Ende des Kontinuums hin durchlässig sind, aber in sich stabil, weil sie direkt kommunikative Funktionen und perzeptuell erreichbare Formseiten haben und stark situativ konnotierte Bedeutungsseiten haben können. In Abschnitt 5 wurde zunächst erläutert, warum es

wünschenswert ist, Chunks auch als grammatische Phänomene zu begreifen: Weil sie sich für die Beschreibung des Sprachgebrauchs und für die psychologische Plausibilität, wie auch aus der Datenlage zu grammatischen Phänomenen und aus der Grammatikalisierungsforschung aufdrängen, und weil sie eine gute Möglichkeit sind, um die Diskussion um die Struktur des Konstruktions zu beginnen. Chunks als zentrale, als Basic Level Category, erlauben daher den Beginn einer Lösung gleich mehrerer theoretischer Probleme der Konstruktionsgrammatik: Sie verbinden Lexikon und Grammatik zu einem tatsächlichen Kontinuum, sie erlauben Aussagen über die Unterschiedlichkeit von schriftlicher und mündlicher Sprache, sie erlauben Interaktionen untereinander und können so erklären, wie es zu „Fehlkonstruktionen“ durch Überlappungen kommt, sie stellen den Bezug zum kommunikativen Aspekt der Sprache her. Und sie sind problemlos auch in formale Modelle integrierbar, da sie auch neurologische Abdrücke, grammatische Fußspuren, sozusagen, hinterlassen, die sich abstrakt modellieren lassen. Einen Aspekt der Konstruktionsgrammatik teilt diese Analyse allerdings nicht: Wenn Chunks zentrale Einheiten sind, dann ist dieses Modell kein striktes Vereinigungsmodell mehr (obwohl sie in geringerem Maße Vereinigungen mit anderen Elementen erlauben). Das muss nicht schaden, wenn sich herausstellt, dass strikte Unifikation kognitiv nicht plausibel ist und natürliche Sprache besser zu erklären ist, wenn nicht zu viel Unifikation und mehr Aneinanderreihung angenommen wird.

Zuletzt wurde anhand von Attribut-Wert-Matrizen, einem Formalismus, wie er in der Sign Based Construction Grammar genutzt wird, und einer schematischen Darstellung erläutert, wie die theoretische Modellierung von Chunks aussehen kann, und angerissen, in welcher Beziehung die chunkbasierte Analyse, wie sie hier vorgestellt ist, zu exemplarbasierten und frame-semantischen Ansätzen steht.

Ein Nachteil bei der analytischen Einbeziehung von Chunks ist, dass die kombinatorischen Möglichkeiten um ein vielfaches im Verhältnis zu Wörtern + Konstruktionen gesteigert werden, und die Streitbarkeit jeder Analyse erhöht wird. Diese Problematik ist ernstzunehmen, sie ist jedoch unumgänglich, wenn sich die Grundannahmen als zutreffend erweisen sollten. Dazu müssen die Annahmen in dieser Arbeit an geeigneten Testhypothesen geprüft werden.

