Head-Driven Phrase Structure Grammar: The handbook

Edited by

Stefan Müller

Anne Abeillé

Robert D. Borsley

Jean-Pierre Koenig

Empirically Oriented Theoretical Morphology and Syntax 99

language science press

Empirically Oriented Theoretical Morphology and Syntax

In this series:

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

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

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

ISSN: 2366-3529

Head-Driven Phrase Structure Grammar: The handbook

Edited by

Stefan Müller

Anne Abeillé

Robert D. Borsley

Jean-Pierre Koenig

language science press

Stefan Müller , Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (ed.). 2018. *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax 99). Berlin: Language Science Press.

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

ISBN: no digital ISBN
no print ISBNs!
ISSN: 2366-3529
no DOI
Source code available from www.github.com/langsci/000
Collaborative reading: paperhive.org/documents/remote?type=langsci&id=000

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

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

Storage and cataloguing done by FU Berlin

no logo

Contents

Preface

Acknowledgments

Part I

Introduction

Chapter 1

Basic properties and elements

Bob Borsley
University of Essex

Anne Abeillé
Université Paris Diderot

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

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

1 Introduction

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

Meier (2017)

Bob Borsley & Anne Abeillé. 2018. Basic properties and elements. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 3–5. Berlin: Language Science Press. DOI:??

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

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

Table 1: Frequencies of word classes

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

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

Abbreviations

COP	copula	NEG	negation
FV	final vowel	SM	subject marker

[1] eros condimentum mi consectetur, ac consectetur

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

Acknowledgements

References

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

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

Chapter 2

The evolution of HPSG

Dan Flickinger
Stanford University

Carl Pollard
Ohio State Universtiy

Tom Wasow
Stanford University

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

Introduction

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

Dan Flickinger, Carl Pollard & Tom Wasow. 2018. The evolution of HPSG. in Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 7–38. Berlin: Language Science Press. DOI:??

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

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

1 Precursors

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

1.1 Generalized Phrase Structure Grammar

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

These developments led a few linguists to begin questioning the central role transformations had played in the syntactic research of the preceding two decades (notably, Bresnan (1978)). This questioning of Transformational Grammar (TG) culminated in a series of papers by Gerald Gazdar, which (in those pre-internet days) were widely distributed as paper manuscripts. The project that they laid out was succinctly summarized in one of Gazdar's later publications (Gazdar 1981: 155) as follows:

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

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

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

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

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

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

1.2 Head Grammar

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

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

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

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

2 The HP NLP project

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

The initial stages consisted of the linguists and programmers coming up with a notation that would serve the purposes of both. Once this was accomplished, the linguists set to work writing a grammar of English in Lisp to run on the DEC-20 mainframe computer that they all worked on. The first publication coming out of this project was a 1982 Association for Computational Linguistics paper (Gawron et al. 1982). The paper's conclusion (p. 80) begins:

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

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

This goal was carried over into HPSG.

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

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

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

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

3 The emergence of HPSG

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

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

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

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

4 Early HPSG

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

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

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

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

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

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

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

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

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

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

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

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

(3) $$\frac{\mathrm{A}\quad \mathrm{A\backslash B}}{\mathrm{B}} \qquad \frac{\mathrm{B/A}\quad \mathrm{A}}{\mathrm{B}}$$

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

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

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

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

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

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

(5) $$\frac{\mathrm{B[SUBCAT} <\ldots, \mathrm{A}>\mathrm{,SLASH}\ \Gamma] \quad \mathrm{A[SLASH}\ \Delta]}{\mathrm{B[SUBCAT} <\ldots>][\mathrm{SLASH}\ \Gamma,\Delta]}$$

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

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

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

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

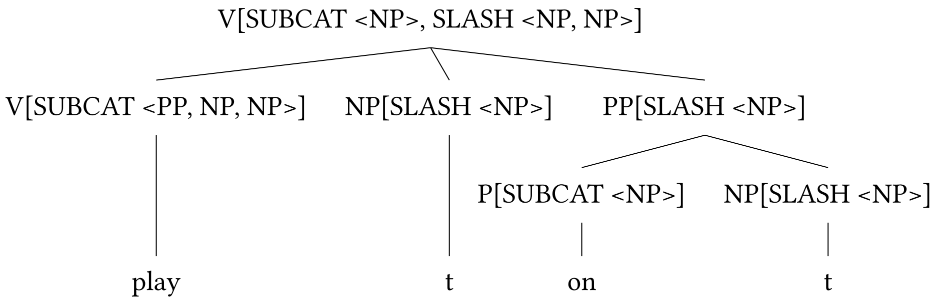


Figure 1: *play on* as part of *Violins this finely crafted, even the most challenging sonatas are easy to play on.*

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

An overview and elaboration of early HPSG is provided by the two books Pollard & Sag (1987) and Pollard & Sag (1994). Confusingly, the former is called *Information-Based Syntax and Semantics, Volume 1: Fundamentals*, and the second simply *Head-Driven Phrase Structure Grammar* (not *Information-Based Syntax and Semantics, Volume 2*). The reason for the title change had to do with a change in the underlying mathematical theory of feature structures. In the first book, following work in theoretical computer science by Rounds & Kasper (1986) and Moshier & Rounds (1987), feature structures were treated as data structures that supplied partial information about the linguistic objects being theorized about; this perspective in turn was based on Scott's (1982) mathematical theory of computation in terms of what he called information systems. Subsequently, Paul King persuaded Pollard and Sag that it was more straightforward to distinguish between feature structures, thought of as formal models of the linguistic objects, and feature descriptions or formulas of feature logic, which provided partial information about them, as described in his Manchester dissertation (King 1989). Although the formal issues involved in distinguishing between the two approaches are of interest in their own right, they seem not to have had a lasting effect on how theoretical linguists used HPSG, or on how computational linguists implemented it. As for subject matter, Pollard & Sag (1987) was limited to the most basic notions, including syntactic features and categories (including the distinction between head features and binding features); subcategorization and the distinction between arguments and adjuncts (the latter of which necessitated one more rule schema beyond the three proposed by Pollard (1985)); basic principles of grammar (especially the Head Feature Principle and the Subcategorization Principle); the obliqueness order and constituent ordering; and the organization of the lexicon by means of a multiple inheritance hierarchy and lexical rules. Pollard & Sag (1994) used HPSG to analyze a wide range of phenomena that had figured prominently in the syntactic literature of the 1960s–1980s, including agreement, expletive pronoun constructions, raising, control, filler-gap constructions (including island constraints and parasitic gaps); so-called binding theory (the distribution of reflexive pronouns, nonreflexive pronouns, and non-pronominal NPs), and scope of quantificational NPs.

5 The LinGO Project

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

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

The second four years of the Verb*mobil* project emphasized development of the generation capabilities of the ERG, along with steady expansion of linguistic coverage, and elaboration of the MRS framework. LinGO contributors in this phase, in addition to Sag, Wasow, Flickinger, Malouf, and Copestake, included Stanford Linguistics graduate students Emily Bender and Susanne Riehemann, along with a regular visitor and steady contributor from the DFKI, Stephan Oepen. Verb*mobil* had meanwhile added Japanese alongside German (Müller & Kasper 2000) and English (Flickinger, Copestake & Sag 2000) for more translation pairs, giving rise to another relatively broad-coverage HPSG grammar, JaCY, authored by Melanie Siegel at the DFKI (Siegel 2000). Work continued at the DFKI, of course, on the German HPSG grammar, written by Stefan Müller, adapted from his earlier Babel grammars (Müller 1999), and with semantics contributed by Walter Kasper.

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

6 Research and Teaching Networks

The first international conference on HPSG was held in 1993 in Columbus, Ohio, in conjunction with the Linguistic Society of America's Summer Institute. The conference has been convened every year since then, with locations in Europe, Asia, and North America. Two of these annual meetings have been held jointly with the annual Lexical Functional Grammar conference, in 2000 in Berkeley and in 2016 in Warsaw. Proceedings of these conferences since 2000 are available on-line from CSLI Publications.[3] Since 2003, HPSG researchers in Europe have frequently held a regional workshop in Bremen, Berlin, Frankfurt, or Paris, to foster informal discussion of current work in HPSG. These follow in the footsteps of European HPSG workshops starting with one on German grammar, held in Saarbrücken in 1991, and including others in Edinburgh and Copenhagen in 1994, and in Tübingen in 1995.

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

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

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

[3] http://csli-publications.stanford.edu/HPSG/, 2018-08-21.
[4] Its archives can be found at https://hpsg.hu-berlin.de/HPSG/MailingList.
[5] http://www.sfs.uni-tuebingen.de/~gazette, 2018-08-21.

7 Implementations and Applications of HPSG

The first implementation of a grammar in the HPSG framework emerged in the Hewlett-Packard Labs natural language project, for English, with a lexical type hierarchy (Flickinger, Pollard & Wasow 1985), a set of grammar rules that provided coverage of core syntactic phenomena including unbounded dependencies and coordination, and a semantic component called Natural Language Logic (Laubsch & Nerbonne 1991). The corresponding parser for this grammar was implemented in Lisp (Proudian & Pollard 1985), as part of a system called HP-NL (Nerbonne & Proudian 1987) which provided a natural language interface for querying relational databases. The grammar and parser were shelved when HP Labs terminated their natural language project in 1991, leading Sag and Flickinger to begin the LinGO project and development of the English Resource Grammar at Stanford.

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

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

In 1999, research groups at the DFKI, Stanford, and Tokyo set up a consortium called DELPH-IN (Initiative for Deep Linguistic Processing in HPSG), to foster broader development of both grammars and platform components, described in Oepen, Flickinger, Tsujii & Uszkoreit (2002). Over the next two decades, substantial DELPH-IN grammars were developed for Norwegian, Portuguese, and Spanish, along with moderate-coverage grammars for Bulgarian, Greek, Hausa, Hebrew, Indonesian, Mandarin Chinese, Thai, and Wambaya, all described at http://delph-in.net. Several of these grammars are based on the Grammar Matrix (Bender, Flickinger & Oepen 2002), a starter kit generalized from the ERG and Jacy for rapid prototyping of HPSG grammars, along with a much larger set of coursework grammars.[6]

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

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

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

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

& Malouf 2001), include analyses which the grammar may not yet be able to produce automatically.

8 Prospects

As we noted early in this chapter, HPSG's origins are rooted in the desire simultaneously to address the theoretical concerns of linguists and the practical issues involved in building a useful natural language processing system. In the decades since the birth of HPSG, the mainstream of work in both theoretical linguistics and NLP developed in ways that could not have been anticipated at the time. NLP is now dominated by statistical methods, with almost all practical applications making use of machine learning technologies. It is hard to see any influence of research by linguists in most NLP systems. Mainstream grammatical theory, on the other hand, is now dominated by the Minimalist Program (MP), a vaguely formulated[7] variant of Categorial Grammar (see Retoré & Stabler (2004), Berwick & Epstein (1995), and Müller (2013b) for comparison). Concern with computational implementation plays virtually no role in MP research; see Müller (2016) for a discussion.

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

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

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

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

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

More interesting are the connections that have been forged between research in HPSG and work in Construction Grammar (CxG). Fillmore (1988: 35) describes this approach as follows:

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

This (admittedly vague) description would include HPSG as a form of CxG (see Müller (2018), Chapter 36 of this volume for a much more detailed discussion). Nevertheless, work that labels itself CxG tends to look very different from HPSG. This is in part because of the difference in their origins: many proponents of CxG come from the tradition of *cognitive grammar* or typological studies, whereas HPSG's roots are in computational concerns. Hence, most of the CxG literature is not precise enough to allow a straightforward comparison with HPSG, though the variants called Embodied Construction Grammar and Fluid Construction Grammar have more in common with HPSG; see Müller (2017) for a comparison.

In the last years of his life, Ivan Sag sought to bridge this gap through collaboration with construction grammarians at the University of California at Berkeley, particularly Chuck Fillmore, Paul Kay, and Laura Michaelis. He developed a theory he called *Sign-Based Construction Grammar* (SBCG), which would combine the insights of CxG with the explicitness of HPSG. Sag (2012) wrote, "To readers steeped in HPSG theory, SBCG will no doubt seem like a minor variant of constructional HPSG." Indeed, despite the name change, the main feature of SBCG that differs from HPSG is that it posits an inheritance hierarchy of constructs, which includes feature structure descriptions for such partially lexicalized multiword expressions as Ved – X's – *way* – PP, instantiated in such VPs as *ad-libbed his way through a largely secret meeting*. While this is a non-trivial extension to HPSG, there is no fundamental change to the technical machinery. Indeed, it has been a part of the LinGO implementation for years. For a more detailed discussion, see the chapter on SBCG in this volume.

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

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

References

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

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

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

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

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

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

Berwick, Robert C. & Samuel David Epstein. 1995. On the convergence of the 'Minimalist' Syntax and Categorial Grammar. In Anton Nijholt, Giuseppe Scollo & Rene Steetskamp (eds.), *Algebraic methods in language processing*, 143–148. Enschede: University of Twente. http://eprints.eemcs.utwente.nl/9555/01/twlt10.pdf, accessed 2018-2-25.

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

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

Borsley, Robert D. 1996. *Modern Phrase Structure Grammar* (Blackwell Textbooks in Linguistics 11). Oxford, UK/Cambridge, USA: Blackwell Publishing Ltd.

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

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

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

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

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

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

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

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

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

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

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 4(3). 281–332.

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

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

Culy, Christopher. 1985. The complexity of the vocabulary of Bambara. *Linguistics and Philosophy* 8. 345–351.

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

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

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

Dowty, David R. 1982a. Grammatical relations and Montague Grammar. In Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic representation* (Synthese Language Library 15). Dordrecht: D. Reidel Publishing Company.

Dowty, David R. 1982b. More on the categorial analysis of grammatical relations. In Annie Zaenen (ed.), *Subjects and other subjects: Proceedings of the Harvard*

Conference on the Representation of Grammatical Relations, 115–153. Bloomington: Indiana University Linguistics Club.

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

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

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

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

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

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

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

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

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

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

Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Müller, Stefan. 2013a. *Grammatiktheorie*. 2nd edn. (Stauffenburg Einführungen 20). Tübingen: Stauffenburg Verlag. https://hpsg.hu-berlin.de/~stefan/Pub/grammatiktheorie.html, accessed 2018-3-20.

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

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

Müller, Stefan. 2016. *Grammatical theory: From Transformational Grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

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

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

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

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

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

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

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

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

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

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

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

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

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

Peters, Stanley & R. W. Ritchie. 1973. On the generative power of Transformational Grammar. *Information Sciences* 6(C). 49–83.

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

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

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

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

Pollard, Carl J. 2013. Traces exist (hypothetically)! https://web.stanford.edu/dept/linguistics/structures_evidence/website/docs/slides/pollard.pdf, accessed 2018-9-10. Presented at the Workshop on Structure and Evidence in Linguistics, Stanford.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Chapter 3

Formal Background

Frank Richter
Goethe-Universität Frankfurt

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

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

1 Introduction

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

Frank Richter. 2018. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 39–40. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 4

The nature and role of the lexicon in HPSG

Anthony Davis
Southern Oregon University

Jean-Pierre Koenig
University at Buffalo

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

1 Introduction

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

Anthony Davis & Jean-Pierre Koenig. 2018. The nature and role of the lexicon in HPSG. in Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 41–80. Berlin: Language Science Press. DOI:??

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

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

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

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

Finally, we discuss lexical rules and their alternatives. Along with the "vertical" relationships between classes of lexical entries modeled by types and their subtypes in the hierarchical lexicon, there is a perceived need for "horizontal" relationships between lexical entries that are based on a single root or stem, such as forms of inflectional paradigms. Yet formalizing lexical rules adequately within HPSG has proven tricky; specifying just what information is preserved and what is changed by a lexical rule is one prominent issue. We conclude this chapter by describing alternatives to lexical rules. One is implicational statements on partially specified feature structures; this might be thought of as a kind of "online lexical rule". The second augments the type hierarchy via online type construction, extending the predefined lexical types specified in the hierarchy to include "virtual types" that combine the information from multiple predefined types.

2 Lexicalism

2.1 Lexicalism and the origins of HPSG

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

Lexicalism, at least within HPSG, embodies two distinct ideas. First is the idea that parts of words are invisible to syntactic operations (*lexical integrity*, see Bresnan & Mchombo 1995), so that relations between stems and between word forms cannot be the result of or follow syntactic operations, as in distributed morphology, or other linguistic models that assign no special status to the notion of word. Relations between words are therefore not modeled via syntactic operations (hence the appeal to Jackendoff's lexical rules). Second is the idea that the occurrence of a lexical head in distinct syntactic contexts arises from distinct variants of words. For instance, the fact that the verb *expect* can occur both with a finite clause and an NP+VP sequence (see (1) vs. (2)) means that there are two variants of the verb *expect*, one that subcategorizes for a finite clause and one where it subcategorizes for an NP+VP sequence.[1] Not all lexicalist theories, though, cash out these two distinct ideas the same way. The net effect of lexicalism within HPSG is that words and phrases are put together via distinct sets of constructions and that words are syntactic atoms. These two assumptions justify positing two kinds of signs, *phrasal-sign* and *lexical-sign* and go hand in hand with the surface-oriented character of HPSG and what one might call a principle of surface combinatorics: If expression A consists of B⊕C, then all grammatical

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

constraints that make reference to B and C are circumscribed to A.

(1) I expected to leave yesterday

(2) I expected that I would leave yesterday.

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

2.2 Recent challenges to lexicalism

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

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

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

We can provisionally accept for the sake of argument Bruening's contention that *don't-you-dare!* is a word in (3), despite its reliance on the (unjustified) assumption that the secondary object in (3) involves N-N compounding rather than an AP N structure (we refer readers to Bresnan & Mchombo (1995) or Müller (2010) for counter-arguments to Bruening's claim). Crucially, though, examples such as (3) have no bearing on HPSG's model of lexical knowledge, as HPSG-style lexicalism does not preclude constructions that form words from phrases. Nothing, as far as we know, rules out constructions of the form *phrase* → *stem/word* in HPSG. The two assumptions underlying HPSG brand of lexicalism we mentioned above do not preclude a *lexical-sign* having a *phrasal-sign* as sole daughter (although we do not know of any HPSG work that exploits this possibility) and examples such as (3) are simply irrelevant to whether HPSG's lexicalist stance is empirically correct.

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

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

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

Haspelmath (2011) similarly challenges the view that syntactic processes may not access the internal structure of words, although Haspelmath's point is merely that what is a word is cross-linguistically unclear. So-called suspended affixation in Turkish (see (5)) also shows that word parts can be elided. We cannot discuss here whether Chaves' analysis can be extended to cases like (5) where suffixes are seemingly elided or whether lexical sharing (where a single word can be the daughter of two c-structure nodes à la McCawley 1982), as proposed in Broadwell (2008) is needed. What is important for current purposes is that these putative challenges to lexical integrity such as (4) or (5) do not necessarily render a substantive version of it implausible. The same is true of another potential challenge to lexical integrity which neither Bruening nor Haspelmath discuss, endoclitics, which we discuss next.

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

Endoclitics are clitics that at least appear to be situated within a word, rather than immediately preceding or following it, as clitics often do. (cross-reference to Abeillé & Penn chapter on clitics) In many cases, endoclitics appear at morphological boundaries, as in the well-studied pronominal clitics of European Portuguese (Crysmann 2001). An approach similar to what we have referenced above for composite words and elided morphology may well extend to these as well. But some trickier cases have also come to light, in which the clitic appears within a morpheme, not at a boundary. Two of the best documented cases from the Northeast Caucasian language Udi (Harris 2000) and from Pashto (Tegey 1977; Roberts 2000; Dost 2007). Here are examples from Udi (6) and Pashto (7), where the clitics appear in the middle of verbs.

(6) q'ačaɣ-ɣ-on bez tänginax baš=**q'un**-q'-e
thief-PL-ERG my money.DAT $steal_1$-3PL-$steal_2$-AORII
'Thieves stole my money.' (root *bašq'*, 'steal')

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

b. ṭə́l=**me**-wɑhə
$push_1$-cl.1SG-$push_2$.PF.PST.3SG
'I pushed it.' (from Tegey 1977; Dost 2007)

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

If clitics are viewed as a syntactic phenomenon ("phrasal affixes", as Anderson (2005) puts it), these endoclitics must have "visibility" into the internal structure of words (be it morphological, prosodic, or something else), thereby seemingly violating lexical integrity. Anderson's brief account invokes a reranking of optimality theoretic constraints from their typical ordering, whereby the clitic's positional requirements outrank lexical integrity requirements. Crysmann (2015) proposes an analysis, paralleling in many respects his account of European Portuguese clitics in Crysmann (2001), using Kathol's topological fields (Kathol 1999). The "morphosyntactic paradox in Udi" is effectively "resolved on the basis of discontinuous lexical items"; this account then "parallels HPSG's representation of syntactic discontinuity." (cross reference to Müller's chapter on word order here)

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

Still, these accounts of endoclitics in Udi and Pashto appear to breach the wall of the strictest kind lexical integrity, requiring that they have access to some of the internal structure of lexical entries through a partial decomposition of their morphophonology into distinct order domains. Yet we would not wish to advocate models that permit unconstrained violations of lexical integrity, either. The troublesome cases we have noted here are relatively marginal or cross-linguistically rare, and limited in scope to prosodic or morphophonological information and seem to only pertain to phonological interactions (ellipsis, insertion). As Broadwell (2008) points out when comparing possible analyses of Turkish suspended affixation, rejecting lexicalism altogether may lead to an unconstrained theory of the interaction between words/stems and phrases and incorrect predictions (e.g., that all affixes in Turkish can be suspended). Likewise, we would not expect to find a language in which endoclitics positioning is utterly unconstrained, and thus we would not wish to see grammatical theories abandon lexical integrity altogether.

3 Lexical entries in HPSG

3.1 What are lexical entries?

A consequence of HPSG's lexicalist stance is that there will be many lexical entries where one might at first glance expect a single entry. We will see below how HPSG handles multiple entries and classes of entries while avoiding redundancy, but it is important at the outset to clarify what a lexical entry is in HPSG. One of the misunderstandings about lexical knowledge is that it confuses descriptions and entities being described, or the distinction between constructions and constructs (lexical entry vs. fully instantiated lexeme). As the chapter on the formal foundations of HPSG discusses, grammars in HPSG consist of *descrip-*

tions of structures, and the lexicon thus consists of descriptions of what are fully specified lexemes. What is stored in the lexicon is descriptions of fully instantiated lexemes. To see the importance of the distinction between descriptions (stored entries) and the fully instantiated entries that are being described, consider HPSG's model of subcategorization and consider the relevant portion of the tree for sentence (2). HPSG's model of the dependency between heads and complements involves identity between the syntactic and semantic information of each complement (the value of the SYNSEM attribute) and a member of the list of complements the head subcategorizes for. Since there are infinitely many SYNSEM values, on the assumption that there are infinitely many clausal meanings (a point Jackendoff (1990) emphasizes), there are, in principle, infinitely many fully instantiated entries for the verb *see* subcategorizing for a clausal complement (as in (2)). But each of these fully instantiated entries for *expect*, one for each clausal sentence that corresponds to the tree in (8) corresponds to a single abstract description, and it is this description that the lexicon contains.

(8)

$$\left[\textsc{comps}\ \langle \boxed{1} \rangle\right] \quad \left[\textsc{synsem}\ \boxed{1}\right]$$

The formal status of lexical entries has engendered a fair amount of theoretical work and some debate, particularly over the question of whether lexical entries must be fully specified. We will touch on some aspects of this further below, in connection with online type construction. For further discussion of these kinds of issues, see the chapters on Basic Properties and Elements and on Formal Background.

3.2 What information is in lexical entries?

Because lexical items play a critical role in accounting for the syntax of natural languages, lexical entries are informationally rich in HPSG. Aside from the expected phonological and semantic information, specific to each lexeme, they include morphological and combinatorial potential information. Morphological information serves as input to inflectional rules, but is also used to select the appropriate types of phrases (via their projection through the Head-Feature Principle), as shown in (9). Some verbs, for instance, select for a PP headed by a particular preposition; others select for VPs whose verb is a gerund, or a bare infinitive, and so forth. Lexical entries thus include as much morphological information as both (inflectional) morphology and syntactic selection require.

(9) a. John conceived *of/*about* the world's tastiest potato chip..

b. John regretted *going/*(to) go* to the party.

We illustrate the second leading idea behind HPSG or LFG's lexicalism (that there are different variants of lexical heads for different contexts in which heads occur) with the French examples in (10). The verb *aller* 'go' in (10a) combines with a PP headed by *à* that expresses its goal argument and a subject that expresses its theme argument. The same verb in (10b) combines with the so-called non-subject clitic *y* that expresses its goal argument. We follow Miller & Sag (1997) and assume here that French non-subject clitics are prefixes. Since the context of occurrence of the head of the sentence, *aller*, differs across these two sentences (NP____PP[*à*] and NP *y*____ , respectively and informally), there will be two distinct entries for *aller* for both sentences, shown in (11) and (12) (we simplify the entries' feature geometry for expository purposes).

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

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

(11)
$$\left[\begin{array}{ll}
\text{MORPH} & \left[\begin{array}{ll} \text{FORM} & \boxed{5} \\ \text{I-FORM} & \boxed{5}\ \textit{va} \\ \text{STEM} & \textit{v-} \end{array}\right] \\
\text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{verb}} \\ \text{VFORM} & \left[\begin{array}{ll} \text{MOOD} & \textit{indic} \\ \text{TNS} & \textit{pres} \\ \text{AGR} & \textit{3rdsing} \end{array}\right] \end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2} \rangle \end{array}\right] \\ \text{ARG-ST} & \left\langle \boxed{1}\text{NP}[\textit{3rdsg}]_{\boxed{3}},\ \boxed{2}\text{PP}[\textit{à}]_{\boxed{4}} \right\rangle \end{array}\right] \\
\text{CONT} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{go-rel}} \\ \text{THEME} & \boxed{3} \\ \text{GOAL} & \boxed{4} \end{array}\right]
\end{array}\right]$$

(12)
$$\left[\begin{array}{ll} \textsc{morph} & \left[\begin{array}{ll} \textsc{form} & \boxed{\textit{y-va}} \\ \textsc{i-form} & \textit{va} \\ \textsc{stem} & \textit{v-} \end{array}\right] \\ \textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \left[\begin{array}{l} \textit{verb} \\ \textsc{vform} \left[\begin{array}{ll} \textsc{mood} & \textit{indic} \\ \textsc{tns} & \textit{pres} \\ \textsc{agr} & \textit{3rdsing} \end{array}\right] \end{array}\right] \\ \textsc{val} & \left[\begin{array}{ll} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \boxed{\langle\rangle} \end{array}\right] \\ \textsc{arg-st} & \left\langle \boxed{1}\text{NP}[\textit{3rdsg}]_{\boxed{3}}, \boxed{\text{PP}[\textit{p-aff,loc}]_{\boxed{4}}} \right\rangle \end{array}\right] \\ \textsc{cont} & \left[\begin{array}{ll} \textit{go-rel} & \\ \textsc{theme} & \boxed{3} \\ \textsc{goal} & \boxed{4} \end{array}\right] \end{array}\right]$$

CATegory information in both entries include part of speech information (including morphologically relevant features of verb forms), ARGument-STructure information and VALence information. MORPH information includes both stem form information, inflected form information (I-FORM) and, in case so-called clitics are present, the combination of the clitic and inflected form information. Both entries illustrate how informationally rich lexical entries are in HPSG. But, postulating informationally rich entries does not mean stipulating all of the information within every entry. In fact, only the stem form and the relation denoted by the semantic content of the verb *aller* need be stipulated within either entry. All the other information can be inferred once it is known which classes of verbs these entries belong to. In other words, most of the information included in the entries in (11) and (12) is not specific to these individual entries, an issue we take up in Section 4. The entry-specific information in (11) and (12) is in black font while the shared information is in gray font; the informational difference between the two entries for *va* and *y va* is included in shadowed boxes in the respective entries. The first difference between the two variants of *va* 'goes' is in the list of complements: the entry for *y va* does not subcategorizes for a locative PP since the affix *y* satisfies the relevant argument structure requirement. This difference in the realization of syntactic arguments (via phrases and pronominal affixes) is recorded in the type of the PP members of ARG-ST, *p-aff* in (12) but not in (11). Finally, the two entries differ in the FORM of the verb, which is the same as the inflected form of the verb in (11) (as indicated by the identically numbered $\boxed{5}$), but not in (12) whose FORM includes the prefix *y*.

One other question arises with regard to the information in lexical entries. Are there attributes or values that occur solely within lexical signs, and not in phrasal ones? If so, they would provide a diagnostic for distinguishing lexical signs from others. Specific phonological information, for instance, is something we would expect to be introduced by lexical entries, and not elsewhere. Some information is claimed to be specific to lexical signs, such as phonological information (cross reference to Tseng here) and the ARG-ST list, on the premise that lexical items alone specify combinatorial requirements (but see Przepiórkowski (2001) for a contrary view, and see the chapter on Construction Grammar for other views questioning this assumption). But HPSG researchers have generally not typically explored this question in depth, and we will leave this issue here.

3.3 The role of the lexicon in HPSG

As we hope is evident by now, the lexicon plays a critical role in HPSG's explanatory mechanisms, as words encode their distributional potential, as well as their idiosyncratic phonological and semantic charateristics. Much of the information contained in lexical entries is geared to modeling the combinatorial potential of words. As detailed in the chapter on Argument Structure, their combinatorial potential is recorded using two kinds of information, a list of syntactic arguments or syntactic requirements to be satisfied, and distinct lists that indicate how these requirements are to be satisfied (as local dependents, as non-local dependents, as clitics/affixes). Not only are syntactic arguments recorded; so is their relative obliqueness (in terms of grammatical function), as per the partial hierarchy in (13) from Pollard and Sag 1992.

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

We illustrate this explanatory role by alluding to the role of the lexicon in HPSG's approach to binding, as described in Pollard & Sag (1992) (see the chapter on Binding for details). As lexical entries of heads record both syntactic and semantic properties of their dependents, constraints between properties of heads and properties of dependents, e.g. subject-verb agreement, or between dependents, e.g. binding constraints illustrated in (15), can be stated as constraints on classes of lexical entries. The principle in (15) is such a constraint.

(14) a. Mathilda$_i$ saw herself$_i$ in the mirror.

b. *Mathilda$_i$ saw her$_i$ in the mirror.

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

Principle (15) is, formally, a constraint on lexical entries that makes use of the fact that an entry's argument structure records the syntactic and semantic properties of a word's dependents. The three argument structures in (16) illustrate permissible and ungrammatical entries. (16a) illustrates exempt anaphors as there is no less oblique syntactic argument than the anaphoric NP; (16b) illustrates a non-exempt anaphor properly bound by a less oblique, co-indexed non-anaphor; (16c) illustrates an ungrammatical lexical entry that selects for an anaphoric syntactic argument that is not co-indexed by a less oblique syntactic argument, despite not being an exempt anaphor (i.e., not being the least oblique syntactic argument).

(16) a. $\left[\textsc{arg-st}\ \langle \mathrm{NP}_{i,+ana}, \ldots\rangle\right]$

b. $\left[\textsc{arg-st}\ \langle \mathrm{NP}_{i,-ana}, \ldots, \mathrm{NP}_{i,+ana}, \ldots\rangle\right]$

c. * $\left[\textsc{arg-st}\ \langle \mathrm{XP}_{j}, \ldots, \mathrm{NP}_{i,ana}, \ldots\rangle\right]$

Our purpose here is not to argue in favor of the specific approach to binding just outlined. Rather, we wish to illustrate that in a theory like HPSG where much of syntactic distribution is accounted for by properties of lexical entries, co-occurrence restrictions treated traditionally as constraints on trees (via some notion of command) are modeled as constraints on the argument structure of lexical entries. It is tempting to think of such a lexicalization of binding principles as a notational variant of tree-centric approaches. Interestingly, this is not the case, as argued in Wechsler (1999). Wechsler argues that the difference between argument structure and valence is critical to a proper model of binding in Balinese. Summarizing briefly, voice alternations in Balinese (e.g., objective or agentive voices) do not alter a verb's argument structure but do alter its valence, which is the subject and object it subcategorizes for. As binding is sensitive to relative obliqueness within ARG-ST, binding possibilities are not affected by voice alternations within the same clause, which are represented with different valence values. In the case of raising, on the other hand, the argument structure of the raising verb and the valence of the complement verb interact, as the subject of the complement verb is part of the argument structure of the raising verb. An HPSG approach to binding therefore predicts that voice alternations within the embedded clause will not affect binding of co-arguments of the embedded verb, but will affect binding of the raised NP and an argument of the embedded verb. This prediction seems to be borne out, as the examples in (17) show.

(17) a. Ia$_i$ nawang awakne$_i$/Ia$_{*i}$ lakar tangkep polisi.
3rd AV.know self/3rd FUT OV.arrest police
'He$_i$ knew that the police would arrest self$_i$./him$_{*i}$.'

b. Cang ngaden ia$_i$ suba ningalin awakne$_i$/ia$_{*i}$
1sg AV.think 3rd already AV.see self/3rd
'I believe him$_i$ to have seen himself$_i$/ him$_{*i}$.'

c. Cang ngaden awakne$_i$ suba tingalin=a$_i$.
1sg AV.think self$_i$ already OV.see=3
'I believe him to have seen himself.'

Sentence (17a) shows that the proto-agent (the first element of ARG-ST) of the subject-to-object raising verb *nawang* 'know' can bind the raised subject (which in this case corresponds to the proto-patient of the complement verb *tangkep* 'arrest' since that verb is in the objective voice). Sentence (17b) shows that the raised (proto-agent) subject of the complement verb can bind its proto-patient argument. Critically, sentence (17c) shows that the raised proto-patient (second) argument of the complement verb can be bound by the complement verb's proto-agent. The contrast between sentences (17b) and (17c) illustrates that while binding is insensitive to valence alternations (the same proto-agent binds the same proto-patient argument in both sentences), raising is not (the proto-agent argument is raised in (17b) and the proto-patient argument in (17c)). As Wechsler argues, this dissociation between valence subjects and less oblique argumentson the ARG-ST list is hard to model in a configurational approach to binding that equates the two notions in terms of c-command or the like. What is important for our purposes is that a 'lexicalization' of argument structure, valence, and binding has explanatory power beyond tree configurations, illustrating some of analytical possibilities informationally rich lexical entries create.

3.4 Lexical vs. constructional explanations

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

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

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

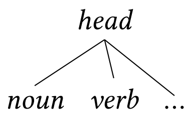
4 The hierarchical lexicon

We have seen that lexicalism demands that lexical entries be information rich, in order to encode what might otherwise be represented as syntactic rules. To avoid massive and redundant stipulation within each lexical entry, we need mechanisms to represent regularities within the lexicon. Two main mechanisms have been used in HPSG to represent these regularities. The first mechanism is the organization of information shared by lexical entries or parts of entries into a hierarchy of types in a way quite similar to semantic networks within Knowledge Representation systems (see among others Brachman & Schmolze 1985). This hierarchy of types (present in HPSG since the beginning, Pollard & Sag (1987) and the seminal work of Flickinger (1987)) ensures that individual lexical entries only specify information that is unique to them. The second mechanism is lexical rules, which relates variants of entries, and more generally, members of a lexeme's morphological family (which consists of a root or stem as well as all stems derived from that root or stem). In this section, we discuss the hierarchical organization of the lexicon into cross-cutting classes of lexical entries at various levels of generality.

4.1 Inheritance

All grammatical frameworks classify lexical entries to some extent, of course. Basic part of speech information is one obvious case. This high-level classification is present in HPSG, too, as part of the hierarchy of types of heads. That information is recorded in the value of the HEAD feature. A simple hierarchy of types of heads is depicted in (18).

(18)



Each of these types is a partial specification of a lexical entry's head properties. Typing of HEAD information allows the ascription of appropriate properties to different classes of lexical entries. For example, case information is only relevant to nouns, and whether a verb is an auxiliary or not is only relevant to verbs. Each type in (18) includes in its definition a specification of which features are appropriate for it, as shown in (19). (19) specifies what it means to be a noun or a verb in a particular language (of course, there will strong similarities in these properties across languages).

(19) a. $noun \Rightarrow \begin{bmatrix} \textsc{case} & case \end{bmatrix}$

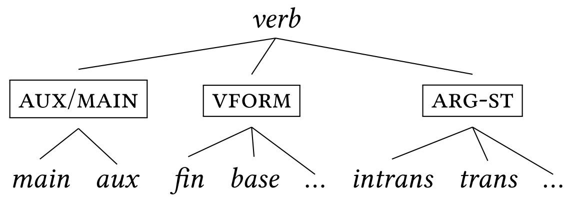
b. $verb \Rightarrow \begin{bmatrix} \textsc{aux} & boolean \\ \textsc{tense} & tns \\ \textsc{aspect} & asp \end{bmatrix}$

More technically, each type of head imposes some constraint on lexical entries of that type. Thus, (19a) requires all noun lexemes to be eligible for case information. Here, the constraint on each type specifies the value of SYNSEM|LOC|CAT|HEAD.

as the atomic value *noun*, *verb*, etc. Since these atomic values are disjoint, and since the HEAD value is unique for each lexical entry, the types in (18) are also disjoint. If there's a type corresponding to each possible HEAD value, then they constitute a partition of lexical entries as well. Lexical entries for particular lexemes make use of the definitions of types like (19) to abstract information that is shared across classes of entries. Thus, the pronoun *him* need only include the fact that its HEAD is of type *noun*; the fact that it might bear case can be inferred. Similarly, the entry for the verb *can* need only include information that its head information include the specification [AUX +] for us to be able to infer that it is a *verb*.

So far, this is merely an HPSG implementation of a part of speech taxonomy, but once we consider subtypes with additional constraints the utility of the hierarchical lexicon within a lexicalist framework becomes apparent. There are interesting generalizations to be made about more specific classes, such as transitive verbs, or past participles, or predicators denoting caused motion (regardless of their part of speech). In the hierarchical lexicon, we can represent these "interesting" classes as types. Which classes are worth instantiating in the grammar of a given language depends on its grammar; thus we expect lexical classes to specify a mix of cross-linguistically common (maybe, in some cases, universal) and language-particular constraints. Consider some of the subtypes of verbs shown in (20) adapted from Bouma et al. (2000):

(20)



Again, each subtype specifies additional information constraining the lexical entries belonging to it. The boxed labels have no independent formal status (although they play a role in the framework of online type construction, discussed below), indicating simply that the parent type, here *verb*, is partitioned by the subtypes under each box. Typically, this means that each subtype specifies an atomic value for a particular attribute, out of a set of mutually disjoint values, as in the part of speech types above. Thus *main-verb* and *aux-verb* are disjoint subtypes of *verb*, with the values + and − for the attribute SYNSEM|LOC|CAT|HEAD|AUX.

More specific verb subtypes can combine the constraints of the types depicted in (20), through multiple inheritance. Infinitive forms of transitive verbs, for example, inherit the constraints of both *infinitive-verb* and *transitive-verb*. Provided that the constraints do not conflict (i.e., the descriptions of the two types unify), such a type can exist and have members. Whether it is useful to reify such a type is another matter; not all possible combinations of constraints yield linguistically interesting classes of lexical entries (this is another issue we address in the discussion of online type construction). It seems desirable, for instance, to avoid a proliferation of types for each form of a verb in a paradigm in an inflectionally complex language, as the number of forms, and thus types, would be extremely large (there are, for example, 2,494 combinations of inflectional prefixes in Oneida, a Northern Iroquoian language, Karin Michelson, p.c.). While an economical type-based description of regular morphological paradigms may prove descriptively adequate, it is implausible in assuming that each form of every fully regular verb is reified as a lexical entry. We will discuss mechanisms (lexical rules and online type construction) that offer better accounts of morphologically regular and productive word formation below.

In general, types are posited in the hierarchy when there is some additional constraint to state about them. We now briefly examine some of the lower levels of the lexical hierarchy; that is, some more specific lexical types that illustrate how types is one way to reduce the amount of information that needs to be stipulated in individual lexical entries and are one of the tools HPSG employs to represent lexical generalizations. We begin with the *transitive-verb* type (*trans-vb* for short). Apart from requiring its ARG-ST list to contain two NPs, *trans-vb* is further constrained, at least in English, to be a main verb rather than an auxiliary verb (see the value of HEAD *main* in (21) and the hierarchy of verbal head information in (20)): there are no transitive auxiliaries in English. So, *trans-vb* includes information constraining the feature values of transitive verbs that goes beyond simply specifying the nature of the ARG-ST list.

The partial representation of the type *transitive-verb* in (21).

(21) $\begin{bmatrix} \textit{trans-vb} & \\ \textsc{head} & \textit{main} \\ \textsc{arg-st} & \langle \text{NP, NP, ...} \rangle \end{bmatrix}$

A more specific subtype of *trans-vb* is *caused-motion-transitive-verb*, which states information about the semantics of verbs in the class as well as their subcategorization, as in (22). (↑ indicates that the type that follows it is a supertype of the type indicated in the feature structure; in this case that *caused-mot-trans-vb* is a subtype of *trans-vb*.)

(22)
$$\begin{bmatrix} \multicolumn{2}{l}{\textit{caused-mot-trans-vb}(\uparrow \textit{trans-vb})} \\ \textsc{content} & \begin{bmatrix}\textit{caused-motion-rel} & \\ \textsc{causer} & \boxed{1}\\ \textsc{moved} & \boxed{2}\end{bmatrix} \\ \textsc{arg-st} & \left\langle \mathrm{NP}_{\boxed{1}}, \mathrm{NP}_{\boxed{2}}\right\rangle \end{bmatrix}$$

The information in each type constitutes constraints on objects of that type. With the types situated in a hierarchy, each type inherits all the constraints of its supertypes. Thus constraints will be inherited from supertypes. But additional constraints can be added at the level of that type as well; this is the principal fashion in which generalizations about classes of lexical entries can be stated. For example, following Davis & Koenig (2000) and Koenig & Davis (2003) we might state a contraint on argument realization on the type *caused-motion-transitive-verb*, to ensure that the causer is linked to the subject and the entity that is caused to move to the direct object, as in (23). (cross-reference to chapter on argument structure and linking here) Here, we make use of Richter's logic (Richter 1999) to encode constraints on information that is included in lexical entries. The constraint in (23) says that if a verb's semantic content is a cause relation, the causer arguments corresponds to the index of the first NP on the ARG-ST list and that if a verb's semantic content is a motion relation, the moved entity is realized as an NP. Implicational constraints such as (23) relieve some of the burden of encoding generalizations over lexical entries exclusively through the lexical type hierarchy, and can lead to a simpler model of lexical generalizations in some cases, as Koenig & Davis (2003) point out. When it is preferable to use a hierarchy of lexical types or conditional constraints on the information included in lexical types remains an open issue.

(23)
$$\begin{bmatrix}\textsc{content} & \textit{cause-rel}\end{bmatrix} \Rightarrow \begin{bmatrix}\textsc{cont} & \begin{bmatrix}\textsc{causer} & \boxed{1}\end{bmatrix}\\ \textsc{arg-st} & \langle \mathrm{NP}{:}\boxed{1}, \ldots\rangle\end{bmatrix}$$

$$\begin{bmatrix}\textsc{content} & \textit{move-rel}\end{bmatrix} \Rightarrow \begin{bmatrix}\textsc{cont} & \begin{bmatrix}\textsc{moved} & \boxed{1}\end{bmatrix}\\ \textsc{arg-st} & \langle \ldots \mathrm{NP}{:}\boxed{1}, \ldots\rangle\end{bmatrix}$$

More specific classes of transitive caused-motion verbs, such as the *spray* verbs in English that exhibit locative alternations, inherit the additional constraints in (22) and further specify additional semantic constraints that characterize these alternating verbs. The hierarchical organization of lexical types allows us to state these additional restrictions, which are often language-particular, in the appropriate place without additional formal mechanisms. The range of ditransitive constructions, to take one such case, varies across languages, with some lacking them entirely and others freely allowing them in, e.g., morphologicallly productive causatives of any transitive verb. For those languages, like English, in between these extremes, semantic (and possibly other) constraints can be placed on the type *ditransitive-verb*, limiting such verbs to those involving, e.g., transfer of possession.

We now illustrate how the organization of the lexicon in a hierarchy of lexical types minimizes the information that needs to be specified within individual entries, such as those for the forms of the French verb *va* we provided earlier (see (11)). We start with semantics and how it links to the argument structure. We can infer that the use of *va* illustrated in (10a) includes two arguments, a theme and a goal, from the hierarchy of semantic relations, which ensures that all types of directed motion events, of which *go-rel* is a subtype, includes these two arguments (see Davis (2001) for such an approach to semantic relations). The linking of these arguments to an NP and PP follows either from linking types, as in Davis & Koenig (2000) or Davis (2001), or from constraints similar to those we show above in (23) for English caused-motion verbs. The relation between the argument structure of *va* and its subcategorization requirements for a subject and PP complement follows from general constraints on words and a general type for intransitive verbs, analogous to (21) for transitive verbs. The inflectional features of this form are instantiations of possible combinations of values of mood, tense, and agreement information within French verbs. Finally, the expression of these inflectional features is the result of either general lexical rules (see Miller & Sag (1997) for some examples) or, as in more recent work in HPSG, a network of associations between morphosyntactic features and forms at various positions in the word (see Crysmann & Bonami 2016). In the end, nothing but the meaning of this use of *va* and the fact that the stem form is *v-* need be stipulated in the entry.

4.2 The lexicon as repository of generalizations at various levels

The hierarchical lexicon makes it possible to specify constraints on classes of lexical entries at any level, not just, e.g., all nouns, or a single word. An illustrative example, drawn from Ackerman & Webelhuth (1998), involves German passives, which come in several varieties, each with its own constraints. Each passive construction uses a different auxiliary: (*werden*, *sein*, or *bekommen*) and two of these constructions require a participial form of the verb, while the *sein* passive requires *zu* followed by an infinitive VP. Additionally, passives appear attributively, as NP modifiers, as well as predicatively. Here are two examples of the *zu* + infinitive passive, the first attributive, the second predicative:

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

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

Ackerman & Webelhuth's account of German passives posits a multiple inheritance hierarchy of lexical types in German, a portion of which is shown in Figure 1.

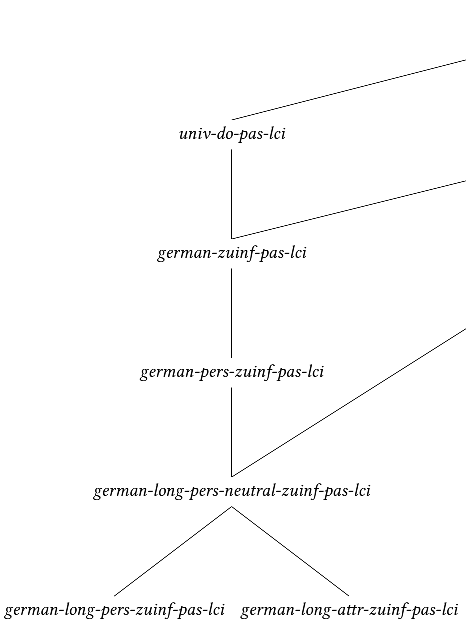


Figure 1: A portion of the hierarchy of passive lexical types according to Ackermann and Webelhuth, p.244

While all passives share the constraint that a logical subject is demoted, as stipulated on a general *univ-pas-bas-lci* passive type, the other requirements for each kind of passive are stated on various subtypes. The *zu*+infinitive passive, for instance, requires not only that *sein* is the auxiliary and that the main verb is infinitive, but that the semantics involves necessity or obligation. This differs from the other passives, which simply maintain the semantics of their active counterparts. However, the types of the passive verb *schenken(den)* in (24) both inherit from several passive verb supertypes. As mentioned, at a general level, there is information common to all German passives, or indeed to passives universally, namely that the "logical subject" (first element of the basic verb's ARG-ST list) is realized as an oblique complement of the passive verb, or not at all. A very common subtype, which Ackerman & Webelhuth also regard as universal, rather than specific to German, specifies that the base verb's direct object is realized as the subject of its passive counterpart; this defines personal passives. Once in the German-specific realm, an additional subtype specifies that the logical subject, if realized, is the object of a *von*-PP; this holds true of all three types of German personal passives. Among its subtypes is one that requires *zu* and the infinitive form of the verb; moreover, although Ackerman & Webelhuth do not spell this out in detail, this subtype specifies the modal force characteristic of this passive construction but not of the others. Finally, both the predicative and attributive forms are subtypes of all the preceding, but these inherit also from distinct supertypes for predicative and attributive passives of all kinds. The supertype for predicative passives constrains them to occur with an auxiliary; its subtype for *zu* + infinitive passives further specifies that the auxiliary is *sein*. The attributive passive type, on the other hand, inherits from modifier types generally, which do not allow auxiliaries, but do require agreement in person, number, and case with the modified noun. In summary, the hierarchical lexicon is deployed here to factor out the differing properties of the various German passive constructions, each of which obtains its particular combination of properties via multiple inheritance.

The most specific types of the lexical hierarchy, where individual lexical entries reside, is where constraints pertaining solely to a given word or root – its phonological form, inflectional class, specific semantics, register, and so forth – are stated. Specific information about a word needs to be spelled out somewhere in any grammatical framework. In a hierarchically organized lexicon we can view this as just the narrowest, most particular case of specifying information about a class of linguistic entities. But where information is shared across a broader set of lexical entries, it need not be stated separately for each one. Thus, the phonology of the word *spray* and the precise manner of motion of the particles or liquid caused to move in a spraying event are unique to this lexical entry. However, much of its syntactic and semantic behavior— it is a regular verb, participating in a locative alternation, involving caused ballistic motion of a liquid or other dispersable material— is shared with other English verbs such as *splash*, *splatter*, *inject*, *squirt*, and *drizzle*. To the extent that these "narrow conflation classes," as Pinker (1989) terms them, are founded on clear semantic criteria, we can readily state syntactic and semantic constraints at the appropriate level in the hierarchical lexicon (some, however, such as Briscoe & Copestake (1999), cast doubt on the feasibility of formulating such constraints for dative and other alternations in English, suggesting that lexical rules might be a better alternative). Given this semantic similarity, it may be that much of the semantics of a verb like *spray* need not be specified at the level of that individual lexical entry. Apart from the broad semantics of caused motion, shared by numerous verbs, the verbs in the narrow conflation class containing *spray* share the selectional restriction, noted above, that their objects are set in motion by an initial impulse and that they are liquid or particulate material. We might therefore posit a subtype of the type *caused-motion-rel* to represent this shared semantics triggering the locative alternation, with further subtypes of that for the semantics of the individual verbs. Note that not all these constraints apply to precisely the same class (there are other verbs with somewhat different semantics, like *load* and *wrap*, exhibiting the locative alternation, for example), so a multitude of types in the hierarchy is crucial.

4.3 Default inheritance in the lexicon

So far, we have assumed rigid, monotonic inheritance of all information in supertypes to their subtypes; none of the inherited information can be overridden. This runs into difficulties when dealing with lexical entries that appear to be exceptional in some way, the obvious examples being morphological irregularities. How can productive regular forms such as **childs* be blocked, and only *children* allowed as a lexical entry?

While several approaches to exceptions have been proposed, we will focus here on *default unification*; that is, weakening monotonic inheritance in some circumstances. Then, although the plural of *child* might inherit the information from the pertinent lexical entry and from the *plural-noun* type, which would entail the phonology for **childs*, this regular plural form is overridden. Various complex issues arise in attempting to formulate a workable system of default unification and inheritance. See, e.g., Briscoe & Copestake (1999) for a brief overview of various ways that default unification might be defined. Lascarides & Copestake (1999) list several desirable criteria, including:

- Non-default information is always preserved
- Default unification behaves like monotonic unification whenever possible
- Default unification is order-independent

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

As another example of the use of default, nonmonotonic inheritance, outside of morphology, consider the account of the syntax of gerunds in various languages developed by Malouf (2000). Gerunds exhibit both verbal and nominal characteristics, and furnish a well-known example of seemingly graded category membership, which does not accord well with the categorical assumptions of mainstream syntactic frameworks. Roughly speaking, English gerunds, and their counterparts in other languages, act much like verbs in their "internal" syntax, allowing direct objects and adverbial modifiers, but function distributionally ("externally") as NPs. To take but a couple of pieces of evidence (see Malouf, op.cit. p.27 et seq. for more details), gerunds can be the complement of prepositions when finite clauses cannot (see (25)); conversely, adverbs, but not adjectives can modify gerunds, but adjectives must modify deverbal nouns (see (26)).

(25) a. Pat is concerned about Sandy('s) getting arrested.
b. *Pat is concerned about (that) Sandy got arrested.

(26) a. Pat disapproved of (me/my) *quiet/quietly leaving before anyone noticed.
b. Pat disapproved of my quiet/*quietly departure.

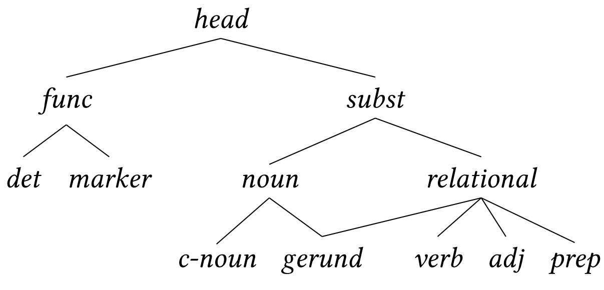
In contrast to accounts that attempt to model this dichotomy directly, via syntactic rules that allow an NP to be expanded as a constituent internally headed by a verb, Malouf posits a lexical rule, which converts the lexical category of a verb to *noun*, but otherwise preserves its verbal properties, such as subcategorization. This would pose problems with strictly monotonic inheritance, however, as it would force us to abandon generalizations about nouns other than gerunds (e.g., they do not take direct object complements, as many verbs and their gerunds do). Default inheritance provides one way to model the observed phenomena, without weakening the constraints on parts of speech to the point where no meaningful constraints distinguish them.

Malouf notes that some possible combinations of noun-like and verb-like attributes are frequently attested cross-linguistically in gerunds and their equivalents, while others are rare or unattested. Cross-linguistically, gerunds vary in their subcategorization possibilities: some allow subjects and complements, while some allow only complements and no subjects. But there appear to be no cases of gerund-like lexical items that can take a subject but cannot take complements. Malouf invokes default inheritance (Lascarides & Copestake 1999) as a mechanism to represent these generalizations. In his account, there are both “hard” constraints – a verb lexical entry, for example, must have a HEAD value of type *relational* (encompassing verbs, adjectives, and adpositions) – and “soft”, overridable constraints – a verb lexical entry by default has a HEAD value of type *verb*. In addition, following Bouma et al. (2001), he posits the types *ext-subj* and *ext-spr*. The former constrains the HEAD value to *relational* and the first element of the ARG-ST list to be the SUBJ (only adjective, adpositions, and verbs have subjects), while the latter constrains the HEAD value to *noun* and the first element of the ARG-ST list to be the SPR (only nouns have specifiers), as shown in (27).

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

Malouf then specifies default HEAD values for the lexical classes *n* and *v* (see (29) for the latter’s definition). As gerunds have both properties of nominal and relational heads, they are subtypes of both, as shown in the multiple inheritance hierarchy in (28). The *v* type, which concerns us here, has a default HEAD value *verb*, as shown in (29) in addition to the non-default, more general type *relational* it also includes (default information follows /).

(28)



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

However, the default value *verb* is overridden in the subtype *vger*, in which the HEAD value is *gerund*, which is a subtype of both *noun* and *relational*, but not of *verb*. The type *vger* is shown in (30); where *f-ing* is a function the produces the *-ing* form of an English verb from its root.

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

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

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

Malouf further notes a key difference between gerunds and exceptions like **childs*/*children*: English gerunds are productive (and completely regular morphologically). If the same mechanisms of default unification are involved in both, what accounts for this difference? His answer is that productive and predictable processes involve online type construction (see Section 5.2 for details). The irregular form *children* must of course be learned and stored, not generated online. The default mechanisms described above, however, are employed at higher levels of the lexical hierarchy, and the individual gerunds forms *are* productively generated online. Note that, in contrast to the morphological and syntactic consistency among gerunds, English nominalizations display some idiosycrasies that suggest at least some of them must be stored as distinct lexical items. Thus, as Malouf emphasizes, modeling prototypicality in the lexicon within HPSG can draw on both default inheritance and online type construction; together, they make "the connection between prototypicality, and productivity."

5 Lexical rules

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

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

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

In a slight variant of the representation of lexical rules proposed by Copestake & Briscoe and Meurers, the OUT attribute can be dispensed with; the information in the lexical rule type not within the IN value then constitutes the output of the rule. In this variant, lexical rules could alternatively be viewed as be subtypes of a *derived-word* type, which could combine with other types in the lexical hierarchy, merely adding the derivational source via the IN value. Formulated in either fashion, lexical rules are essentially equivalent to unary syntactic rules, with the IN attribute corresponding to the daugther and the OUT attribute (or the rest of the information in the rule, if the OUT attribute is done away with) to the mother. This is the way lexical rules are implemented in the English Resource Grammar (see http://www.delph-in.net/erg/ for demos and details about this large-scale implemented grammar of English). (cross-reference to Bender &

Emerson's chapter on computational linguistics and language engineering here)

5.1 Phenomena accounted for by lexical rules

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

We begin with the Complement Extraction Lexical Rule (hereafter, CELR) proposed in Pollard & Sag (1994) shown in (31). The input to the rule is any lexeme that selects for a syntactic argument ($\boxed{3}$) that the lexeme requires be expressed as a complement (as indicated, this syntactic argument is also a member of the COMPS list). The ouput stipulates that this same syntactic argument is no longer a member of the COMPS list; however, the SLASH set now includes a new element, which is the local information of this syntactic argument ($\boxed{1}$). Informally stated, the input entry specifies that a syntactic argument must be realized as a complement, whereas the output entry specifies that the same syntactic argument must be realized by a non-local dependent (see Pollard & Sag (1994) for why only LOCAL information is shared between syntactic arguments and fillers that realize them).

(31)
$$
\begin{bmatrix} \textsc{arg-st} & \langle \ldots, \boxed{3}, \ldots \rangle \\ \textsc{comps} & \left\langle \ldots, \boxed{3}\begin{bmatrix}\textsc{loc} & \boxed{1}\end{bmatrix}, \ldots \right\rangle \\ \textsc{slash} & \boxed{2} \end{bmatrix} \mapsto \begin{bmatrix} \textsc{arg-st} & \left\langle \ldots, \boxed{4}\begin{bmatrix}\textsc{loc} & \boxed{1} \\ \textsc{slash} & \boxed{1}\end{bmatrix}, \ldots \right\rangle \\ \textsc{comps} & \langle \ldots \rangle \\ \textsc{slash} & \{\boxed{1}\} \cup \boxed{2} \end{bmatrix}
$$

A similar use of lexical rules to model alternative realizations of arguments can be found in Monachesi (1993), who analyzes alternations between complements and so-called object clitics in Italian in a way that parallels the French examples in (10). In the output of her lexical rule, in (32), a subset of the list of complements in the input ($\boxed{2}$) corresponds to a list of clitic SYNSEMS, realized as prefixes through inflectional rules not shown here.

(32) $$\begin{bmatrix} \textit{word} & \\ \textsc{head} & \textit{verb} \\ \textsc{val}|\textsc{comps} & \boxed{1} \bigcirc \boxed{2} \\ \textsc{clts} & \textit{elist} \end{bmatrix} \mapsto \begin{bmatrix} \textit{word} & \\ \textsc{val}|\textsc{comps} & \boxed{1} \\ \textsc{clts} & \boxed{2}\,\textit{list(cl-ss)} \end{bmatrix}$$

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

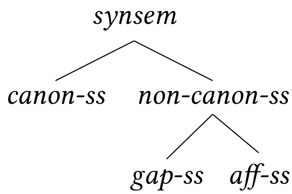
Before turning to the alternatives, we note in passing that lexical rules are inherently "directional", with an input and an output. This seems intuitively correct in the cases we've discussed, but might not always be so. Is there inherent directionality, for example between the causative and inchoative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

5.2 Alternatives to lexical rules

In this section we briefly examine two alternatives to lexical rules, each involving underspecification. The types of members of the ARG-ST list might be underspecified so that a lexical entry accounts for more than one subcategorization. Or the type of the entry itself may be underspecified, so that it subsumes multiple inflectional or derivations forms. In both cases, the intent is that sufficiently underspecified information covers multiple entries that would otherwise have to be specified and related by lexical rules. We begin with alternatives to the complement extraction and clitic lexical rules in (31) and (32), proposed in Bouma et al. (2001) and Miller & Sag (1997).

In both cases, the idea is to distinguish between "canonical" and "non-canonical" realizations of syntactic arguments, as shown in the hierarchy of *synsem*s in (33). "Canonical" means local realization as a complement or subject/specifier, and "non-canonical" means realization as an affix or filler of an unbounded dependency. Linking constraints between semantic roles (values of argument positions) and syntactic arguments (members of ARG-ST) do not specify whether the realization is canonical or not; thus they retain their original form. Only canonical members of ARG-ST must be structured-shared with members of valence lists. The two constraints that determine the non-canonical realization of fillers are shown in (34). (34a) specifies what it means to be a *gap-ss*), namely that the argument is extracted (its local information is "slashed,") whereras (34b) prohibits any *gap-ss* member from being a member of the COMPS list. As these two constraints are compatible with either a canonical or extracted object, there is no need for the lexical rule is (31).

(33)



(34) a. *gap-ss* $\Rightarrow \begin{bmatrix} \textsc{loc} & \boxed{1} \\ \textsc{slash} & \boxed{1} \end{bmatrix}$

b. *word* $\Rightarrow \begin{bmatrix} \textsc{subj} & \boxed{1} \\ \textsc{comps} & \boxed{2} \ominus \mathit{list(gap\text{-}ss)} \\ \textsc{deps} & \boxed{1} \oplus \boxed{2} \end{bmatrix}$

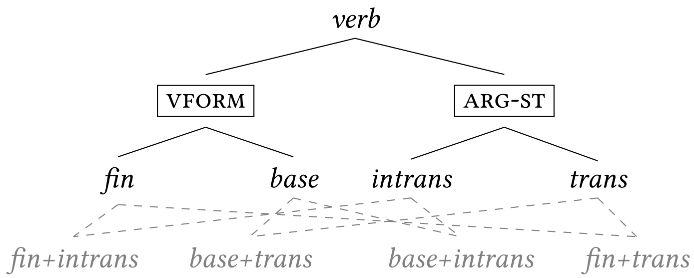
Miller & Sag (1997) make a similar use of non-canonical relations between the ARG-ST list and the valence lists, eschewing lexical rules to model French clitics and proposing instead the constraint in (35), where a subset of ARG-ST members, those that are realized as affixes (of type *aff*) are not also subcategorized for as complements.

(35)
$$\begin{bmatrix} \textsc{morph} & \begin{bmatrix} \textsc{form} & F_{PRAF}(\boxed{1}, ...) \\ \textsc{i-form} & \boxed{1} \end{bmatrix} \\ \textsc{synsem} & \begin{bmatrix} \textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head} & \mathit{verb} \\ \textsc{val} & \begin{bmatrix} \textsc{subj} & \boxed{2} \\ \textsc{comps} & \boxed{3}\mathit{list(non\text{-}aff)} \end{bmatrix} \\ \textsc{arg-st} & (\boxed{2} \oplus \boxed{3}) \bigcirc \mathit{nelist(aff)} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

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

In (36) is a simplified version of the hierarchy in (20), where the boxed labels in small caps VFORM and ARG-ST are mnemonic names of orthogonal dimensions of classification of subcategories of verbs (and are not themselves labels of subcategories). Inheritance links to the predictable subtypes are dashed and their names grayed out; this indicates that these types can be inferred, and need not be declared explicitly as part of the grammar. A grammar of English would include statements to the effect that head information about verbs includes a classification of verbs into finite or base forms (of course, there would be more types of verb forms in a realistic grammar of English) as well a classification into intransitive and transitive verbs (again, a realistic grammar would include many more types).

(36)



Crysmann & Bonami (2016) have shown how this *online type construction*, where predictable combinations of types of orthogonal dimensions of classification are not reified in the grammar, is useful when modeling productive inflectional morphology. Consider, for example, exponents of morphosyntactic features whose shape remains constant, but whose position within a word's template (to speak informally here) varies. One case like this is the subject and object markers of Swahili, which can occur in multiple slots in the Swahili verb template. For reasons of space we illustrate the usefulness of this dynamic approach to type creation, the Type Underspecified Hierarchical Lexicon (TUHL) with an example from Koenig (1999), the cross-cutting classification of syntactic/semantic information and stem form in the entry for the French verb *aller* (see Bonami & Boyé (2001) for a much more thorough discussion of French stem allomorphy along similar lines; Crysmann and Bonami's much more developed approach to stem allomorphy would model the same phenomena differently and we use Koenig's simplified presentation for expository purposes only). The forms of *aller* are based on four different suppletive stems: *all-* (1^{st} and 2^{nd} person plural of the indicative and imperative present, infinitive, past participle, and imperfective past), *i-* (future and conditional), *v-* (1^{st}-3^{rd} person singular and 3^{rd} person plural of the indicative present), and *aill-* (subjunctive present). These four suppletive stems are shared by all entries (i.e., senses) of the lexeme *aller*: the one which means 'to fit' as well as the one which means 'to leave', as shown in (37) (see Koenig, op.cit, p.40-41). The cross-cutting generalizations over lexemes and stems are represented in Figure 2. Any *aller* stem combines one entry and one stem form. In a traditional HPSG type hierarchy, each combination of types (grayed out in Figure 2), would have to be stipulated. In a TUHL, these combinations can be dynamically created when an instance of *aller* needs to be produced or comprehended.

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

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

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

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

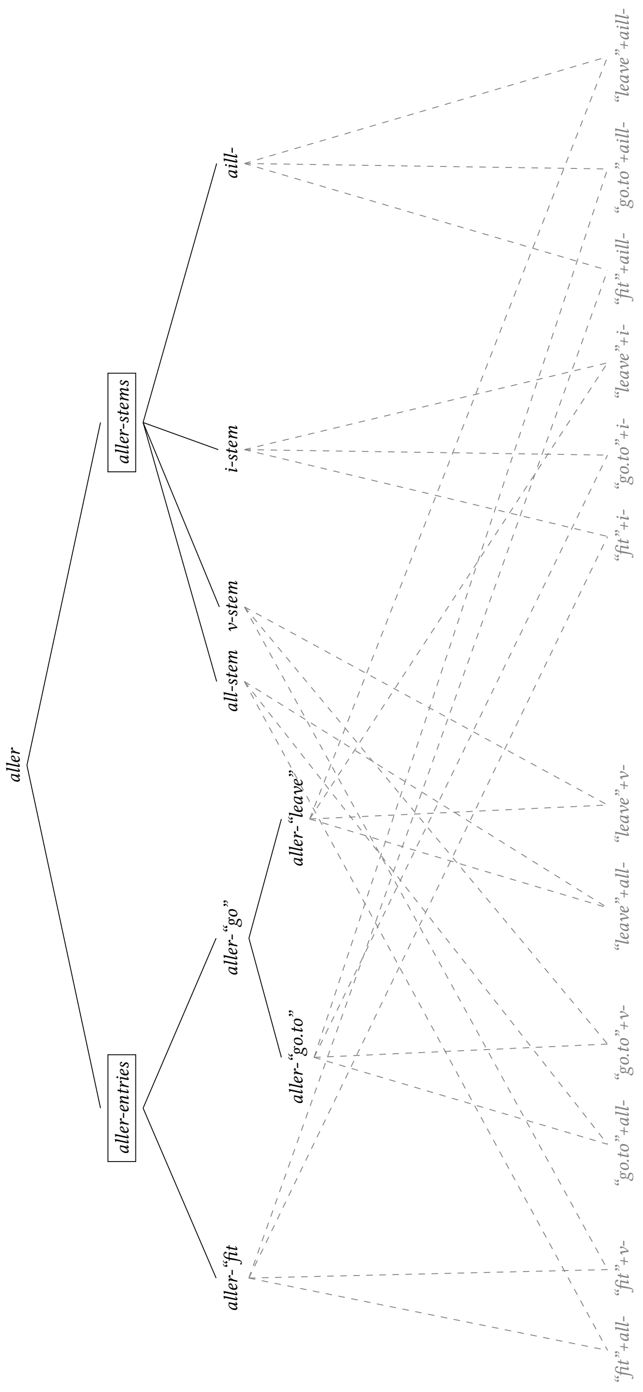


Figure 2: A hierarchy of lexical entries and stem-forms for the French verb *aller*, from Koenig (1999)

Both *synsem* and type underspecification avoid conflict between the information specified in the variants of words based on a single lexeme (e.g., conflicts on how syntactic arguments are realized); they abstract over the relevant pieces of conflicting information. Underspecifying information included in lexical entries or lexical types allows a single entry or type to stand for the two distinct entries or types that would be related as input and output by lexical rules. The third alternative to lexical rules eschews informational conflict by adding internal structure to stems and words. This is the approach to derivational morphology taken by Riehemann (1998). Example (38) (Riehemann's (1)) illustrates *-bar* suffixation in German, a process by which an adjective that includes a modal component can be derived from verb stems (similar to English *-able* suffixation). A lexical rule approach would posit a verb stem input and derive an adjective output. As Riehemann stresses, though, there are many different subtypes of *-bar* suffixation, some productive, some unproductive, all sharing some information. This combination of productive and unproductive variants of a lexical process is exactly what the type hierarchy is meant to capture and what Riehemann's *Type-Based Derivational Morphology* capitalizes on. (39) presents the relevant information of Riehemann's type for regular *-bar* adjectives (see p.68 for more details). Critically, *-bar* adjectives include a singleton-list base (the value of MORPH-B) that records the information of the adjective's verbal base (what would be the lexical rule's input). Because of this extra layer, the local information in the base (*local*1) and the *-bar* adjective (*local*$_2$) can differ without being in conflict.

(38)

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

(39)
$$\begin{bmatrix} \textit{reg-bar-adj} & \\ \textsc{phonology} & \boxed{1} + \textit{bar} \\ \textsc{morph-b} & \left\langle \begin{bmatrix} \textit{trans-verb} & \\ \textsc{phon} & \boxed{1} \\ \textsc{local} & \textit{local}_1 \end{bmatrix} \right\rangle \\ \textsc{synsem}|\textsc{local} & \textit{local}_2 \end{bmatrix}$$

Lexical rules played a critical role in the rise of lexicalist approaches to syntax. But the three alternative analytical tools we discussed in this section (which, of course, can be combined in an analysis) have chipped away at their use in HPSG. Inflectional morphology is now dealt with through lexical types associating morphosyntactic features with forms/positions and constraints on words (ensuring that all morphosyntactic features are realized). (cross-reference the chapter on Morphology) Derivational morphology is handled via lexical types too, but ones that add an extra internal layer (the MORPH-BASE in Riehemann's analysis and (39)). Non-canonical realization of syntactic arguments as affixes or fillers in unbounded dependencies is now modeled by distinguishing kinds of members of the ARG-ST list and constraints on words that relate valence, argument structure, and dependents lists. So, what remains of the case for lexical rules now? Müller (2006; 2010) argues that diathesis phenomena, broadly speaking, favor a lexical rules approach over a phrase-structural constructional approach à la Goldberg (1995) or an online type construction approach suggested in Kay (2002). The arguments are convincing, but it should be noted that some of the data involves derivational morphology (e.g., causatives) or passive morphemes, which, arguably, could be handled via a Type-Based Derivational Morphology of the kind Riehemann argues for (such an approach was suggested in Koenig (1999: Chapter 4)). It is unclear to us whether there are always motivated types for morphologically derived stems to dispense entirely with lexical rules of the kind Müller argues for. On the other hand, if one adopts the version of lexical rules in which the OUT attribute is eliminated and lexical rules are subtypes of *derived-lexeme*, little will be at stake formally, as lexical rules and derivational processes "look the same."

Abbreviations

Acknowledgements

References

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

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

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

Baldridge, Jason. 2002. *Lexically specified derivational control in combinatory categorial grammar*. Edinburgh: University of Edinburgh dissertation.

Bochner, Harry. 1993. *Simplicity in generative morphology*. Berlin: Mouton de Gruyter.

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

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

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

Bouma, Gosse, Frank Van Eynde & Dan Flickinger. 2000. Constraint-based lexicons. In Frank Van Eynde & Dafydd Gibbon (eds.), *Lexicon development for speech and language processing* (Text, Speech and Language Technology). Dordrecht: Kluwer Academic Publishers.

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

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

Bresnan, Joan & Sam Mchombo. 1995. The lexical integrity principle: evidence from Bantu. *Natural Language and Linguistic Theory* 13. 181–254.

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

Broadwell, George Aaron. 2008. Turkish suspended affixation is lexical sharing. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2008 conference*, 198–213. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/13/, accessed 2018-2-25.

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

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

Carpenter, Bob. 1992b. *The logic of typed feature structures* (Tracts in Theoretical Computer Science). Cambridge: Cambridge University Press.

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

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

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

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

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

Crysmann, Berthold. 2001. Clitics and coordination in linear structure. In Birgit Gerlach & Janet Grijzenhout (eds.), *Clitics in phonology, morphology and syntax* (Linguistik Aktuell/Linguistics Today 36), 121–159. Amsterdam: John Benjamins Publishing Co.

Crysmann, Berthold. 2015. On the placement and morphology of Udi subject agreement. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 Conference, Norwegian University of Science and Technology, Trondheim*. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2001/, accessed 2015-12-23.

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

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

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

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

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

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

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

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

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

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

Haspelmath, Martin. 2011. The indeterminacy of word segmentation1 and the nature of morphology and syntax. *Folia Linguistica* 45. 31–80.

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

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

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

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

Koenig, Jean-Pierre. 1999. *Lexical relations* (Stanford Monographs in Linguistics). Stanford, CA: CSLI Publications.

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

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

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

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

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

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

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

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

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

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

Monachesi, Paola. 1993. Object clitics and clitic climbing in Italian HPSG grammar. In Steven Krauwer, Michael Moortgat & Louis des Tombe (eds.), *Sixth Conference of the European Chapter of the Association for Computational Linguistics. Proceedings of the conference*, 437–442. Uetrecht: Association for Computational Linguistics.

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

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

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

Pollard, Carl & Ivan Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry*. 2nd ser. 23. 261–303.

Pollard, Carl & Ivan Sag. 1994. *Head-driven phrase-structure grammar*. Chicago, IL: Chicago University Press.

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

Reape, Mike. 1994. Domain union and word order variation in german. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in head-driven phrase structure grammar*, 151–197. Stanford University: CSLI Publications.

Richter, Frank. 1999. RSRL for HPSG. In Valia Kordoni (ed.), *Tübingen studies in head-driven phrase structure grammar* (Arbeitsberichte des SFB 340 132), 74–115. Tübingen: Universität Tübingen.

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

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

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

Wechsler, Stephen. 1999. HPSG, GB, and the Balinese bind. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 179–195. Stanford, CA: CSLI Publications.

Chapter 5

Understudied languages

Doug Ball

Aron Broadwell

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

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

1 Introduction

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

Doug Ball & Aron Broadwell. 2018. Understudied languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 81–82. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Part II

Syntactic phenomena

Chapter 6

Agreement

Stephen Wechsler
The University of Texas at Austin

Put abstract here with \abstract.

1 Introduction

Agreement is the systematic covariation between a semantic or formal property of one element (called the agreement *trigger*) and a formal property of another (called the agreement *target*). In the sentences *I am here* and *They are here*, the subjects (*I* and *they*, respectively) are the triggers; the target verb forms (*am* and *are*, respectively) covary with them. Research on agreement systems within HPSG has been devoted to describing and explaining a number of observed aspects of such systems. Regarding the grammatical relationship between the trigger and the target, we may first of all ask how local that relationship is, and in what grammatical terms it is defined. Having determined the prevailing locality conditions on agreement in a given language, we attempt to explain observed exceptions, that is, cases of apparent 'long-distance agreement', as well as cases of superficial agreement defined on string adjacency. Agreement features across languages include person, number, and gender (known as *phi* features), as well as deictic features and case, but various different subsets of those features are involved in particular agreement relations. How can we explain the distribution of features? How are locality and feature distribution related to the diachronic origin of agreement systems? Also, as indicated in the definition of agreement provided in the first sentence of this paper, the features of the target are sometimes determined by the trigger's form and sometimes by its meaning. What regulates this choice? In some cases a single trigger in a sentence determines different features on two different targets. Why does such 'mixed agreement'

Stephen Wechsler. 2018. Agreement. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 85–106. Berlin: Language Science Press. DOI:??

exist, and what does its existence tell us about the grammatical representation of agreement? This chapter reviews HPSG approaches to these questions of locality, grammatical representation, feature distribution, diachrony, semantic versus formal agreement, and mixed agreement.

HPSG offers an integrated account of these phenomena. In most cases the analysis of agreement phenomena does not involve any special formal devices dedicated for agreement, comparable to the *probe* and *goal*, or the AGREE relation, found in GB/Minimalist accounts. Instead the observed agreement phenomena arise as a side effect of other grammatical processes such as valence saturation, the semantics of modification, coreference, and feature defaults.

2 Agreement as unification

Constraint-based formalisms such as HPSG are uniquely well-suited for modeling agreement. Within such formalisms agreement occurs when multiple feature sets arising from distinct elements of a sentence specify information about a single abstract object, so that the information must be mutually consistent (Kay 1984). The two forms are said to agree when the values imposed by the two constraints are compatible, while ungrammaticality results when they are incompatible. For example the English verb *is* in (1) specifies that its initial ARG-ST list item, which is identified with the SUBJ list item, has third person, singular features. In the process of valence saturation, the NP list item in the value of SUBJ unifies with the feature structure representing the subject NP. The features specified by the verb for its subject and by the subject NP must be compatible; otherwise the representation for the resulting sentence is ill-formed, predicting ungrammaticality as in (3a).

(1) Simplified lexical sign for the verb *is*:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{is} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \boxed{2} \rangle \end{bmatrix} \\ \textsc{arg-st} & \left\langle \boxed{1}\text{NP}\begin{bmatrix} \textsc{pers} & \mathit{3rd} \\ \textsc{num} & \mathit{sg} \end{bmatrix}, \boxed{2}\text{XP} \right\rangle \end{bmatrix}$$

(2) Simplified lexical signs for *I* and *she*:

$$\begin{bmatrix} \text{PHON} & \langle \textit{I} \rangle \\ \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{PERS} & \textit{1st} \\ \text{NUM} & \textit{sg} \end{bmatrix} \end{bmatrix} \qquad \begin{bmatrix} \text{PHON} & \langle \textit{she} \rangle \\ \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{PERS} & \textit{3rd} \\ \text{NUM} & \textit{sg} \\ \text{GEN} & \textit{fem} \end{bmatrix} \end{bmatrix}$$

(3) a. *I is sober.
    b. She is sober.

The features supplied by the trigger and target must be consistent, but there is no general minimum requirement on how many features they specify. Both of them can be, and typically are, underspecified for some agreement features.

Since unification is commutative, the representation of an agreement construction is the same regardless of whether a feature originates from the trigger or the target. This immediately accounts for common agreement behavior observed when triggers are underspecified (Barlow 1988). For example, Serbo-Croatian is a grammatical gender language, where common nouns are assigned to the masculine, feminine, or neuter gender. The noun *knjiga* 'book' in (4) is feminine, so the modifying determiner and adjective appear in feminine form, as does the participle.

(4) Ov-a star-a knjig-a stalno pad-a.
    this-NOM.F.SG old-NOM.F.SG book-NOM.SG always fall-3SG
    'This old book keeps falling.' (Wechsler & Zlatić 2003: p. / 4, ex. 1)

However, some nouns are unspecified for gender, such as *sudija* 'judge'. Interestingly, the gender of an agreeing adjective actually adds semantic information, indicating the sex of the judge:

(5) a. Taj stari sudija je dobro sudio.
       that.M old.M judge AUX well judged.M
       'That old (male) judge judged well.'
    b. Ta stara sudija je dobro sudila.
       that.F old.F judge AUX well judged.F
       'That old (female) judge judged well.'

Here the gender feature comes from the targets instead of the trigger. This illustrates an advantage of constraint-based theories like HPSG over transformational accounts in which a feature is copied from the trigger, where it originates, to the target, where it is then realized. The usual source of the feature (the noun) lacks it in (5), a problem for the feature-copying view.

The same problem occurs even more dramatically in *pro*-drop. Many languages allow subject pronouns to drop, and distinguish person, number, and/or gender on the verb. If those features originate from the null subject, then there would have to be distinct null pronouns, one for each verbal inflection (Pollard & Sag 1994: p. 64). In HPSG a *pro*-dropped argument usually appears on the ARG-ST list but not a VALENCE list (see Ch. ARGST CHAPTER). In the context given in (6) a Serbo-Croatian speaker could omit the subject pronoun.

(6) Context: Speaker comes home to find her bookcase mysteriously empty.

Gde	su	(one)	nestale?
Where	did	(they.F.PL)	disappear.F.PL

'Where did they (i.e. the books) go?'

The sign for the inflected participle specifies feminine plural features on the initial item in its ARG-ST list. The SUBJ list item is optional:

(7) Simplified lexical sign for the participle form *nestale*:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{nestale} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \left\langle \left(\boxed{1}\right) \right\rangle \\ \textsc{comps} & \langle\ \rangle \end{bmatrix} \\ \textsc{arg-st} & \left\langle \boxed{1}\textsc{np}\begin{bmatrix}\textsc{num} & pl \\ \textsc{gen} & fem\end{bmatrix} \right\rangle \end{bmatrix}$$

The feminine plural features are specified regardless of whether the subject pronoun appears. When the pronoun is dropped we have the usual underspecification, only in this case the trigger does not exist, so it is effectively fully underspecified, realizing no features at all.

3 Locality in agreement

3.1 Argument and modifier agreement

In HPSG, the grammatical agreement of a predicator with its subject or object, or an adjective, determiner, or other modifier with its head noun, piggy-backs on the processes of valence saturation and modification. Agreement is encoded in the grammar by adding features of person, number, gender, case and deixis to the existing feature structures involved in syntactic and semantic composition. This simple assumption is sufficient to explain the broad patterning of distribution of agreement, in contrast to the transformational approach where complex locality conditions must be stipulated.

In HPSG predicate-argument agreement arises directly from the process of valence saturation, as illustrated already in (1) above. Thus the locality conditions on the trigger-target relation follow from the conditions on the specifier-head or complement-head relation. Similarly, attributive adjectives agree with nouns directly through the composition of the modifier with its head, which takes place with the MOD feature. For example, the Serbo-Croatian feminine adjective form *stara* 'old.F' in (5b) specifies feminine singular features for the common noun phrase that it modifies.