## 7. Literatur

- Aitchison, J. (2012): *Words in the mind: An introduction to the mental lexicon*. Hoboken.
- Antonopoulou, E. and K. Nikiforidou (2011): Construction grammar and conventional discourse: A construction-based approach to discursal incongruity. In: *Journal of Pragmatics* 43 (2011). Pp. 2594–2609 .
- Barðdal, J. (to appear): Construction-Based Historical-Comparative Reconstruction. To appear in: G. Trousdale and T. Hoffmann (eds): *Oxford Handbook of Construction Grammar*. <http://org.uib.no/iecastp/barddal/Construction-Based%20Historical-Comparative%20Reconstruction.pdf> (07.03.2012).
- Bencini, G. and Goldberg, A. (2000): The contribution of argument structure constructions to sentence meaning. In: *Journal of Memory and Language*, 43. Pp. 640-651.
- Bergen, B. and N. Chang (2005): Embodied Construction Grammar in simulation-based language understanding. In: J.-O. Östman and M. Fried (eds.): *Construction Grammars: Cognitive grounding and theoretical extensions*. *Constructional Approaches to Language*, vol. 3. Amsterdam. Pp. 147-190.
- Biber, D. and J. Jones (2007): Vocabulary-based discourse units in biology research articles. In: D. Biber, U. Connor and T. Upton (eds.): *Discourse on the Move: Using corpus analysis to describe discourse structure*. *Studies in Corpus Linguistics*, vol. 28. Amsterdam/Philadelphia. Pp. 175-212.
- Bladas, Ó. (2012): Conversational routines, formulaic language and subjectification . In: *Journal of Pragmatics* 44 (2012). Pp. 929-957 .
- Boas, H. C. (2003): *A Constructional Approach to Resultatives*. Stanford.
- Boas, H. C. (2005): Determining the Productivity of Resultatives: A Reply to Goldberg and Jackendoff . In: *Language*, Vol. 81, No. 2 (Jun., 2005). Pp. 448-464 .
- Boas, H. C. (2010): The syntax lexicon continuum in Construction Grammar: A case study of English communication verbs. In: *Belgian Journal of Linguistics*, 24(1). Pp. 54-82.
- Boas, H.C. (2011): Zum Abstraktionsgrad von Resultativkonstruktionen. In: S. Engelberg, A. Holler und K. Proost (Hgg.): *Sprachliches Wissen zwischen Lexikon und Grammatik*. Berlin/Boston. S. 37-69.
- Boroditsky, L. (2000): Metaphoric structuring: Understanding time through spatial metaphors . In: *Cognition* 75 (2000). Pp. 1-28 .
- Bybee, J. (2006): From Usage to Grammar: The Mind's Response to Repetition. In: *Language* 82(4). Pp. 711-773.
- Bybee, J. (to appear): Usage-based Theory and Exemplar Representations of Constructions. To appear in: G. Trousdale and T. Hoffmann (eds): *Oxford Handbook of Construction Grammar*.
- Bybee, J. and R. Cacoulios (2009): The role of prefabs in grammaticization: How the particular and the general interact in language change. In: R. Corrigan, E. Moravcsik, H. Ouali and K. Wheatly (eds.): *Formulaic Language, Vol. 1: Distribution and historical change*. Amsterdam/Philadelphia. Pp. 187-218.

- Cacoullos, R. and J. Walker (2009): The present of the English future: Grammatical variation and collocations in discourse. In: *Language*, 85(2). Pp. 321-354.
- Calude, A. (2009). Formulaic tendencies of demonstrative clefts in spoken English. In: R. Corrigan, E. Moravcsik, H. Ouali and K. Wheatly (eds.): *Formulaic Language*, Vol. 1: Distribution and historical change. Amsterdam/Philadelphia. Pp. 55-76.
- Carroll, S. (2012): When is input salient? An exploratory study of sentence location and word length effects on input processing . In: *IRAL* 50 (2012). Pp. 39–67 .
- Casenhiser, D. And A. Goldberg (2005): Fast mapping between a phrasal form and meaning. In: *Development Science*, 8(6). Pp. 500-508.
- Cheng, W., C. Greaves, J. Sinclair and M. Warren (2009): Uncovering the Extent of the Phraseological Tendency: Towards a Systematic Analysis of Concrgrams In: *Applied Linguistics* 30/2. Pp. 236–252.
- Chomsky, N. (1965): *Aspects of the Theory of Syntax*. Cambridge, MA.
- Coxhead, A. (2012): Academic Vocabulary, Writing and English for Academic Purposes: Perspectives from Second Language Learners. IN: *RELC Journal* 43 (2012), 1. Pp. 137-145.
- Croft, W. (2001): *Radical Construction Grammar: syntactic theory in typological perspective*. Oxford.
- Croft, W. (2004): Logical and typological arguments for Radical Construction Grammar. In: J.-O. Östman and M. Fried (eds.): *Construction Grammars: Cognitive grounding and theoretical extensions*. *Constructional Approaches to Language*, vol. 3. Amsterdam/Philadelphia. Pp. 273-314.
- Crossley, S. and. T. Salsbury (2011): The development of lexical bundle accuracy and production in English second language speakers. In: *IRAL* (50). Pp. 1-26.
- Csomay, E. (2007): Vocabulary-based discourse units in university class sessions. In: D. Biber, U. Connor and T. Upton (eds.): *Discourse on the Move: Using corpus analysis to describe discourse structure*. *Studies in Corpus Linguistics*, vol. 28. Amsterdam/Philadelphia. Pp. 213-238.
- Csomay, E. (2012): Lexical Bundles in Discourse Structure: A Corpus-Based Study of Classroom Discourse. In: *Applied Linguistics* 2012. Pp. 1–21 .
- Deppermann, A, (2011): Konstruktionsgrammatik und Interaktionale Linguistik: Affinitäten, Komplementaritäten und Diskrepanzen. In: A. Lasch und A. Ziem (Hgg.): *Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze*. Tübingen. S. 205-238.
- Dobrovolskij, D. (2011): Phraseologie und Konstruktionsgrammatik. In: A. Lasch und A. Ziem (Hgg.): *Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze*. Tübingen. S. 111-130.
- Ellis, N. (1996): Sequencing in SLA: Phonological Memory, Chunking and Points of Order. In: *SSLA*, 18. Pp. 91-126.
- Ellis, N. (1998): Emergentism, connectionism and language learning. In: *Language Learning* 48. Pp. 631-664.

- Ellis, N. (2003): Constructions, chunking, and connectionism: The emergence of second language structure. In: C. Doughty and M. H. Long (eds.): *Handbook of second language acquisition*. Oxford. Pp. 33-68.
- Ellis, N. (2008). *Phraseology: The periphery and the heart of language*. In F. Meunier and S. Granger (eds.): *Phraseology in Foreign Language Learning and Teaching*. Amsterdam/Philadelphia. Pp. 1–13.
- Ellis, N. (2011): Frequency-based accounts of SLA. In: S. Gass and A. Mackey (Eds.): *Handbook of Second Language Acquisition*. London. Pp. 193-210.
- Ellis, N. and L. Collins (2009): *Input and Second Language Acquisition: The Roles of Frequency, Form, and Function*. Introduction to the Special Issue. In: *The Modern Language Journal*, 93, iii.
- Ellis, N. and D. Larsen-Freeman (2009): *Constructing a Second Language: Analyses and Computational Simulations of the Emergence of Linguistic Constructions from Usage*. In: *Language Learning* 59: Suppl.1. Pp. 90–125.
- Ellis, N., R. Simpson-Vlach and C. Maynard (2008): Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. In: *Tesol Quarterly* 42 (3). Pp. 375-396.
- Enikő N. and B. Károly (2010): Interaction between grammar and pragmatics: The case of implicit arguments, implicit predicates and co-composition in Hungarian. In: *Journal of Pragmatics*, 42(2). Pp. 501-524.
- Erman, B. and B. Warren (2000): The idiom principle and the open choice principle. In: *Text* 20(1). Pp. 29-62.
- Evans, V. and M. Green (2006); *Cognitive linguistics: An introduction*. Edinburgh.
- Faulhaber, S. (2011): *Verb Valency Patterns: A challenge for semantics-based accounts*. *Topics in English Linguistics*, vol. 71. Berlin/New York.
- Fillmore, C. (1982/2011): *Frame Semantics*. In: Linguistic Society of Korea (ed.): *Linguistics in the Morning Calm*. Seoul. Pp. 373-400. Reprinted in: F. Stjernfelt and P. Bundgaard (eds.): *Semiotics. Vol 2: Linguistics. Critical Concepts in Language Studies*. London/New York. Pp.187-213.
- Fillmore, C. (1988), *The Mechanisms of ‘Construction Grammar’*. In: *Proceedings of the Berkeley Linguistics Society* 14. Pp. 35–55.
- Fillmore, C., P. Kay and C. O'Connor (1988): *Regularity and Idiomaticity in Grammatical Constructions*. In: *Language* 64. Pp. 501–38.
- Fischer, K. und A. Stefanowitsch (Hgg.) (2006a): *Konstruktionsgrammatik: Von der Anwendung zur Theorie*. Tübingen.
- Fischer, K. und A. Stefanowitsch (2006b): *Konstruktionsgrammatik: Ein Überblick*. In: Kerstin Fischer and Anatol Stefanowitsch (Hgg.): *Konstruktionsgrammatik: Von der Anwendung zur Theorie*. Tübingen. S. 3-15.
- Fried, M. and H.C. Boas (eds.)(2005): *Grammatical Constructions: Back to the roots*. *Constructional Approaches to Language* vol. 4. Amsterdam/Philadelphia.