(8) Simplified lexical sign for *stara*:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{stara} \rangle \\ \textsc{mod} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{noun} & \\ \textsc{num} & \textit{sg} \\ \textsc{gend} & \textit{fem} \end{bmatrix} \\ \textsc{comps} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$$

In head-adjunct phrases, the MOD value of the adjunct daughter is token-identical with the *synsem* value of the head daughter. So *stara*'s feminine singular features cannot conflict with the features of the noun it modifies.

The predicted locality conditions are also affected by the percolation of features from words to phrasal nodes, and this depends on the location of the features within the feature structure. Agreement features of the *trigger* appear either within the HEAD field or the CONTENT field (these give rise to CONCORD and INDEX agreement, respectively; see Section 4.2). In either case these features percolate from the trigger's head word to its maximal phrasal projection, due to the Head Feature Principle in the former case and the Semantics Principle in the latter. For example the the noun phrase *the books* inherits its [NUM *plur*] feature from the head word *books*. This determines plural agreement on a verb: *These books are/*is interesting*. Apparent exceptions, where a target seems to fail to agree with the head of the trigger, are discussed below.

However, agreement features of the *target* appear in neither the HEAD nor the CONTENT fields of the target form, but rather appear embedded in an ARG-ST list item or MOD features. So agreement features of the target do not project to the target's phrasal projection such as VP, S, or AP. This is a welcome consequence. If the subject agreement features of the verb projected to the VP, for example, we would expect to find VP-modifying adverbs that agree with them, but we do not.

4 Varieties of agreement target

4.1 Anaphoric agreement

In anaphoric agreement, an anaphoric pronoun agrees in person, number, and gender with its antecedent. Since Pollard & Sag (1994; 1992), anaphoric agreement has been analyzed in HPSG by assuming that person, number, and gender are formal features of the referential index associated with an NP. Anaphoric binding in HPSG is modeled as coindexation, i.e. sharing of the INDEX value, between the binder and bindee. Thus any specifications for agreement features of the INDEX arising from the binder and bindee must be mutually consistent. In (9) Principle A of the binding theory requires the reflexive pronoun to be coindexed with an o-commanding item, here the subject pronoun:

(9) a. She admires herself.

b. $$\textit{admire}: \begin{bmatrix} \text{ARG-ST} \left\langle \text{NP:}\begin{bmatrix} \textit{ppro} \\ \text{INDEX } \boxed{1}\begin{bmatrix}\text{PERS} & \textit{3rd}\\ \text{NUM} & \textit{sg} \\ \text{GEN} & \textit{fem}\end{bmatrix}\end{bmatrix}, \text{NP:}\begin{bmatrix}\textit{ana}\\ \text{INDEX } \boxed{1}\begin{bmatrix}\text{PERS} & \textit{3rd}\\ \text{NUM} & \textit{sg} \\ \text{GEN} & \textit{fem}\end{bmatrix}\end{bmatrix}\right\rangle\end{bmatrix}$$

The agreement features are formal features and not semantic ones, but the semantic correlates of person (speaker, addressee, other), number (cardinality), and gender (male, female, inanimate, etc.) are invoked under certain conditions (described in Section 5). Thus index agreement is distinct from *pragmatic agreement* whereby semantic features of two coreferential expressions must be semantically consistent in order for them to refer to a single entity. Index agreement is enforced only within the syntactic domain defined by binding theory, while pragmatic agreement applies everywhere. For example, feminine pronouns are sometimes used for ships, in addition to neuter pronouns. Whichever gender is chosen, it must be consistent in binding contexts (example based on Pollard & Sag (1994: p. 79, ex. 46a)):

(10) a. The ship lurched, and then it righted itself. She is a fine ship.
b. The ship lurched, and then she righted herself. It is a fine ship.
c. *The ship lurched, and then she righted itself.
d. *The ship lurched, and then it righted herself.

The bound reflexive must agree formally with its antecedent, while other coreferential pronouns need not agree, as they are not coarguments of the antecedent and not subject to the structural binding theory.

In grammatical gender languages, where common nouns are conventionally assigned to a gender, an anaphoric pronoun appearing outside the binding domain of its antecedent can generally agree with that antecedent either formally or, if it is semantically appropriate (such as an animate, sexed entity), it can alternatively agree pragmatically. In most situations pronouns allow either pragmatic or INDEX agreement with their antecedents. For example, pronouns coreferential with the Serbian/Croatian grammatically neuter diminutive noun *devojče* 'girl' can appear in either neuter or feminine gender (from Wechsler & Zlatić (2003: p. 198)):

(11)

Ovo	malo	devojče$_i$	je	ušlo.NT.SG
this.NT.SG	little.NT.SG	girl.NT.SG	AUX.3SG	entered.

a.

Ono$_i$	je	htelo	da	telefonira.
it.NT.SG	AUX.SG	wanted.NT.SG	that	telephone

b.

Ona$_i$	je	htela	da	telefonira.
she.F.SG	AUX.SG	wanted.F.SG	that	telephone

'This little girl$_i$ came in. She$_i$ wanted to use the telephone.'

The neuter pronoun in (11a) reflects INDEX agreement with the antecedent while the feminine pronoun (11b) reflects its reference to a female (pragmatic agreement). But when a reflexive pronoun is locally bound by a nominative subject, agreement in formal INDEX features is preferred:

(12)

Devojče	je	volelo	samo/?*samu	sebe.
girl.NOM.NT.SG	AUX3.SG	liked.NT.SG	own.ACC.NT.SG/ACC.F.SG	self.ACC

'The girl liked herself.'

Again, this illustrates INDEX agreement in the domain defined by the structural binding theory.

4.2 Grammatical agreement: INDEX and CONCORD

As noted above, in HPSG agreement effectively piggy-backs on other independently justified grammatical processes. Anaphoric agreement is a side-effect of binding (Section 4.1) while grammatical agreement is a side-effect of valence saturation and modification (Section 3.1). The formal HPSG analysis of a particular agreement process mainly consists of positing agreement features somewhere in the feature structure; the observed properties follow from the location of those agreement features. With regard to the location of the features, grammatical agreement bifurcates into two types, INDEX and CONCORD.[1] (The attribute name CONCORD was introduced by Wechsler & Zlatić (2000; 2003); precursors to the idea were treated as HEAD features in Pollard & Sag (1994), and called AGR by Kathol (1999).) The best way to understand this bifurcation of agreement, and indeed the operation of grammatical agreement systems generally, is by considering their diachronic origin. Although our primary goal is the description of synchronic grammar, a look at diachrony can help explain the forms that the grammar takes, and can also provide clues as to the best formalization of it.

Within the diachronic literature on agreement there are thought to be two different lexical sources for agreement inflections: (i) incorporated pronouns; and (ii) incorporated noun classifiers (Greenberg 1978). Going back further, the noun classifiers derive from so-called 'generic' common nouns, i.e. semantically superordinate common nouns meaning 'animal', 'man', 'woman', and so on. These two sources, ultimately traced to pronouns and common nouns, give rise to Index and Concord target inflections, respectively, as explained next.

4.2.1 INDEX agreement

Taking pronouns first, many grammatical agreement systems evolve historically from the incorporation of pronominal arguments into the predicates selecting those arguments, such as verbs and nouns (Bopp 1842; Givón 1976; Wald 1979, inter alia). When a nominal topic serving as antecedent to the incorporated pronoun is reanalyzed as the true subject or object of the predicate, the pronominal affix effectively becomes an agreement marker. With this reanalysis the only change in the affix is that it loses its ability to refer: it no longer functions as a pronoun. The affix retains its agreement features, and what was formerly anaphoric agreement with the topic becomes grammatical agreement with the subject or object. This explains why the features of grammatical agreement match those of pronominal anaphora: typically person, number, and gender, with occasional deictic features (Bresnan & Mchombo 1987: p. / 752).

As explained above, structural anaphoric binding involves identifying (structure-sharing) the referential indices of the pronoun and its binder. Therefore grammatical agreement derived from it is also INDEX agreement. For example, the signs for English *is* and *I* in (1) and (2) above should be rewritten as follows:

[1]The INDEX / CONCORD theory is sketched in Pollard & Sag (1994: ch. 2) and Kathol (1999), and developed in detail in Wechsler & Zlatić (2000; 2003), all in the HPSG framework. It has since been adopted into LFG (King & Dalrymple 2004: inter alia) and GB/Minimalism (Danon 2009).

(13) Sign for *is*, illustrating INDEX agreement

$$\begin{bmatrix} \textsc{phon} & \langle \textit{is} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \left\langle \textsc{np} \begin{bmatrix} \textsc{content}|\textsc{index} & \begin{bmatrix}\textsc{pers} & \textit{3rd} \\ \textsc{num} & \textit{sg}\end{bmatrix}\end{bmatrix}\right\rangle \\ \textsc{comps} & \langle \textsc{xp}\rangle \end{bmatrix}\end{bmatrix}$$

(14) Sign for *I*, illustrating INDEX features:

$$\begin{bmatrix} \textsc{phon} & \langle \textit{I} \rangle \\ \textsc{content}|\textsc{index} & \boxed{1}\begin{bmatrix} \textit{ppro} & \\ \textsc{pers} & \textit{1st} \\ \textsc{num} & \textit{sg} \end{bmatrix} \\ \textsc{context} & \textit{speaker}(\boxed{1}) \end{bmatrix}$$

This finite verb form specifies third person singular features of its subject's referential index.

One salient distinguishing characteristic of INDEX agreement is that it includes the PERSON feature. The only known diachronic source of the person feature is from pronouns. Therefore, the other type of agreement, CONCORD, lacks the person feature (as we will see below).

By modeling verb agreement in a way that reflects its historical origin, we are able to explain an array of facts concerning particular agreement systems. Some of these facts and explanations are presented in Section 6 below.

4.2.2 CONCORD

The agreement inflections on modifiers of nouns, such as adjectives and determiners, are thought to derive historically, not from pronouns, but from noun classifiers (Greenberg (1978); Reid (1997); Seifart (2009); Grinevald & Seifart (2004), Corbett (2006: p. 268–9)). The classifier morphemes in turn derive historically from lexical common nouns denoting superordinate categories like animal, woman, man, etc. For example Reid (1997) posits the following historical development of Ngan'gityemerri (southern Daly; southwest of Darwin, Australia), a language where the historical stages continue to cooccur in the current synchronic grammar. Originally the language had general-specific pairings of nouns as a common syntactic construction, such as *gagu wamanggal* 'animal wallaby' in (15a) (from Reid (1997, 216) example 162-5). The specific noun can be omitted when reference to it is established in discourse, leaving the general noun and modifier, to form NPs like *gagu kerre*, literally 'animal big' but functioning roughly like nominal ellipsis 'big one'. Then, where the specific noun is also included, both noun and modifier attract the generic term (15b). The gender markers then reduce phonologically and incorporate, producing modifier gender agreement (15c).

(15) a. Stage I:
Gagu wamanggal kerre ngeben-da.
animal wallaby big 1SG.SB.AUX-shoot
'I shot a big wallaby.'

b. Stage II:
Gagu wamanggal gagu kerre ngeben-da.
animal wallaby animal big 1SG.SB.AUX-shoot
'I shot a big wallaby.' (Reid (1997, 216) example 164)

c. Stage III:
wa=ngurmumba wa=ngayi darany-fipal-nyine.
male=youth male=mine 3SG.AUX-return-FOC
'My initiand son has just returned.'

If the same affix is retained on the modifiers and the noun they modify, then the result is symmetrical agreement (also known as alliterative agreement), like the feminine *-a* endings in Spanish *zona rosa* (Corbett 2006, 87-88). But often an asymmetry between the affixes on the noun and the modifiers develops: the noun affix becomes obligatory and is subject to morphophonological processes that do not affect the modifier affix (Reid 1997, 216). This process may further progress to 'prefix absorption' into the common noun, as evidenced by 'gender prefixed nominal roots being interpreted as stems for further gender marking.' (Reid 1997, 217)

CONCORD is agreement marked with inflections from such nominal sources. What is the proper HPSG formalization of this type of agreement, given its provenance? The last stages of the diachronic development, described in the previous paragraph, imply that the *form* of the trigger (the noun) is influenced by the agreement features. That is, noun declension classes correlate with gender assignment (and more generally, phonological and morphological characteristics of nouns correlate with gender assignment); and number is marked on nouns as well. (This close relation between declension class and CONCORD is demonstrated in detail in Wechsler & Zlatić (2003: ch. 2).) Thus the agreement features must appear both on the head noun (to inform its form and/or its gender selection and number value) and on the phrasal projection of that noun (to trigger agreement via the MOD feature of the agreement targets). Ergo CONCORD is a HEAD feature of the trigger.

Along with number and gender, CONCORD often includes the case feature, since case is a feature of NPs that can be realized on both the head noun and its modifying adjectives or determiner. CONCORD lacks the person feature, since common nouns, from which the agreement inflections derive, lack the person feature. Meanwhile, INDEX agreement preserves the pronominal features of person, number, and gender, reflecting its origins. In the usual case the number and gender values found in CONCORD match those found in INDEX. The Serbo-Croatian noun form *knjiga* triggers feminine singular nominative CONCORD on its adjectival possessive specifier and modifier, and third person singular INDEX agreement on the finite auxiliary. (The status of the participle is discussed below.)

(16) Moja stara knjiga je pala.
my.F.NOM.SG old.F.NOM book.NOM.SG AUX.3.SG fall.PPRT.F.SG
'My old book fell.' (from Wechsler & Zlatić (2003: p. 18))

The nominative singular noun form *knjiga* specifies its agreement features in both CONCORD (in the HEAD field) and INDEX, with the respective values for number and gender shared:

(17) Lexical sign for *knjiga* 'book' (from Wechsler & Zlatić (2003: p. 18)):

$$\begin{bmatrix} \text{PHONOLOGY} & \langle \textit{knjiga} \rangle \\ \text{SYNSEM} & \begin{bmatrix} \text{CATEGORY} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{CONCORD} & \boxed{3}\begin{bmatrix}\text{CASE} & \textit{nom}\\ \text{NUM} & \boxed{1}\,\textit{sing} \\ \text{GEN} & \boxed{2}\,\textit{fem}\end{bmatrix}\end{bmatrix} \\ \text{VALENCE}|\text{SPR} & \left\langle \left(\text{AP}\begin{bmatrix}\textit{poss};\ \text{CONCORD}\ \boxed{3}\end{bmatrix}_{\boxed{j}}\right) \right\rangle \end{bmatrix} \\ \text{CONTENT} & \begin{bmatrix} \text{INDEX} & \boxed{i}\begin{bmatrix}\text{PERS} & \textit{3rd}\\ \text{NUM} & \boxed{1}\\ \text{GEN} & \boxed{2}\end{bmatrix} \\ \text{RESTR} & \left\{ \textit{BOOK(I)} \left(, \textit{possess}(j,i)\right)\right\}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

The features in the overlap between the two agreement fields are normally shared as in this example. But in some cases features can be asymmetrically specified in only one field (with no reentrancy linking them, of course). This leads to mismatches between CONCORD and INDEX targets, discussed in Section 6 below.

4.3 Conclusion

To summarize this section, we have seen the two main historical paths to agreement, and shown how HPSG formalizes these two types of agreement so as to capture the syntactic and semantic properties that follow directly from their origins. Agreement that descends from anaphoric agreement of pronouns with their antecedents, through the incorporation of personal pronouns into verbs and other predicators, inherits the INDEX matching process found in the anaphoric agreement from which it descends. Agreement that descends from the incorporation of noun classifiers involves features located in the HEAD field that connect a trigger noun form to its phrasal projection. The feature sets differ for the same reason; person is a feature only of the first type, and case only of the second. CONCORD correlates strongly with declension class, while INDEX agreement need not correlate as strongly (for evidence see Wechsler & Zlatić (2003: ch. 2)). The differences in feature sets and morphology further correlate with systematic syntactic differences, described in the following section.

5 Syntactic, semantic, and default agreement

This article has so far focused mainly on formal agreement, as opposed to semantic agreement. But this is one of three different ways in which the form of an agreement target may be determined by a grammar:

(18) Formal, semantic, and default determinants of target form.
   a. Formal agreement: The target form depends on the trigger's formal phi features.
   b. Semantic 'agreement': The target form depends on the trigger's meaning.
   c. Failure of agreement: The target fails to agree and hence takes its default form.

In formal agreement the trigger is grammatically specified for certain features as a consequence of the words making up the trigger phrase: for example a nominal may be marked for a gender as a consequence of the lexical gender of the head noun. In semantic agreement the target is sensitive to the meaning of the trigger instead of its formal features. English number agreement can be formal as in (19) or semantic as in (20):

(19) a. His clothes are/*is dirty.
   b. His clothing is/*are dirty.
   (from Wechsler (2013: p. 92))

(20) a. That the position will be funded and that Mary will be hired now seems/??seem likely.

b. That the president will be reelected and that he will be impeached are/??is equally likely at this point.
(from McCloskey (1991: p. 92))

Regarding (20) McCloskey (1991:564-5) observes that singular is used for 'a single complex state of affairs or situation-type', while plural is possible for 'a plurality of distinct states of affairs or situation-types'. The latter sort of interpretation is facilitated by the use of the adverb *equally*. Formal and semantic gender agreement are illustrated by the French examples in (21):

(21) a. La sentinelle à la barbe a été {prise/*pris} en otage.
the.F sentry bearded AUX been taken.F.SG/taken.M hostage
'The bearded sentry was taken hostage.'

b. Dupont est {compétent / compétente}.
Dupont is competent.M.SG / competent.F.SG
'Dupont {a man / a woman} is competent.'

The grammatically feminine noun *sentinelle* triggers feminine agreement regardless of the sex of the sentry; but in (21b) feminine agreement indicates that Dupont is female.

How does the grammar negotiate between formal and semantic agreement? In HPSG syntactic and semantic representations are composed in tandem, making the framework well suited to address this question. It was addressed in early HPSG work, including Pollard & Sag (1994: ch. 1). The specific approach due to Wechsler (2011) exploits the underspecification of agreement features (see Section 2). He posits the Agreement Marking Principle (AMP), which states that target agreement features are semantically interpreted whenever the trigger is underspecified for the formal grammatical features to which the target would normally be sensitive. The subject phrases in (19) are specified for number due to the formal features of the head nouns; but those in (20) are not, as a (coordinate) clause has no grammatical source for those features. Consequently, by the AMP the verb's number feature is semantically interpreted in (20). Similarly, *sentinelle* in (21a) gives its formal feminine gender feature to the subject, while *Dupont* lacks a gender specification, triggering the semantic interpretation of the target adjectives in (21b): feminine is interpreted as 'female'.

Agreement targets generally have a default form for use when there is no trigger or the normal agreement relation is blocked for some reason. Blocking of agreement comes about in various situations; here we consider a case where the trigger is interpreted metonymically, apparently resulting in a reassignment of the referential index. Swedish predicate adjectives normally agree with their subjects in number (either singular or plural) and grammatical gender, either neuter (NT) or 'common' gender (COM), the gender held in common between masculine and feminine:

(22) a. Hus-et är gott.
house-DEF.NT.SG is good.NT.SG
'The house is good.'

b. Pannkaka-n är god.
pancake-DEF.COM.SG be.PRES good.COM.SG
'The pancake is good.'

c. Hus-en/ Pannkak-orna är god-a.
house-PL.DEF/ pancake-PL.DEF be.PRES good-PL
'The houses/ The pancakes are good.'

As shown in (22), a predicate adjective is inflected for number, and, in the singular, for gender, and agrees with its subject. But in sentences like (23), the adjective appears in the neuter singular form, regardless of the number and gender features of the subject. Note that *pannkakor* is the plural form of a common gender noun (Faarlund 1977; Enger 2004; Josefsson 2009):

(23) Pannkak-or är gott.
pancake-PL be.PRES good.NT.SG
'Situations involving pancakes are good.' (e.g. 'Eating pancakes is good.')

In general Swedish predicate adjectives appear in neuter singular when there is no triggering NP, such as with clausal subjects (see (25a) below). Wechsler and Zlatić (2003) posit the index type *unm* ('unmarked') for referential indices that lack phi features, such as those introduced by verbs. So *gott* has a SUBJ list item whose index is disjunctively specified for either neuter singular, or type *unm*.

The lack of agreement in (23) then arises because the subject phrase refers, not to the pancakes, but to a situation involving them; hence its referential index is distinct from the one lexically introduced by the noun *pannkakor*. A rule shifts the index and encodes the metonymic relation between the entity and the situation involving it. This is implemented with a non-branching phrasal construction rule in Wechsler (2013: p. 82, ex. 20):

(24) *metonymy-ctx:*

$$\begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{SYN} & \textit{NP} \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & s_{unm} \\ \text{RESTR} & \left\{\textit{involve}(s,i)\right\} \cup \boxed{1} \end{bmatrix} \end{bmatrix} \\ \text{DTRS} & \begin{bmatrix} \text{SYN} & \textit{NP} \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & i \\ \text{RESTR} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The noun *pannkakor* has an index marked with the features [PERSON *3rd*], [GENDER *com*], and [NUMBER *pl*], which, by the Semantics Principle, are therefore shared with the index of the daughter NP node in rule (24). But the mother NP node's index is unmarked for those features, thus explaining the neuter singular adjective.

On the alternative ellipsis analysis, sentence (23) has an elliptical clausal or infinitival subject, with a structure like (25a) except that *att äta* is silent (Faarlund 1977; Enger 2004; Josefsson 2009):

(25) a. Att äta pannkakor är gott.
to eat pancakes be.PRES good.NT.SG
'Eating pancakes is good.'

b. Det är gott att äta pannkakor.
it be.PRES good.NT.SG to eat pancakes
'It is good to eat pancakes.'

c. *Det är gott pannkakor.
it be.PRES good.NT.SG pancakes
('It is good to eat pancakes')

But the metonymic subject behaves in all respects like an NP, and unlike a clause or infinitival phrase. For example, unlike an infinitival it resists extraposition, as shown in (25b,c). The metonymy analysis captures the fact that the subject has a clause-like meaning but not clause-like syntax.

6 Mixed agreement

The two-feature (INDEX/CONCORD) theory of agreement was originally motivated by *mixed agreement*, where a single phrase triggers different features on distinct targets (Pollard & Sag (1994: ch. 2); Kathol (1999)). For example, the French second person plural pronoun *vous* refers to multiple addressees, and also has an honorific or polite use for a single (or multiple) addressee. When used to refer politely to one addressee, *vous* triggers singular on a predicate adjective but plural on the verb, as in (26a):

(26) a. Vous êtes loyal.
you.PL be.2PL loyal.M.SG
'You (singular, formal, male) are loyal.'

b. Vous êtes loyaux.
you.PL be.2PL loyal.PL
'You (plural) are loyal.'

Wechsler (2011) analyzes this by adopting the following suppositions: (i) *vous* has a second person plural marked referential INDEX; (ii) *vous* lacks phi features for CONCORD; (iii) finite verbs agree with their subjects in INDEX; and (iv) predicate adjectives agree with their subjects in CONCORD. Suppositions (i) and (iii) need not be stipulated, as they follow from the theory: the pronoun must have INDEX phi features since it shows anaphoric agreement (when it serves as binder or bindee); and the verb must agree in INDEX since it includes the PERSON feature. By the Agreement Marking Principle (see Section 5), the (CONCORD) number and gender features of the predicate adjective are interpreted semantically, which is what is shown by example (26).

'Polite plural pronouns' of this kind are found in many languages of the world (Head 1978). The cross-linguistic agreement patterns observed in typological studies (Comrie 1975; Wechsler 2011) confirm the predictions of the theory. Taken together, suppositions (i) and (iii) from the previous paragraph entail that any person agreement targets agreeing with polite pronouns should show formal, rather than semantic, agreement. Targets lacking person, meanwhile, can vary across languages. This pattern is confirmed for all languages with polite plurals that have been surveyed, including Romance languages, Modern Greek, Germanic (Icelandic), West, South and East Slavic, Hindi, Gbaya (Niger-Congo), Kobon and Usan (Papuan), and Sakha (Turkic) (see Comrie (1975) and Wechsler (2011)).

The Index/Concord distinction plays a crucial role in this account of mixed agreement. An earlier hypothesis, proposed by Kathol (1999), is that French predicate adjectives are grammatically specified for semantic agreement with their subjects, while finite verbs show formal agreement. But plurale tantum noun such as *ciseaux* 'scissors', triggers syntactic agreement on the predicate adjective:

(27) Ces ciseaux sont géniaux! (*génial!)
these.PL scissors(M.PL) are.PL brilliant.M.PL (*brilliant.M.SG)
'These scissors are cool!'

As far as the syntax is concerned, *ciseaux* 'scissors' is an ordinary common noun with masculine plural CONCORD features, so it triggers those features on the adjective. More generally, agreement target types cannot be split into 'formal' and 'semantic' agreement targets; both formal and semantic agreement are found across all target types. Which of the two is observed depends, according to the Index/Concord theory, on the feature markedness of the trigger, together with the Index versus Concord status of the target.

7 Agreement defined on other structures

So far our look at grammatical agreement has focused primarily on agreement defined on local grammatical relations like subject, object, and modifier. In this section we look at HPSG analyses of two other types of agreement, namely long-distance and superficial agreement.

7.1 Long-distance agreement

The simple picture of locality in the previous sections is challenged by the phenomenon of long-distance agreement, where trigger appears within a clause subordinate to the one headed by the target verb. Long-distance agreement has been observed in a number of languages, including Tsez (Nakh-Dagestanian; Polinsky & Potsdam (2001), Hindi-Urdu (Bhatt 2005), and Passamaquoddy (Athabaskan; Bruening (2001); LeSourd (2018)).

Passamaquoddy long-distance agreement is illustrated by this sentence (LeSourd (2018: ex. 5)):

(28) N-kosicíy-a-**k** [eli- Píyel -litahási-t
1-know-DIR-PROX.PL thus- Peter -think-3AN
[eli-kis-ankum-í-hti-t **nìkt** **ehpíc-ik**
thus-PAST-sell-3/1-PROX.PL-3AN those.PROX woman-PROX.PL
posonúti-yil]]
basket-IN.PL
'I know that Peter thinks that those women sold me the baskets.'

The *-k* suffix on the matrix verb *kosicíy* 'know' marks plural, deictically proximate agreement with the phrase *nìkt ehpícik* 'those women' in the doubly embedded subordinate clause. LeSourd (2018) analyzes Passamaquoddy long distance agreement in the HPSG framework. He notes that Passamaquoddy long distance agreement is parallelled by long-distance raising, in which an NP in the matrix clause is coreferential with an implicit argument of a subordinate clause (LeSourd (2018: ex. 4)):.

(29)

N-kosicíy-a-**k**	**nìkt**	**ehpíc-ik**$_i$	[eli-	Píyel	-litahási-t
1-know-DIR-PROX.PL	those.PROX	woman-PROX.PL	thus-	Peter	-think-3AN

[eli-kis-ankum-í-hti-t	**e**$_i$	posonúti-yil]]
thus-PAST-sell-3/1-PROX.PL-3AN		basket-IN.PL

'I know about those women$_i$ that Peter thinks that they$_i$ sold me the baskets.'

Passamaquoddy speakers report that sentences (28) and (29) suggest the subject of 'know' (the speaker) is familiar with the women. This provides evidence that the phrase 'those women' in (29) is an argument of the matrix verb 'know', as implied by the translation. (In other words, this is a prolepsis construction.) Similarly, the *-k* suffix in (28) is an incorporated object pronoun cataphoric to 'those women'.[2] What the long-distance agreement and raising constructions share is simply that the matrix object is coreferential with some argument contained in the subordinate clause. The following lexical entry for *kosicíy* 'know' captures that:

(30) *kosicíy* 'know':

$$\begin{bmatrix} \textsc{phon} & \langle \textit{kosicy} \rangle \\ \textsc{arg-st} & \left\langle \textsc{np}_i\ ,\ \textsc{np}_j\ ,\ \textsc{s}{:}\begin{bmatrix}\textsc{restr}\left\langle \ldots \begin{bmatrix}\textsc{prd}|\textsc{arg} & j\end{bmatrix} \ldots \right\rangle\end{bmatrix} \right\rangle \end{bmatrix}$$

By the Semantic Inheritance Principle (Sag et al. 2003), the semantic restrictions (RESTR) list inherits all the restrictions of the daughter nodes. Thus every semantic argument contained within the S complement, whether overt or null, will correspond to some argument (ARG) of an elementary predication (PRD) in S's RESTR list. The lexical entry in (30) stipulates that the matrix object NP corefers with some such argument. In conclusion, Passamaquoddy long distance agreement is really anaphoric agreement in a prolepsis construction.

[2]LeSourd notes that Passamaquoddy lacks Principle C effects, so cataphora of this kind is permitted.

7.2 Superficial agreement

In some languages string adjacency of the trigger and target, rather than a grammatical relation such as subject or modifier, is a grammatical condition on agreement. This may arise because person agreement derives historically from pronoun incorporation, and a basic syntactic precondition for incorporation is string adjacency between the pronoun and the head into which it incorporates (Givón 1976; Ariel 1999; Wechsler et al. 2010; Fuss 2005). If the trigger occupies the syntactic position that the pronoun occupied prior to incorporation (for example because the trigger is itself a pronoun) then the result is that trigger and target are adjacent. For example, West Flemish complementizers agree with an immediately following subject, even though the complementizer and subject are not related by any grammatical relation (Haegeman 1992). To take another example, Borsley (2009) analyzes Welsh superficial agreement in the HPSG framework, citing examples like the following:

(31) a. Gwelon nhw ddraig.
see.PAST.3PL they dragon
'They saw a dragon.'

b. arno fo
on.3SG.M he
'on him'

c. Gweles i a Megan geffyl.
see.PAST.1SG I and Megan horse
'Megan and I saw a horse.'

The trigger is the subject in (31a), object in in (31b), first conjunct of a coordinate subject in in (31c). But in every case, 'An agreeing element agrees with an immediately following noun phrase if and only if the latter is a pronoun' (Borsley 2009: ex. 48). Borsely (2009, ex. 99) expresses this as an HPSG implicational constraint using the DOMain feature from linearization theory (Reape 1994; Kathol 2000):

(32) $[\text{DOM} < [\text{AGR } \boxed{1}], \text{NP: } \textit{ppro}\boxed{2}, \ldots >] \rightarrow \boxed{1} = \boxed{2}$

The DOMain list encodes linear precedence between constituents that are not necessarily sisters. In (32) the AGR value is the phi feature field of the target; the rule states that when a constituent bearing this attribute is immediately followed by a personal pronoun (content of type *ppro*), then this value is identified with the pronoun's index (shown here as $\boxed{2}$), that is, it agrees with a right-adjacent pronoun.

8 Conclusion

HPSG analysis of agreement consists primarily of assigning phi features to specific fields in the feature structures representing the grammar. Anaphoric agreement results from phi features appearing on the referential indices. Verbal agreement with subjects and objects results when phi features appear on the verb's ARG-ST list items. Modifier agreement with heads occurs when phi features appear on the MOD value of the modifier. According to the Index/Concord theory, those features within the ARG-ST list or MOD items are located on the referential index for agreement historically descended from anaphoric agreement of incorporated pronouns; while otherwise they are head features (collected in the CONCORD field). The locality conditions on agreement follow from the normal operation of the grammar in which those phi features are embedded. Some cases of agreement seem to exist outside those conditions. Long-distance agreement has been analyzed as a kind of anaphoric agreement within a prolepsis construction; and superficial agreement has been defined on string adjacency and precedence, within linearization theory.

Abbreviations

Acknowledgements

References

Ariel, Mira. 1999. The development of person agreement markers: from pronouns to higher accessibility markers. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 197–260. Stanford: CSLI Publications.

Barlow, Michael. 1988. *A Situated Theory of Agreement*. Stanford University PhD thesis.

Bhatt, Rajesh. 2005. Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23(4). 757–807.

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

Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63. 741–82.

Bruening, Benjamin. 2001. *Syntax at the edge: cross-clausal phenomena and the syntax of Passamaquoddy*. MIT dissertation.

Comrie, Bernard. 1975. Polite Plurals and Predicate Agreement. *Language* 51(2). 406–418.

Corbett, Greville. 2006. *Agreement*. Cambridge: Cambridge University Press.

Danon, Gabi. 2009. Agreement with quantified nominals: Implications for feature theory. In *Colloque de Syntaxe et Sémantique à Paris*.

Enger, Hans-Olav. 2004. Scandinavian pancake sentences as semantic agreement. *Nordic Journal of Linguistics* 27(01). 5–34.

Faarlund, Jan Terje. 1977. Embedded Clause Reduction and Scandinavian Gender Agreement. *Journal of Linguistics* 13(2). 239–257.

Fuss, Eric. 2005. *The rise of agreement: a formal approach to the syntax and grammaticalization of verbal inflection*. Amsterdam: John Benjamins.

Givón, Talmy. 1976. Topic, pronoun and grammatical agreement. In Charles N. Li (ed.), *Subject and topic*, 149–88. New York: Academic Press.

Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Morovcsik (eds.), *Word structure*, vol. III (Universals of Human Language). Stanford: Stanford University Press.

Grinevald, Colette & Frank Seifart. 2004. Noun classes in African and Amazonian languages: Towards a comparison. *Linguistic Typology* 8(2). 243–285.

Haegeman, Liliane. 1992. *Theory and description in generative syntax: a case study in West Flemish*. Cambridge, England: Cambridge University Press.

Head, Brian F. 1978. Respect degrees in pronominal reference. In Joseph H. Greenberg (ed.), *Universals of Human Language: Vol. 3, Word Structure*, 151–212. Stanford, California: Stanford University Press.

Josefsson, Gunlög. 2009. Peas and Pancakes: On apparent disagreement and (null) light verbs in Swedish. *Nordic Journal of Linguistics* 32(01). 35–72.

Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In R. Levine & G. Green (eds.), *Studies in contemporary phrase-structure grammar*. Cambridge: Cambridge University Press.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University Press.

Kay, Martin. 1984. *Functional unification grammar: a formalism for machine translation*. Association for Computational Linguistics. 75–78.

King, Tracy Halloway & Mary Dalrymple. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40. 69–104.

LeSourd, Philip S. 2018. Raising and long distance agreement in Passamaquoddy: a unified analysis. *Journal of Linguistics*.

McCloskey, James. 1991. There, it, and agreement. *Linguistic Inquiry* 22(3). 563–567.

Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19(19). 583–646.

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

Pollard, Carl J. & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.

Reape, Mike. 1994. Domain Union and Word Order Variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford: CSLI Publications.

Reid, Nicholas. 1997. Class and classifier in Ngan'gityemerri. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in aboriginal Australia*, 165–228. Amsterdam: John Benjamins.

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

Seifart, Frank. 2009. Multidimensional typology and Miraña class markers. In Patience Epps & Alexandre Arkhipov (eds.), *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, 365–385. Berlin: Walter de Gruyter.

Wechsler, Stephen. 2011. Mixed agreement, the person feature, and the Index/Concord distinction. *Natural Language and Linguistic Theory* 29(4). 999–1031.

Wechsler, Stephen. 2013. The structure of Swedish *pancakes*. In Elisabeth Norcliffe & Philip Hofmeister (eds.), *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, 71–98. Stanford: CSLI publications.

Wechsler, Stephen, Patience Epps & Elizabeth Coppock. 2010. The origin and distribution of agreement features: a typological investigation. Project description, National Science Foundation grant application.

Wechsler, Stephen & Larisa Zlatić. 2000. A theory of agreement and its application to Serbo-Croatian. *Language* 76. 799–832.

Wechsler, Stephen & Larisa Zlatić. 2003. *The many faces of agreement*. Stanford: CSLI Publications.

Chapter 7

Case

Adam Przepiórkowski
Polish Academy of Sciences

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

HPSG is not widely known for its approach to grammatical case. For example, it is only mentioned in passing in the 2006 monograph *Theories of Case* (Butt 2006: 225) and in the 2009 *Oxford Handbook of Case* (Malchukov & Spencer 2009: 43), which features separate articles on GB/Minimalism, LFG, Optimality Theory, and other grammatical frameworks. As most of the HPSG work on case was carried out in 1990s and early 2000s, this perception is unlikely to have changed since the publication of these 2 volumes.

The aim of this chapter is to provide an overview of HPSG work on grammatical case and to show that it does offer novel solutions to some of the problems related to case. Two main research areas are presented in the two ensuing sections: structural case assignment is discussed in Section 1 and case neutralisation ('indeterminacy') and underspecification – in Section 2. Some of the other HPSG work on case, including implementational work, is outlined in Section 3.

Adam Przepiórkowski. 2018. Case. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 107–133. Berlin: Language Science Press. DOI:??

1 Structural case assignment

Pollard & Sag (1994) did not envisage a separate theory of case:[1] "Nominative case assignment takes place directly within the lexical entry of the finite verb," while "the subject SUBCAT element of a nonfinite verb… does not have a case value specified" (30). However, they added in a footnote on the same page that "for languages with more complex case systems, some sort of distinction analogous to the one characterized in GB work as 'inherent' vs. 'structural' is required."

In the transformational Government and Binding theory of 1980s (GB; Chomsky 1981; 1986), *inherent* – or *lexical* – case is understood as rigidly assigned by the head and independent of syntactic environment, while 'structural' case varies with the structural context (e.g., Haider 1985: 70). This difference can be illustrated on the basis of the following examples from German (Przepiórkowski 1999a: 63, based on data from Heinz & Matiasek 1994):

(1) a. Der Mann unterstützt *den Installateur.*
the man.NOM supports the plumber.ACC
'The man is supporting the plumber.'

b. *Der Installateur* wird unterstützt.
the plumber.NOM AUX supported
'The plumber is supported.'

c. das Unterstützen *des Installateurs*
the supporting the plumber.GEN
'the support for/from the plumber'

(2) a. Der Mann hilft *dem Installateur.*
the man.NOM helps the plumber.DAT
'The man is helping the plumber.'

b. *Dem Installateur* wird geholfen.
the plumber.DAT AUX helped
'The plumber is helped.'

c. das Helfen *des Installateurs*
the helping the plumber.GEN
'the help from/*for the plumber'

[1] This section is to some extent based on Przepiórkowski 1999a, Section 3.4 and Chapter 4; see also Müller 2013, Chapter 14.

In (1), both arguments of the verb UNTERSTÜTZEN 'support' receive structural case: the patient argument occurs in the accusative in (1a), in the nominative in (1b), and in the genitive in (1c). Similarly, the agent argument is in the nominative in (1a), but it may only occur in the genitive in (1c); hence, the single argument marked as genitive in (1c) is ambiguous between the agent and the patient. In the case of (2), the agent argument of HELFEN 'help' is similarly assigned structural case, but the patient argument receives a rigid inherent case: it is always the dative, so, e.g., the genitive in (2c) may only be understood as marking the agent.

Examples such as above may still be handled without any general principles of case assignment. For example, lexical rules responsible for forming passive participles (as in the b. examples above) and gerunds (as in the c. examples) might be responsible for manipulating case values of arguments, e.g., for translating nominative and accusative – but not dative – to genitive in the case of gerunds. However, the interaction of the structural/inherent case dichotomy with raising (and – in some languages – with control) motivates a more comprehensive approach to case assignment.

Consider Icelandic raising verbs (all Icelandic data is taken from Sag et al. 1992: 304–305):

(3) a. *Hann* virðist elska hana.
he.NOM seems love.INF her.ACC
'He seems to love her.'

b. Þeir telja *Maríu* hafa skrifað ritgerðina.
they believe Mary.ACC have.INF written the-thesis
'They believe Mary to have written her thesis.'

As in other languages, the subject of the infinitival verb raised to the higher subject position, as in (3a), normally receives the nominative case there, while – in the case it is raised to the object position, as in (3b) – it normally receives the accusative case. This could be easily modelled in accordance with the suggestion of Pollard & Sag (1994: 30) that infinitival verbs do not assign case to their subjects, while finite verbs – in this case finite raising verbs – normally assign the nominative to their subjects and the accusative to their objects. But, as is well known (Andrews 1982; Zaenen & Maling 1983; Zaenen et al. 1985), some Icelandic verbs idiosyncratically assign specific "quirky" cases to their subjects, and when they do, the higher raising verbs must honour this assignment:

(4) a. *Hana* virðist vanta peninga.
her.ACC seems lack.INF money
'She seems to lack money.'

b. Hann telur *mig* vanta peninga.
he.NOM believes me.ACC lack.INF money
'He believes that I lack money.'

(5) a. *Barninu* virðist hafa batnað veikin.
the-child.DAT seems have.INF recovered-from the-disease
'The child seems to have recovered from the disease.'

b. Hann telur *barninu* hafa batnað veikin.
he believes the-child.DAT have.INF recovered-from the-disease
'He believes the child to have recovered from the disease.'

(6) a. *Verkjanna* virðist ekki gæta.
the-pains.GEN seems not be-noticeable.INF
'The pains don't seem to be noticeable.'

b. Hann telur *verkjanna* ekki gæta.
he believes the-pains.GEN not be-noticeable.INF
'He believes the pains to be not noticeable.'

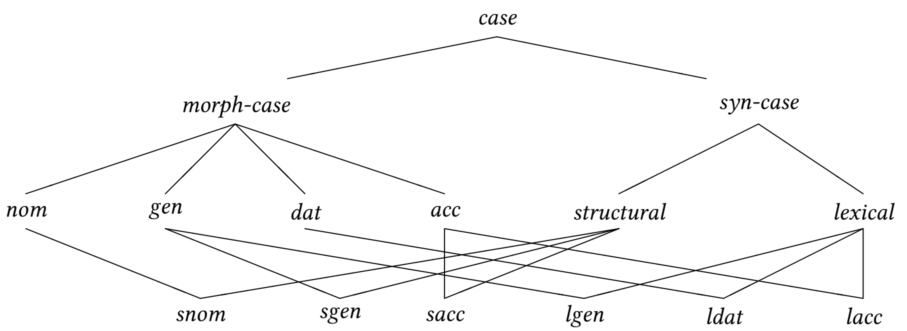
Thus, in (4), the understood subject of the infinitival VANTA 'lack' must be in the accusative, whether it is raised the object position, as in (4b), where the accusative would be expected anyway, or to the subject position, as in (4a), where normally the nominative case would be expected. Similarly in the case of verbs idiosyncratically assigning their subject the dative case, as in (5), or the genitive case, as in (6).

The difficulty presented by such examples is this. If the finite raising verbs were assumed to assign case to the raised subjects – nominative in the case of raising to subject and accusative in the case of raising to object – than this would clash with 'quirky' cases assigned to their subjects by some verbs: (4a), (5) and (6) would be predicted to be ungrammatical. If, on the other hand, such raising verbs did not assign case to the raised arguments, instead relying on the lower verbs to assign appropriate cases to their subjects, then it is not clear what case should be assigned to their subjects by the usual – not 'quirky' – verbs: it cannot always be the nominative, as the accusative case is witnessed when the subject is raised to the object position, as in (3b); similarly, it cannot always be the accusative, as the nominative case surfaces when the subject is raised to the subject position, as in (3a).

The intuition of the analysis proposed in Sag et al. 1992 relies on the distinction between structural and inherent case assignment, although these terms do not appear in this paper. Verbs such as those in (4)–(6) assign their subjects specific inherent cases (accusative in (4), dative in (5) and genitive in (6)), while the usual verbs, as in (3), only mark their subjects as structural, to be assigned case elsewhere. Finite raising verbs are, in a way, sensitive to this distinction, and only assign the nominative (in the case of raising to subject) or accusative (in the case of raising to object) to such structural arguments. While Sag et al. (1992) represent this distinction between structural and inherent case implicitly, via the interaction of two attributes, CASE (realised case) and DCASE (default case), later HPSG work assumes explicit representation of the two kinds of case as two subtypes of *case* in the type hierarchy: *str*(uctural) and *lex*(ical). Such a *case* type hierarchy is, apparently independently, alluded to in Pollard 1994 and introduced in detail in Heinz & Matiasek 1994, to which we turn presently.

On the basis of German examples such as (1)–(2), Heinz & Matiasek (1994) argue that out of 4 morphological cases in German – nominative, accusative, genitive and dative — the first three (i.e., without the dative) may be assigned structurally, by general case assignment principles. Similarly, they argue that the last three (i.e., without the nominative) may also be assigned lexically, in which case they are stable across various syntactic environments. These empirical observations are translated into the following *case* hierarchy:

(7)



Particular verbs may assign specific lexical cases to their arguments, e.g., *ldat*.

They may also specify arguments as bearing structural case, in which case only the *str*(*uctural*) supertype is mentioned in the lexicon. For example, the lexical entries for UNTERSTÜTZEN 'support' and HELFEN 'help' contain the following subcategorisation requirements:

(8) a. UNTERSTÜTZEN: [SUBCAT ⟨ NP[*str*], NP[*str*] ⟩]
b. HELFEN: [SUBCAT ⟨ NP[*str*], NP[*ldat*] ⟩]

Assuming a similar *case* hierarchy for Icelandic, the difference between the usual verbs, such as ELSKA 'love' in (3a), and 'quirky' subject verbs, such as VANTA 'lack' in (4), could be represented as below (omitting non-initial arguments):

(9) a. ELSKA: [SUBCAT ⟨ NP[*str*], ...⟩]
b. VANTA: [SUBCAT ⟨ NP[*lacc*], ...⟩]

Since Pollard 1994 and Heinz & Matiasek 1994, such representations of case requirements are generally adopted in HPSG,[2] with the only difference that SUBCAT is currently replaced with ARG-ST. The point where different approaches diverge is how exactly structural case is resolved to a specific morphological case.

The simplest principle would resolve the case of the first *str* argument of a pure (non-gerundial) verb to nominative, i.e., to *snom*, the case of any subsequent *str* argument of a pure verb to accusative, i.e., to *sacc*, and the case of any *str* argument of a gerund to *sgen*. Unfortunately, this simple principle would not work in various cases of raising, e.g., in the case of the Icelandic data above. While the 'quirky' cases in (4)–(6) would be properly taken care of by this approach – once the subject is assigned a specific lexical case it is outside of the realm of a principle resolving structural cases – structural subjects raised to a higher verb would be assigned specific case twice (or more times, in the case of longer raising chains): on the SUBCAT (or ARG-ST) of the lower verb and on the SUBCAT (or ARG-ST) of the raising verb. This would not necessarily lead to problems in the case of raising to subject verbs, as in (3a), as the structural argument would be the subject in both subcategorisation frames, so it would be resolved to *snom* twice, but it would create a problem in the case of raising to object verbs, as in (3b), as the raised argument would be resolved to the nominative on the lower subcategorisation frame and to the accusative on the higher frame. So, the problem is not limited to Icelandic, but may be observed in any language with raising to object (also known as Exceptional Case Marking or Accusativus cum Infinitivo), including German (cf., e.g., Heinz & Matiasek 1994: 231). Obviously, even if a structural argument occurs on a number of SUBCAT or ARG-ST lists, it should be assigned

[2] A very recent example being Machicao y Priemer & Fritz-Huechante 2018.

specific morphological case according to its position on just one of them – the highest one.

Both Pollard 1994 and Heinz & Matiasek 1994 account for such facts via configurational case principles, e.g. (Heinz & Matiasek 1994: 209):

(10) Case Principle (for German):

In a *head-complement-structure* whose head has category
verb[*fin*] the external argument has a CASE value of *snom*,
verb the internal argument has a CASE value of *sacc*,
noun the internal argument has a CASE value of *sgen*.
These are the only saturated or almost saturated *head-complement-structure*s with structural arguments.

(11) *Syntactically External Argument ('Subject')*:

If the first element of the SUBCAT list of a sign is an NP[*str*], it is called the *(syntactically) external argument* of that sign.

(12) *Syntactically Internal Argument ('Direct Object')*:

If the second element of the SUBCAT list of a sign is an NP[*str*], it is called the *(syntactically) internal argument* of that sign.

Heinz & Matiasek (1994: 209–210) formalize this Case Principle by giving the following constraints:

(13)
$$\begin{bmatrix} \textit{phrase} \\ \text{SYNSEM|LOC|CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \textit{verb} \\ \text{VFORM} \ \textit{fin} \end{bmatrix} \\ \text{SUBCAT} \ \langle\rangle \end{bmatrix} \\ \text{DTRS} \begin{bmatrix} \textit{h-c-str} \\ \text{HEAD-DTR|...|SUBCAT} \ \langle \ \text{NP}[\textit{str}] \ \rangle \end{bmatrix} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} \text{DTRS|HEAD-DTR|...|SUBCAT} \ \langle \ \text{NP}[\textit{snom}] \rangle, \ldots \end{bmatrix}$$

(14)
$$\begin{bmatrix} \textit{phrase} \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{verb} \\ \textsc{vform} \ \textit{fin} \end{bmatrix} \\ \textsc{subcat} & \langle\rangle \vee \left\langle \textit{synsem} \right\rangle \end{bmatrix} \\ \textsc{dtrs} \begin{bmatrix} \textit{h-c-str} \\ \textsc{head-dtr}|\ldots|\textsc{subcat} \ \left\langle \textit{synsem}, \text{NP}[\textit{str}], \ldots \right\rangle \end{bmatrix} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} \textsc{dtrs}|\textsc{head-dtr}|\ldots|\textsc{subcat} \ \left\langle \textit{synsem}, \text{NP}[\textit{sacc}] \right\rangle, \ldots \end{bmatrix}$$

(15)
$$\begin{bmatrix} \textit{phrase} \\ \textsc{synsem}|\textsc{loc}|\textsc{cat} \begin{bmatrix} \textsc{head} & \textit{noun} \\ \textsc{subcat} & \langle\rangle \vee \left\langle \textit{synsem} \right\rangle \end{bmatrix} \\ \textsc{dtrs} \begin{bmatrix} \textit{h-c-str} \\ \textsc{head-dtr}|\ldots|\textsc{subcat} \ \left\langle \textit{synsem}, \text{NP}[\textit{str}], \ldots \right\rangle \end{bmatrix} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} \textsc{dtrs}|\textsc{head-dtr}|\ldots|\textsc{subcat} \ \left\langle \textit{synsem}, \text{NP}[\textit{sgen}] \right\rangle, \ldots \end{bmatrix}$$

Note that the locus of this Case Principle is *phrase* and that it makes reference to *head-complement-structure* values of the DAUGHTERS (DTRS) attribute. In this sense, this principle is configurational. Similar principles were proposed for Korean (Yoo 1993; Bratt 1996), English (Grover 1995) and Polish (Przepiórkowski 1996a), *inter alia*.

This configurational approach to case assignment is criticised in Przepiórkowski 1996b; 1999b,a on the basis of conceptual and theory-internal problems. The conceptual problem is that a configurational analysis is employed to what is usually considered an essentially local phenomenon, one concerned with the relation between a head and its dependents (Blake 1994). The – more immediate – theory-internal problem is that such configurational case principles are restricted to locally realised arguments, and are not directly compatible with those – dominant since Pollard & Sag 1994: Chapter 9 – HPSG analyses of extraction which do not assume traces and with those HPSG approaches to cliticisation in which the clitic is realised as an affix rather than a tree-configurational constituent (cf., e.g., Miller & Sag 1997 on French and Monachesi 1999 on Italian).

The solution proposed in Przepiórkowski 1996b; 1999b,a is to resolve structural cases directly within ARG-ST, via local principles operating at the level of *category* of a word (where both head information and argument structure information – but not constituent structure – is available) rather than at the level of *phrase*. This seems to bring back the problem, discussed in the connection of Icelandic data above, of raised arguments, which occur on a number of ARG-ST lists. The innovation of Przepiórkowski 1996b; 1999b,a is the proposal to mark, within ARG-ST, whether a given argument is realised locally (either tree-configurationally, or as a gap to be extracted higher on, or as an affix) or not. If it is realised locally, it may be assigned appropriate case; if it is not (because it is raised), its structural case must be resolved higher up. On this setup, the above constraints (13)–(14) responsible for the assignment of structural nominative and accusative are replaced with the following two constraints (and similarly for the structural genitive):

(16) $$\begin{bmatrix} \textit{cat} & \\ \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{ARG NP}[\textit{str}] \\ \text{REALIZED} + \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{ARG-ST} \left\langle \begin{bmatrix} \text{ARG NP}[\textit{snom}] \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

(17) $$\begin{bmatrix} \textit{cat} & \\ \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \boxed{1}\ \textit{nelist} \oplus \left\langle \begin{bmatrix} \text{ARG NP}[\textit{str}] \\ \text{REALIZED} + \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} \text{ARG-ST}\ \boxed{1} \oplus \left\langle \begin{bmatrix} \text{ARG NP}[\textit{snom}] \end{bmatrix} \right\rangle \oplus \boxed{2} \end{bmatrix}$$

Obviously, for such constraints to work, values of ARG-ST must be lists of slightly more complex objects than *synsem* (these are now values of ARG within such more complex objects), and additional principles must make sure that values of REALIZED are instantiated properly (see Przepiórkowski 1999a: 78–79 for details).

While this approach seems to be sufficient to account for almost all known structural case phenomena, German presents additional difficulties, as discussed in Müller 1997a; 2001 and Meurers 1999a,b. In brief, Müller (1997a; 2001) recalls arguments from Höhle 1983; 2018 that in German controlled (not: raised) structural subjects bear the nominative case. Since such subjects are never realised locally (as such), the above case principle stated in terms of REALIZED would not resolve their case. The problem with the particular approach of Przepiórkowski 1996b; 1999b,a is the assumption that an argument is locally realised – and hence may be assigned structural case – if and only if it is not raised to a higher argument structure. The kind of data discussed in Höhle 1983; 2018 and Müller 1997a; 2001 shows that this equivalence does not always hold and suggests that structural case should be assigned to arguments on the basis of whether they are raised or not, and not whether they are locally realised or not.

The same conclusion may be reached on the basis of different data, discussed in Meurers 1999a,b on the basis of empirical observations in Haider 1990, Grewendorf 1994 and Müller 1997b:

(18) a. [*Ein* *Außenseiter* gewinnen] wird hier nie.
an.NOM outsider win.INF will here never
'An outsider will never win here.'

b. [*Einen* *Außenseiter* gewinnen] läßt Gott hier nie.
an.ACC outsider win.INF lets god here never
'God never lets an outsider win here.'

Assuming that fronted fragments, marked with square brackets, are single constituents,[3] the subject of *gewinnen* 'win' forms a constituent with this verb, i.e., it has the same configurational realisation in both examples. Hence, configurational case assignment principles should assign it the same case in both instances, contrary to facts: *ein Außenseiter* occurs in the nominative in (18a) and *einen Außenseiter* bears the accusative case in (18b). As argued by Meurers 1999a,b, the reason is that – although the subject is realised locally to its infinitival head – it is in some sense raised further to the subject position of the auxiliary *wird* in (18a) and to the object position of the AcI verb *läßt* in (18b); hence, the difference in cases. This, again, suggests that structural case should be assigned not where the argument is realised, but on the highest ARG-ST on which it occurs. A corresponding modification of the non-configurational case assignment approach of Przepiórkowski 1996b; 1999b,a – replacing the [REALIZED +] with [RAISED −] in constraints such as (16)–(17) and providing appropriate constraints on values of RAISED – is proposed in Przepiórkowski 1999a: 93–95; see also Müller 2013, Section 17.4 (and references therein), for further improvements.

While this non-configurational approach to syntactic case assignment was motivated largely by theory-internal technical considerations, it turns out to formalise sometimes apparently contradictory intuitions expressed in various approaches to case. First of all, it preserves the common intuition that case is a local phenomenon, an intimate relation between a head and its dependents.

[3] This assumption is not completely uncontroversial; see Kiss (1994: 100–101) for potential counterexamples.

Second, it successfully formalises the distinction between structural and inherent/lexical case known from the transformational literature of 1980s, and non-configurationally encodes the apparently configurational principles of structural case assignment. Third, while most HPSG literature on case is concerned with syntactic phenomena in European languages, this approach has been extended to case stacking known, e.g., from languages of Australia and case attraction observed, e.g., in Classical Armenian and in Gothic (Malouf 2000). Fourth, by allowing antecedents of implicational constraints such as (16)–(17) to be *local* objects, not just syntactic *categories*, semantic factors influencing case assignment may also be taken into account, as in differential case marking, repeatedly considered in Lexical Functional Grammar (cf., e.g., Butt & King 2003 and references therein), but apparently not (so far) in HPSG. Fifth, as pointed out in Przepiórkowski 1999b,a, the above approach to case formalises the 'case tier' intuition of Zaenen et al. 1985, Yip et al. 1987 and Maling 1993 (see also Maling 2009).

Let us illustrate the last point with some Finnish data from Maling 1993:

(19) a. Liisa muisti matkan vuoden.
Liisa.NOM remembered trip.ACC year.ACC
'Liisa remembered the trip for a year.'

b. Lapsen täytyy lukea kirja kolmannen kerran.
child.GEN must read book.NOM [third time].ACC
'The child must read the book for a 3rd time.'

c. Kekkoseen luotettiin yksi kerta.
Kekkonen.ILL trust.PASSP [one time].NOM
'Kekkonen was trusted once.'

d. Kekkoseen luotettiin yhden kerran yksi vuosi.
Kekkonen.ILL trust.PASSP [one time].ACC [one year].NOM
'Kekkonen was trusted for one year once.'

Maling (1993) argues at length that some adjuncts (adverbials of measure, duration and frequency) behave just like objects with respect to case assignment and, in particular, notes the following generalization about syntactic case assignment: only one NP dependent of the verb receives the nominative, namely the one which has the highest grammatical function; other dependents receive the accusative.[4] Thus, if none of the arguments bears inherent case, the subject is in

[4] See also Zaenen & Maling 1983 and Zaenen et al. 1985 for a similar generalisation with respect to Icelandic.

the nominative and other dependents are in the accusative, cf. (19), but if the subject bears an idiosyncratic case, it is the object that gets the nominative, cf. (19b). Furthermore, if all arguments (if any) bear inherent case, the next 'available' grammatical function is that of an adjunct, thus one of the adjuncts receives the nominative, cf. (19c)–(19d).

Given such facts, Maling (1993) claims that syntactic case is assigned in Finnish on the basis of grammatical hierarchy and that (at least some) adjuncts belong to this hierarchy. Moreover, as evidenced by (19c)–(19d), adjuncts do not form a single class in this hierarchy: although the multiplicative adverbial *yksi kerta* is nominative in (19c), this case is won over by the duration adverbial in (19d). Taking into consideration also the partitive of negation facts (measure adverbials, but not duration or frequency adverbials, behave like direct objects in the sense that they take partitive case under sentential negation), Maling (1993) extends the grammatical function hierarchy for Finnish in the following way:

(20) SUBJ > OBJ > MEASURE > DURATION > FREQUENCY

While these generalisations are developed in the context of Lexical Functional Grammar, it is not clear how they could be encoded in LFG: there are no formal mechanisms for stating such a hierarchy of grammatical functions and, additionally, all adjuncts are assumed to be elements of an unordered set. On the other hand, given the 'adjuncts as complements' approach of Bouma et al. 2001 and others, upon which at least some adjuncts are added to ARG-ST (perhaps renamed to DEPS), and assuming – as standard in HPSG – that ARG-ST elements satisfy the obliqueness hierarchy, formalisation of the 'case tier' approach is easy and consists of two implicational principles similar to (16)–(17). The first principle resolves the first structurally-cased element of extended ARG-ST to nominative, whether this element is the first element of ARG-ST or not (it is not in the case of (19b)–(19d)), and whether it corresponds to the subject, the direct object or an adjunct. The second principle resolves the structural case of all subsequent elements, if any, to accusative.

2 Case syncretism

Another important strand of HPSG work on case concerns situations in which a single syncretic form seems to simultaneously bear two (or more) case values, as in the following examples involving coordination, free relatives and parasitic gaps:

(21) Polish coordination (Dyła 1984: 701–702):

a. Kogo Janek lubi a Jerzy
who.ACC/GEN Janek.NOM likes(OBJ.ACC) and Jerzy.NOM
nienawidzi?
hates(OBJ.GEN)
'Who does Janek (John) like and Jerzy (Jerry) hate?'

b. * Co Janek lubi a Jerzy
what.NOM/ACC Janek.NOM likes(OBJ.ACC) and Jerzy.NOM
nienawidzi?
hates(OBJ.GEN)
'What does Janek (John) like and Jerzy (Jerry) hate?' (putative)

(22) German coordination (Pullum & Zwicky 1986: 764–765):

a. Er findet und hilft
he.NOM finds(OBJ.ACC) and helps(OBJ.DAT)
Frauen.
women.NOM/ACC/GEN/DAT
'He finds and helps women.'

b. * Sie findet und hilft Männer.
she.NOM finds(OBJ.ACC) and helps(OBJ.DAT) men.NOM/ACC/GEN
'She finds and helps men.' (putative)

c. * Sie findet und hilft Männern.
she.NOM finds(OBJ.ACC) and helps(OBJ.DAT) men.DAT
'She finds and helps men.' (putative)

(23) German free relatives (Groos & van Riemsdijk 1981: 212):

Was du mir gegeben hast, ist
what.NOM/ACC you.NOM me.DAT given(OBJ.ACC) have is(SUBJ.NOM)
prächtig.
wonderful
'What you have given to me is wonderful.'

(24) English parasitic gaps (Hukari & Levine 1996: 482, Levine et al. 2001: 205):

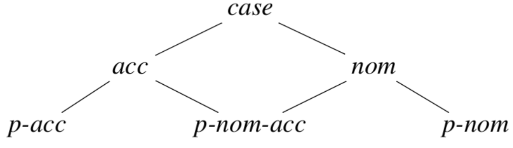
Robin is someone who$_i$.NOM/ACC even good friends of e_i.ACC believe e_i.NOM should be closely watched.

In (21a), the fronted syncretic accusative/genitive form *kogo* 'who' satisfies the requirements of the two coordinated verbal constituents: in one, *lubi* 'likes' requires an accusative object, and in the other, *nienawidzi* 'hates' expects a genitive object. A form which is not syncretic between (at least) these two cases cannot occur in the place of *kogo*; this is illustrated in (21b), where the element putatively shared by the two verbal constituents is syncretic between accusative and nominative, rather than accusative and genitive. The well-known example (22) illustrates essentially the same phenomenon in German: the form *Frauen* 'women', which is fully syncretic with respect to case, simultaneously satisfies the accusative requirement of *findet* 'finds' and the dative requirement of *hilft* 'helps'. By contrast, this joint requirement is not satisfied either by *Männer*, which is accusative (among other cases) but not dative, or by *Männern*, which is dative but not accusative. The other two examples show that this phenomenon is not restricted to coordination. In (23), the syncretic form *was* 'what' simultaneously satisfies the constraint that the object of *gegeben* 'given' is accusative and that the subject of *ist* 'is' is nominative. Similarly, the extracted *who* in (24) seems to simultaneously bear the accusative case assigned by the preposition *of* and the nominative case of the subject of *should.*

Such examples were at one point considered as problematic not only for HPSG, but for unification-based theories in general (Ingria 1990). The reason is that, on the straightforward approach to case, they should all be ungrammatical. For example, in the case of (22a), the assignment of the accusative to the object of *findet* 'finds' should clash with the assignment of the dative to the object of *hilft* 'helps', as both objects are realised by the same noun *Frauen* 'women'. In other words, the attempt to unify accusative and dative should fail.

The solution first proposed by Levine et al. (2001: 207–208) is to enrich the *case* hierarchy in such a way that the unification of two different morphological cases does not necessarily result in failure. Specifically, assuming that nominative and accusative are structural cases in English, they propose the following part of the structural case hierarchy:[5]

(25)



Particular nominal forms are specified in the lexicon as either pure accusative (*p-acc*), pure nominative (*p-nom*) or syncretic between the two (*p-nom-acc*):

[5]Type names follow the convention in Daniels 2002, for increased uniformity with the remainder of this section.

(26) *he* [case *p-nom*]
him [case *p-acc*]
whom [case *p-acc*]
who [case *p-nom-acc*]
Robin [case *p-nom-acc*]

On the other hand, heads – or constraints of a case principle of the kind presented in the previous section – specify particular arguments as *nom* or *acc*. So, in the case of the parasitic gap example (24), the *acc* requirement associated with the preposition *of* and the *nom* requirement on the subject of *should* are not incompatible: their unification results in *p-nom-acc* and the shared dependent may be any form compatible with this case value, e.g., *who* (but not *whom*). Examples (21)–(23) can be handled in a similar way.

A situation often perceived as dual to such case neutrality, sometimes called 'case underspecification', occurs when a head specifies the case of its dependent disjunctively and may combine with a coordinate structure containing phrases in both cases, e.g.:

(27) Polish (Przepiórkowski 1999a: 175):

Dajcie	wina	i	całą	świnię!
give	wine.gen	and	whole.acc	pig.acc

'Serve (some) wine and a whole pig!'

(28) Russian (Levy 2001: 11):

Včera	vec'	den'	on	proždal	svoju	podrugu	Irinu	i
yesterday	all	day	he	expected	self's.acc	girlfriend.acc	Irina.acc	and

zvonka	ot	svoego	brata	Grigorija.
call.gen	from	self's	brother	Grigory

'Yesterday he waited all day for his girlfriend Irina and for a call from his brother Grigory.'

In Polish, the object of the verb *dajcie* 'give' is normally in the accusative, but may also be realised as the genitive, when its meaning is partitive; in (27), the object is a coordination of such a genitive noun *wina* '(some) wine' and the accusative *całą świnię* 'whole pig'. Similarly, according to Levy 2001, the Russian verb *proždal* 'awaited' may combine with accusative or genitive, and in (28) it happily combines with a coordinate phrase containing both.

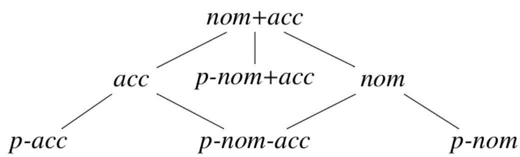
If such 'accusative and genitive' coordinate phrases bear case at all, the value of this grammatical category must be something like *acc+gen*. Note that this situation differs from case neutrality discussed above: a neutral case such as *p-acc-gen* intuitively corresponds to intersection: a nominal bearing this case is accusative and genitive at the same time. On the other hand, the intuition behind *acc+gen* is that of union: a (coordinated) nominal with this case value has accusative elements (perhaps purely accusative) and genitive elements (perhaps purely genitive). In particular, *acc+gen* coordinate phrases cannot fill either purely accusative position, or purely genitive positions, or positions in which both the accusative and the genitive is expected, as in (21) above.[6]

This duality is a feature of the categorial grammar approach to case and coordination of Bayer 1996 (see also Bayer & Johnson 1995) and the corresponding HPSG analyses were presented in Levy 2001 and Levy & Pollard 2002, as well as in Daniels 2002. As noted in Levy & Pollard 2002: 233, the two approaches are isomorphic. The main technical difference is that the relevant case hierarchies are construed outside of the usual HPSG type hierarchy in the approach of Levy 2001 and Levy & Pollard 2002, but they are fully integrated in the approach of Daniels 2002. For this reason, and also because it is the basis of some further HPSG work (e.g., Crysmann 2005), this latter approach is presented below.

Intuitively, just as the common subtype of *acc* and *nom*, i.e., *p-nom-acc* in (25), represents forms which are simultaneously accusative and nominative, the common supertype, i.e., *case*, which should perhaps be renamed to *nom+acc*, should represent coordinate structures involving nominative and accusative conjuncts. However, given that all objects are assumed to be sort-resolved in standard HPSG, saying that the case of a coordinate structure is *case* (or *nom+acc*) is paramount to saying that it is either *p-acc* (pure accusative), or *p-nom-acc* (syncretic nominative/accusative), or *p-nom* (pure nominative). One solution is to "make a simple change to the framework's foundational assumptions" (Sag 2003: 268) and to allow linguistic objects to bear non-maximal types. This is proposed and illustrated in detail in Sag 2003. A more conservative solution, proposed in Daniels 2002, is to add dedicated maximal types to all such non-maximal types; for example, the above hierarchy (25) becomes (29):

[6] To the extent that coordinate phrases may fill such a neutralised accusative/genitive position, all conjuncts – and, hence, also the whole coordinate structure – should bear the neutralised *p-acc-gen* case.

(29)



Apart from the trivial renaming of *case* to the more explicit *nom+acc*, a maximal type corresponding to this renamed non-maximal type is added here, namely, *p-nom+acc*.

Let us illustrate this approach with the two Polish examples (21a) and (27), repeated below as (30a) and (30b):

(30) a.

Kogo	Janek	lubi	a	Jerzy	nienawidzi?
who.ACC/GEN	Janek.NOM	likes(OBJ.ACC)	and	Jerzy.NOM	hates(OBJ.GEN)

'Who does Janek like and Jerzy hate?'

b.

Dajcie	wina	i	całą	świnię!
give	wine.GEN	and	whole.ACC	pig.ACC

'Serve (some) wine and a whole pig!'

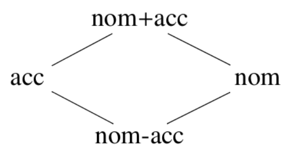
As these examples involve accusative and genitive, we will assume that the complete case hierarchy contains a subhierarchy such as (29) above, with all occurrences of *nom* replaced by *gen*.

First of all, heads subcategorise for (or relevant case principles specify) 'non-pure' cases, i.e., *acc*, *gen*, *gen+acc*, etc., but not *p-acc*, *p-gen*, *p-gen+acc*, etc. For example, *lubi* 'likes' and *nienawidzi* 'hates' in (30a) expect their objects to have the case values: *acc* and *gen*, respectively. Moreover, *dajcie* 'give' in (30b) specifies the case of its object as *gen+acc*. On the other hand, nominal dependents bear 'pure' cases. For example, *kogo* 'who' in (30a) is lexically specified as *p-gen-acc*. Similarly to the analysis of the English parasitic gap example above, this neutralised case is compatible with both specifications: *acc* and *gen*.

The analysis of (30b) is a little more complicated, as a new principle is needed to determine the case of a coordinate structure. The two conjuncts, *wina* 'wine' and *całą świnię* 'whole pig', have — by virtue of lexical specifications of their head nouns – the case values: *p-gen* and *p-acc*, respectively. Now, the case value of the coordination is determined as follows: take the 'non-pure' versions of the cases of all conjuncts (here: *gen* and *acc*), find their (lowest) common supertype (here: *gen+acc*), and assign to the coordinate structure the 'pure' type corresponding to this common supertype (here: *p-gen+acc*). This way the coordinate structure in (30b) ends up with the case value *p-gen+acc*, which is compatible with the *gen+acc* requirement posited by the verb *dajcie* (or by an appropriate principle of structural case assignment). Obviously, a purely accusative, purely genitive or accusative/genitive neutralised object would also satisfy this requirement.

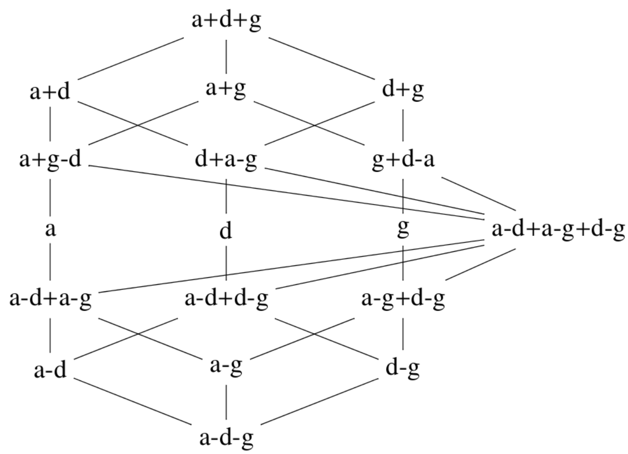
One often perceived – both within and outside of HPSG – problem with this approach is that it leads to very complex type hierarchies for *case* and rather inelegant constraints (Sag 2003: 272, Dalrymple et al. 2009: 63–66). Let us, following Daniels 2002, simplify the presentation of type hierarchies such as (29) in the following way, by removing all those 'pure' types which are only needed to represent some non-maximal types as maximal:

(31)



Hence, the above representation corresponds to 7 types shown explicitly in (29) (each non-maximal type in (31) has an additional *p-* type, while the maximal *nom-acc* in (31) is the same as *p-nom-acc* in (29)). What would a similar hierarchy for three morphological cases look like? Daniels 2002: 143 provides the following visualisation, involving 18 nodes, corresponding to 35 types in the full type hierarchy:

(32)



As mentioned in Levy & Pollard 2002: 225, the size of such a type hierarchy grows double exponentially with the number of grammatical cases, so it would already be next to impossible to visualise such a hierarchy for German, with its four cases, not to mention Polish with its 7 cases or Finno-Ugric languages with around 15 cases. And matters are further complicated by the fact that sometimes form syncretism simultaneously involves a number of grammatical categories, so perhaps such type hierarchies should combine case information with person, gender and number (Daniels 2002: 145, Crysmann 2005), and by the fact that coordinated elements may be specified for different categories (e.g., an NP specified for case may be coordinated with a sentence), in which case it is not clear what categories should be borne by the coordinate structure as a whole (see, e.g., the inconclusive fn.10 in Sag 2003: 277).

After early 2000s, such complex *case* hierarchies seem to have been abandoned in HPSG. A possible reason for this is the increasing popularity of 'conjunction reduction', i.e., ellipsis-based, accounts of various coordinate constructions, including unlike category coordination cases, of which the 'case underspecification' examples (27)–(28) may be seen as special cases.[7] Such 'conjunction reduction' accounts are usually formulated within the linearisation approach of Reape 1992; 1994 and Kathol 1995, and they have been claimed to deal with some of the cases discussed in this section, e.g., by Crysmann 2008, Beavers & Sag 2004, and Chaves 2006; 2008. However, such linearisation-based approaches to coordination have more recently come under attack: see Levine 2011 and Kubota & Levine 2015, as well as Yatabe 2012; 2016 and, especially, Yatabe & Tam 2018 for a defence. Hence, it is difficult to predict at the moment whether 'conjunction reduction' analyses will permanently remove the need for complex type hierarchies modelling neutralisation and underspecification. But even if they do, some of the examples given at the beginning of this section, namely, (23)–(24), demonstrated that feature neutrality is not limited to coordinate structures, but also occurs at least in free relatives and multiple gapping, so case hierarchies of the kind illustrated in (25), with separate types representing syncretic cases, are still needed in contemporary HPSG, regardless of the analysis of coordination; an example of a more recent analysis which does assume such a case hierarchy (to account for gapping and resumptive pronouns in Modern Standard Arabic) is Alotaibi & Borsley 2013.[8]

[7] Another HPSG approach to unlike category coordination which obviates the need for such complex hierarchies is that of Yatabe 2004, according to which the – perhaps disjunctive or underspecified – requirements of the head independently distribute to all conjuncts, in a manner similar to distributivity within coordinate structures assumed in LFG (Dalrymple & Kaplan 2000; Dalrymple et al. 2009; Przepiórkowski & Patejuk 2012).

[8] But see Crysmann 2017 for a reanalysis which does not need to refer to such a case hierarchy.

3 Other HPSG work on case

Apart from the two clearly identifiable strands of HPSG work described in the two preceding sections, there are also single papers concerned with various theoretical and implementational aspects of grammatical case. Of these, the report by Drellishak 2008 on modelling complex case phenomena in the Grammar Matrix (Bender et al. 2002) has the widest typological scope. It describes the treatment of various case systems in the multilingual platform for implementing HPSG grammars: not only the pure nominative-accusative, ergative-absolutive and tripartite systems, but also systems with various types of split ergativitity, systems – known from Austronesian languages, including Tagalog – in which case marking interact with focus marking, and so-called 'direct-inverse' systems, exemplified by Algonquian languages, in which case marking partially depends on the hierarchies – or scales – of nominal phrases, e.g., based on person and/or animacy. Similarly to the non-configurational case assignment principles discussed in Section (1) above, such systems are described – via constraints on specific lexical types – by specifying case values of elements on ARG-ST.

Two other works mentioned here are concerned with two very different aspects of case systems of particular languages. Ryu 2013 investigates the issue of case spreading from an argument of a verb to certain nominal dependents of this argument. He investigates the semantic relations that must hold between the two nominals for such 'case copying' to occur and proposes a repertoire of 16 semantic relations (collected in five coherent groups, further classified into two general classes) which make the spreading of the nominative possible, 10 of which (three of the five groups, one of the two classes) license the spreading of the accusative. On the syntactic side, the dependents of such nominal arguments are raised to become valency elements of the governing verbs. In particular, dependents of the subject are raised to the VAR|SUBJ list, resulting in multiple valency subjects. Configurational case assignment rules constrain the value of case of each valency subject to nominative, and of each valency complement – to accusative. The paper does not discuss the (im)possibility of formulating such case assignment rules non-configurationally, within local ARG-ST (or DEPS), but the challenge for the non-configurational case assignment seems to be the fact that multiple argument structure elements may correspond to valency subjects (and multiple – to valency complements), so – looking at the argument structure alone – it is not immediately clear how many initial elements of this list should be assigned the nominative case, and which final elements should get the accusative.

Finally, a very different aspect of Hungarian case is investigated in Thuilier 2011, namely, whether case affixes should be distinguished from postpositions and, if so, where to draw the line. In Hungarian, postpositions behave in some respect just like case affixes (e.g., they do not allow any intervening material between them and the nominal phrase) which led some researches to deny the existence of this distinction. Thuilier 2011 shows that, in this case, the traditional received wisdom is right, and that case affixes and postpositions differ in a number of morphological and syntactic ways. The proposed tests suggest that the essive element *ként*, normally considered to be a case affix, should be reanalysed as a postposition, thus establishing the number of Hungarian cases as 16. The resulting analysis of Hungarian case affixes and postpositions is couched within Sign-based Construction Grammar (Boas & Sag 2012).

References

Alotaibi, Mansour & Robert D. Borsley. 2013. Gaps and resumptive pronouns in Modern Standard Arabic. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 6–26. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2013/alotaibi-borsley.pdf, accessed 2018-2-25.

Andrews, Avery D. 1982. The representation of case in Modern Icelandic. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 427–503. Cambridge, MA/London: MIT Press.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72(3). 579–616.

Bayer, Samuel & Mark Johnson. 1995. Features and agreement. In Hans Uszkoreit (ed.), *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 70–76. Cambridge, MA: Association for Computational Linguistics.

Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar, Center for Computational Linguistics, Katholieke Universiteit Leuven*, 48–69. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2004/, accessed 2018-2-25.

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

Blake, Barry J. 1994. *Case*. Cambridge, UK: Cambridge University Press.

Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based Construction Grammar* (CSLI Lecture Notes 193). Stanford, CA: CSLI Publications.

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

Bratt, Elizabeth Owen. 1996. *Argument composition and the lexicon: lexical and periphrastic causatives in Korean*. Stanford University dissertation.

Butt, Miriam. 2006. *Theories of case*. Cambridge, UK: Cambridge University Press.

Butt, Miriam & Tracy Holloway King. 2003. Case systems: beyond structural distinctions. In Ellen Brandner & Heike Zinsmeister (eds.), *New perspectives on Case Theory* (CSLI Lecture Notes 156), 53–87. Stanford, CA: CSLI Publications.

Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. In Stefan Müller (ed.), *Proceedings of the HPSG 2006 conference*, 102–122. Stanford, CA: Center for the Study of Language & Information.

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

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

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use* (Convergence). New York/Westport, Connecticut/London: Praeger.

Crysmann, Berthold. 2005. Syncretism in German: A unified approach to underspecification, indeterminacy, and likeness of case. In Stefan Müller (ed.), *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, 91–107. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2005/, accessed 2018-2-25.

Crysmann, Berthold. 2008. An asymmetric theory of peripheral sharing in HPSG: conjunction reduction and coordination of unlikes. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2003, Vienna, Austria*, 47–62. Stanford, CA: CSLI Publications.

Crysmann, Berthold. 2017. Resumption and case: A new take on Modern Standard Arabic. In Stefan Müller (ed.), *Proceedings of the HPSG 2017 conference*, 120–240. Stanford, CA: Center for the Study of Language & Information.

Dalrymple, Mary & Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 4(74). 759–798.

Dalrymple, Mary, Tracy Holloway King & Louisa Sadler. 2009. Indeterminacy by underspecification. *Journal of Linguistics* 45. 31–68.

Daniels, Michael W. 2002. On a type-based analysis of feature neutrality and the coordination of unlikes. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 137–147. Stanford, CA: CSLI Publications.

Drellishak, Scott. 2008. Complex case phenomena in the Grammar Matrix. In Stefan Müller (ed.), *Proceedings of the HPSG 2008 conference*, 67–86. Stanford, CA: Center for the Study of Language & Information.

Dyła, Stefan. 1984. Across-the-board dependencies and case in Polish. *Linguistic Inquiry* 15(4). 701–705.

Grewendorf, Günther. 1994. Kohärente Infinitive und Inkorporation. In Anita Steube & Gerhild Zybatow (eds.), *Zur Satzwertigkeit von Infinitiven und Small Clauses* (Linguistische Arbeiten 315), 31–50. Tübingen: Max Niemeyer Verlag.

Groos, Anneke & Henk van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar. In A. Belletti, L. Brandi & L. Rizzi (eds.), *Theory of markedness in Generative Grammar*, 171–216. Pisa: Scuola Normale Superiore.

Grover, Claire. 1995. *Rethinking some empty categories: missing objects and parasitic gaps in HPSG*. Department of Language & Linguistics University of Essex dissertation.

Haider, Hubert. 1985. The case of German. In Jindřich Toman (ed.), *Studies in German grammar* (Studies in Generative Grammar 21), 23–64. Dordrecht: Foris Publications.

Haider, Hubert. 1990. Topicalization and other puzzles of German syntax. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and Barriers* (Linguistik Aktuell/Linguistics Today 5), 93–112. Amsterdam: John Benjamins Publishing Co.

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

Höhle, Tilman N. 1983. Topologische Felder. Köln, ms, Published as Höhle (2018).

Höhle, Tilman N. 2018. Topologische Felder. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5). Berlin: Language Science Press. First circulated as draft in 1983. DOI:10.5281/zenodo.1145680

Hukari, Thomas E. & Robert D. Levine. 1996. Phrase Structure Grammar: The next generation. *Journal of Linguistics* 32. 465–496.

Ingria, Robert J. P. 1990. The limits of unification. In *28th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 194–204. Pittsburgh, Pennsylvania: Association for Computational Linguistics.

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

Kiss, Tibor. 1994. Obligatory coherence: The structure of German modal verb constructions. USenglish. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 71–108. Stanford, CA: CSLI Publications.

Kubota, Yusuke & Robert Levine. 2015. Against ellipsis: Arguments for the direct licensing of 'noncanonical' coordinations. *Linguistics and Philosophy* 38. 521–576.