- Fried, M. and J.-O. Östman (2005): Construction Grammar and spoken language: The case of pragmatic particles . In: Journal of Pragmatics 37 (2005). Pp. 1752–1778 .
- Goldberg, A. (1995): Constructions: A construction grammar approach to argument structure. Chicago.
- Goldberg, A. (2006): Constructions at work: The nature of generalization in language. Oxford.
- Granger, S. And F. Meunier (eds.)(2007a): Phraseology: An interdisciplinary perspective. Amsterdam/Philadelphia.
- Granger, S. And F. Meunier (2007b): Introduction: The many faces of phraseology. In: S. Granger and F. Meunier (eds.): Phraseology: An interdisciplinary perspective. Amsterdam/Philadelphia. Pp. XIV-XXVIII.
- Granger, S. and M. Paquot (2007): Disentangling the phraseological web. In: In: S. Granger and F. Meunier (eds.): Phraseology: An interdisciplinary perspective. Amsterdam/Philadelphia. Pp. 28-50.
- Gries, S. (2007): Phraseology and linguistic theory. In: S. Granger and F. Meunier (eds.): Phraseology: An interdisciplinary perspective. Amsterdam/Philadelphia. Pp. 3-25.
- Gries, S. (2011): Phonological similarity in multi-word units. In: Cognitive Linguistics 22–3 (2011). Pp. 491–510.
- Gries, S. and A. Stefanowitsch (2004): Extending collocation analysis: a corpus-based perspective on 'alternations'. In: International Journal of Corpus Linguistics 9(1). Pp. 97-129.
- Gries, S. and S. Wulff (2005): Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. In: Annual Review of Cognitive Linguistics 3. 182-200.
- Guckelsberger, S. (2008): Diskursive Basisqualifikation. In: K. Ehlich, Konrad, U. Bredel und H. Reich (Hgg.): Referenzrahmen zur altersspezifischen Sprachaneignung – Forschungsgrundlagen. Bildungsforschung Band 29/II. Berlin/Bonn. S. 103-134.
- Hampe, B. and D. Schönefeld (2006): Syntactic leaps or lexical variation? - More on „Creative Syntax“. In: S. Gries and A. Stefanowitsch (eds.): Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis. Berlin. Pp. 127-158.
- Handwerker, B. und K. Madlener (2009): Chunks für DaF: Theoretischer Hintergrund und Prototyp einer multimedialen Lernumgebung. Hohengehren.
- Helasvuo, M. (2004): Shared syntax: the grammar of co-constructions. In: Journal of Pragmatics 36 (2004). Pp. 1315–1336 .
- Höche, S. und A. Shahrokny-Prehn (2011): Arbeitsteilung zwischen Konstruktionen? Zur diachronen Entwicklung von *have a V* vs. *take a V*. In: A. Lasch und A. Ziem (Hgg.): Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze. Tübingen. S. 95-110.
- Hoey, M. (2004): Textual colligation: a special kind of lexical priming. In: K. Aijmer and B. Altenberg (eds.): Language and Computers, Advances in Corpus Linguistics. Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23) Göteborg 22-26 May 2002. Pp. 171-194.