Levine, Robert D. 2011. Linearization and its discontents. In Stefan Müller (ed.), *Proceedings of the HPSG 2011 conference*, 126–146. Stanford, CA: Center for the Study of Language & Information.

Levine, Robert D., Thomas E. Hukari & Mike Calcagno. 2001. Parasitic gaps in English: Some overlooked cases and their theoretical consequences. In Peter W. Culicover & Paul M. Postal (eds.), *Parasitic gaps* (Current Studies in Linguistics 35), 181–222 Check. London, England/Cambridge, MA: MIT Press.

Levy, Roger. 2001. Feature indeterminacy and the coordination of unlikes in a totally well-typed HPSG. http://www.mit.edu/~rplevy/papers/feature-indet.pdf, accessed 2018-7-29. Unpublished manuscript.

Levy, Roger & Carl Pollard. 2002. Coordination and neutralization in HPSG. In Frank van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG 2001 conference*, 221–234. Stanford, CA: CSLI Publications.

Machicao y Priemer, Antonio & Paola Fritz-Huechante. 2018. Unmarked word order in Korean and Spanish psych-verbs: Interaction of case, theta-roles, and event structures in HPSG. http://phiz.c.u-tokyo.ac.jp/~hpsg2018/MachicaoYPriemer-Fritz.pdf, accessed 2018-7-25. Abstract for HPSG 2018.

Malchukov, Andrej & Andrew Spencer (eds.). 2009. *The Oxford handbook of case*. Oxford: Oxford University Press.

Maling, Joan. 1993. Of nominative and accusative: the hierarchical assignment of grammatical case in Finnish. In Anders Holmberg & Urpo Nikanne (eds.), *Case and other functional categories in Finnish syntax*, 51–76. Dordrecht: Mouton.

Maling, Joan. 2009. The case tier: A hierarchical approach to morphological case. In Andrej Malchukov & Andrew Spencer (eds.), 72–87. Oxford: Oxford University Press.

Malouf, Robert. 2000. A head-driven account of long-distance case assignment. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical interfaces in HPSG* (Studies in Constraint-Based Lexicalism 8), 201–214. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar. 1999a. *Lexical generalizations in the syntax of German non-finite constructions.* Tübingen: Eberhard-Karls-Universität dissertation.

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

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

Monachesi, Paola. 1999. *A lexical approach to Italian cliticization* (CSLI Lecture Notes 84). Stanford, CA: CSLI Publications.

Müller, Stefan. 1996. Yet another paper about partial verb phrase fronting in German. In *Proceedings of COLING 96*, 800–805. Copenhagen.

Müller, Stefan. 1997a. *Complement extraction lexical rules and argument attraction.* Research Report RR-97-08. Saarbrücken: Deutsches Forschungszentrum für Künstliche Intelligenz. Eine geringfügig andere Fassung ist in *Natural Language Processing and Speech Technology. Results of the 3rd KONVENS Conference, Bielefeld, October 1996* erschienen.

Müller, Stefan. 1997b. *Yet another paper about partial verb phrase fronting in German.* Research Report RR-97-07. Saarbrücken: DFKI. A shorter version appeared as Müller (1996).

Müller, Stefan. 2001. Case in German: Towards an HPSG analysis. In Walt Detmar Meurers & Tibor Kiss (eds.), *Constraint-based approaches to Germanic syntax* (Studies in Constraint-Based Lexicalism 7), 217–255. Stanford, CA: CSLI Publications.

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

Pollard, Carl J. 1994. Toward a unified account of passive in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 273–296. Stanford, CA: CSLI Publications.

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

Przepiórkowski, Adam. 1996a. Case assignment in Polish: towards an HPSG analysis. In Claire Grover & Enric Vallduví (eds.), *Studies in HPSG*, vol. 12 (Edinburgh Working Papers in Cognitive Science), 191–228. Centre for Cognitive Science, University of Edinburgh.

Przepiórkowski, Adam. 1996b. Non-configurational case assignment in HPSG. Paper delivered at the 3rd International Conference on HPSG, 20–22 May 1996, Marseille, France.

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

Przepiórkowski, Adam. 1999b. On case assignment and "adjuncts as complements". In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 231–245. Stanford, CA: CSLI Publications.

Przepiórkowski, Adam & Agnieszka Patejuk. 2012. On case assignment and the coordination of unlikes: The limits of distributive features. In Miriam Butt & Tracy Holloway King (eds.), *The proceedings of the LFG'12 conference*, 479–489. Stanford, CA: Center for the Study of Language & Information. http://cslipublications.stanford.edu/LFG/17/lfg12.html.

Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 62(4). 751–773.

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

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Ryu, Byong-Rae. 2013. Multiple case marking as case copying: A unified approach to multiple nominative and accusative constructions in Korean. In Stefan Müller (ed.), *Proceedings of the HPSG 2013 conference*, 182–202. Stanford, CA: Center for the Study of Language & Information.

Sag, Ivan A. 2003. Coordination and underspecification. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the HPSG 2002 conference*, 267–291. Stanford, CA: Center for the Study of Language & Information.

Sag, Ivan A., Lauri Karttunen & Jeffrey Goldberg. 1992. A lexical analysis of Icelandic case. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical matters* (CSLI Lecture Notes 24), chap. 11, 301–318. Stanford, CA: CSLI Publications.

Thuilier, Juliette. 2011. Case suffixes and pospositions in Hungarian. In Stefan Müller (ed.), *Proceedings of the HPSG 2011 conference*, 209–226. Stanford, CA: Center for the Study of Language & Information.

Yatabe, Shûichi. 2004. A comprehensive theory of coordination of unlikes. In Stefan Müller (ed.), *Proceedings of the HPSG 2004 conference*, 335–355. Stanford, CA: Center for the Study of Language & Information.

Yatabe, Shûichi. 2012. Comparison of the ellipsis-based theory of non-constituent coordination with its alternatives. In Stefan Müller (ed.), *Proceedings of the HPSG 2012 conference*, 454–474. Stanford, CA: Center for the Study of Language & Information.

Yatabe, Shûichi. 2016. Medial left-node raising in Japanese. In Doug Arnold, Miriam Butt, Berthold Crysmann & Tracy Holloway King (eds.), *The proceedings of the HeadLex16 conference*, 681–701. Stanford, CA: Center for the Study of Language & Information. http://cslipublications.stanford.edu/.

Yatabe, Shûichi & Wai Lok Tam. 2018. In defense of an HPSG-based theory of non-constituent coordination: A reply to Kubota and Levine. http://ling.auf.net/lingbuzz/003152. Unpublished manuscript, version of April 16, 2018.

Yip, Moira, Joan Maling & Ray S. Jackendoff. 1987. Case in tiers. *Language* 63(2). 217–250. DOI:10.2307/415655

Yoo, Eun Jung. 1993. Subcategorization and case marking in Korean. In Andreas Kathol & Carl J. Pollard (eds.), *Papers in syntax* (OSU Working Papers in Linguistics 42), 178–198. Ohio State University: Department of Linguistics.

Zaenen, Annie & Joan Maling. 1983. Passive and oblique case. In Lori S. Levin, Malka Rappaport & Annie Zaenen (eds.), *Papers in Lexical-Functional Grammar*, 159–191. Bloomington, IN: Indiana University Linguistics Club.

Zaenen, Annie, Joan Maling & Höskuldur Thráinsson. 1985. Case and grammatical functions: The Icelandic passive. *Natural Language and Linguistic Theory* 3(4). 441–483.

Chapter 8

Argument structure and linking

Stephen Wechsler
The University of Texas

Jean-Pierre Koenig
University at Buffalo

Anthony Davis
Southern Oregon University

What we cannot speak about, we must pass over in silence.

1 Introduction

When a verb or other predicator is composed with the phrases or pronominal affixes expressing its semantic arguments, the grammar must specify the mapping between the semantic participant roles and syntactic dependents of that verb. For example, the grammar of English indicates that the subject of *eat* fills the eater role and the object of *eat* fills the role of the thing eaten. In HPSG this mapping is usually broken down into two simpler mappings by positing an intermediate representation called ARG-ST ('argument structure'). The first mapping connects the participant roles within the semantic CONTENT with the elements of the ARG-ST feature; here we will call the theory of this mapping *linking theory* (see Section 4). The second mapping connects those ARG-ST list elements to the elements of the VAL lists, namely COMPS ('complements') and SUBJ ('subject'; or SPR, 'specifier'); we will refer to this second mapping as *argument realization* (see Section 2).[1] These two mappings are illustrated with the simplified lexical sign

[1]Some linguists, such as Levin & Rappaport Hovav (2005), use the term 'argument realization' more broadly, to encompass linking as well.

Stephen Wechsler, Jean-Pierre Koenig & Anthony Davis. 2018. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 135–175. Berlin: Language Science Press. DOI:??

for the verb *eat* in (1).

(1) Lexical sign for the verb *eat*

$$\begin{bmatrix} \textsc{phon} & \langle \text{eat} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \boxed{2} \rangle \end{bmatrix} \\ \textsc{arg-st} & \langle \boxed{1}\text{NP}_i \text{ , } \boxed{2}\text{NP}_j \rangle \\ \textsc{content} & \boldsymbol{eat}(i, j) \end{bmatrix}$$

In (1), 'NP' abbreviates a feature structure representing syntactic and semantic information about a nominal phrase. The variables i and j are the referential indices for the eater and eaten arguments, respectively, of the **eat** relation. The semantic information in NP_i semantically restricts the value or referent of i.

The ARG-ST feature plays an important role in HPSG grammatical theory. In addition to regulating the mapping from semantic arguments to grammatical relations, ARG-ST is the locus of the theories of anaphoric binding and other construal relations such as control and raising. (This chapter focuses on the function of ARG-ST in semantic mapping, with some discussion of binding and other construal relations only insofar as they interact with that mapping. A more detailed look at binding is presented in Chapter **??**.)

In HPSG, verb diathesis alternations, voice alternations, and derivational processes such as category conversions are all captured within the lexicon (see Section 5 and Chapter **??**). The different variants of a word are grammatically related either through lexical rules or by means of the lexical type hierarchy. HPSG grammars explicitly capture paradigmatic relations between word variants, making HPSG a *lexical approach to argument structure*, in the sense of Müller & Wechsler (2014a). This fundamental property of lexicalist theories contrasts with many transformational approaches, where such relationships are treated as syntagmatically related through operations on phrasal structures representing sentences and other syntactic constituents. Arguments for the lexical approach are reviewed in Section 8.

Within the HPSG framework presented here, we will formulate and address a number of empirical questions:

- We know that a verb's meaning influences its valence requirements, (via the ARG-ST list, on this theory). What are the principles governing the mapping from CONTENT to ARG-ST? Are some aspects of ARG-ST idiosyncratically stipulated for individual verbs? What aspects of the semantic CONTENT bear on the value of ARG-ST, and what aspects do not? (For example, what is the role of modality?)

- How are argument alternations defined with respect to our formal system? For each alternation we may ask which of the following it involves: a shuffling of the ARG-ST list; a change in the mapping from ARG-ST to VAL; or a change in the CONTENT, with a concomitant change in the ARG-ST?

These questions will be addressed below in the course of presenting the theory. We begin by considering ARG-ST itself (Section 2), followed by the mapping from ARG-ST to VAL (Section 3) and the mapping from CONTENT to ARG-ST (Section 4). The remaining sections address further issues relating to argument structure: the nature of argument alternations, extending the ARG-ST attribute to include additional elements, whether ARG-ST is a universal feature of languages, and a comparison of the lexicalist view of argument structure presented here with phrasal approaches.

2 The representation of argument structure in HPSG

In the earliest versions of HPSG, the selection of dependent phrases was specified in the SUBCAT feature of the head word (Pollard & Sag (1987), Pollard & Sag (1994: ch. 1–8)). The value of SUBCAT is a list of items, each of which corresponds to the SYNSEM value of a complement or subject. Following are SUBCAT features for an intransitive verb, a transitive verb, and a transitive verb with obligatory PP complement:

(2) a. *laugh*: [SUBCAT ⟨ NP ⟩]
    b. *eat*: [SUBCAT ⟨ NP, NP ⟩]
    c. *put*: [SUBCAT ⟨ NP, NP, PP ⟩]

Phrase structure rules in the form of immediate dominance schemata identify a certain daughter node as the head daughter (HEAD-DTR) and others as complement daughters (COMP-DTRS). In keeping with the *Subcategorization Principle*, here paraphrased from Pollard & Sag (1994: 34), list items are effectively 'cancelled' from the SUBCAT list as complement phrases are joined with the selecting head:

(3) Subcategorization Principle: In a headed phrase, the SUBCAT value of the HEAD-DTR ('head daughter') is the concatenation of the phrase's SUBCAT list with the list of SYNSEM values of the COMPS-DTRS ('complement daughters').

Phrasal positions are distinguished by their saturation level: 'VP' is defined as a verbal projection whose SUBCAT list contains a single item, corresponding to the subject; and 'S' is defined as a verbal projection whose SUBCAT list is empty.

The 'subject' of a verb, a distinguished dependent with respect to construal processes such as binding, control, and raising, was then defined as the first item in the SUBCAT list, hence the last item with which the verb combines. However, defining 'subject' as the last item to combine with the head proved inadequate (Pollard & Sag 1994: Ch. 9). There are many cases where the dependent displaying subject properties need not be the last item added to the head projection. For example, in German the construal subject is a nominal in nominative case (Reis 1982), but the language allows subjectless clauses containing only a dative or genitive non-subject NP. If that oblique NP is the only NP dependent to combine with the verb then it is *ipso facto* the last NP to combine.

Consequently, the SUBCAT list was split into two valence lists, a SUBJ list of length zero or one for subjects, and a COMPS list for complements. Nonetheless, certain grammatical phenomena, such as binding and other construal processes, must still be defined on a single list comprising both subject and complements (Manning & Sag 1999). Additionally, some syntactic arguments are unexpressed or realized by affixal pronouns, rather than as subject or complement phrases. The new list containing all the syntactic arguments of a predicator was named ARG-ST ('argument structure').

In clauses without implicit or affixal arguments, the ARG-ST is the concatenation of SUBJ and COMPS respectively. For example, the SUBCAT list for *put* in (2c) is replaced with the following:

(4) $$\begin{bmatrix} \text{PHON} & \langle \text{put} \rangle \\ \text{VALENCE} & \begin{bmatrix} \text{SUBJ} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \boxed{1}\text{NP}, \boxed{2}\text{NP}, \boxed{3}\text{PP} \rangle \end{bmatrix}$$

The idealization according to which ARG-ST is the concatenation of SUBJ and COMPS is canonized as the *Argument Realization Principle* (ARP) (Sag et al. 2003: PAGE). Exceptions to the ARP, that is, dissociations between VALENCE and ARG-ST, are discussed in Section 3.2 below.

A predicator's VALENCE lists indicate its requirements for syntactic concatenation with phrasal dependents (Section 3). ARG-ST, meanwhile, provides syntactic information about the expression of semantic roles and is related, via linking theory, to the lexical semantics of the word (Section 3.2). The ARG-ST list contains specifications for the union of the verb's syntactic local phrasal dependents (the subject and complements, whether they are semantic arguments, raised phrases, or expletives) and its syntactic arguments that are not realized locally, whether they are unbounded dependents, affixal, or unexpressed.

Figure 1 provides a schematic representation of linking and argument realization in HPSG. Linking principles govern the mapping of of participant roles in a predicator's CONTENT to syntactic arguments on ARG-ST. Argument realization is shown in this figure only for mapping to VAL; affixal and null arguments are not depicted. Here, the semantic roles are just arbitrary labels, but we discuss in Section (4) how they can be systematically related to lexical entailments of predicators. The ARG-ST and VAL lists in this figure contain only arguments linked to participant roles, but in Section ?? we note motivations for extending ARG-ST to include additional elements. And in Section (3), we examine cases where the relationship between ARG-ST and VAL violates the ARP.

Semantics (CONTENT)	$\begin{bmatrix} \textit{sem-rel-1} & \\ \text{SEM-ROLE-1} & \boxed{1} \\ \text{SEM-ROLE-2} & \boxed{2} \\ \text{SOA} & \begin{bmatrix} \textit{sem-rel-2} & \\ \text{SEM-ROLE-3} & \boxed{3} \\ \text{SEM-ROLE-4} & \boxed{4} \end{bmatrix} \end{bmatrix}$	Structured representation of semantic relations, their roles and the participants filling them
	$\Updownarrow$ *Linking principles*	
Argument structure (ARG-ST)	$\begin{bmatrix} \text{ARG-ST} \ \langle \text{XP}_1, \text{XP}_2, \ldots \text{XP}_n \rangle \end{bmatrix}$	List of syntactic arguments
	$\Updownarrow$ *Argument Realization Principle*	
Syntax (VALENCE)	$\begin{bmatrix} \text{SPR} & \langle\ \rangle \\ \text{SUBJ} & \langle \text{XP}_1 \rangle \\ \text{COMPS} & \langle \text{XP}_2, \ldots \text{XP}_n \rangle \end{bmatrix}$	Lists of locally realized dependents

Figure 1: How linking works in HPSG

3 Argument realization: The mapping from ARG-ST to VAL lists

3.1 Variation in the expression of arguments

The VAL feature is responsible for composing a verb with its phrasal dependents, but this is just one of the ways that semantic arguments of a verb are expressed in natural language. Semantic arguments can be expressed in various linguistic forms: as local syntactic dependents (SUBJ and COMPS), as affixes, or displaced in unbounded dependency constructions (SLASH).

Affixal arguments can be illustrated with the first person singular Spanish verb *hablo* ‘speak.1SG’, as in 5.

(5) a. Habl-o español.
speak-1SG Spanish
‘I speak Spanish.’

b. *hablo*‘speak.1SG’:

$$\left[\begin{array}{ll} \textsc{phon} & \langle \textsc{hablo} \rangle \\ \textsc{valence} & \left[\begin{array}{ll} \textsc{subj} & \langle\ \rangle \\ \textsc{comps} & \langle \boxed{2} \rangle \end{array}\right] \\ \textsc{arg-st} & \left\langle \textsc{np}{:}\left[\begin{array}{ll} \textit{ppro} & \\ \textsc{index} & \left[\begin{array}{ll} \textsc{pers} & \textit{1st} \\ \textsc{num} & \textit{sg} \end{array}\right] \end{array}\right], \boxed{2}\textsc{np} \right\rangle \end{array}\right]$$

The *-o* suffix contributes the first person singular pronominal subject content to the verb form (the morphological process is not shown here; see Chapter MORPHOLOGY-CHAPTER). The pronominal subject appears on the ARG-ST list and hence is subject to the binding theory. But it does not appear in SUBJ, if no subject NP appears in construction with the verb.

A lexical sign whose ARG-ST list that is just the concatenation of its SUBJ and COMPS lists conforms to the Argument Realization Principle (ARP); such signs are called *canonical signs* by Bouma et al. (2001). Non-canonical signs, which violate the ARP, have been approached in two ways. In one approach, a lexical rule takes as input a canonical entry and derives a non-canonical one by removing items from the VAL lists, while adding an affix or designating an item as an unbounded dependent by placement on the SLASH list. In the other approach, a feature of each ARG-ST list item specifies whether the item is subject to the ARP (hence mapped to a VAL list), or ignored by it (hence expressed in some other way). See the chapter on the lexicon for more detail and Miller & Sag (1997) for a treatment of French clitics as affixes.

A final case to consider is null anaphora, in which a semantic argument is simply left unexpressed and receives a definite pronoun-like interpretation. Japanese *mi-* 'see' is transitive but the object NP can be omitted as in (6).

(6) Naoki-ga mi-ta.
Naoki-NOM see-PAST
'Naoki saw it/him/her/*himself.'

Null anaphors of this kind typically arise in discourse contexts similar to those that license ordinary weak pronouns, and the unexpressed object often has the (Principle B) obviation effects characteristic of overt pronouns, as shown in (6). But HPSG eschews the use of silent formatives like 'small *pro*' when there is no evidence for such items, such as local interactions with the phrase structure. Instead, null anaphors of this kind are present in ARG-ST but absent from VAL lists. ARG-ST is directly linked to the semantic CONTENT and is the locus of binding theory, so the presence of a syntactic argument on the ARG-ST list but not a VAL list accounts for null anaphora. To account for obviation, the ARG-ST list item, when unexpressed, receives the binding feature of ordinary (non-reflexive) pronouns, usually *ppro*. This language-specific option can be captured in a general way by VAL and ARG-ST defaults in the lexical hierarchy for verbs.

3.2 The syntax of ARG-ST and its relation to VALENCE

The ARG-ST ordering represents a preliminary syntactic structuring of the set of argument roles. In that sense it functions as an interface between the lexical semantics of the verb, and the expressions of dependents as described in Section 3. Its role thus bears some relation to the initial stratum in Relational Grammar, *argument structure* (including intrinsic classifications) in LFG Lexical Mapping Theory, D-structure in Government/Binding theory, and the Merge positions of arguments in Minimalism, assuming in the last two cases the Uniform Thematic Alignment Hypothesis (Baker 1988) or something similar. However, it also differs from all of those in important ways.

Semantic constraints on ARG-ST are explored in Section 4 below. But ARG-ST is not only structured by semantic distinctions between the arguments but also by syntactic ones. Specifically, the list ordering represents relative syntactic *obliqueness* of arguments. The least oblique argument is the subject (SUBJ), followed by the complements (COMPS). Following Manning (1996) term arguments (direct arguments, i.e. objects) are assumed to be less oblique than 'oblique' arguments (adpositional and oblique case marked phrases), followed finally by predicate and clausal complements. The transitive ordering relation on the ARG-ST list is called *o-command* ('obliqueness command'): the subject list item o-commands those of the complements; an object list item o-commands those of any obliques; and so on.

Voice alternations like the passive, which are defined on the ARG-ST list, illustrate the ordering of terms before obliques on the ARG-ST list. Passivization alters the syntactic properties of ARG-ST list items: the initial item of the active, normally mapped to SUBJ of the active, is an oblique (*by* phrase) or unexpressed argument in the passive. Given that terms precede obliques in the list order, any term arguments must o-command the passive oblique, so passive effectively reorders the initial item in ARG-ST to a list position following any terms.

(7) a. Susan gave Mary a book.
b. Mary was given a book by Susan.

(8) a. *give* (active): $[\textsc{arg-st} < \textsc{np}_i, \textsc{np}_j, \textsc{np}_k >]$
b. *given* (passive): $[\textsc{arg-st} < \textsc{np}_j, \textsc{np}_k, \textsc{pp}[by]_i >]$

Relative obliqueness conditions a number of syntactic processes and phenomena, including anaphoric binding. The o-command relation replaces the c-command in the Principles A, B, and C of Chomsky's (1981) configurational theory of binding. For example, HPSG's Principle B states that an ordinary pronoun cannot be o-commanded by its coargument antecedent, which accounts for the pronoun obviation observed in the English sentence *Naoki*$_i$ *saw him*$_{*i/j}$, and also accounts for obviation in the Japanese sentence (6) above.

Relative obliqueness also conditions the accessibility hierarchy of Keenan & Comrie (1977), according to which a language allowing relativization of some type of dependent also allows relativization of any less oblique than it. Hence if a language has relative clauses at all, it has subject relatives; if it allows obliques to relativize then it also allows subject and object relatives; and so on. Similar implicational universals apply to verb agreement with subjects, objects, and obliques Greenberg (1966).

Returning now to argument realization, we saw above that the rules for the selection of the subject from among the verb's arguments are also stated on the ARG-ST list. In a 'canonical' realization the subject is the first list item, o-commanding all of its coarguments.

3.3 Syntactic ergativity

The autonomy of ARG-ST from the VAL lists is further illustrated by cross-linguistic variation in the mapping between them. As just noted, in English and many other languages the initial item in ARG-ST maps to the subject. However, languages with so-called *syntactically ergative* clauses have been analyzed as following a different mapping rule. Crucially, the ARG-ST ordering in those languages is still supported by independent evidence from properties such as binding and NP versus PP categorial status of arguments. Balinese (Austronesian), as analyzed by Wechsler & Arka (1998), is such a language. In the morphologically unmarked, and most common voice, called 'Objective voice' (OV), the subject is any term *except* the ARG-ST-initial one.

Balinese canonically has SVO order, regardless of the verb's voice form (Artawa 1994; Wechsler & Arka 1998). The preverbal NPs in (9) are the surface subjects and the postverbal ones are complements. When the verb appears in the unmarked OV verb, a non-initial term is the subject, as in (9a). But verbs in 'Agentive Voice' (AV) select as their subject the ARG-ST-initial item, as in (9b).

(9) a. Bawi adol ida.
pig OV.sell 3sg
'He/She sold a pig.'

b. Ida ng-adol bawi.
3sg AV-sell pig
'He/She sold a pig.'

A ditransitive verb such as the benefactive applied form of *beli* 'buy' in (10), has three term arguments on its ARG-ST list. The subject can be either term that is non-initial in ARG-ST:

(10) a. Potlote ento beli-ang=a I Wayan.
pencil-DEF that OV.buy-APPL=3 Art Wayan
'(s)he bought Wayan the pencil.'

b. I Wayan beli-ang=a potlote ento.
Art Wayan OV.buy-APPL=3 pencil-DEF that
'(s)he bought Wayan the pencil.'

Unlike the passive voice, which reorders the ARG-ST list, the Balinese OV does not affect ARG-ST list order. Thus the agent argument can bind a coargument reflexive pronoun (but not vice versa), regardless of whether the verb is in OV or AV form:

(11) a. Ida ny-ingakin ragan idane.
3sg AV-see self
'(s)he saw himself/herself'

b. Ragan idane cingakin ida.
self OV.see 3sg
'(s)he saw himself/herself'

The 'seer' argument o-commands the 'seen', with the AV versus OV voice forms regulating subject selection:

(12) Agentive Voice form of 'see':

$$\begin{bmatrix} \textsc{phon} & \langle \textit{nyinkagin} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{1} \rangle \\ \textsc{comps} & \langle \boxed{2} \rangle \end{bmatrix} \\ \textsc{arg-st} & \langle \boxed{1}\textsc{np}_i\ ,\ \boxed{2}\textsc{np}_j \rangle \\ \textsc{content} & \begin{bmatrix} \textit{see-rel} & \\ \textsc{seer} & i \\ \textsc{seen} & j \end{bmatrix} \end{bmatrix}$$

(13) Objective Voice form of 'see':

$$\begin{bmatrix} \textsc{phon} & \langle \textit{cinkagin} \rangle \\ \textsc{valence} & \begin{bmatrix} \textsc{subj} & \langle \boxed{2} \rangle \\ \textsc{comps} & \langle \boxed{1} \rangle \end{bmatrix} \\ \textsc{arg-st} & \langle \boxed{1}\textsc{np}_i\ ,\ \boxed{2}\textsc{np}_j \rangle \\ \textsc{content} & \begin{bmatrix} \textit{see-rel} & \\ \textsc{seer} & i \\ \textsc{seen} & j \end{bmatrix} \end{bmatrix}$$

Languages like Balinese illustrate the autonomy of ARG-ST. Although the agent binds the patient in both (11)a and b, the binding conditions cannot be stated directly on the thematic hierarchy. For example, in HPSG a raised argument appears on the ARG-ST list of the raising verb even though that verb assigns no thematic role to that list item. But a raised subject can bind a coargument reflexive in Balinese (this is comparable to English *John seems to himself to be ugly*). Anaphoric binding in Balinese raising constructions thus behaves as predicted by the ARG-ST based theory (Wechsler 1999). In conclusion, neither VAL nor CONTENT provides the right representation for defining binding conditions, but ARG-ST fills the bill.

Syntactically ergative languages that have been analyzed as using an alternative mapping between ARG-ST and VAL include Tagalog, Inuit, some Mayan languages, Chukchi, Toba Batak, Tsimshian languages, and Nadëb (Manning 1996; Manning & Sag 1999).

Interestingly, while the GB/Minimalist configurational binding theory may be defined on analogues of VAL or CONTENT, those theories lack any analogue of ARG-ST. This leads to special problems for such theories in accounting for binding in many Austronesian languages like Balinese. In transformational theories since Chomsky (1981), anaphoric binding conditions are usually stated with respect to the A-positions ('argument positions'). A-positions are analogous to HPSG VAL list items, with relative c-command in the configurational structure corresponding to relative list ordering in HPSG, in the simplest cases. Meanwhile, to account for data similar to (11), where agents asymmetrically bind patients, Austronesian languages like Balinese were said to define binding on the 'thematic structure' encoded in d-structure or Merge positions, where agents asymmetrically c-command patients regardless of their surface positions (Guilfoyle et al. 1992). But the interaction with raising shows that neither of those levels is appropriate as the locus of binding theory (Wechsler 1999).[2]

3.4 Symmetrical objects

We have thus far tacitly assumed a total ordering of elements on the ARG-ST list, but Ackerman et al. (2018), Ackerman et al. (2017) propose a partial ordering for certain so-called 'symmetrical object' languages. In Moro (Kordofanian), the two term complements of a ditransitive verb have exactly the same object properties. Relative linear order of the theme and goal arguments is free, as shown by the two translations of (14) (from Ackerman et al. 2017: 9):

(14) é-g-a-natʃ-ó óráŋ ŋeɾá
1SB.SM-CLg-MAIN-give-PFV CLg.man CLŋ.girl
'I gave the girl to the man.' / 'I gave the man to the girl.'

More generally, the two objects have identical object properties with respect to occurrence in post-predicate position, case marking, realization by an object marker, and ability to undergo passivization (Ackerman et al. 2017: 9).

[2] To account for (11b) under the configurational binding theory, the subject position must be an A-bar position; but to account for binding by a raised subject, it must be an A-position. See Wechsler (1999).

Ackerman et al. (2017) propose that the two objects are unordered on the ARG-ST list. This allows for two different mappings to the COMPS list, as shown here:

(15) a. Goal argument as primary object:

$$\begin{bmatrix} \textsc{valence} & \begin{bmatrix} \textsc{subj} & \left\langle \boxed{1} \right\rangle \\ \textsc{comps} & \left\langle \boxed{2}, \boxed{3} \right\rangle \end{bmatrix} \\ \textsc{arg-st} & \left\langle \boxed{1}\textsc{np}_i, \left\{ \boxed{2}\textsc{np}_j, \boxed{3}\textsc{np}_k \right\} \right\rangle \\ \textsc{content} & \begin{bmatrix} \textit{give-rel} & \\ \textsc{agent} & i \\ \textsc{goal} & j \\ \textsc{theme} & k \end{bmatrix} \end{bmatrix}$$

b. Theme argument as primary object:

$$\begin{bmatrix} \textsc{valence} & \begin{bmatrix} \textsc{subj} & \left\langle \boxed{1} \right\rangle \\ \textsc{comps} & \left\langle \boxed{3}, \boxed{2} \right\rangle \end{bmatrix} \\ \textsc{arg-st} & \left\langle \boxed{1}\textsc{np}_i, \left\{ \boxed{2}\textsc{np}_j, \boxed{3}\textsc{np}_k \right\} \right\rangle \\ \textsc{content} & \begin{bmatrix} \textit{give-rel} & \\ \textsc{agent} & i \\ \textsc{goal} & j \\ \textsc{theme} & k \end{bmatrix} \end{bmatrix}$$

The primary object properties, which are associated with the initial term argument of COMPS, can go with either the goal or theme argument.

To summarize this section, while the relationship between ARG-ST, SUBJ, and COMPS lists was originally conceived as a straightforward one, enabling binding principles to maintain their simple form by defining ARG-ST as the concatenation of the other two, the relationship was soon loosened. Non-canonical relationships between ARG-ST and the VAL lists are invoked in accounts of several core syntactic phenomena. Arguments not realized overtly in their canonical positions, due to extraction, cliticization, or pro-drop (null anaphora), appear on ARG-ST but not in any VAL list. Accounts of syntactic ergativity in HPSG involve variations in the mapping between ARG-ST and VAL lists; in particular, the element of SUBJ is not, in such languages, the first element of ARG-ST. Modifications of ARG-ST play a role in treatments of passivization, where its expected first element is suppressed, and in languages with multiple, symmetric objects, where a partial rather than total ordering of ARG-ST elements has been postulated. Thus ARG-ST has now acquired an autonomous status within HPSG, and is not merely a predictable rearrangement of information present elsewhere in lexical entries.

4 Linking: the mapping between semantics and ARG-ST

4.1 HPSG approaches to linking

The term *linking* refers to the mapping specified in a lexical entry between participant roles in the semantics and their syntactic representations on the ARG-ST list. Early HPSG grammars stipulated the linking of each verb: semantic CONTENT values with predicator-specific attributes like DEVOURER and DEVOURED were mapped to the subject and object, respectively, of the verb *devour*. But linking observes prevailing patterns, e.g. if one argument of a transitive verb in active voice has an agentive role, it will map to the subject, not the object. Thus these early accounts were unsatisfying, as they lead to purely stipulative accounts of linking, specified verb by verb. Beginning with Wechsler (1995b) and Davis (1996), researchers formulated linking principles stated on more general semantic properties holding across verbs.

Within the history of linguistics there have been three general approaches to modeling the lexico-semantic side of linking: thematic role types (Pāṇini ca 400 B.C., Fillmore 1968); lexical decomposition (Foley & Van Valin 1984; Rappaport Hovav & Levin 1998); and the proto-roles approach (Dowty 1991). In developing linking theories within the HPSG framework Wechsler (1995b) and Davis (1996) employed a kind of lexical decomposition that also incorporated some elements of the proto-roles approach. The reasons for preferring this over the alternatives are discussed in Section 4.4 below.

Wechsler's 1995b linking theory constrains the relative order of pairs arguments on the ARG-ST list according to semantic relations entailed between them. For example, his *notion rule* states that if one participant in an event is entailed to have a mental notion of another, then the first must precede the second on the ARG-ST list. The *conceive-pred* type is defined by the following type declaration (based on Wechsler (1995b: 127), with formal details adjusted for consistency with current usage):

$$\text{(16)}\quad \textit{conceive-pred:}\ \begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_i, \text{NP}_j \rangle \\ \text{CONTENT} & \begin{bmatrix} \textit{conceive-rel} & \\ \text{CONCEIVER} & i \\ \text{CONCEIVED} & j \end{bmatrix} \end{bmatrix}$$

This accounts for a host of linking facts in verbs as varied as *like*, *enjoy*, *invent*, *claim*, and *murder*, assuming these verbs belong to the type *conceive-pred*. It explains the well known contrast between experiencer-subject *fear* and experiencer-object *frighten* verbs: *fear* entails that its subject has some notion of its object, so

The tourists feared the lumberjacks entails that the tourists are aware of the lumberjacks. But the object of *frighten* need not have a notion of its subject: in *The lumberjacks frightened the tourists (by cutting down a large tree that crashed right in front of them*, the tourists may not be aware of the lumberjacks' existence.

Two other linking rules appeared in Wechsler (1995). One stated that 'affected themes,' that is, participants that are entailed to undergo a change, map to the object, rather than subject, of a transitive verb. Another pertained to stative transitive verbs entailing a part-whole relation between the two participants, such as *include* and *contain*: the whole maps to the subject and the part to the object.

The linking constraints do not rely on a total ordering of thematic roles, nor on an exhaustive asssignment of thematic role types to every semantic role in a predicator. Instead, a small set of partial orderings of semantic roles, based on lexical entailments, suffices to account for the linking patterns of a wide range of verbs. This insight was adopted in a slightly different guise in work by Davis (1996; 2001) and Davis & Koenig (2000), who develop a more elaborated representation of lexical semantics, with which simple linking constraints can be stated. The essence of this approach is to posit a small number of dyadic semantic relations such as *act-und-rel* ('actor-undergoer relation') with attributes ACT(OR) and UND(ERGOER) that serve as intermediaries between semantic roles and syntactic arguments (akin to the notion of Generalized Semantic Roles discussed in Van Valin 1999).

What are the truth conditions of *act-und-rel*? Following Fillmore (1977), Dowty (1991), and Wechsler (1995b), Davis and Koenig note that many of the pertinent lexical entailments come in related pairs. For instance, one of Dowty's entailments is that one participant causally affects another, and of course the other is entailed to be causally affected. Another involves the entailments in Wechsler's notion rule (16); one participant is entailed to have a notion of another. These entailments of paired participant types characterize classes of verbs (or other predicators), and can then be naturally represented as dyadic relations in CONTENT. Collecting those entailments we arrive at a disjunctive statement of truth conditions:

(17) **act-und-rel**(x, y) is true iff x causes a change in y, or x has a notion of y.

We can designate the x participant in the pair as the value of ACTOR (or ACT) and y as the value of UNDERGOER (or UND), in a relation of type *act-und-rel*. Semantic arguments that are ACTOR or UNDERGOER will then bear at least one of the entailments characteristic of ACTORS or UNDERGOERS (Davis & Koenig 2000: 72). This then simplifies the statement of linking constraints for all of these paired participant types. Davis (1996) and Koenig & Davis (2001) argue that this obviates counting the relative number of proto-agent and proto-patient entailments, as advocated by Dowty (1991).

The linking constraints 18 and 19 state that a verb whose semantic CONTENT is of type *act-und-rel* will be constrained to link the ACT participant to the the first element of the verb's ARG-ST list (its subject), and the UND participant to the second element of the verb's ARG-ST list (this is analogous to Wechsler's constraints based on partial orderings).

These linking constraints can be viewed as parts of the definition of lexical types, as in Davis (2001), where (18) defines a particular class of lexemes (or words).[3]

(18) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix}\textsc{actor} & \boxed{1}\end{bmatrix} \\ \textsc{arg-st} & \left\langle \textsc{np}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$$

(19) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix}\textsc{undergoer} & \boxed{2}\end{bmatrix} \\ \textsc{arg-st} & \left\langle \ldots, \textsc{np}_{\boxed{2}}, \ldots \right\rangle \end{bmatrix}$$

(20) $$\begin{bmatrix} \textsc{content}|\textsc{key} & \begin{bmatrix} \textit{cause-possess-rel} \\ \textsc{soa} \begin{bmatrix}\textsc{actor} & \boxed{3}\end{bmatrix} \end{bmatrix} \\ \textsc{arg-st} & \left\langle \textit{synsem} \right\rangle \oplus \left\langle \textsc{np}_{\boxed{3}}, \ldots \right\rangle \end{bmatrix}$$

The first constraint, in (18), links the value of ACT (when not embedded within another attribute) to the first element of ARG-ST. The second, in (19), merely links the value of UND (again, when not embedded within another attribute) to some NP on ARG-ST. Given this understanding of how the values of ACTOR and UNDERGOER are determined, these constraints cover the linking patterns of a wide range of transitive verbs: *throw* (ACT causes motion of UND), *slice* (ACT causes change of state in UND), *frighten* (ACT causes emotion in UND), *imagine* (ACT has a notion of UND), *traverse* (ACT "measures out" UND as an incrememtnal theme), and *outnumber* (ACT is superior to UND on a scale).

[3] Alternatively, (18) (and other linking constraints) can be recast as implicational constraints on lexemes or words (Koenig & Davis 2003). (i) is an implicational constraint indicating that a word whose semantic content includes an ACTOR role must map that role to the initial item in the ARG-ST list.

(i) $$\begin{bmatrix} \textsc{content}|\textsc{key} \begin{bmatrix}\textsc{actor} & \boxed{1}\end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \textsc{arg-st} \left\langle \textsc{np}_{\boxed{1}}, \ldots \right\rangle \end{bmatrix}$$

The third constraint, in (20), links the value of an ACT attribute embedded within a SOA attribute to an NP that is second on ARG-ST. This constraint accounts for the linking of the (primary) object of ditransitives. In English, these verbs (*give*, *hand*, *send*, *earn*, *owe*, etc.) involve (prospective) causing of possession (Pinker 1989; Goldberg 1995), and the possessor is represented as the value of the embedded ACT in (20). There could be additional constraints of a similar form in langages with a wider range of ditransitive constructions; conversely, such a constraint might be absent in languages that lack ditransitives entirely. As mentioned earlier in this section, the range of subcategorization options varies somewhat from one language to another.

The KEY attribute in (18) – (20) also requires explanation. The formulation of linking constraints here employs the architecture used in Koenig & Davis (2006), in which the semantics represented in CONTENT values is expressed as a set of *elementary predications*, formalized within Minimal Recursion Semantics (Copestake et al. 2001; 2005). Each elementary predication is a simple relation, but the relationships among them may be left unspecified. For linking, one of the elementary predications is designated the KEY, and it serves as the locus of linking. This allows us to indicate the linking of participants that play multiple roles in the denoted situation. The KEY selects one relation as the "focal point," and the other elementary predications are then irrelevant as far as linking is concerned. The choice of KEY then becomes an issue demanding consideration; we will see in the discussion of argument alternations in Section 5 how this choice might account for some alternation phenomena.

Note too that these linking constraints are treated as constraints on classes in the lexical hierarchy (see Chapter **??**). One consequence of this fact merits brief mention. Constraint (19), which links the value of UND to some NP on ARG-ST, is a specification of one class of verbs. Not all verbs (and certainly not all other predicators, such as nominalizations) with a CONTENT value containing an UND value realize it as an NP. Verbs obeying this constraint include the transitive verbs noted above, and intransitive "unaccusative" verbs such as *fall* and *persist*. But some verbs with both ACT and UND attributes in their CONTENT are intransitive, such as *impinge (on)*, *prevail (on*, and *tinker (with)*. Interactions with other constraints, such as the requirement that verbs (in English, at least) have an NP subject, determine the range of observed linking patterns.