- Hoey, M. (2005): *Lexical Priming: A new theory of words and language*. Oxon/New York.
- Hudson, R. (2007): *Language networks: the new Word Grammar*. Oxford, USA.
- Hunston, S. (2001). Colligation, lexis, pattern, and text. In: M. Scott and G. Thompson (eds.): *Patterns of Text: In Honour of Michael Hoey*. Pp. 13-33.
- Hunston, S. and G. Francis (2000): *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. *Studies in Corpus Linguistics*, vol. 4. Amsterdam/Philadelphia.
- Kanoksilapatham, B. (2007a): Introduction to move analysis. In: D. Biber, U. Connor and T. Upton (eds.): *Discourse on the Move: Using corpus analysis to describe discourse structure*. *Studies in Corpus Linguistics*, vol. 28. Amsterdam/Philadelphia. Pp. 23-42.
- Kanoksilapatham, B. (2007b): Rhetorical moves in biochemistry research articles. In: D. Biber, U. Connor and T. Upton (eds.): *Discourse on the Move: Using corpus analysis to describe discourse structure*. *Studies in Corpus Linguistics*, vol. 28. Amsterdam/Philadelphia. Pp. 73-120.
- Kawabe, K. And Y. Matsumoto (1998): Acquisition of normal lexical knowledge based on basic level category. In: *Information Processing Society of Japan, SIG note*, 125(9). Pp. 87-92.
- Kerz, E. And F. Haas (2009): The aim is to analyze NP: The function of prefabricated chunks in academic texts. In: R. Corrigan, E. Moravcsik, H. Ouali and K. Wheatly (eds.): *Formulaic Language, Vol. 1: Distribution and historical change*. Amsterdam/Philadelphia. Pp. 97-116.
- Kuiper, K. (2004): Formulaic performance in conventionalised varieties of speech. In: N. Schmitt (ed.): *Formulaic Sequences: Acquisition, processing and use*. *Language Learning and Language Teaching*, vol. 9. Amsterdam/Philadelphia. Pp. 37-54.
- Langacker, R. (1987): *Foundations of Cognitive Grammar, Volume 1, Theoretical Prerequisites*. Stanford.
- Langacker, R. (2008): *Cognitive Grammar: A Basic Introduction*. Oxford/New York.
- Lasch, A. und A. Ziem (Hgg.)(2011): *Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze*. Tübingen.
- Lieven, E., H. Behrens, J. Speares and M. Tomasello (2003): Early syntactic crativity: a usage-based approach. In: *Journal of Child Language* 30 (2). Pp. 333-370.
- Li, J. And N. Schmitt (2009): The acquisition of lexical phrases in academic writing: A longitudinal case study . In: *Journal of Second Language Writing* 18 (2009). Pp. 85–102 .
- Maclagan, M., B. Davis and R. Lunsford (2007): Fixed expressions, extenders and metonymy in the speech of people with Alzheimer's disease. S. Granger and F. Meunier (eds.): *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia. Pp. 175-187.
- Martin, W. (2007): A unified approach to semantic frames and collocational patterns. In: S. Granger and F. Meunier (eds.): *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia. Pp. 51-65.
- Mauranen, A. (2009): Chunking in ELF: Expressions for managing interaction. In: *Intercultural Pragmatics* 6-2 (2009). Pp. 217–233.
- Melinger, A. and C. Döbel (2005): Lexically-driven syntactic priming . In: *Cognition* 98 (2005). P. B11–B20 .