These linking constraints also assume that the proto-role attributes ACTOR, UNDERGOER, and SOA are appropriately matched to entailments, as described above. Other formulations are possible, such as that of Koenig & Davis (2003), where the participant roles pertinent to each lexical entailment are represented in CONTENT by corresponding, distinct attributes.

In addition to the linking constraints, there may be some very general well-formedness conditions on linking. We rarely find verbs that obligatorily map one semantic role to two distinct members of the ARG-ST list that are both expressed overtly. A verb meaning 'eat', but with that disallowed property, could appear in a ditransitive sentence like (4.1), with the meaning that Pat ate dinner, and his dinner was a large steak.

(21) *Pat ate dinner a large steak.

Typically semantic arguments map to at most one (overtly expressed) ARG-ST list item (Davis 2001: 262-268).

4.2 Linking oblique arguments

In this section we discuss linking of oblique arguments, that is, PP's and oblique case marked NP's. In some instances, a verb's selection of a particular preposition appears at least partly arbitrary; it is hard to explain why we *hanker after* and *yearn for*, but we don't **yearn after*. In these cases, the choice of preposition may be stipulated by the individual lexical entry. But as Gawron (1986) and Wechsler (1995a) have shown, many prepositions are semantically meaningful. *For* in the above-mentioned cases, and in *look for*, *wait for*, and *aim for* is surely not a lexical accident. And in the cases like *cut with*, *with* is used in an instrumental sense, denoting a *use-rel* relation, as with verbs that either allow (*eat*) or require (*cut*) an instrument. Davis (1996; 2001) adopts the position of Gawron and Wechsler in his treatment of linking to PPs As an example of this kind of account, the linking type in (22) characterizes a verb selecting a *with*-PP. The PP argument is linked from the RELS list rather the KEY.

(22)
$$\left[\begin{array}{ll} \text{CONTENT} & \left[\begin{array}{ll} \text{KEY} & \boxed{1} \\ \text{RELATIONS} & \left\langle \boxed{1}, \boxed{2}\left[\begin{array}{ll} \multicolumn{2}{l}{\textit{use-rel}} \\ \text{ACT} & a \\ \text{UND} & u \\ \text{SOA} & s \end{array}\right], \ldots \right\rangle \end{array}\right] \\ \text{ARG-ST} & \left\langle \ldots, \text{PP}_{with}{:}\boxed{2} \ldots \right\rangle \end{array}\right]$$

Apart from the details of individual linking constraints, we have endeavored here to describe how linking can be modeled in HPSG using the same kinds of constraints used ubiquitously in the framework. Within the hierarchical lexicon, constraints between semantically defined classes and syntactically defined ones, can furnish an account of linking patterns, and there is no resort to additional mechanisms such as a thematic hierarchy or numerical comparison of entailments.

4.3 To what extent does meaning predict linking?

The framework outlined above allows us to address the following question: how much of linking is strictly determined by semantic factors, and how much is left open to lexically arbitrary subcategorization specifications, or perhaps subject to other factors?

Subcategorization— the position and nature of ARG-ST elements, in HPSG terms— is evidently driven to a great extent by semantics, but debate continues about how much, and which components of semantics are involved. Views have ranged from the strict, highly constrained relationship in which lexical semantics essentially determines syntactic argument structure, to a looser one in which some elements of subcategorization may be stipulated. Among the first camp are those who espouse the Uniformity of Theta Assignment Hypothesis proposed in Baker (1988: 46) or Baker (1997), which maintains that "identical thematic relationships between items are represented by identical structural relationships" in the syntax. With regard to the source of diathesis alternations, Levin (1993: 12-13) notes that "studies of these properties suggest that argument structures might in turn be derivable to a large extent from the meaning of words", and accordingly "pursues the hypothesis of semantic determinism seriously to see just how far it can be taken."

Others, including Pollard & Sag (1987) (Section 5.3) and Davis (2001) (Section 5.1), have expressed caution, pointing out cases where subcategorization and diathesis alternations seem to be at least partly arbitrary. Pollard & Sag note contrasts like these:

(23) a. Sandy spared/*deprived Kim a second helping.
 b. Sandy *spared/deprived Kim of a second helping. (Pollard & Sag 1987: ex. 214–215)

And Davis provides these pairs of semantically similar verbs with differing subcategorization requirements:

(24) a. Few passengers waited for/awaited the train.
 b. Homer opted for/chose a chocolate frosted donut.
 c. The music grated on/irritated the critics. (Davis 2001: ex. 5.4)

Other cases where argument structure seems not to mirror semantics precisely include raising constructions, in which one of a verb's direct arguments bears no semantic role to it at all. Similarly, overt expletive arguments cannot be seen as deriving from some participant role in a predicator's semantics. Like the examples above, these phenomena suggest that some aspects of subcategorization are specified independently of semantics.

Another point against strict semantic determination of argument structure comes from cross-linguistic observations of subcategorization possibilities. It is evident, for example, that not all languages display the same range of direct argument mappings. Some lack ditransitive constructions entirely (Halkomelem), some allow them across a limited semantic range (English), some quite generally (Georgian), and a few permit tritransitives (Kinyarwanda and Moro). Gerdts (1992) surveys about twenty languages and describes consistent patterns like these. The range of phenomena such as causative and applicative formation in a language is constrained by what she terms its "relational profile;" this includes, in HPSG terms, the number of direct NP arguments permitted on its ARG-ST lists. Again, it is unclear that underlying semantic differences across languages in the semantics of verbs meaning 'give' or 'write' would be responsible for these general patterns.

Summarizing, there is much evidence tempting us to derive the contents of ARG-ST solely from lexical semantics. If this ultimately proves feasible, then ARG-ST serves more as a convenient interface notion with little possibility of independently expressing strictly syntactic aspects of subcategorization. This view, however satisfying it might be, does not accord with our current best understanding of the syntactic and semantic evidence. In the following sections we delve into some of the nuances that make linking more than a simple rendering of lexical semantics. We begin by noting a point on which HPSG accounts of linking differ from many others— the absence of traditional thematic roles.

4.4 HPSG and thematic roles

The ARG-ST list constitutes the syntactic side of the mapping between semantic roles and syntactic dependents. As ARG-ST is merely an ordered list of arguments, without any semantic "labels," it contains no counterparts to thematic roles, such as AGENT, PATIENT, THEME, or GOAL. Thematic roles like these, however, have been a mainstay of linking in generative grammar since Fillmore (1968) and have antecedents going back to (Pāṇini. Ranking them in a *thematic hierarchy*, and labeling each of a predicator's semantic roles with a unique thematic role, then yields an ordering of roles analogous to the ordering on the ARG-ST list. Indeed, it would not be difficult to import this kind of system into HPSG, as a means of determining the order of elements on the ARG-ST list. However, HPSG researchers have generally avoided using a thematic hierarchy, for reasons we now briefly set out.

Fillmore (1968) and many others thereafter have posited a small set of disjoint, thematic roles, with each semantic role of a predicator assigned exactly one thematic role. Thematic hierarchies depend on these properties for a consistent linking theory. But they do not hold up well to formal scrutiny. Jackendoff (1987) and Dowty (1991) note (from somewhat different perspectives) that numerous verbs have arguments not easily assigned a thematic role from the typically posited inventory (e.g., the objects of *risk*, *blame*, and *avoid*), that more than one argument might sensibly be assigned the same role (e.g., the subjects and objects of *resemble*, *border*, and some alternants of commercial transaction verbs), and that multiple roles can be sensibly assigned to a single argument (the subjects of verbs of volitional motion are like both an AGENT and a THEME). In addition, consensus on the inventory of thematic roles has proven elusive, and some, notoriously THEME, have resisted clear definition. Work in formal semantics, including Ladusaw & Dowty (1988), Dowty (1989), Landman (2000), and Schein (2002), casts doubt on the prospects of assigning formally defined thematic roles to all of a predicator's arguments, at least in a manner that would allow them to play a crucial part in linking. Thematic role types seem to pose problems, and there are alternatives that avoids those problems. As Carlson (1998) notes about thematic roles, "It is easy to conceive of how to write a lexicon, a syntax, a morphology, a semantics, or a pragmatics without them."

4.5 CONTENT decomposition and ARG-ST

Instead of thematic role types, lexical decomposition is typically used in HPSG to model the semantic side of the linking relation. The word meaning represented by the CONTENT value is decomposed into a set of elementary predications that share arguments, as described in Section 4 above. Lexical decompositions cannot be directly observed, but the decompositions are justified indirectly by the roles they play in the grammar. Decompositions play a role in at least the following processes:

- *Linking*. As described in Section 4, linking constraints are stated on semantic relations like *act-und-rel* ('actor-undergoer relation'), so those relations must be called out in the CONTENT field.

- *Sublexical scope.* Certain modifiers can scope over a part of the situation denoted by a verb (Dowty 1979).

  (25) John sold the car, and then he bought it again.

  In this sentence the adverb *again* either adds the presupposition that John bought it before, or, in the more probable interpretation, it adds the presupposition that *the result of buying the car* obtained previously. The result of buying a car is owning it, so this sentence presupposes that John previously owned the car. Thus the decomposition of the verb *buy* in (36) below includes the *possess-rel* ('possession relation') holding between the buyer and the goods. This is available for modification by adverbials like *again*.

- *Argument alternations.* Some argument alternations can be modeled as highlighting of different portions of a single lexical decomposition. See Section 5.

In general, sublexical decompositions are included in the CONTENT field only insofar as they are visible to the grammar for processes like these.

The ARG-ST feature lies at the syntax side of the linking relation. Much like the CONTENT field, the ARG-ST items are justified only insofar as they are visible to the syntax. Many ARG-ST list items are obviously justified by being explicitly expressed as subject and complement phrases or as affixal pronouns. Certain implicit arguments appear if they are subject to the binding theory as applied to the ARG-ST list (as discussed in Section 3.1 above).

Implicit arguments can also participate in the syntax, and therefore appear on the ARG-ST list, by acting as controllers of adjunct clauses. For example, English rationale clauses like the infinitival phrase in (26a) are controlled by the agent argument in the clause, *the hunter* in this example. The implicit agent of a short passive can likewise control the rationale clause as shown in(26b). But the middle in (26c) lacks an implicit agent that is capable of controlling, even though native speakers assume that some agent must have caused the gun to load. This contrast was observed by Keyser & Roeper (1984) and confirmed in experimental work by Mauner & Koenig (2000).

(26) a. The shotgun was loaded quietly by the hunter to avoid the possibility of frightening off the deer.
     b. The shotgun was loaded quietly to avoid the possibility of frightening off the deer.

c. *The shotgun had loaded quietly to avoid the possibility of frightening off the deer.

If the syntax of control is specified such that the controller of the rationale clause is an (agent) argument on the ARG-ST list of the verb, then this contrast is captured by assuming that the agent appears on the ARG-ST list of the passive verb but not the middle.

4.6 Modal transparency

Another observation concerning lexical entailments and linking was developed by Koenig & Davis (2001), who point out that linking appears to ignore modal elements of lexical semantics, even when those elements invalidate entailments (expanding on an observation implicit in Goldberg 1995). For instance, there are various English verbs displaying linking patterns like the ditransitive verbs of possession transfer *give* and *hand*, but which denote situations in which the transfer need not, or does not, take place. Thus, *offer* describes a situation where the transferor is willing to effect the transfer, *owe* one in which the transferor should effect the transfer but has not yet, *promise* describes a situation where the transferor commits to effect the transfer, and *deny* one in which the transferor does not effect the contemplated transfer. Koenig and Davis argue that modal elements should be clearly separated in CONTENT values from the representations of predicators and their arguments. (27) exemplifies this factoring out of sublexical modal information from core situational information. This pattern of linking functioning independently of sublexical modal information applies not only to these ditransitive cases, but also to verbs involving possession (cf. *own* and *obtain*, vs. *lack*, *covet* and *lose*), perception (*see* vs. *ignore* and *overlook*), and carrying out an action (*manage* vs. *fail* and *try*). Whatever the role of lexical entailments in linking, then, the modal information should be factored out, since the entailments canonically driving, e.g., the ditransitive linking patterns of verbs like *give* and *hand*, do not hold of *offer*, *owe*, or *deny*. The constraints in (18)-(20) need only been minimally altered to target the value of SIT-CORE within the representation of relation.

(27) The lexical semantic representation of *promise* (Koenig & Davis 2001: 101)

$$\begin{bmatrix} \textit{promise-sem} \wedge \textit{cause-possess-sem} & \\ \text{SIT-CORE} & \boxed{3}\begin{bmatrix}\textit{cause-possess-rel} & \\ \text{ACTOR} & \boxed{1} \\ \text{UNDERGOER} & \boxed{2} \\ \text{SOA} & \begin{bmatrix} \text{sit-core} & \boxed{5}\begin{bmatrix}\textit{have-rel} & \\ \text{ACTOR} & \boxed{2}\\ \text{UNDERGOER} & \boxed{4}\end{bmatrix}\end{bmatrix}\end{bmatrix} \\ \text{MODAL-BASE} & \left\langle \begin{bmatrix}\textit{deontic-mb} \wedge \textit{condit-satis-mb} & \\ \text{SOA} & \boxed{3}\end{bmatrix}\right\rangle \end{bmatrix}$$

5 The semantics and linking of argument alternations

A verb can often occur in varied syntactic contexts, as *find* does in (28); these are termed *valence alternations* or *diathesis alternations*, in reference to their different argument structuress. Levin (1993) lists around 50 kinds of alternations in English, and there are still more, including the alternation illustrated in (28).

(28) a. I found that the chair was comfortable
b. I found the chair to be comfortable
c. I found the chair comfortable

Another well studied alternation, the locative alternation, is exemplified by the two uses of *spray* in (29).

(29) a. $spray_{loc}$: Joan sprayed the paint onto the statue.
b. $spray_{with}$: Joan sprayed the statue with paint.

It is typically assumed that these two different uses of *spray* in (29) have slightly different meanings, with the statue being in some sense more affected in the *with* alternant. This exemplifies the “holistic” effect of direct objecthood, which we will return to. Here, we will examine how semantic differences between alternant relate to their linking patterns. The semantic side of linking has often been devised with an eye to syntax (e.g., Pinker (1989), and see Koenig & Davis (2006) for more examples). There is a risk of stipulation here, without independent evidence for these semantic differences. In the case of locative alternations, though, the meaning difference between (29a) and (29b) is easily stated (and Pinker had the right intuition), as (29b) entails (29a), bot not conversely. Informally, (29a) describes a particular kind of caused motion situation, while (29b) describes a situation in which this kind of caused motion additionally results in a caused change of state. The difference is depicted in the two structures in (30).

(30) a. CAUSE (JOAN, GO (PAINT, TO (STATUE)))

b. ACT-ON (JOAN, STATUE, BY (CAUSE (JOAN, GO (PAINT, TO (STATUE)))))

This description of the semantic difference between sentences (29a) and (29b) provides a strong basis for predicting their different argument structures. But we still need to explain how linking principles give rise to this difference. Pinker's account rests on semantic structures like (30), in which depth of embedding reflects sequence of causation, with ordering on ARG-ST stemming from depth of semantic embedding, a strategy adopted in Davis (1996; 2001). This is one reasonable alternative, although the resulting complexity of some of the semantic rpresentations raises valid questions about what independent evidence supports them. An alternative appears in Koenig & Davis (2006), who borrow from Minimal Recursion Semantics (see the chapter on Semantics for an introduction to MRS). MRS "flattens" semantic relationss, rather than embedding them in one another, so the arrangment of these *elementary predications*, as they are temed, is of less import. They posit a RELATIONS (or RELS) attribute that collects a set of elementary predications, each representing some part of the predicator's semantics. A KEY attribute specifies a particular member of RELS as the relevant one for linking (of direct syntactic arguments). In the case of (29b) the KEY is the caused change of state description. These MRS-style representations of the two alternants of *spray*, with different KEY values, are shown in (31) and (32).

(31)
$$\begin{bmatrix} \text{KEY} & \boxed{5} \\ \text{RELATIONS} & \left\langle \boxed{5}\begin{bmatrix} \textit{spray-ch-of-loc-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-loc-rel} & \\ \text{FIG} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

(32)
$$\begin{bmatrix} \text{KEY} & \boxed{3}\begin{bmatrix} \textit{spray-ch-of-st-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-st-rel} & \\ \text{UND} & \boxed{2} \end{bmatrix} \end{bmatrix} \\ \text{RELATIONS} & \left\langle \boxed{3}, \begin{bmatrix} \textit{use-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \boxed{3} \end{bmatrix}, \begin{bmatrix} \textit{spray-ch-of-loc-rel} & \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \begin{bmatrix} \textit{ch-of-loc-rel} & \\ \text{FIG} & \boxed{4} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Generalizing from this example, one possible characterization of valence alternations, implicit in Koenig & Davis (2006), is as systematic relations between two sets of lexical entries in which the RELS of any pair of related entries are in a subset/subset relation (a weaker version of that definition would merely require an overlap between the RELS values of the two entries). Consider another case; (33) illustrates the causative-inchoative alternation, where the intransitive alternant describes only the change of state, while the transitive one ascribes a explicit causing agent.

(33) a. John broke the window.
b. The window broke.

Under a MRS representation, the change of state relation is a separate member of RELS; it is also included in the RELS of the transitive alternant, which contains a cause relation as well. Again, the RELS value of one member of each pair of related entries is a subset of the RELS value of the other.

Many other alternations linvolve one argument shifting from direct to oblique. Some English examples include conative, locative preposition drop, and *with* preposition drop alternations, as shown in (34):

(34) a. Rover clawed (at) Spot.
b. Bill hiked (along) the Appalachian Trail.
c. Burns debated (with) Smithers.

It is well known that the direct objects in these alternations seem to be "affected" more than their oblique counterparts. So if Rover clawed Spot, we infer that Spot was subjected to direct contact with Rover's claws and may have been injured by them, while if Rover merely clawed *at* Spot, no such inference can be made. Similarly, to say the one has hiked the Appalachian Trail suggests that one has hiked its entire length, not merely hiked along some portion of it. This holistic effect is not so evident in cases like (34c), though the direct object variant suggests that a formally organized debate took place, while the *with* variant could just describe an informal discussion. How might these varying intuitions related to "affectedness" relate to lexical semantic representations like those in (31) and (32)? Beavers (2010) provides one analytical advance in this direction, similar to the subset relationship between RELS values described above. He generalizes from affectedness to strength of entailments, where one semantic role's entailments are stronger than another's if and only if the set of entailments characterizing the second role also hold of the first. That is, what is true for any particpant that bears the first role will be true for any particpant that bears the second, but not necessarily the converse. His *Morphosyntactic Alignment Principle* then relates this to linking, as stated in (35), where an "L-thematic role" is a linguistically relevant semantic role:

(35) When participant x may be realized as either a direct or oblique argument of verb V, it bears L-thematic role R as a direct argument and L-thematic role $Q \subseteq_M R$ as an oblique. (Beavers 2010: 848)

Here, $Q \subseteq_M R$ means that Q is a "minimally weaker" role than R; in other words, there is no role P in the predicator such that $Q \subset P \subset R$. Thus, the substantive claim is essentially that the MAP rules out "verbs where the alternating participant has MORE lexical entailments as an oblique than the corresponding object realization" (Beavers 2010: 849).

The entailments Beavers employs differ somewhat from those we have discussed here, involving quantized change, nonquantized change, potential for change (where change can refer to change in location, possession, state, or something more abstract), furnishing the clear ordering by strength that is central to his proposal. But they do resemble entailments of semantic relations we have represented as elementary predications, such as incremental theme, change of state, and possession, along with the modal effects described in Koenig & Davis (2001). Thus the notion of a stronger role in Beavers' analysis has a rough analog in terms of whether a particular elementary predication is present in the semantics of a particular alternant. And only if an elementary predication is present, can it be designated as the KEY, and its roles linked directly. For example, in (31), there is nothing representing affectedness of the location, while in (32), there is, and it is designated as the KEY. As noted earlier, the semantics in (32) represents this additional entailment borne by the location argument. However, we are not aware of any simple, general way to represent Beavers' MAP within the EP-based model of Koenig & Davis (2006). Indeed, there is an aspect of Beavers' view that seems more in accord with numerical comparison approaches such as those of Dowty (1991) and Ackerman & Moore (2001), in that role strength is determined by the number of entailments that hold of it relative to others.

Having outlined the semantic basis of the different linking patterns of alternating verbs, we briefly take up three other issues. First is the question of how the alternants are related to one another. Second is how KEY selection has been used to account not just for alternants of the same verb, but for (nearly) synonymous verbs whose semantics contain the same set of elementary predications. Third is whether passives, which arguably do not differ semantically from their active counterparts, should be assimilated with other alternations or treated distinctly, as a kind of non-canonical lexical item.

The hypothesis pursued in Davis (1996; 2001) is that most alternations are the consequence of classes of lexical entries having two related meanings. This follows researchers such as Pinker (1989) and Levin (1993) in modeling subcategorization alternations as underlyingly meaning alternations. This change in meaning is crucial in Koenig & Davis (2006) KEY shifts as well. In some cases, the value of the RELS attribute of the two valence alternates differ (as in the two alternates of *spray* in *spray/load* alternation we discussed earlier). In some cases, the alternation might be different construals of the same event for some verbs, but not others, as Rappaport Hovav & Levin (2008) claim for the English ditransitive alternations, which adds the meaning of transfer for verbs like *send*, but not for verbs like *promise*; a KEY change would be involved (with the addition of a *cause-possess-rel*) for the first verb only. But KEY shifts and diathesis alternations do not always involve a change in meaning. The same elementary predications can be present in as the CONTENT values of two alternants, with each alternant designating a different elementary predication as the KEY.

Koenig & Davis propose this not only for cases in which there is no obvious meaning difference betwen two alternants of a verb, but also for different verbs that appear to be truth-conditionally equivalent, one famous example being the verbs of commercial exchange *buy* and *sell* (but see Van Valin (1999: 387-388), Levin & Rappaport Hovav (2005: 20), and Wechsler (2005a) for arguments that *buy* and *sell* are not equivalent). Koenig and Davis argue that a commercial event involves two reciprocal actions, an exchange of goods (which involves giving goods and obtaining goods) and an exchange of money (which involves giving money and obtaining money). Individual verbs might select one or the other these four relations, thus accounting for the differences in subject and object selection. As shown in (36) and (37), each of these verbs contains four elementary predications: one *exch-give-rel* and one *obtain-rel* for the transfer of goods, and one of each for the counter-transfer of money. *Buy* designates the *obtain-rel* representing the transfer of goods as the KEY, while *sell* designates the *exch-give-rel* representing the transfer of goods as the KEY. Other verbs, such as *pay* or *charge*, choose elementary predications representing the counter-transfer as the KEY. In all cases, the same linking constraints apply between the KEY and the ARG-ST list, yielding the different argument realizations of these verbs while preserving their underlying semantic commonality. The relevant portions of the entries for *buy* and *sell* in (36) and (37) below illustrate: critically, the KEY relation for *buy* is not the same as that for *sell*.

(36) A representation of the relevant parts of the lexical entry for *buy*:

$$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \text{KEY} & \boxed{7} \\ \text{RELATIONS} & \left\langle \begin{array}{l} \boxed{7}\begin{bmatrix} \text{ACT} & \boxed{1}\ \textit{(buyer)} \\ \text{UND} & \boxed{2} \\ \text{SOA} & \boxed{5}\begin{bmatrix} \textit{possess-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2}\ \textit{(goods)} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{ACT} & \boxed{1}\ \textit{(buyer)} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \boxed{6}\begin{bmatrix} \textit{possess-rel} \\ \text{ACT} & \boxed{3}\ \textit{(seller)} \\ \text{UND} & \boxed{4}\ \textit{(money)} \end{bmatrix} \end{bmatrix} \\ \begin{bmatrix} \textit{obtain-rel} \\ \text{ACT} & \boxed{3}\ \textit{(seller)} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \boxed{6}\begin{bmatrix} \textit{possess-rel} \\ \text{ACT} & \boxed{3}\ \textit{(seller)} \\ \text{UND} & \boxed{4}\ \textit{(money)} \end{bmatrix} \end{bmatrix}, \boxed{8}\begin{bmatrix} \textit{exch-give-rel} \\ \text{ACT} & \boxed{3}\ \textit{(seller)} \\ \text{UND} & \boxed{2} \\ \text{SOA} & \boxed{5}\begin{bmatrix} \textit{possess-rel} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{2}\ \textit{(goods)} \end{bmatrix} \end{bmatrix} \end{array} \right\rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \text{NP}:\boxed{1}, \text{NP}:\boxed{2}, \text{PP(FROM)}:\boxed{8} \rangle \end{bmatrix}$$

(37) A representation of the relevant parts of the lexical entry for *sell*:

$$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \text{KEY} & \boxed{7} \\ \text{RELATIONS} & \left\langle \boxed{7}\begin{bmatrix} \text{ACT} & \boxed{1}\ \textit{(seller)} \\ \text{UND} & \boxed{2} \\ \text{SOA} & \boxed{6}\begin{bmatrix} \textit{POSSESS-REL} \\ \text{ACT} & \boxed{3}\textit{(buyer)} \\ \text{UND} & \boxed{2}\ \textit{(goods)} \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \text{ACT} & \boxed{1}\ \textit{(seller)} \\ \text{UND} & \boxed{4} \\ \text{SOA} & \boxed{6}\begin{bmatrix} \textit{POSSESS-REL} \\ \text{ACT} & \boxed{1} \\ \text{UND} & \boxed{4}\ \textit{(money)} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{ARG-ST} & \langle \text{NP}:\boxed{1}, \text{NP}:\boxed{2}, \text{PP(TO)}:\boxed{7} \rangle \end{bmatrix}$$

As a final example of semantic alternations in the fine-grained meaning possibilities of verbs, we consider here the source-final product alternation exemplified in (38) where the direct object can be either the final product or the material source of the final product. Davis proposes that the (38a) sentences involve an alternation between the two meanings of entries represented in (39). We adapt Davis (2001) to make it consistent with Koenig & Davis (2006) and also treat the alternation as an alternation of meaning of *entries.* Note that in the meaning alternation described in (39), we use, informally, a double-headed arrow. One of the potential drawbacks of a lexical rule approach to valence alternations is that it requires selecting one or the other alternant as basic and the other as derived (e.g., is the inchoative or the causative basic?). This is not always an easy decision, as Goldberg (1995) or Levin & Rappaport Hovav (1994) have pointed out. Sometimes, morphology provides a clue, although in different languages the clues may point in different directions. French, and other Romance languages, use a "reflexive" clitic as a detransitivizing affix. In English, though, there is no obvious "basic" form or directionality. It is to avoid committing ourselves to a directionality in the meaning relation described in (39) that we eschews treating it as a lexical rule.

(38) a. Kim made/carved/sculpted/crafted a toy (out of the wood).

b. Kim made/carved/sculpted/crafted the wood into a toy.

(39)
$$\left[\begin{array}{ll}\textsc{content} & \left[\begin{array}{ll}\textsc{key} & \boxed{3}\left[\begin{array}{ll}\multicolumn{2}{l}{\textit{affect-incr-th-rel}}\\ \textsc{act} & \boxed{1}\\ \textsc{soa} & \boxed{2}\textit{(final product)}\end{array}\right]\\ \textsc{rels} & \left\langle \boxed{3}, \boxed{4} \right\rangle\end{array}\right]\end{array}\right]$$

$$\longleftrightarrow$$

$$\left[\begin{array}{ll}\textsc{content} & \left[\begin{array}{ll}\textsc{key} & \boxed{4}\left[\begin{array}{ll}\multicolumn{2}{l}{\textit{affect-incr-th-rel}}\\ \textsc{act} & \boxed{1}\\ \textsc{und} & \textit{(source material)}\\ \textsc{soa} & \left[\begin{array}{ll}\multicolumn{2}{l}{\textit{affect-incr-th-rel}}\\ \textsc{act} & \boxed{1}\\ \textsc{soa} & \boxed{2}\textit{(final product)}\end{array}\right]\end{array}\right]\\ \textsc{rels} & \left\langle \boxed{3}, \boxed{4} \right\rangle\end{array}\right]\end{array}\right]$$

Although most diathesis alternations can be modeled as alternations in meaning and as KEY shifts, some arguably cannot. We discuss the active/passive alternation here, but impersonals, as well as raising structures exemplified in (28) are good candidates too. The semantic relations of actives and long passives, as in (40), are practically identical and the difference between the two alternates is pragmatic in nature. Arguably, then, passives are a degenerate case of the subset relationship between RELS attributes, where the RELS values of the two entries are identical and so are the two entries' KEY. But this raises the question of whether linking in passives violates the constraints in (18)–(20), especially (18, which links the value of ACT to the first element of ARG-ST.

(40) a. Fido dug a couple of holes.

b. A couple of holes were dug by Fido.

One typical HPSG method for modeling valence alternations like passives is through lexical rules (see Chapter ??) with one alternant serving as input and the other as output; the main effect of the lexical rule in such an approach is to alter the ARG-ST of the input, going from the ARG-ST of *give* to that of *given* in (8). Critically, we must assume that the output cannot be subject to linking constraint (18), since the actor argument is not linked to the first member of the ARG-ST list. A simplified representation of what such a rule would look like is provided in (41) where we assume that the input to the rule must be transitive.

(41) $\begin{bmatrix} \textit{trans-vb} & \\ \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{1}}, \text{NP}_{\boxed{2}} \right\rangle \oplus \boxed{3}\textit{list} \end{bmatrix} \mapsto \begin{bmatrix} \text{HEAD} & \begin{bmatrix}\text{VFORM} & \textit{pass}\end{bmatrix} \\ \text{ARG-ST} & \left\langle \text{NP}_{\boxed{2}} \right\rangle \oplus \boxed{3} \oplus \left\langle \text{PP}[\textit{by}]_{\boxed{1}} \right\rangle \end{bmatrix}$

To sum up, in contrast to most meaning-driven alternations, valence alternations like the active/passive are modeled through the use of lexical rules that alter the ARG-ST of "base" entries. Which alternations pattern with active/passive and require positing lexical rules that alter a "base" entry's ARG-ST list is, as of yet, not settled. We turn now to roles that are putatively present semantically, but not realized syntactically at all.

6 Extended ARG-ST

Most of this chapter focuses on cases where semantic roles linked to the ARG-ST list are arguments of the verb's core meaning. But in quite a few cases, complements (or even subjects) of a verb are not part of this basic meaning; consequently, the ARG-ST list must be extended to include elements beyond the basic meaning. We consider three cases here, illustrated in (42)-(44).

Resultatives, illustrated in (42), express an effect, which is caused by an action of the type denoted by the basic meaning of the verb. The verb *fischen* 'to fish' is a simple intransitive verb (42a) that does not entail that any fish were caught, or any other specific effect of the fishing.

(42) a. dass er fischt
that he fishes
'that he is fishing'

b. dass er ihn leer fischt
that he it empty fishes
'that he is fishing it empty'

c. wegen der Leerfischung der Nordsee
because.of the empty.fishing of.the North.See.GEN
'because of the North Sea being fished empty'

In (42b) we see a resultative construction, with an object NP and a secondary predicate AP. The meaning is that he is fishing, causing it (the body of water) to become empty of fish. Müller (2002) posits a German lexical rule applying to the verb that augments the ARG-ST list with an NP and AP, and adds the causal semantics to the CONTENT (see Wechsler (2005b) for a similar analysis of English resultatives). The existence of deverbal nouns like *Leerfischung* 'fishing empty', which takes the body of water as an argument in genitive case (see 42c) confirms that the addition of the object is a lexical process, as noted by Müller (2002).

Romance clause-union structures as in (43) have long been analyzed as cases where the complements of the complements of a clause-union verb (*faire* in (43)) are complements of the clause-union verb itself (Aissen 1979/1974).

(43) Johanna a fait manger les enfants.
Johanna have made eat the children
'Johanna had the children eat.'

Within HPSG, the "union" of the two verbs' dependents is modeled via the composition of ARG-ST lists of the clause union verb, following Hinrichs & Nakazawa (1994) (this is a slight simplification, see Chapter ?? for details).

Sentence (44) illustrates a slightly different point, namely that some semantic modifiers, such as *souvent* in (44), can be realized as complements, and thus should be added as members of ARG-ST (or members of the DEPS list if one countenances such an additional list).

(44) Mes amis m' ont souvent aidé.
My friends me have often helped
'My friends often helped me.'

Abeillé & Godard (1997) have argued that many adverbs including negative adverbs and negation in French are complements of the verb and Kim & Sag (2002) extended that view to some uses of negation in English. In contrast to resultatives, which affect the meaning of the verb, or to clause union, where one verb co-opts the argument structure of another verb, what is added to the ARG-ST list in these cases is typically considered a semantic adjunct and a modifier in HPSG (thus it selects the verb or VP via the MOD attribute).

Another case of an adjunct that behaves like a complement is found in (45). The clitic *en* expressing the cause of death is not normally an argument of the verb *mourir* 'die', but rather an adjunct (Koenig & Davis 2006):

(45) Il en est mort
He of.it is dead.PERF.PAST
'He died of it' (Koenig & Davis 2006, ex. 12a)

On the widespread assumption (at least within HPSG) that pronominal clitics are verbal affixes (see Miller & Sag 1997), the adjunct cause of the verb *mourir* must be represented within the entry for *mourir*, so as to trigger affixation by *en*. Bouma, Malouf & Sag (2001) discuss such cases and other cases where "adverbials" as they call them, can be part of a verb's lexical entry. To avoid mixing those adverbials with the argument structure list (and have to address their relative obliqueness with syntactic arguments of verbs), they introduce yet an additional list, the dependents list (abbreviated as DEPS) which includes the ARG-ST list but also a list of adverbials. Each adverbial selects for the verb on whose DEPS list it appears as argument, as shown in (46). But, of course, not all verb modifiers can be part of the DEPS list and Bouma, Malouf & Sag discuss at length some of the differences between the two kinds of "adverbials."

(46) $$verb \Rightarrow \begin{bmatrix} \text{CONT}|\text{KEY} & \boxed{2} \\ \text{HEAD} & \boxed{3} \\ \text{DEPS} & \boxed{1} \oplus list\left(\begin{bmatrix}\text{MOD} & \begin{bmatrix}\text{HEAD} & \boxed{3} \\ \text{key} & \boxed{2}\end{bmatrix}\end{bmatrix}\right) \\ \text{ARG-ST} & \boxed{1} \end{bmatrix}$$

Although the three cases we have outlined result in an extended ARG-ST, the ways in which this extension arises differ. In the case of resultatives, the extension results partly or wholly from changing the meaning in a way similar to Rappaport Hovav & Levin (1998): by adding a causal relation, the effect argument of this causal relation is added to the membership in the base ARG-ST list (see Section 5 for a definition of the attributes KEY and RELS; suffice it to say for now that a *cause-rel* is added to the list of relations that are the input of the rule). The entries of the clause union verbs are simply stipulated to include on their ARG-ST lists the syntatic arguments of their (lexical) verbal arguments. Finally, (negative) adverbs that select for a verb (VP) are added to the ARG-ST of the verb they select. A simplified representation of all three processes is provided in (47)-(49).

(47) $$\begin{bmatrix} \text{KEY} & \boxed{2} \\ \text{RELS} & \boxed{1}\langle \ldots \boxed{2} \ldots \rangle \end{bmatrix} \mapsto \begin{bmatrix} \text{KEY} & \boxed{3} cause\text{-}rel \\ \text{RELS} & \boxed{1} \oplus \boxed{3} \end{bmatrix}$$

(48) $$\begin{bmatrix} \text{ARG-ST} & \left\langle \ldots, \begin{bmatrix} \text{HEAD} & verb \\ \text{ARG-ST} & \boxed{1} \end{bmatrix} \right\rangle \bigcirc \boxed{1} \end{bmatrix}$$

(49) $$\begin{bmatrix} \text{ARG-ST} & \boxed{1} \end{bmatrix} \mapsto \begin{bmatrix} \text{ARG-ST} & \boxed{1} \bigcirc \langle \text{ADV}_{neg} \rangle \end{bmatrix}$$

7 Is ARG-ST universal?

In this section, we briefly consider the question of whether something akin to Figure 1 offers a satisfying account of the grammatical encoding of semantic arguments across all languages. Because of its role in accounting for the syntax of basic clauses, the presence of an ARG-ST list on lexical entries comes with expectations about the syntactic realization of semantic roles. In recent work, Koenig & Michelson (2015) argue these expectations are not universally borne out, based on data from Oneida (Norther Iroquoian). The only grammatical reflex of semantic arguments in Oneida, they argue, is inflectional: the referencing of semantic arguments by so-called pronominal prefixes, which are better thought of as agreement markers à la Evans (2002). Koenig and Michelson distinguish between grammatical and syntactic arguments. Grammatical arguments include not only syntactic arguments (that is, those on ARG-ST) but also inflectional referencing of semantic roles. Some ordering analogous to linking in other languages is present in Oneida, because the semantic roles are not arbitrarily associated with agreement morphemes, but this can be captured in an ordered list of semantic indices , called INFL-STR in this proposal. INFL-STR is part of the morphological information relevant to word-internal inflectional processes, what Anderson (1992) calls 'Morphosyntactic Representation' in his treatment of Georgian agreement markers. The ordering of semantic indices on INFL-STR insures that the predicator is properly inflected. For example, the prefix *lak-* occurs if a third singular masculine proto-agent argument is acting on a first singular proto-patient argument as in *lak-hlo·lí-heʔ* 'he tells me' (habitual aspect), whereas the prefix *li-* occurs if a first singular proto-agent argument is acting on a third masculine singular argument, as in *li-hlo·lí-heʔ* 'I tell him' (habitual aspect).

In (50) and (51) we show the distinction between grammatical arguments realized as syntactic arguments (as in most languages) and those that are not (as in Oneida). Here, Koenig and Michelson follow the encoding of linking constraints as implicational constraints, as in Koenig & Davis (2003), although nothing critical hinges on that choice.

(50) $$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} \ \boxed{1} \end{bmatrix} \\ \text{ARG-ST} & \langle \text{NP}, \ldots \rangle \end{bmatrix} \Rightarrow \begin{bmatrix} \text{ARG-ST} \ \langle \text{NP:}\boxed{1}, \ldots \rangle \end{bmatrix}$$

(51) $$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \textit{cause-rel} \\ \text{CAUSER} \ \boxed{1}\textit{anim} \end{bmatrix} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{INFL-STR} \ \langle \boxed{1}, \ldots \rangle \end{bmatrix}$$

(50) constrains the association between a cause and a *synsem* member of the ARG-ST list; (51) constrains the semantic *index* of the cause to be the first member of INFL-STR.

If a language like Oneida (Northern Iroquoian) only includes an ordering of semantic indices for inflectional purposes and constraints such as (51) and no

ARG-ST list, a number of predictions follow, which Koenig and Michelson claim are borne out. Briefly summarizing their evidence, the relation between semantic arguments and external phrases, when they occur, is not necessarily one of co-indexing, no binding constraints exist between external phrases (e.g., condition C violations can be found), there are no valence alternations, and no syntactic constraints on extraction. In other words, Oneida contains none of the evidence supporting the presence of an ARG-ST list and an ordering of syntactic arguments along an obliqueness hierarchy we have discussed in this chapter. The ARG-ST list may thus not a universal attribute of words, though present in the overwhelming majority of languages. Linking, understood as constraints between semantic roles and members of the ARG-ST list, is then but one possibility; constraints that relate semantic roles to an INFL-STR list of semantic indices is also an option. In languages that exclusively exploit that latter possibility, syntax is indeed simpler.

8 The lexical approach to argument structure

We end this chapter with a necessarily brief comparison between the approach to argument structure we describe in this chapter with other approaches to argument structure that have developed since the 1990's. This chapter describes a *lexical approach to argument structure*, which is typical of research in HPSG. The basic tenet of such approaches is that lexical items include argument structures, which represent essential information about potential argument selection and expression, but abstract away from the actual local phrasal structure. In contrast, *phrasal approaches*, which are common both in Construction Grammar and in transformational approaches such as Distributed Morphology, reject such lexical argument structures. Let us briefly review the reasons for preferring a lexical approach. (This section is drawn from Müller & Wechsler (2014b), which may be consulted for more detailed and extensive argumentation).

In phrasal approaches to argument structure, components of a verb's apparent meaning are actually 'constructional meaning' contributed directly by the phrasal structure. The linking constraints of the sort discussed above are then said to arise from the interaction of the verb meaning with the constructional meaning. For example, agentive arguments tend to be realized as subjects, not objects, of transitive verbs. On the theory presented above, that generalization is captured by the linking constraint (18), which states that the ACTOR argument of an *act-und-rel* ('actor-undergoer relation') is mapped to the initial item in the ARG-ST list. In a phrasal approach, the agentive semantics is directly associated with the subject position in the phrase structure. In transformational theories, a silent 'light verb' (usually called 'little *v*') heads a projection in the phrase structure and assigns the agent role to its specifier (the subject). In constructional theories, the phrase structure itself assigns the agent role. In either type of phrasal approach, the agentive component of the verb meaning is actually expressed by the phrasal structure into which the verb is inserted.

The lexicalist's predicate argument structure provides essential information for a verb's potential combination with argument phrases. If a given lexical entry could only combine with the particular set of phrases specified in a single VAL feature, then the lexical and phrasal approaches would be difficult to distinguish: whatever information the lexicalist specifies for each VAL list item could, on the phrasal view, be specified instead for the phrases realizing those list items. But crucially, the verb need not immediately combine with its specified arguments. Alternatively it can meet other fates: it can serve as the input to a lexical rule; it can combine first with a modifier in an adjunction structure; it can be coordinated with another word with the same predicate argument structure; instead of being realized locally, one or more of its arguments can be effectively transferred to another head's valence feature (raising or argument transfer); or arguments can be saved for expression in some other syntactic position (partial fronting). Here we consider two of these, lexical rules and coordination.