- Messenger K., H. Branigan, J. McLean and A. Sorace (2012): Is young children's passive syntax semantically constrained? Evidence from syntactic priming . In: *Journal of Memory and Language* 66 (2012). Pp. 568–587.
- Moon, R. (1999): Needles and haystacks, idioms and corpora: Gaining insights into idioms, using corpus analysis. In: Th. Herbst and K. Popp (Hg.): *The Perfect Learners' Dictionary*. Tübingen. Pp. 265-281.
- Müller, S. (2008): *Head-Driven Phrase Structure Grammar: Eine Einführung*. Tübingen. 2., überarbeitete Auflage.
- Myles, F., J. Hooper and R. Mitchell (1998): Rote or rule? Exoring the role of formulaic language in classroom foreign language learning. In: *Language Learning* 48(3). Pp. 323-363.
- Myles, F., R. Mitchell and J. Hooper (1999): Interrogative chunks in French L2: a basis for creative construction? In: *Studies in Second Language Acquisition* 21(1). Pp. 49-80.
- Östman, J.-O. and M. Fried (eds.) (2005): *Construction Grammars: Cognitive grounding and theoretical extensions*. *Constructional Approaches to Language*, vol. 3. Amsterdam/Philadelphia.
- Pawley, A. And F. Syder (1983): Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: *Language and communication* 191. P. 225.
- Pickering, M. and S. Garrod (2006): Alignment as the Basis for Successful Communication. In: *Research on Language and Computation*, vol. 4, issue 2-3. Pp 203-228.
- Römer, U. (2005): *Progressives, Patterns, Pedagogy: A corpus-driven approach to English progressive forms, functions, contexts and didactics*. *Studies in Corpus Linguistics*, vol. 18. Amsterdam/Philadelphia.
- Rosch, E. (1978): Principles of Categorization. In: E. Rosch and B. Lloyd (eds.): *Cognition and Categorization*. Hillsdale, NJ. Pp. 27-48.
- Sag, I., H. Boas and P. Kay (2012): Introducing Sign-Based Construction Grammar. In: H. Boas and I. Sag (eds.): *Sign-based construction grammar*. CSLI lecture notes no. 193. Pp. 1-30.
- Sag, I. (2012): *Sign-Based Construction Grammar: An Informal Synopsis*. In: H. Boas and I. Sag (eds.): *Sign-based construction grammar*. CSLI lecture notes no. 193. Pp. 69-202.
- Santesteban, M., M. Pickering and J. McLean (2010): Lexical and phonological effects on syntactic processing: Evidence from syntactic priming . In: *Journal of Memory and Language* 63 (2010). Pp. 347–366 .
- Saussure, F. (1916/1967): *Grundfragen der allgemeinen Sprachwissenschaft*. Berlin. Zuerst: Lausanne (1916).
- Schmitt, N. (ed.)(2004): *Formulaic Sequences: Acquisition, processing and use*. *Language Learning and Language Teaching*, vol. 9. Amsterdam/Philadelphia.
- Schmitt, N. and R. Carter (2004): Formulaic sequences in action: An introduction. In: N. Schmitt (ed.): *Formulaic Sequences: Acquisition, processing and use*. *Language Learning and Language Teaching*, vol. 9. Amsterdam/Philadelphia. Pp. 1-22.
- Schmitt, N., S. Grandage and S. Adolphs (2004): Are corpus-derived recurrent clusters



- psycholinguistically valid? In: N. Schmitt (ed.): *Formulaic Sequences: Acquisition, processing and use. Language Learning and Language Teaching*, vol. 9. Amsterdam/Philadelphia. Pp. 127-151.
- Simpson-Vlach, R. and N. Ellis (2010): *An Academic Formulas List (AFL)*. In: *Applied Linguistics*, 31. Pp. 487-512.
- Sinclair, J. (1991): *Corpus, concordance, collocation*. Oxford.
- Sinclair, J. (2007): Preface. In: S. Granger and F. Meunier (eds.): *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia. Pp. XV-XIII.
- Sinclair, J. and A. Mauranen (2006): *Linear Unit Grammar: Integrating Speech and Writing*. *Studies in Corpus Linguistics* 25. Amsterdam/Philadelphia.
- Staffeldt, S. (2011): In der Hand von Konstruktionen: Eine Fallstudie zu bestimmten Phraseologismen mit *in ... Hand*. In: A. Lasch und A. Ziem (Hgg.): *Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze*. Tübingen. S. 131-148.
- Stefanowitsch, A. (2006): *Konstruktionsgrammatik und Korpuslinguistik*. In: K. Fischer und A. Stefanowitsch (Hgg.): *Konstruktionsgrammatik: Von der Anwendung zur Theorie*. Tübingen. S. 151-176.
- Stefanowitsch, A. (2011): *Konstruktionsgrammatik und Grammatiktheorie*. In: A. Lasch und A. Ziem (Hgg.): *Konstruktionsgrammatik III: Aktuelle Fragen und Lösungsansätze*. Tübingen. S. 11-26.
- Stefanowitsch, A. und K. Fischer (Hgg.) (2008): *Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik*. Tübingen.
- Stefanowitsch, A. and S. Gries (2003): *Collostructions: investigating the interaction between words and constructions*. In: *International Journal of Corpus Linguistics* 8(2). Pp. 209-243.
- Stefanowitsch, A. and S. Gries (2005): *Covarying collexemes*. In: *Corpus Linguistics and Linguistic Theory* 1(1). Pp. 1-43.
- Stengers, H., F. Boers, A. Housen and J. Eyckmans (2011): *Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association?* In: *IRAL* 49 (2011). Pp. 321-343 .
- Svensson, M.(2007): *A very complex criterion of fixedness: Non-compositionality*. In: S. Granger and F. Meunier (eds.): *Phraseology: An interdisciplinary perspective*. Amsterdam/Philadelphia. Pp. 81-93.
- Taguchi, N. (2008). *Building Language blocks in L2 Japanese: Chunk learning and the development of complexity and fluency in spoken production*. *Foreign Language Annals*, 41, 130-154.
- Tomasello, M. (1995): *Language is not an instinct*. In: *Cognitive development*, 10(1). Pp. 131-156.
- Tomasello, M. (2003): *Constructing a language: A usage-based account of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., M. Carpenter, J. Call, T. Behne and H. Moll (2005/2011): *Understanding and Sharing Intentions*. In: *Behavioral and Brain Sciences*, 18 (2005), Pp. 675-691, 727-735.