The predicate argument structure is abstract: it does not directly encode the phrase structure or precedence relations between this verb and its arguments. This abstraction captures the commonality across different syntactic expressions of the arguments of a given root.

(52) a. The rabbits were nibbling the carrots.
  b. The carrots were being nibbled (by the rabbits).
  c. a large, partly nibbled, orange carrot
  d. the quiet, nibbling, old rabbits
  e. the rabbit's nibbling of the carrots
  f. The rabbit gave the carrot a nibble.
  g. The rabbit wants a nibble (on the carrot).
  h. The rabbit nibbled the carrot smooth.

Verbs undergo morpholexical operations like passive (52d), as well as antipassive, causative, and applicative in other languages. They have cognates in other parts of speech such as adjectives (52c,d) and nouns (52e,f,g). Verbs have been argued to form complex predicates with resultative secondary predicates (52h), and with serial verbs in other languages.

The same root lexical entry *nibble*, with the same meaning, appears in all of these contexts. The effects of lexical rules together with the rules of syntax dictate the proper argument expression in each context. For example, if we call the first two arguments in an ARG-ST list (such as the one in (52) above) Arg1 and Arg2, respectively, then in an active transitive sentence Arg1 is the subject and Arg2 the object; in the passive, Arg2 is the subject and the referential index of Arg1 is optionally assigned to a *by*-phrase. The same rules of syntax dictate the position of the subject, whether the verb is active or passive. When adjectives are derived from verbal participles, whether active (*a nibbling rabbit*) or passive (*a nibbled carrot*), the rule is that whichever role would have been expressed as the subject of the verb is assigned by the participial adjective to the referent of the noun that it modifies, see Bresnan (1982) and Bresnan et al. (2015: Chapter 3). The phrasal approach, in which the agent role is assigned to the subject position, is too rigid.

Nor could this be solved by associating each syntactic environment with a different meaningful phrasal construction: an active construction with agent role in the subject position; a passive construction with agent in the *by*-phrase position; etc. The problem for that view is that that one lexical rule can feed another. In the example above, the output of the verbal passive rule (see (52d)) feeds the adjective formation rule (see (52e)).

A verb can also be coordinated with another verb with the same valence requirements. The two verbs then share their dependents. This causes problems for the phrasal view, especially when a given dependent receives different semantic roles from the two verbs. For example, in an influential phrasal analysis, Hale & Keyser (1993) derived denominal verbs like *to saddle* through noun incorporation out of a structure akin to [PUT a saddle ON x]. Verbs with this putative derivation routinely coordinate and share dependents with verbs of other types:

(53) Realizing the dire results of such a capture and that he was the only one to prevent it, he quickly [saddled and mounted] his trusted horse and with a grim determination began a journey that would become legendary.[4]

Under the phrasal analysis the two verbs place contradictory demands on a single phrase structure. But on the lexical analysis, this is simple V^0 coordination.

To summarize, a lexical argument structure is an abstraction or generalization over various occurrences of the verb in syntactic contexts. To be sure, one key use of that argument structure is simply to indicate what sort of phrases the verb must (or can) combine with, and the result of semantic composition; if that were

[4] http://www.jouetthouse.org/index.php?option=com_content&view=article&id=56&Itemid=63, 21.07.2012

the whole story then the phrasal theory would be viable. But it is not. As it turns out, this lexical valence structure, once abstracted, can alternatively be used in other ways: among other possibilities, the verb (crucially including its valence structure) can be coordinated with other verbs that have a similar valence structure; or it can serve as the input to lexical rules specifying a new word bearing a systematic relation to the input word. The phrasal approach prematurely commits to a single phrasal position for the realization of a semantic argument. In contrast, a lexical argument structure gives a word the appropriate flexibility to account for the full range of expressions found in natural language.

References

Abeillé, Anne & Danièle Godard. 1997. The syntax of French negative adverbs. In Paul Hirschbühler & F. Marineau (eds.), *Negation and polarity: syntax and semantics*, 1–17. Amsterdam: John Benjamins.

Ackerman, Farrell, Robert Malouf & John Moore. 2017. Symmetrical objects in Moro: challenges and solutions. *Journal of Linguistics* 53. 3–50.

Ackerman, Farrell, Robert Malouf & John Moore. 2018. Symmetric objects in Moro. http://csli-publications.stanford.edu/HPSG/2013/, accessed 2018-7-5. Paper presentend at the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin.

Ackerman, Farrell & John Moore. 2001. *Proto-properties and grammatical encoding: a correspondence theory of argument selection*. Stanford: CSLI Publications.

Aissen, Judith. 1979/1974. *The syntax of causative constructions*. New York: Garland.

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

Artawa, Ketut. 1994. *Ergativity and Balinese syntax*. Bundoora, Australia: La Trobe University dissertation.

Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press.

Baker, Mark. 1997. Thematic roles and syntactic structures. In Liliane Haegeman (ed.), *Elements of grammar*, 73–137. Kluwer.

Beavers, John. 2010. The structure of lexical meaning: why semantics really matters. *Language* 86. 821–864.

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

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (ed.), *The mental representation of grammatical relations* (MIT Press Series on Cognitive Theory and Mental Representation), 3–86. Cambridge, MA/London: MIT Press.

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

Carlson, Greg. 1998. Thematic roles and the individuation of events. In Susan Rothstein (ed.), *Events and grammar*, 35–51. Dordrecht: Kluwer.

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

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3. 281–332. DOI:10.1007/s11168-006-6327-9

Copestake, Ann, Alex Lascarides & Dan Flickinger. 2001. An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th ACL meeting*. Toulouse, France.

Davis, Anthony. 1996. *Lexical semantics and linking in the hierarchical lexicon*. Stanford University dissertation.

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

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

Dowty, David. 1989. On the semantic content of the notion of 'thematic role'. In Gennaro Chierchia, Barbara Partee & Raymond Turner (eds.), *Properties, types, and meaning*, vol. 2, 69–129. Dordrecht: Kluwer.

Dowty, David R. 1979. *Word meaning and montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ*. Vol. 7 (Studies in Linguistics and Philosophy). Dordrecht, Netherlands: Reidel. DOI:10.1007/978-94-009-9473-7

Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Evans, Nicholas. 2002. The true status of grammatical object affixes: evidence from Bininj Gun-Wok. In Nicholas Evans & Hans-Jürgen Sasse (eds.), *Problems of polysynthesis*, 15–50. Berlin: Akademie Verlag.

Fillmore, Charles. 1968. The case for case. In Emmon Bach & Robert Harms (eds.), *Universals in linguistic theory*, 1–87. New York: Holt, Rinehart & Winston.

Fillmore, Charles. 1977. The case for case reopened. In Peter Cole & Jerold Sadock (eds.), *Grammatical relations*, 59–81. New York: Academic Press.

Foley, William & Robert Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Gawron, Jean Mark. 1986. Situations and prepositions. *Linguistics and Philosophy* 9. 327–382.

Gerdts, Donna. 1992. Morphologically-mediated relational profiles. In Laura Buszard-Welcher et al. (ed.), *Proceedings of the 18th annual meeting of the berkeley linguistics society*, 322–337. Berkeley, CA: Berkeley Linguistics Society.

Goldberg, Adele. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: Chicago University Press.

Greenberg, Joseph H. 1966. *Language universals*. The Netherlands: Mouton.

Guilfoyle, Eithne, Henrietta Hung & Lisa Travis. 1992. Spec of IP and Spec of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10(3). 375–414.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 53–109. Cambridge, MA: MIT Press.

Hinrichs, Erhard & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-driven Phrase-Structure Grammar*, 11–37. Stanford: CSLI Publications.

Jackendoff, Ray. 1987. The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18. 369–411.

Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. 63–99.

Keyser, Samuel & T. Roeper. 1984. On the middle and ergative constructions in English. *Linguistic Inquiry*. 2nd ser. 15. 381–416.

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

Koenig, Jean-Pierre & Anthony Davis. 2001. Sublexical modality and the structure of lexical semantic representations. *Linguistics and Philosophy* 24. 71–124.

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

Koenig, Jean-Pierre & Anthony Davis. 2006. The KEY to lexical semantic representations. *Journal of Linguistics* 42. 71–108.

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

Ladusaw, William & David Dowty. 1988. Toward a nongrammatical account of thematic roles. In Wendy Wilkins (ed.), *Syntax and semantics 21: thematic relations*, 62–74. New York: Academic Press.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures*. Dordrecht: Kluwer.

Levin, Beth. 1993. *English verb classes and alternations*. Chicago, IL: Chicago University Press.

Levin, Beth & Malka Rappaport Hovav. 1994. A preliminary analysis of causative verbs in English. In Lila Gleitman & Barbara Landau (eds.), *The acquisition of the lexicon*, 35–77. Cambridge, MA: MIT Press.

Levin, Beth & Malka Rappaport Hovav. 2005. *Argument realization*. Cambridge, UK: Cambridge University Press.

Manning, Christopher. 1996. *Ergativity: argument structure and grammatical relations*. Stanford, CA: CSLI publications.

Manning, Christopher & Ivan A Sag. 1999. Dissociations between argument structure and grammatical relations. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation*, 63–78. Stanford: CSLI Publications.

Mauner, Gail & Jean-Pierre Koenig. 2000. Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language* 43. 110–134.

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

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

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

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

Pinker, Steven. 1989. *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.

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

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

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments*, 97–134. Stanford: CSLI Publications.

Rappaport Hovav, Malka & Beth Levin. 2008. The English dative alternation: the case for verb sensitivity. *Journal of Linguistics* 44. 129–167.

Reis, Marga. 1982. Zum Subjektbegriff im Deutschen. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 171–211. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag.

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

Schein, Barry. 2002. Events and the semantic content of thematic relations. In Gerhard Preyer & Georg Peter (eds.), *Logical form and language*, 263–344. Oxford: Oxford University Press.

Van Valin, Robert. 1999. Generalized semantic roles and the syntax-semantics interface. In *Empirical issues in formal syntax and semantics 2*, 373–388. The Hague, The Netherlands: Thesus.

Wechsler, Stephen. 1995a. Preposition selection outside the lexicon. In Raul Aranovich, William Byrne, Susanne Preuss & Matha Senturia (eds.), *Proceedings of WCCFL XIII*, 416–431. Stanford: CSLI Publications.

Wechsler, Stephen. 1995b. *The semantic basis of argument structure.* Stanford, CA: CSLI Publications.

Wechsler, Stephen. 1999. HPSG, GB, and the Balinese bind. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 179–195. Stanford, CA: CSLI Publications.

Wechsler, Stephen. 2005a. What is right and wrong about little v. In Mila Vulchanova & Tor A. Åfarli (eds.), *Grammar and beyond – essays in honour of Lars Hellan.* Novus Press.

Wechsler, Stephen Mark. 2005b. Resultatives under the event-argument homomorphism model of telicity. In Nomi Erteschik-Shir & Tova Rapoport (eds.), *The syntax of aspect—Deriving thematic and aspectual interpretation*, 255–273. Oxford University Press.

Wechsler, Stephen & I Wayan Arka. 1998. Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory* 16. 387–442.

Chapter 9

Constituent order

Stefan Müller

Humboldt-Universität zu Berlin

1 Introduction

This chapter deals with constituent order with a focus on local order variants. English is the language that is treated most thoroughly in theoretical linguistics but it is also the most boring language as far as the possibilities of reordering constituents is concerned: the order of subject verb and object is fixed in sentences like (1):

(1) Kim likes bagels.

Of course there is the possibility to front the object as in (2) but this is a special, non-local construction that is not the topic of this chapter but is treated in Borsley & Crysmann (2018), Chapter 13 of this volume.

(2) Bagels, Kim likes.

This chapter deals with scrambling (the local reordering of arguments) and with alternative placements of heads (called head movement in some theories). Examples of the former are the sentences in (3) and an example of the latter is given in (4):

(3) a. [weil] der Mann der Frau das Buch gibt (German)
because the.NOM man the.DAT woman the.ACC book gives

b. [weil] der Mann das Buch der Frau gibt
because the.NOM man the.ACC book the.DAT woman gives

Stefan Müller. 2018. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 177–207. Berlin: Language Science Press. DOI:??

c. [weil] das Buch der Mann der Frau gibt
because the.ACC book the.NOM man the.DAT woman gives

d. [weil] das Buch der Frau der Mann gibt
because the.ACC book the.DAT woman the.NOM man gives

e. [weil] der Frau der Mann das Buch gibt
because the.DAT woman the.NOM man the.ACC book gives

f. [weil] der Frau das Buch der Mann gibt
because the.DAT woman the.ACC book the.NOM man gives

(4) Gibt der Mann der Frau das Buch? (German)
gives the.NOM man the.DAT woman the.ACC book
'Does the man give the woman the book?'

(3) shows that in addition to the unmarked order in (3a) (see Höhle (1982) on the notion of unmarked order), five other argument orders are possible in sentences with three-place verbs.

(4) shows that the verb is placed in initial position in questions in German. This contrasts with the verb final order in (3a). This alternation of verb placement is usually treated as head movement in the transformational literature (Bach 1962; Bierwisch 1963: 34; Reis 1974; Thiersch 1978: Chapter 1). Declarative main clauses in German are V2 clauses and the respective fronting is usually treated as a non-local dependency (see Borsley & Crysmann (2018), Chapter 13 of this volume). Hence, these sentences will not be handled here.

The following sections deal with the theoretical options within the HPSG framework for dealing with these phenomena. I first discuss the separation of grammar rules in the immediate dominance part and a linear precedence component in Section 2 and then flat vs. binary branching structures (Section 3). While flat structures allow verbs to be ordered finally or initially, this is not the case for binary branching structures. Proposals with binary branching structures are usually paired with so-called head-movement approaches. These are explained in Section 4. Section 5 introduces an extension to standard HPSG developed by Reape (1994): constituent order domains. Such constituent order domains allow for discontinuous constituents and have been used to account for languages like Warlipri (Donohue & Sag 1999). Section 6 shows how such languages can be analyzed without admitting discontinuous constituents.

2 ID/LP format

HPSG was developed out of GPSG and Categorial Grammar. The ideas concerning linearization of daughters in a local tree were taken over from GPSG (Gazdar, Klein, Pullum & Sag 1985). In GPSG a separation between immediate dominance and linear precedence is assumed. So, while in classical phrase structure grammar a phrase structure rule like (5) states that the NP[nom], NP[dat] and NP[acc] have to appear in exactly this order, this is not the case in GPSG and HPSG:

(5) S → NP[nom], NP[dat], NP[acc], V

The HPSG schemata corresponding to the phrase structure rule in (5) do not express information about ordering. Instead there are separate linearization rules. A schema like (5) licenses 24 different orders: the six permutations of the three arguments that were shown in (3) and all possible placements of the verb (to the right of NP[acc], between NP[dat] and NP[acc], between NP[nom] and NP[dat], to the left of NP[nom]). Orders like NP[nom], NP[dat], V, NP[acc] are not attested in German and hence these linearizations have to be filtered out. This is done by linearization rules, which can refer to features or to the function of a daughter in a schema. (6) shows some examples of linearization rules:

(6) a. X < V
b. X < V[INI−]
c. X < Head [INI−]

The first rule says that all constituents have to precede a V in the local tree. The second rule says that all constituents have to precede a V that has the INITIAL value −. One option to analyze German would be the one that was suggested by Uszkoreit (1987) within the framework of GPSG: one could allow for two linearization variants of finite verbs. So in addition to the INI− variant there could be a INI+ variant and this variant would be linearized initially. The LP rule in (6c) is more general than (6b) in that it does not mention the part of speech but instead refers to the function of the constituent. The rule says that a head that has the INI value − has to be linearized to the right of all other elements in the local tree.

This treatment of constraint on linearization has an advantage that was already pointed out by researchers working in GPSG: it captures the generalizations regarding linearization. For instance the order of verbs and their arguments is the same in embedded sentences in German independent of the finiteness of the verb:

(7) a. dass er dem Mann das Buch gab
that he the man the book gave
'that he gave the man the book'

b. dass er versucht, [dem Mann das Buch zu geben]
that he tried the man the book to give
'that he tried to give the man the book'

This is also true for the relative order of dative and accusative object in (7). The constraints regarding linearization hold across rules. By factoring these constraints out, the generalizations can be captured.

3 Flat and binary branching structures

The previous section discussed LP rules and used flat phrase structure rules for illustration. The corresponding flat structures are also used in HPSG. Schema 1 shows a Head-Complement schema that combines a head with all the complements selected via the COMPS-Liste.

Schema 1 (Head-Complement Schema)
head-complement-phrase ⇒

$$\begin{bmatrix} \ldots|\textsc{comps} & \langle\rangle \\ \textsc{head-dtr} & \begin{bmatrix} \ldots|\textsc{comps}\ \boxed{1} \end{bmatrix} \\ \textsc{non-head-dtrs} & \text{synsem2sign}(\boxed{1}) \end{bmatrix}$$

`synsem2sign` is a relational constraint mapping *synsem* objects as they are contained in the COMPS-Liste onto objects of type *sign* as they are contained in daughters (Pollard & Sag 1994). [1] How this schema can be used to analyze VPs like the one in (8) is shown in Figure 1.

(8) Kim gave Sandy a book.

Researchers working on English usually assume a flat structure but assuming binary branching structures would be possible as well, as is clear from analyses in Categorial Grammar, where binary combinatory rules are assumed (Ajdukiewicz 1935; Steedman 2000). For languages like German it is usually assumed that structures are binary branching (but see Reape (1994: 156) and Bouma & van Noord

[1] In Sign-Based Construction Grammar the objects in valence lists are of the same type as the daughters. A relational constraint would not be needed in this variant of the HPSG theory.

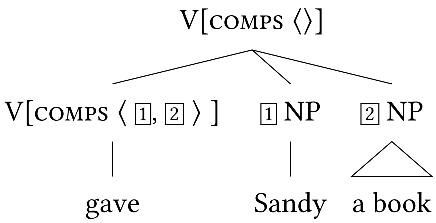


Figure 1: Analysis of the VP *gave Sandy a book* with a flat structure

(1998: 51)). The reason for this is that adverbials can be placed anywhere between the arguments as the following example from Uszkoreit (1987: 145) shows:

(9) *Gestern* hatte *in der Mittagspause* der Vorarbeiter *in der Werkzeugkammer* dem Lehrling *aus Boshaftigkeit langsam* zehn schmierige Gußeisenscheiben *unbemerkt* in die Hosentasche gesteckt.
yesterday had during the lunch.break the foreman in the tool.shop the apprentice maliciously slowly ten greasy cast.iron.disks unnoticed in the pocket put
'Yesterday during lunch break, the foreman maliciously and unnoticed, put ten greasy cast iron disks slowly into the apprentice's pocket.'

A straightforward analysis of adjunct placement in German and Dutch is to assume that adjuncts can attach to any verbal projection.

Binary branching structures with attachment of adjuncts to any verbal projection also accounts for recursion and hence the fact that arbitrarily many adjuncts can attach to a verbal projection. Of course it is possible to formulate analyses with flat structures that involve arbitrarily many adjuncts (Kasper 1994; van Noord & Bouma 1994; Bouma et al. 2001), but these analyses involve relational constraints in schemata or in lexical items. In Kasper's analysis the relational constraints walk through lists of daughters of unbounded length in order to compute the semantics. In the other two analyses adjuncts are treated as valents, which may be problematic because of scope issues. This cannot be dealt with in detail here but see Levine & Hukari (2006) and Chaves (2009) for discussion.

The following schema licenses binary branching head-complement phrases:

Schema 2 (Head-Complement Schema (binary branching))
head-complement-phrase ⇒

$$\begin{bmatrix} \textsc{comps} & \boxed{1} \oplus \boxed{2} \\ \textsc{head-dtr} & \begin{bmatrix} \textsc{comps}\ \boxed{1} \oplus \langle\ \boxed{3}\ \rangle \oplus \boxed{2} \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle \begin{bmatrix} \textsc{synsem}\ \boxed{3} \end{bmatrix} \right\rangle \end{bmatrix}$$

$\oplus$ (append) is a relational constraint that concatenates two lists. The COMPS-Liste of the head daughter is split into three lists: a beginning ($\boxed{1}$), a list containing $\boxed{3}$ and a rest ($\boxed{2}$). $\boxed{3}$ is identified with the SYNSEM value of the non-head daughter. All other elements of the COMPS-Liste of the head daughter are concatenated and the result of this concatenation ($\boxed{1} \oplus \boxed{2}$) is the COMPS-Liste of the mother node. This schema is very general. It works for languages that allow for scrambling since it allows to take an arbitrary element out of the COMPS-Liste of the head daughter and realize it in a local tree. The schema can also be "parametrized" to account for languages with fixed word order. For head final languages with fixed order $\boxed{2}$ would be the empty list and for head-initial languages with fixed order (e.g., English) $\boxed{1}$ would be the empty list.

The alternative to using relational constraints as in Schema 2 is to use sets rather than lists for the representation of valence information (Gunji 1986; Hinrichs & Nakazawa 1989; Pollard 1996; Engelkamp, Erbach & Uszkoreit 1992). The Head-Complement Schema would combine the head with one of its complements. Since the elements of a set are not ordered, any complement can be taken and hence all permutations of complements are accounted for.

The disadvantage of set-based approaches is that sets do not impose an order on their members but an order is needed for various subtheories of HPSG (see Przepiórkowski (2018), Chapter 7 of this volume on case assignment, and Branco (2018), Chapter 20 of this volume on Binding Theory). In the approach proposed above and in Müller (2003b; 2015a,b), the valence lists are ordered but the schema allows for combination with any element of the list. For valence representation and the order of elements in valence lists see Wechsler, Koenig & Davis (2018), Chapter 8 of this volume.

4 Head movement vs. constructional approaches assuming flat structures

The Germanic languages signal the clause type by verb position. All Germanic languages with the exception of English are V2 languages: the finite verb is in second position in declarative main clauses. The same holds for questions with *wh* phrases. Yes/no questions are formed by putting the verb in initial position. English is a so-called *residual V2 language*. While declarative clauses are in base order (SVO), questions follow the pattern that is known from other Germanic languages.

(10) What$_i$ did Kim read $\_{i}$?

Analyses assuming flat structures (or flat linearization domains, see Section 5) usually treat alternative orders of verbs in Germanic languages as linearization variants (Reape 1994; Kathol 2001; Müller 1995; 2003b; Bjerre 2006), but this is not necessarily so as Bouma and van Noord's analysis of Dutch clauses show (Bouma & van Noord 1998: 62, 71). The alternative to verb placement as linearization is something that is similar to verb-movement in Government & Binding: a trace takes the position of the verb in its canonical position and the verb is realized in initial or second position. The following subsection deals with such approaches in more detail. Subsection 4.2 deals with a constructional approach.

4.1 Head movement approaches

Borsley (1989) showed that in addition to the analysis of auxiliary inversion in English that was suggested in GPSG (Gazdar et al. 1985) an analysis that is similar to the movement-based analysis in GB is possible in HPSG as well. The technique that is used in the analysis is basically the same that was developed by Gazdar (1981) for the treatment of nonlocal dependencies in GPSG. A trace is assumed and the information about the missing element is passed up the tree until it is bound off at an appropriate place (that is by the fronted verb). The analysis of (11) is shown in Figure 2.

(11) Did Kim get the job?

A special variant of the auxiliary selects a full clause in which an auxiliary is missing. The fact that the auxiliary is missing is represented as the value of DSL. The value of DSL is a *synsem* object, that is something that contains syntactic and semantic information ($\boxed{2}$ in Figure 2). DSL is a head feature and hence available everywhere along a projection path (see Borsley & Abeillé (2018), Chapter 1 of

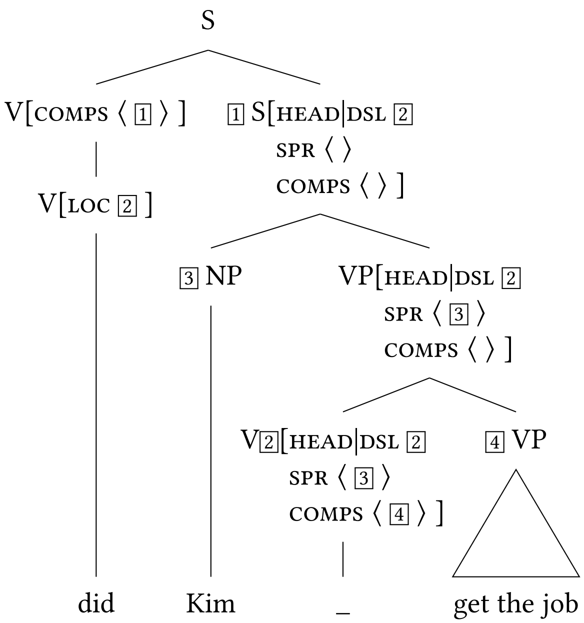


Figure 2: Analysis of English auxiliary constructions as head-movement following (Borsley 1989)

check reference to Chapter 1.

this volume for the Head Feature Principle). The trace for head movement is rather simple:

(12) $\begin{bmatrix} \textit{word} & & \\ \text{PHON} & \langle\rangle & \\ \text{SYNSEM}|\text{LOC} & \boxed{1} & \begin{bmatrix}\text{CAT}|\text{HEAD}|\text{DSL}\ \boxed{1}\end{bmatrix} \end{bmatrix}$

It states that there is an empty element that has the local requirements that correspond to its DSL value. For cases of verb movement it says: I am a verb that is missing itself. The fronted auxiliary is licensed by a lexical rule that maps a non-fronted auxiliary onto one that selects a complete clause from which the input auxiliary is missing.

Such head-movement analyses are assumed by most researchers working on German (Kiss & Wesche 1991: Section 4.7; Oliva 1992; Netter 1992; Kiss 1993; Frank 1994; Kiss 1995; Feldhaus 1997, Meurers 2000; Müller 2005a; 2017) and also by (Bouma & van Noord 1998: 62, 71) in their work on Dutch, by Müller & Ørsnes (2015) in their grammar of Danish and by Müller (2018) for Germanic in general.

4.2 Constructional approaches

The alternative to head-movement-based approaches is a flat analysis with an alternative serialization of the verb. This was already discussed with respect to German, but I want to discuss English auxiliary constructions here, since the figured prominently in linguistic discussions. In the analysis of (13) shown in Figure 3, the auxiliary *did* selects for the subject *Kim* and a VP *get the job*.

(13) Did Kim get the job?

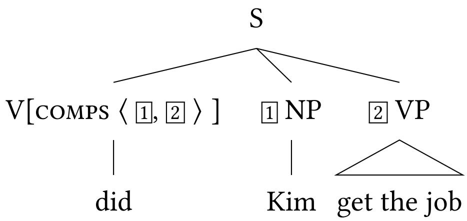


Figure 3: Analysis of English auxiliary constructions according to (Sag et al. 2018)

The tree in Figure 3 is licensed by a schema combining a head with its subject ($\boxed{1}$) and its VP complement ($\boxed{2}$) in one go. As is common in HPSG since 1995 (Sag 1997) phrasal schemata are organized in type hierarchies and the general schema for auxiliary initial constructions has the type *aux-initial-cxt*. Fillmore (1999) and Sag et al. (2018) argue that there are various usages of auxiliary-initial constructions and assign the respective usages to subconstructions of the general auxiliary-initial construction. Technically this amounts to stating subtypes of *aux-initial-cxt*. For example, Sag et al. (2018) posit a subtype *polar-int-cl* for polar interrogatives like (14a) and another subtype *auxinitial-excl-cl* for exclamatives like (14b).

(14) a. Are they crazy?
b. Are they crazy!

Chomsky (2010) compared the various clause types used in HPSG with the – according to him – much simpler Merge-based analysis in Minimalism. Minimalism assumes just one very general schema for combination (External Merge is basically equivalent to our Schema 2 above, see Müller (2013)), so this rule for combining linguistic objects is very simple, but this does not help in any way when considering the facts: there are at least three different meanings associated with auxiliary initial clauses and these have to be captured somewhere in a grammar. One way is to state them in a type hierarchy as is done in some HPSG analyses and in SBCG, another way is to use implicational constraints that assign meaning with respect to actual configurations (see Section 4.3) and a third way is to do everything lexically. The only option for Minimalism is the lexical one. This means that Minimalism has to either assume as many lexical items for auxiliaries as there are types in HPSG or to assume empty heads that contribute the meaning that is contributed by the phrasal schemata in HPSG (Borsley 2006: Section 5; Borsley & Müller 2018). The latter proposal is generally assumed in Cartographic approaches (Rizzi 1997). Since there is a fixed configuration of functional projections that contribute semantics, one could term these Rizzi-style analyses *Crypto-Constructional*.

4.3 Mixed approaches

The situation with respect to clause types is similar in German. Verb first sentences can be yes/no questions (15a), imperatives (15b), conditional clauses (15c), and declarative sentences with topic drop (15d).

(15) a. Kommt Peter? (German)
comes Peter
'Does Peter come?'

b. Komm!
come

c. Kommt Peter, komme ich nicht.
comes Peter come I not
'If Peter comes, I won't come.'

d. Kommt. (Was ist mit Peter?)
comes what is with Peter
'What about Peter?' 'He comes.'

Verb second sentences can be questions (16a), declarative sentences (16b), or imperatives (16c).

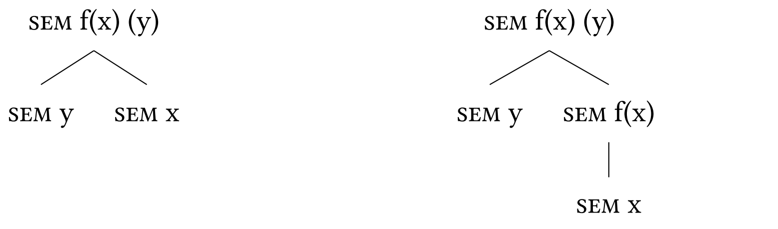
(16) a. Wer kommt?
who comes

b. Peter kommt. (German)
Peter comes

c. Jetzt komm!
now come
'Come now!'

While one could try and capture this situation by assuming surface order-related clause types, such approaches are rarely assumed (but see Kathol (2001) and Wetta (2011). See Section 5.4.2 on why such approaches are doomed to failure). Rather researchers assumed binary branching head-complement structures together with verb movement (I assumed linearization domains (see Section 5) for ten years and then switched to the head-movement approach (Müller 2005a,b; 2017)).

As was explained above, the head movement approaches are based on lexical rules or unary projections. These license new linguistic objects that could contribute the respective semantics. As Borsley (2006) pointed out, this would mean that one needs seven versions of fronted verbs to handle the seven cases in (15 and (16), which would correspond to the seven phrasal types that would have to be stipulated in phrasal approaches. But there is a way out of this: one can assume one lexical item with underspecified semantics. HPSG makes it possible to use implicational constraints referring to a structure in which an item occurs. Depending on the context the semantics contributed by a specific item can be further specified. Figure 4 shows the construction-based and the lexical rule-based analysis for comparison. In the construction-based analysis the daughters



(a) Phrasal construction

(b) Unary construction and implication

Figure 4: Construction-based, phrasal approach and approach with implicational constraint

contribute x and y as semantic values and the whole construction adds the construction meaning f. In the lexical rule- or unary projection-based analysis, the lexical rule/unary projection adds the f and the output of the rule is combined compositionally with the other daughter. Now, implicational constraints can be used to determine the exact contribution of the lexical item (Müller 2016). This is shown with the example of a question in Figure 5. The implication says: when the

QUE ⟨ [ ] ⟩ $\Rightarrow$ int(x)

Figure 5: Implication for interrogative sentences

configuration has the form that there is a question pronoun in the left daughter, the output of the lexical rule gets question semantics. Since HPSG represents all linguistic information in the same AVM, such implicational constraints can refer to intonation as well.

Note that in Constructional HPSG as layed out by Sag (1997) implicational constraints can refer to the structure of a complete utterance. Hence items with a complex internal structure can be seen as contributing a certain meaning. This is ruled out by design in Sign-Based Construction Grammar, where linguistic objects of type *phrase* do not have daughters.

5 Constituent order domains

There is an interesting extension to standard HPSG that opens up possibilities for analyses that are quite different from what is done otherwise in theoretical linguistics: Mike Reape (1991; 1992; 1994) working on German suggested formal tools that allow for the modeling of discontinuous constituents. His original motivation was to account for scrambling of arguments in verbal complexes but this analysis was superseeded by Hinrichs and Nakazawa's analysis (Hinrichs & Nakazawa 1989; 1994) since pureley linearization-based approaches are unable to account for agreement and the so-called remote passive (Kathol 1998: Section 5.1, Section 5.2; Müller 1999: Chapter 21.1). Nevertheless, his work was taken up by others and was used for analyzing German (Kathol & Pollard 1995; Kathol 2000; Müller 1995; 1996; 2004; Wetta 2011; Wetta 2014). As will be discussed below, there were reasons for dropping analyses of German assuming discontinuous constituents (Müller 2005b; 2017) but constituent order domains still play a major role in analyzing ellipsis (Nykiel & Kim 2018, Chapter 19 of this volume) and coordination (Abeillé & Chaves 2018, Chapter 16 of this volume).

5.1 A special representational layer for constituent order

The technique that is used to model discontinuous constituents in frameworks like HPSG goes back to Mike Reape's work on German (1991; 1992; 1994). Reape uses a list called DOMAIN to represent the daughters of a sign in the order in which they appear at the surface of an utterance. (17) shows an example in which the DOM value of a headed-phrase is computed from the DOM value of the head and the list of non-head daughters.

(17) *headed-phrase* $\Rightarrow \begin{bmatrix} \text{HEAD-DTR}|\text{DOM} & \boxed{1} \\ \text{NON-HEAD-DTRS} & \boxed{2} \\ \text{DOM} & \boxed{1} \bigcirc \boxed{2} \end{bmatrix}$

The symbol '$\bigcirc$' stands for the *shuffle* relation. *shuffle* relates three lists A, B and C iff C contains all elements from A and B and the order of the elements in A and the order of the elements of B is preserved in C. (18) shows the combination of two sets with two elements each:

(18)
$$\begin{aligned} \langle\, a, b \,\rangle \bigcirc \langle\, c, d \,\rangle = {} & \langle\, a, b, c, d \,\rangle \vee \\ & \langle\, a, c, b, d \,\rangle \vee \\ & \langle\, a, c, d, b \,\rangle \vee \\ & \langle\, c, a, b, d \,\rangle \vee \\ & \langle\, c, a, d, b \,\rangle \vee \\ & \langle\, c, d, a, b \,\rangle \end{aligned}$$

The result is a disjunction of six lists. *a* is ordered before *b* and *c* before *d* in all of these lists, since this is also the case in the two lists $\langle\, a, b \,\rangle$ and $\langle\, c, d \,\rangle$ that have been combined. But apart from this, *b* can be placed before, between or after *c* and *d*.

Every word comes with a domain value that is a list that contains the word itself:

(19) Domain contribution of single words, here *gibt* 'gives':

$$\boxed{1} \begin{bmatrix} \text{PHON} & \langle\, \textit{gibt} \,\rangle \\ \text{SYNSEM} & \ldots \\ \text{DOM} & \langle\, \boxed{1} \,\rangle \end{bmatrix}$$

The description in (19) may seem strange at first glance, since it is cyclic, but it can be understood as a statement saying that *gibt* contributes itself to the items that occur in linearization domains.

The constraint in (20) is responsible for the determination of the PHON values of phrases:

(20) $$\textit{phrase} \Rightarrow \begin{bmatrix} \textsc{phon} & \boxed{1} \oplus \ldots \oplus \boxed{n} \\ \textsc{dom} & \left\langle \begin{bmatrix} \textit{sign} \\ \textsc{phon}\ \boxed{1} \end{bmatrix}, \ldots, \begin{bmatrix} \textit{sign} \\ \textsc{phon}\ \boxed{n} \end{bmatrix} \right\rangle \end{bmatrix}$$

It states that the PHON value of a sign is the concatenation of the PHON values of its DOMAIN elements. Since the order of the DOMAIN elements corresponds to their surface order, this is the obvious way to determine the PHON value of the whole linguistic object.

Figure 6 shows how this machinery can be used to license binary branching structures with discontinuous constituents. Words or word sequences that are

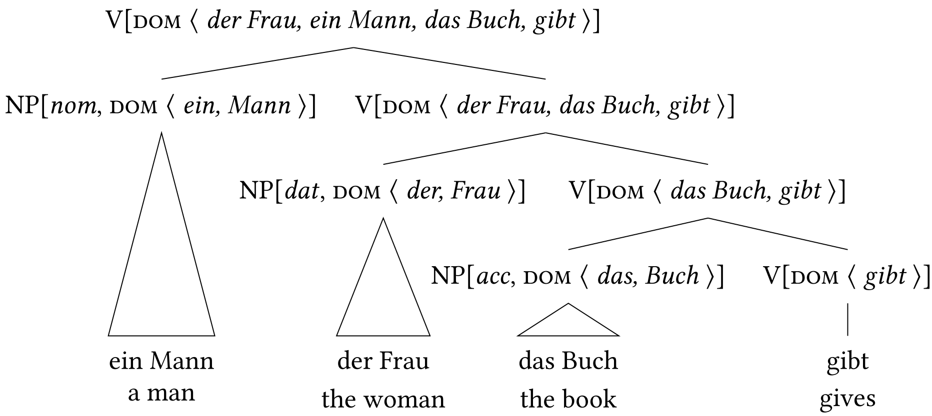


Figure 6: Analysis of *dass der Frau ein Mann das Buch gibt* 'that a man gives the woman the book' with binary branching structures and discontinuous constituents

separated by commas stand for separate domain objects, that is, ⟨ *das, Buch* ⟩ contains the two objects *das* and *Buch* and ⟨ *das Buch, gibt* ⟩ contains the two objects *das Buch* and *gibt*. The important point to note here is that the arguments are combined with the head in the order accusative, dative, nominative, although the elements in the constituent order domain are realized in the order dative, nominative, accusative rather than nominative, dative, accusative, as one would expect. This is possible since the formulation of the computation of the DOM value using the shuffle operator allows for discontinuous constituents. The node for *der Frau das Buch gibt* 'the woman the book gives' is discontinuous: *ein Mann* 'a man' is inserted into the domain between *der Frau* 'the woman' and *das Buch* 'the book'. This is more obvious in Figure 7, which has a serialization of NPs that corresponds to their order.

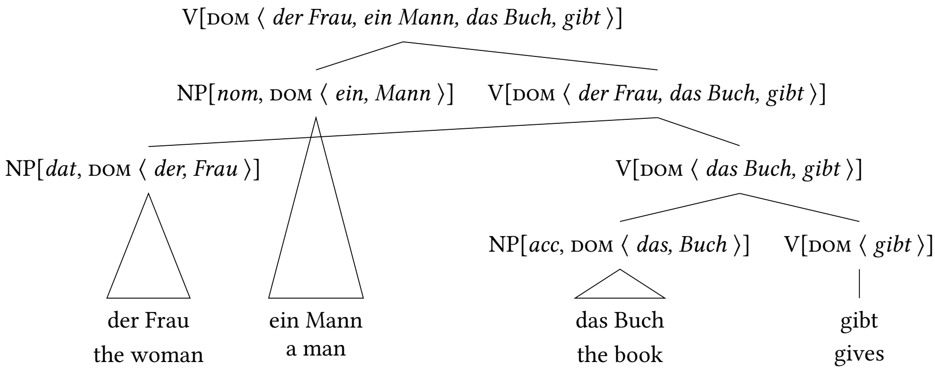


Figure 7: Analysis of *dass der Frau ein Mann das Buch gibt* 'that a man gives the woman the book' with binary branching structures and discontinuous constituents showing the discontinuity

5.2 Absolutely free

While German is more exciting than English in terms of constituent order it is still boring in comparison to languages like Warlpiri which have much freer constituent order. In Warlpiri the auxiliary has to be in first or in second position (Donohue & Sag 1999: 8) and apart from this even parts of what are noun phrases in German and English can appear separated from each other. For example, the two parts of the NP *Kurdu-jarra-rlu wita-jarra-rlu* 'child small' may appear discontinuously since they are marked with the same case:

(21) Kurdu-jarra-rlu ka-pala maliki wajili-pi-nyi wita-jarra-rlu.
child-DU-ERG PRS-3DU.SUBJ dog.ABS chase-NPAST small-DU-ERG
(Warlpiri)

Mistake in glossing. Should *pi-nyi* be *pi.nyi*?

'Two small children are chasing the dog.' or
'Two children are chasing the dog and they are small.'