Reprinted in: F. Stjernfeld and P. F. Bundgaard (eds.): Semiotics. Vol. 4: Logic, Biology, Psychology, Culture and Anthropology. Critical Concepts in Language Studies. London/New York. Pp. 510-553.

Tomasello, M. and D. Stahl (2004): Sampling children's spontaneous speech: How much is enough? In: Journal of Child Language, 31. Pp 101-121.

Underwood, G., N. Schmitt, and A. Galpin (2004): The eyes have it: An eye-movement study into the processing of formulaic sequences. In: N. Schmitt (ed.): Formulaic Sequences: Acquisition, processing and use. Language Learning and Language Teaching, vol. 9. Amsterdam/Philadelphia. Pp. 153-172.

Walsh, M., B. Möbius, T. Wade and H. Schütze (2010): Multilevel exemplar theory. Cognitive science, 34(4). In: 537-582.

Watson, M., M. Pickering and H. Branigan (2004): Alignment of Reference Frames in Dialogue. Proceedings of the 26th Annual Conference of the Cognitive Science Society 2004. Chicago.

Wible, D. (2008): Multiword expressions and the digital turn. In: F. Meunier and S. Granger (eds.): Phraseology in Foreign Language Learning and Teaching. Amsterdam/Philadelphia. Pp. 163-181.

Wray, A. (2002): Formulaic Language and the Lexicon. Cambridge.

Wray, A. and T. Fitzpatrick (2007): Why can't you just leave it alone? Deviations from memorized language as a gauge of nativelike competence. In: F. Meunier and S. Granger (eds.): Phraseology in Foreign Language Learning and Teaching. Amsterdam/Philadelphia. Pp. 124-147.

Wulff, Stefanie (2008). Rethinking Idiomaticity: A Usage-based Approach. Dissertation. London/New York.

Yorio, C. (1980): Conventionalized language forms and the development of communicative competence. In: TESOL Quarterly 14(4). Pp. 433-442.

Yoshikawa, M. (2012): Exemplars supersede abstractions: Presenting an exemplar theory of grammatical constructions. In: ACADEMIC REPORTS Fac. Eng. Tokyo Polytech. Univ. Vol. 35 No.2 (2012)

Zeldes, Amir (to appear). Productivity in Argument Selection: A Usage-Based Approach to Lexical Choice in Syntactic Slots. Dissertation 2011.

## Abbildungsverzeichnis

Abb. 1: Schematische Darstellung der CxG-Analyse von 'Sie streicht die Wand rot'.....	47
Abb. 2: Chunkbasierte CxG-Analyse von 'Sie streicht die Wand rot'.....	48
Abb. 3: Chunkbasierte Analyse von 'Ich habe die Wand rot gestrichen'.....	50
Abb. 4: Chunkanteil in Abhängigkeit von der Modalität und Funktion .....	51
Abb. 5: Sag, Boas und Kay 2012, 17: Darstellung der Äußerungen als vererbende Konstruktionen	52
Abb. 6: Zeichencharakter der Äußerung 'every linguist' .....	53
Abb. 7: Formalisierung einer Resultativkonstruktion in einer Attribut-Wert-Matrix .....	54