Donohue & Sag (1999) developed an analysis for this that simply liberates domain elements and inserts them into the next higher domain. (22) shows how this is formalized:

(22) *liberating-phrase* ⇒

$$\begin{bmatrix} \textsc{dom} & \delta_0 \bigcirc \delta_1 \bigcirc \ldots \bigcirc \delta_n \\ \textsc{head-dtr} & \begin{bmatrix} \textsc{dom} & \delta_0 \end{bmatrix} \\ \textsc{non-head-dtrs} & \left\langle \begin{bmatrix} \textsc{dom} & \delta_1 \end{bmatrix}, \ldots, \begin{bmatrix} \textsc{dom} & \delta_n \end{bmatrix} \right\rangle \end{bmatrix}$$

Rather than inserting the complete daughters into the domain of the mother as in (17), the DOM values of the daughters are shuffled into the domain of the mothers. So instead of having the NPs in the same domain as the verb as in the German example in the previous section one has all the parts of NPs in the next higher domain. Hence, a single nominal element being placed in front of the auxiliary in second position is explained without problems. Figure 8 shows the analysis of Donohue & Sag (1999). *child* and *small* form an NP. They contribute

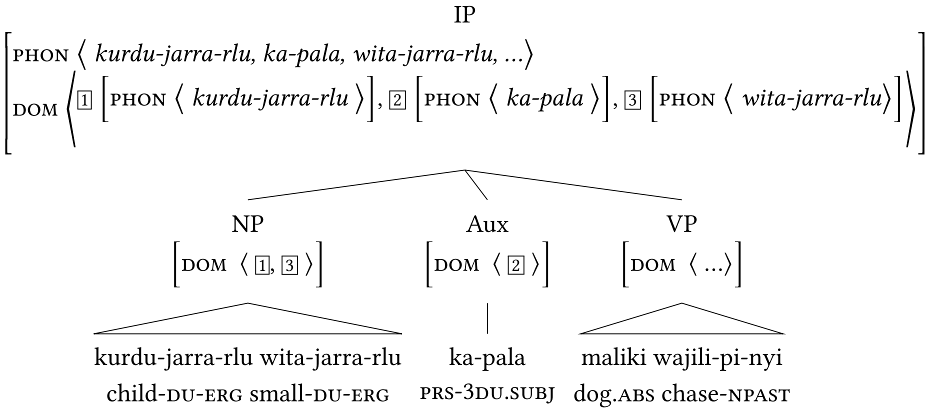


Figure 8: Analysis of free constituent order in Warlpiri according to Donohue & Sag (1999)

two independent domain objects ($\boxed{1}$ and $\boxed{3}$) to the domain of the mother. The second element in this domain has to be the auxiliary ($\boxed{2}$), $\boxed{1}$ is realized initially and $\boxed{3}$ follows the auxiliary.

We have seen so far an analysis that inserts complete objects into the domain of the mother, an analysis that inserts all domain objects of objects into the domain of the mother and in the next subsection I want to look at an intermediate case, so-called *partial compaction*.

5.3 Partial compaction (extraposition)

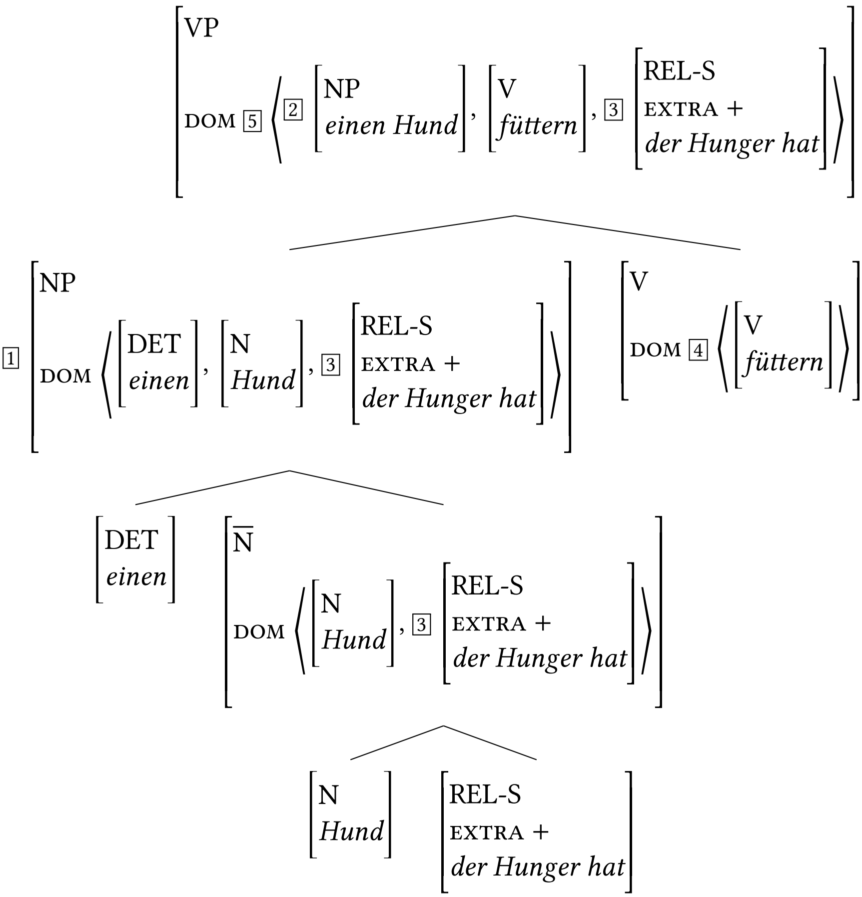
Kathol & Pollard (1995) developed an analysis of extraposition that is a mix of the strategies discussed in the two subsections: most of one NP object is inserted into the domain of the mother as a single object, only those parts that are extraposed are liberated and inserted as individual domain objects into the domain of the mother. Kathol & Pollard's analysis of (23) is given in Figure 9.

(23) einen Hund füttern, der Hunger hat (German)
a dog feed that hunger has
'feed a dog that is hungry'

einen Hund, der Hunger hat 'a dog who is hungry' consists of three domain objects: *einen* 'a', *Hund* 'dog', and *der Hunger hat* 'who hungry is'. The two initial ones are inserted as one object (the NP *ein Hund* 'a dog') into the higher domain and the relative clause is liberated. While the formation of the new domain at the mother node is relatively straight-forward in the cases discussed so far, a complex relational constraint is needed to split the relative clause ($\boxed{3}$) from the other domain objects and construct a new domain object that has the determiner and the noun as constituents ($\boxed{2}$). Kathol and Pollard have a relational constraint called *compaction* that builds new domain objects for insertion into higher domains. *partial compaction* takes an initial part of a domain and forms a new domain object from this returning the remaining domain objects for separate insertion into the higher domain. Due to space limitations, this constraint will not be discussed here but see Müller (1999: 244) for a refined version of Kathol and Pollard's constraint. The effect of partial compaction in Figure 9 is that there is a new object $\boxed{2}$ and a list containing the remaining objects, in the example $\langle \boxed{3} \rangle$. A list containing the new object $\langle \boxed{2} \rangle$, a list containing the remaining objects $\langle \boxed{3} \rangle$ are shuffled with the domain list of the head $\boxed{4}$. Since the relative clause is in the same domain as the verb, it can be serialized to the right of the verb.

5.4 Problems with order domains

Constituent order domains may seem rather straight-forward since linearization facts can be handled easily. I assumed constituent order domains and discontinuous constituents for German myself for over a decade (Müller 1995; 2004). However, there are some problems that seem to suggest that a traditional GB-like head-movement approach is the better alternative. In what follows I want to discuss just two problematic aspects of linearization approaches: spurious ambiguities and apparently multiple frontings.



p-compaction($\boxed{1}$, $\boxed{2}$, $\langle \boxed{3} \rangle$)

$\boxed{5} = \langle \boxed{2} \rangle \bigcirc \langle \boxed{3} \rangle \bigcirc \boxed{4}$

Figure 9: Analysis of extraposition via partial compaction of domain objects according to Kathol & Pollard (1995)

5.4.1 Partial fronting and spurious ambiguities

Kathol (2000) suggests an analysis with binary branching structures in which all arguments are inserted into a linearization domain and can be serialized there in any order provided no LP rule is violated. Normally one would have the elements of the comps-Liste in a fixed order, combine the head with one element from the comps-Liste after the other, and let the freedom in the dom-Liste be responsible for the various attested orders. So both sentences in (24) would have analyses in which the verb *erzählt* 'tells' is combined with *Geschichten* 'stories' first and then *Geschichten erzählt* 'stories tells' is combined with *den Wählern* 'the voters'. Since the verb and all its arguments are in the same linearization domain they can be ordered in any order including the two orders in (24):

(24) a. weil er den Wählern Geschichten erzählt (German)
because he the voters stories tells
'because he tells the voters stories'

b. weil er Geschichten den Wählern erzählt
because he stories the voters tells

The problem with this approach is that examples like (25) show that grammars have to account for combinations of any of the objects to the exclusion of the other:

(25) a. Geschichten erzählen sollte man den Wählern nicht. (German)
stories tell should one the voters not
'One should not tell the voters such stories.'

b. Den Wählern erzählen sollte man diese Geschichten nicht.
the voters tell should one these stories not

Kathol (2000: Section 8.9) accounts for examples like (25) by relaxing the order of the objects in the valence list. He uses the shuffle operator in the valence representation:

(26) $\langle$ NP[*nom*] $\rangle \oplus (\langle$ NP[*dat*] $\rangle \bigcirc \langle$ NP[*acc*] $\rangle)$

This solves the problem with examples like (25) but it introduces a new one: sentences like (24) now have two analyses each. One is the analysis we had before and another one is the one in which *den Wählern* 'the voters' is combined with *erzählt* 'tells' first and the result is then combined with *Geschichten* 'stories'. Since both objects are inserted into the same linearization domain, both orders can be derived. So we have too much freedom: freedom in linearization and freedom in the order of combination. The proposal that I suggested has just the freedom in the order of combination and hence can account for both (24) and (25) without spurious ambiguities.

5.4.2 Surface order, clause types, fields within fields, and empty elements

Kathol (2001) develops an analysis of German that uses constituent order domains and determines the clause types on the basis of the order of elements in such domains. He suggests the topological fields *1*, *2*, *3*, and *4*, which correspond to the traditional topological fields *Vorfeld* 'prefield', *linke Satzklammer* 'left sentence bracket', *Mittelfeld* 'middle field', *rechte Satzklammer* 'right sentence bracket'. Domain objects may assigned to these fields and they are then ordered by linearization constraints stating that objects assigned to *1* have to precede objects of type *2*, type *3*, and type *4*. Objects of type *2* have to precede type *3*, and type *4* and so on. For the Vorfeld and the left sentence he stipulates uniqueness constraints saying that at most one constituent may be of this type. This can be stated in a nice way by using the linearization constraints in (27):

(27) a. *1* < *1*
b. *2* < *2*

This trick was first suggested by Gazdar et al. (1985: 55, Fn. 3) in the framework of GPSG and it works since if there were two objects of type *1* than each one would be required to precede the other one resulting in a violation of the linearization constraint. So in order to avoid such constraint violation there must not be more than one *1*.

Kathol (2001) assumes the following definition for V2 clauses:

(28) $$\textit{V2-clause} \Rightarrow \begin{bmatrix} \text{S}[\textit{fin}] \\ \textsc{dom} \quad \left\langle [1], \begin{bmatrix} 2 \\ \text{V}[\textit{fin}] \end{bmatrix}, \ldots \right\rangle \end{bmatrix}$$

This says that the constituent order domain starts with one element assigned to field *1* followed by another domain object assigned to field *2*. While this is in accordance with general wisdom about German, which is a V2 language, there are problems for entirely surface-based theories: German allows for multiple constituents in front of the finite verb. (29) shows some examples:

(29) a. [Zum zweiten Mal] [die Weltmeisterschaft] errang Clark 1965 … [2]
to.the second time the world.championship won Clark 1965
'Clark won the world championship for the second time in 1965.'

b. [Dem Saft] [eine kräftige Farbe] geben Blutorangen.[3]
the.DAT juice a.ACC strong color give blood.oranges
'Blood oranges give the juice a strong color.'

Müller (2003a) extensively documents this phenomenon. The categories that can appear before the finite verb are almost unrestricted. Even subjects can be fronted together with other material (Bildhauer & Cook 2010: 72; Bildhauer 2011: 371). The empirical side of these apparent multiple frontings was further examined in the Collective Research Center 632, Project A6 and the claim that only constituents depending on the same verb can be fronted together (Fanselow 1993; Hoberg 1997: 1634) was confirmed (Müller 2017: Chapter 3). A further insight is that the linearization properties of the fronted material (NPs, PPs, adverbs, adjectives) correspond to the linearization properties they would have in the Mittelfeld. The example in (30) are even more interesting. It shows that there can be a right sentence bracket (the particle *los*) and an extraposed constituent (something following the particle: *damit*) before the finite verb (*geht* 'goes'):

(30) *Los* damit *geht* es schon am 15. April.[4]
off there.with goes it PRT on 15. April
4 *5* *2* *3* *3* *3*
'The whole thing starts on the 15th April.'

In Kathol's system, *los* would be of type *4* and *damit* would have to be of type *5* (an additional type for extraposed items). Without any modification of the general system, we would get a *4* and a *5* ordered before a *2* (a right sentence bracket and a postfield preceding the left sentence bracket), something that is ruled out by Kathol's linearization constraints.

Müller (2002), still working in a domain-based framework, developed an analysis assuming an empty verbal head to explain the fact that the fronted constituents have to depend on the same verb and that there is a separate topological area that is independent of the remaining clause. So, *los* and *damit* are domain

[2] (Beneš 1971: 162)

[3] Bildhauer & Cook (2010) found this example in the *Deutsches Referenzkorpus* (DeReKo), hosted at Institut für Deutsche Sprache, Mannheim: http://www.ids-mannheim.de/kl/projekte/korpora

[4] taz, 01.03.2002, p. 8.

objects within a larger domain object placed in the prefield. Wetta (2011) suggests an analysis in which two or more constituents are compacted into one domain object, so *los* and *damit* would form one object that is inserted into the domain containing the finite verb. However, this begs the question what kind of object it is that is formed. Section 5.3 dealt with partial compaction of NPs. Some of the elements from an NP domain were liberated and other elements were fused into a new object that had the same category as the object containing all material, namely NP. But the situation with examples like (29) and (30) is quite different. We have a particle and a pronominal adverb in (30) and various other combinations of categories in the examples collected by Müller (2003a) and Bildhauer (2011). It would not make sense to claim that the fronted object is a particle or a pronominal adverb. Note that it is neither an option to leave the category of the fronted object unspecified since HPSG comes with the assumption that models of linguistic objects are total, that is, maximally specific (King 1999, see also Richter (2018), Chapter 3 of this volume). Leaving the category and valence properties of the item in the prefield unspecified would make such sentences infinitely many times ambiguous. Of course Wetta could state that the newly created object is a verbal projection but this would just be stating the effect of the empty verbal head within a relational constraint, which I consider less principled than stating the empty element.

However, the empty verbal head that I stated as part of a linearization grammar in 2002 comes as a stipulation since its only purpose in the grammar of German was to account for apparent multiple frontings. Müller (2005b; 2017) drops the linearization approach and assumes head-movement instead. The empty head that is used for accounting for the verb position in German can also be used to account for apparent multiple frontings. The analysis is sketched in (31):

(31) [$_{\mathrm{VP}}$ [Zum zweiten Mal] [die Weltmeisterschaft] $\_{}_{\mathrm{V}}$]$_i$ errang$_j$ Clark
to.the second time the world.championship won Clark
1965 $\_{}_i$ $\_{}_j$.
1965

The details cannot be explained here but the analysis treats apparent multiple frontings parallel to partial verb phrase frontings. A lexical rule is used for multiple frontings which is a special case of the head-movement rule that was discussed in Section 4. So apparent multiple frontings are analyzed with means that are available to the grammar anyway. This analysis allows to keep the insight that German is a V2 language and it also gets the same-clause constraint and the linearization of elements right. See Müller (2005a,b; 2017) for details.

The paper so far discussed the tools that have been suggested in HPSG to account for constituent order: flat vs. binary branching structures, linearization domains, head-movement via DSL. I showed that analyses of German relying on discontinuous constituents and constituent order domains are not without problems and that head-movement approaches with binary branching and continuous constituents can account for the data. I also demonstrated in Section 5.2 that languages like Warlpiri that allow for much freer constituent order than German can be accounted for in models allowing for discontinuous constituents. The following section discusses a proposal by Bender (2008) showing that even languages like the Australian free constituent order languages can be handled without discontinuous constituents.

6 Free constituent order languages without order domains

Bender (2008) discusses the Australian language Wambaya and shows how phenomena parallel to those treated by Donohue & Sag (1999) can be handled without discontinuous constituents. The trick is that all arguments of a head are projected to higher nodes even when they are combined with the head, that is, arguments are not canceled off from valence lists. See also Meurers (1999); Przepiórkowski (1999) and Müller (2008) for earlier non-cancellation approaches. The example (21) from Section 5.2 can be recast with continuous constituents as is shown in Figure 10. The figure shows that arguments are not removed from the valence representation after combination with the head. Rather they are marked as satisfied. Since they are still in the representation, schemata may refer to them. Bender suggests a schema that identifies the MOD value of an element that could function as an adjunct in a normal head-adjunct structure with an element in the valence representation. In Figure 10 the MOD value of the second ergative nominal *wita-jarra-rlu* 'small' is identified with an argument of the auxiliary verb ($\boxed{1}$). The adjunct hence has access to the referential index of the argument and it is therefore guaranteed that both parts of the noun phrase refer to the same discourse referent. The NP for *kurdu-jarra-rlu* is combined with the projection of the auxiliary to yield a complete sentence. Since $\boxed{1}$ not just contains the semantic index and hence information about number (the dual) but also case information, it is ensured that distributed noun phrases have to bear the same case. Since information about all arguments are projected along the head path, $\boxed{2}$ would also be available for an adjunct referring to it. So in the place of *wita-jarra-rlu* 'small-DU-ERG' we could also have another adjunct referring to *maliki* 'dog.ABS'. This shows that even languages with constituent order as free as the Australian languages

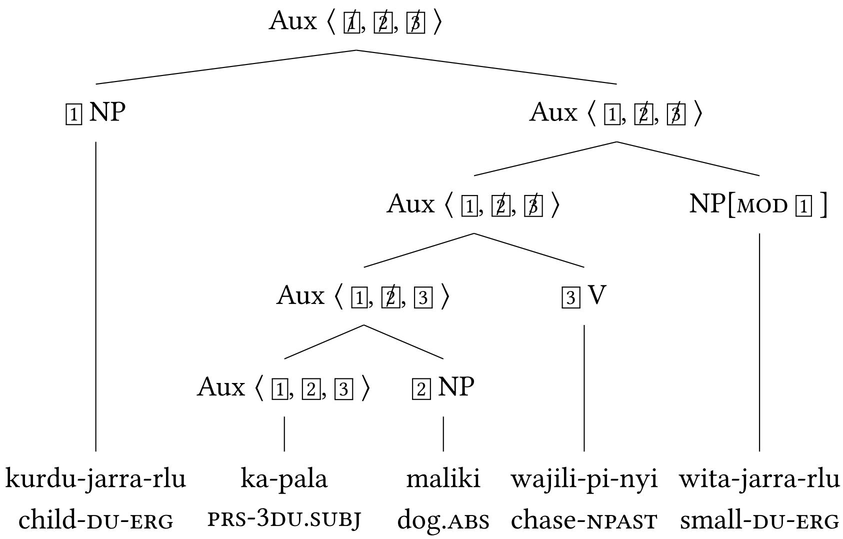


Figure 10: Analysis of free constituent order in Warlpiri using non-cancellation

can be handled within HPSG without assuming discontinuous constituents.

7 Summary

This paper discussed general approaches to constituent order in HPSG. On the one hand there are approaches assuming flat constituent structure allowing permutation of daughters as long as no LP constraints are violated and on the other hand, there are approaches assuming binary branching structures. Approaches that assume flat structures can serialize the head to the left or to the right or somewhere between other daughters in the structure. Approaches assuming binary branching have to use other means. One such means is "head movement", which is analyzed as a series of local dependencies by passing information about the missing head up along the head path. The alternative to head movement is linearization of elements in special linearization domains, allowing for discontinuous constituents. I showed that there are reasons for assuming head-movement for German and how even languages with extremely free constituent order can be analyzed without assuming discontinuous constituents.

Acknowledgements

I thank Bob Borsley for comments.

References

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1. 1–27.

Bach, Emmon. 1962. The order of elements in a Transformational Grammar of German. *Language* 8(3). 263–269.

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

Beneš, Eduard. 1971. Die Besetzung der ersten Position im deutschen Aussagesatz. In Hugo Moser (ed.), *Fragen der strukturellen Syntax und der kontrastiven Grammatik* (Sprache der Gegenwart – Schriften des IdS Mannheim 17), 160–182. Düsseldorf: Pädagogischer Verlag Schwann.

Bierwisch, Manfred. 1963. *Grammatik des deutschen Verbs* (studia grammatica 2). Berlin: Akademie Verlag.

Bildhauer, Felix. 2011. Mehrfache Vorfeldbesetzung und Informationsstruktur: Eine Bestandsaufnahme. *Deutsche Sprache* 39(4). 362–379.

Bildhauer, Felix & Philippa Cook. 2010. German multiple fronting and expected topic-hood. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar, Université Paris Diderot*, 68–79. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2010/, accessed 2018-2-25.

Bjerre, Tavs. 2006. *Object positions in a topological sentence model for Danish: A linearization-based HPSG approach.* http://www.hum.au.dk/engelsk/engsv/objectpositions/workshop/Bjerre.pdf, accessed 2018-2-25. Presentation at Ph.D.-Course at Sandbjerg, Denmark.

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

Borsley, Robert D. 2006. *Syntactic and lexical approaches to unbounded dependencies.* Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.

Borsley, Robert D. & Stefan Müller. 2018. Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. To appear.

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

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press. http://www.let.rug.nl/~vannoord/papers/, accessed 2018-2-25.

Chaves, Rui P. 2009. Construction-based cumulation and adjunct extraction. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 47–67. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2009/, accessed 2018-2-25.

Chomsky, Noam. 2010. *Restricting stipulations: consequences and challenges*. Talk given in Stuttgart.

Donohue, Cathryn & Ivan A. Sag. 1999. Domains in Warlpiri. In *Sixth International Conference on HPSG–Abstracts. 04–06 August 1999*, 101–106. Edinburgh.

Engelkamp, Judith, Gregor Erbach & Hans Uszkoreit. 1992. Handling linear precedence constraints by unification. In Henry S. Thomson (ed.), *30th Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 201–208. Newark, DE: Association for Computational Linguistics. auch als CLAUS-Report, Nr. 19, Universität des Saarlandes erschienen.

Fanselow, Gisbert. 1993. Die Rückkehr der Basisgenerierer. *Groninger Arbeiten zur Germanistischen Linguistik* 36. 1–74.

Feldhaus, Anke. 1997. *Eine HPSG-Analyse ausgewählter Phänomene des deutschen* w*-Fragesatzes*. Working Papers of the Institute for Logic and Linguistics 27. IBM Scientific Center Heidelberg: Institute for Logic & Linguistics.

Fillmore, Charles J. 1999. Inversion and Constructional inheritance. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 113–128. Stanford, CA: CSLI Publications.

Frank, Anette. 1994. *Verb second by lexical rule or by underspecification*. Arbeitspapiere des SFB 340 No. 43. Heidelberg: IBM Deutschland GmbH. ftp://ftp.ims.uni-stuttgart.de/pub/papers/anette/v2-usp.ps.gz, accessed 2018-3-20.

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

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

Gunji, Takao. 1986. Subcategorization and word order. In William J. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, 1–21. Stanford, CA: CSLI Publications.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1989. Subcategorization and VP structure in German. In *Aspects of German VP structure* (SfS-Report-01-93). Tübingen: Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. & Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 11–38. Stanford, CA: CSLI Publications.

Hoberg, Ursula. 1997. Die Linearstruktur des Satzes. In Hans-Werner Eroms, Gerhard Stickel & Gisela Zifonun (eds.), *Grammatik der deutschen Sprache*, vol. 7.2 (Schriften des Instituts für deutsche Sprache), 1495–1680. Berlin: Walter de Gruyter.

Höhle, Tilman N. 1982. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Werner Abraham (ed.), *Satzglieder im Deutschen – Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung* (Studien zur deutschen Grammatik 15), 75–153. Tübingen: originally Gunter Narr Verlag now Stauffenburg Verlag. Republished as Höhle (2018).

Höhle, Tilman N. 2018. Explikationen für „normale Betonung“ und „normale Wortstellung“. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5). Berlin: Language Science Press. DOI:10.5281/zenodo.1145680

Kasper, Robert T. 1994. Adjuncts in the Mittelfeld. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 39–70. Stanford, CA: CSLI Publications.

Kathol, Andreas. 1998. Constituency and linearization of verbal complexes. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 221–270. San Diego: Academic Press.

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

Kathol, Andreas. 2001. Positional effects in a monostratal grammar of German. *Journal of Linguistics* 37(1). 35–66.

Kathol, Andreas & Carl J. Pollard. 1995. Extraposition via complex domain formation. In Hans Uszkoreit (ed.), *33rd Annual Meeting of the Association for Computational Linguistics. Proceedings of the conference*, 174–180. Cambridge, MA: Association for Computational Linguistics.

King, Paul. 1999. Towards truth in Head-Driven Phrase Structure Grammar. In Valia Kordoni (ed.), *Tübingen studies in Head-Driven Phrase Structure Grammar* (Arbeitsberichte des SFB 340 No. 132), 301–352. Tübingen: Universität Tübingen. http://www.sfs.uni-tuebingen.de/sfb/reports/berichte/132/132abs.html, accessed 2018-2-25.

Kiss, Tibor. 1993. *Infinite Komplementation – Neue Studien zum deutschen Verbum infinitum*. Arbeiten des SFB 282 No. 42. Bergische Universität Gesamthochschule Wuppertal.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.

Kiss, Tibor & Birgit Wesche. 1991. Verb order and head movement. In Otthein Herzog & Claus-Rainer Rollinger (eds.), *Text understanding in LILOG* (Lecture Notes in Artificial Intelligence 546), 216–242. Berlin: Springer Verlag.

Levine, Robert D. & Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions* (CSLI Lecture Notes 166). Stanford, CA: CSLI Publications.

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

Meurers, Walt Detmar. 2000. *Lexical generalizations in the syntax of German non-finite constructions*. Arbeitspapiere des SFB 340 No. 145. Tübingen: Eberhard-Karls-Universität. http://www.sfs.uni-tuebingen.de/~dm/papers/diss.html, accessed 2018-2-25.

Müller, Stefan. 1995. Scrambling in German – Extraction into the *Mittelfeld*. In Benjamin K. T'sou & Tom Bong Yeung Lai (eds.), *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*, 79–83. City University of Hong Kong.

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

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

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2003a. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1). 29–62.

Müller, Stefan. 2003b. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI:10.1017/S0022226703002032

Müller, Stefan. 2004. Continuous or discontinuous constituents? A comparison between syntactic analyses for constituent order and their processing systems. *Research on Language and Computation, Special Issue on Linguistic Theory and Grammar Implementation* 2(2). 209–257.

Müller, Stefan. 2005a. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201. 3–39.

Müller, Stefan. 2005b. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203. 297–330.

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

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

Müller, Stefan. 2015a. HPSG – A synopsis. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – Theory and analysis: An international handbook*, 2nd edn. (Handbooks of Linguistics and Communication Science 42.2), 937–973. Berlin: Walter de Gruyter.

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

Müller, Stefan. 2016. Satztypen: Lexikalisch oder/und phrasal. In Rita Finkbeiner & Jörg Meibauer (eds.), *Satztypen und Konstruktionen im Deutschen* (Linguistik – Impulse und Tendenzen 65), 72–105. Berlin, Boston: de Gruyter.

Müller, Stefan. 2017. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Revise and resubmit.

Müller, Stefan. 2018. *Germanic syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin. https://hpsg.hu-berlin.de/~stefan/Pub/germanic.html, accessed 2018-3-20.

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

Netter, Klaus. 1992. On non-head non-movement: An HPSG treatment of finite verb position in German. In Günther Görz (ed.), *Konvens 92. 1. Konferenz „Verarbeitung natürlicher Sprache“. Nürnberg 7.–9. Oktober 1992* (Informatik aktuell), 218–227. Berlin: Springer Verlag.

Oliva, Karel. 1992. *Word order constraints in binary branching syntactic structures.* CLAUS-Report 20. Saarbrücken: Universität des Saarlandes.

Pollard, Carl J. 1996. On head non-movement. In Harry Bunt & Arthur van Horck (eds.), *Discontinuous constituency* (Natural Language Processing 6), 279–305. Berlin: Mouton de Gruyter. Veröffentlichte Version eines Ms. von 1990.

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

Przepiórkowski, Adam. 1999. On case assignment and "adjuncts as complements". In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 231–245. Stanford, CA: CSLI Publications.

Reape, Mike. 1991. *Word order variation in Germanic and parsing*. DYANA Report Deliverable R1.1.C. University of Edinburgh.

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

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Reis, Marga. 1974. Syntaktische Hauptsatzprivilegien und das Problem der deutschen Wortstellung. *Zeitschrift für Germanistische Linguistik* 2(3). 299–327.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publishers.

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

Sag, Ivan A., Rui Chaves, Anne Abeillé, Bruno Estigarribia, Frank Van Eynde, Dan Flickinger, Paul Kay, Laura A. Michaelis-Cummings, Stefan Müller, Geoffrey K. Pullum & Tom Wasow. 2018. Lessons from the English auxiliary system. *Journal of Linguistics*. To appear.

Steedman, Mark J. 2000. *The syntactic process* (Language, Speech, and Communication). Cambridge, MA/London, England: MIT Press.

Thiersch, Craig L. 1978. *Topics in German syntax*. M.I.T. Dissertation. http://hdl.handle.net/1721.1/16327, accessed 2018-2-25.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

van Noord, Gertjan & Gosse Bouma. 1994. The scope of adjuncts and the processing of lexical rules. In Makoto Nagao (ed.), *Proceedings of COLING 94*, 250–256. Kyoto, Japan: Association for Computational Linguistics. http://www.aclweb.org/anthology/C94-1039, accessed 2018-7-5.

Wetta, Andrew C. 2011. A Construction-based cross-linguistic analysis of V2 word order. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 248–268. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2011/, accessed 2018-2-25.

Wetta, Andrew Charles. 2014. *Construction-based approaches to flexible word order*. Buffalo: State University of New York at Buffalo dissertation. http://gradworks.umi.com/36/83/3683106.html, accessed 2016-3-11.

Chapter 10

Clitics

Berthold Crysmann
Université Paris Diderot

Gerald Penn
University of Toronto

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

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

1 Introduction

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

Berthold Crysmann & Gerald Penn. 2018. Clitics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 209–210. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 11

Complex predicates

Danièle Godard
Université Paris Diderot

Pollet Samvelian
Université Sorbonne Nouvelle

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

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

1 Introduction

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

Danièle Godard & Pollet Samvelian. 2018. Complex predicates. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 211–212. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 12

Control and raising

Anne Abeillé
Université Paris Diderot

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

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

1 Introduction

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

Anne Abeillé. 2018. Control and raising. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 213–214. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 13

Unbounded dependencies

Bob Borsley
University of Essex

Berthold Crysmann
Université Paris Diderot

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

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

1 Introduction

finality

Abbreviations

Acknowledgements

Bob Borsley & Berthold Crysmann. 2018. Unbounded dependencies. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 215–215. Berlin: Language Science Press. DOI:??

Chapter 14

Relative clauses

Doug Arnold
University of Essex

Danièle Godard
Université Paris Diderot

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

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

1 Introduction

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

Doug Arnold & Danièle Godard. 2018. Relative clauses. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 217–218. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 15

Island phenomena and related matters

Rui Chaves
University at Buffalo

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

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

1 Introduction

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

Rui Chaves. 2018. Island phenomena and related matters. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 219–220. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 16

Coordination

Anne Abeillé
Université Paris Diderot

Rui Chaves
University at Buffalo

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

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

1 Introduction

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

Anne Abeillé & Rui Chaves. 2018. Coordination. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 221–222. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 17

Idioms

Manfred Sailer
Goethe-Unversität Frankfurt

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

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

1 Introduction

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

Manfred Sailer. 2018. Idioms. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 223–224. Berlin: Language Science Press. DOI:??

ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 18

Negation

Jong-Bok Kim
Kyung Hee University, Seoul

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

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

1 Introduction

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

Jong-Bok Kim. 2018. Negation. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 225–226. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 19

Ellipsis

Joanna Nykiel
University of Silesia

Jong-Bok Kim
Kyung Hee University, Seoul

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

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

1 Introduction

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

Joanna Nykiel & Jong-Bok Kim. 2018. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 227–228. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 20

Binding

António Branco
University of Lisbon

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

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

1 Introduction

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

António Branco. 2018. Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 229–230. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Part III

Other levels of description

Chapter 21

Phonology

Jesse Tseng
Université Paris Diderot

Put abstract here with \abstract.

1 Introduction: PHONOLOGY in the HPSG sign

The PHONOLOGY attribute in (Pollard & Sag 1987) and (Pollard & Sag 1994):

- rudimentary PHON value
- basic Phonology Principle constrained by Linear Precedence rules: corresponds to simple terminal spell-out of the phrase structure tree
- "Phonology-Free Syntax" (Miller et al. 1997): PHON information inaccessible for selection via SYNSEM

There has been relatively little work within HPSG on phonological representation and the analysis of phonological phenomena. Most references to the PHON attribute use it simply as a lexical identifier, or they are dealing with phenomena at the phonology-syntax interface (e.g. constituent order, ellipsis). For such applications, the actual content of the PHON value is unimportant. These topics are covered in other chapters.

2 Phonological representations in HPSG

Proposals for the detailed content of PHON values:

- encoding of phonological constituents (Bird & Klein 1994; Klein 2000; Höhle 1999)

Jesse Tseng. 2018. Phonology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 233–236. Berlin: Language Science Press. DOI:??

- syllable structure Tseng (2008)
- metrical phonology (Klein 2000; Bonami & Delais-Roussarie 2006)

3 Phonological analysis in HPSG

- principles of constraint-based phonology vs derivational phonology (Bird & Klein 1994): compositionality, monotonicity
- compositional construction of prosodic structure in parallel with phrase structure (Klein 2000)

But HPSG is formally compatible with many approaches, and there is as yet no emerging consensus among practitioners.

- Finite state phonology (Bird 1992; 1995)
- need for abstract underlying forms (Skwarski 2009); phonologically empty categories
- OT in HPSG (Orgun 1996)

4 Specific phenomena and case studies

- shape conditions (Asudeh & Klein 2002)
- French (Tseng 2003; Bonami et al. 2004)
- phonological idioms [already covered in Manfred's chapter]
- …

Abbreviations

Acknowledgements

References

Asudeh, Ash & Ewan Klein. 2002. Shape conditions and phonological context. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *Proceedings of the HPSG-2001 Conference, Norwegian University of Science and Technology, Trondheim*, 20–30. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2001/, accessed 2015-12-23.

Bird, Steven. 1992. Finite-state phonology in HPSG. In *Proceedings of coling 92*, 74–80. Nantes, France.

Bird, Steven. 1995. *Computational phonology: A constraint-based approach* (Studies in Natural Language Processing). Cambridge, UK: Cambridge University Press.

Bird, Steven & Ewan Klein. 1994. Phonological analysis in typed feature systems. *Computational Linguistics* 20(3). 455–491.

Bonami, Olivier, Gilles Boyé & Jesse Tseng. 2004. An integrated approach to French liaison. In Gerhard Jäger, Paola Monachesi, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 29–45. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/FG/2004/, accessed 2018-2-25.

Bonami, Olivier & Elisabeth Delais-Roussarie. 2006. Metrical phonology in HPSG. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, 39–59. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2006/, accessed 2018-2-25.

Höhle, Tilman N. 1999. An architecture for phonology. In Robert D. Borsley & Adam Przepiórkowski (eds.), *Slavic in Head-Driven Phrase Structure Grammar*, 61–90. Stanford, CA: CSLI Publications. Republished as Höhle (2018).

Höhle, Tilman N. 2018. An architecture for phonology. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5). Berlin: Language Science Press. Originally published as Höhle (1999). DOI:10.5281/zenodo.1145680

Klein, Ewan. 2000. A constraint-based approach to English prosodic constituents. In *38th Annual Meeting of the Association for Computational Linguistics: Proceedings of the conference*, 217–224. Hong Kong: Association for Computational Linguistics.

Miller, Philip H., Geoffrey K. Pullum & Arnold M. Zwicky. 1997. The principle of Phonology-Free Syntax: Four apparent counterexamples in French. *Journal of Linguistics* 33(1). 67–90.

Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology*. University of California, Berkeley dissertation.

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

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

Skwarski, Filip. 2009. Accounting for underlying forms in HPSG phonology. In Stefan Müller (ed.), *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, 318–337. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2009/skwarski.pdf, accessed 2018-2-25.

Tseng, Jesse. 2003. Edge features and French liaison. In Jongbok Kim & Stephen Mark Wechsler (eds.), *The proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, 313–333. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/3/, accessed 2018-3-20.

Tseng, Jesse. 2008. The representation of syllable structure in HPSG. In Stefan Müller (ed.), *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, 234–252. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2008/, accessed 2018-2-25.

Chapter 22

Morphology

Berthold Crysmann
Université Paris Diderot

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

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

1 Introduction

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

Berthold Crysmann. 2018. Morphology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 237–238. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 23

Semantics

Jean-Pierre Koenig
University at Buffalo

Frank Richter
Goethe-Universität Frankfurt

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

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

1 Introduction

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

Jean-Pierre Koenig & Frank Richter. 2018. Semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 239–240. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 24

Information structure

Kordula de Kuthy
Universität Tübingen

Put abstract here with \abstract.

1 Introduction

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

Abbreviations

Acknowledgements

Kordula de Kuthy. 2018. Information structure. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 241–241. Berlin: Language Science Press. DOI:??

Chapter 25

Pragmatics and dialogue semantics

Andy Lücking
Goethe-Universität Frankfurt

Jonathan Ginzburg
Université Paris Diderot

Robin Cooper
University of Gothenburg

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

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

1 Introduction

Duis pulvinar lacus id gravida ornare. Phasellus eu mauris sed tortor maximus condimentum ultrices in leo. Donec non erat nec nulla ullamcorper ornare sed id ex. Integer risus mauris, aliquet vel aliquam sed, feugiat quis nisi. Suspendisse

Andy Lücking, Jonathan Ginzburg & Robin Cooper. 2018. Pragmatics and dialogue semantics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 243–244. Berlin: Language Science Press. DOI:??

quis nunc a turpis porttitor mollis. In luctus nulla id nunc dapibus, id rhoncus lorem pretium. Nunc eget fringilla velit, semper commodo diam. Suspendisse odio odio, euismod ac ornare sed, tincidunt ac arcu. Pellentesque vitae fringilla orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Part IV

Other areas of linguistics

Chapter 26

Diachronic syntax

Ulrike Demske
Universität Potsdam

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

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

1 Introduction

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

Ulrike Demske. 2018. Diachronic syntax. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 247–248. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 27

Acquisition

Jonathan Ginzburg
Université Paris Diderot

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

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

1 Introduction

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

Jonathan Ginzburg. 2018. Acquisition. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 249–250. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 28

Processing

Tom Wasow

Put abstract here with \abstract.

1 Introduction

Little psycholinguistic research has been guided by ideas from HPSG (but see Konieczny (1996) for a notable exception). This is not so much a reflection on HPSG as on the state of current knowledge of the relationship between language structure and the unconscious processes that underlie language production and comprehension. Other theories of grammar have likewise not figured prominently in theories of language processing, at least in recent decades[1] The focus of this chapter, then, will be on how well the architecture of HPSG comports with available evidence about language production and comprehension.

My argument is much the same as that put forward by Sag et al. (2003: Chapter 9), and Sag & Wasow (2011; 2015), but with some additional observations about the relationship between competence and performance. I presuppose the "competence hypothesis" (see Chomsky 1965: Chapter 1), that is, that a theory of language use (performance) should incorporate a grammar representing the knowledge of language (competence) that is drawn on in everyday comprehension and production, as well as in other linguistic activities, such as language games and the (often artificial) tasks employed in psycholinguistic experiments.

The primary reason for adopting the competence hypothesis is parsimony: a theory of language use is simpler if it does not have to repeat much the same information about about the language in both its production and comprehension

[1]Half a century ago, the Derivational Theory of Complexity (DTC) was an attempt to use psycholinguistic experiments to test aspects of the grammatical theory that was dominant at the time. The DTC was discredited in the 1970s, and the theory it purported to support has longsince been superceded. See Fodor et al. (1974) for discussion.

Tom Wasow. 2018. Processing. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 251–270. Berlin: Language Science Press. DOI:??

components. Such information would include such things as the vocabulary, the preferred word orders, and most of the rest of what linguists encode in their grammars. A performance theory that incorporates a grammar only needs to include such information once.[2]

There is also, however, an empirical reason for preferring a model with a good fit between competence and performance. As noted by Bresnan et al. (2001), preferences that are only statistical tendencies in some languages can show up in others as categorical requirements. The example they discuss in detail is the avoidance of clauses with third-person subjects but first- or second-person objects or obliques. In English, this is a powerful statistical tendency, which they document by showing that the passivization rate in the Switchboard corpus is very significantly lower when the agent is first- or second-person than when it is third-person. In Lummi (a Salish language of British Columbia), this preference is categorical: clauses with third-person subjects but first- or second-person objects or obliques are simply unacceptable. Hawkins (2004; 2014) argues that such examples are by no means exceptional, and formulates the following "Performance-Grammar Correspondence Hypothesis" (PGCH):

> Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by frequency of use and ease of processing.[3]

There are two ways in which a processing model incorporating a grammar might capture this generalization. One is to give up the widespread assumption that grammars provide categorical descriptions, and that any quantitative generalizations must be extra-grammatical; see Francis (In preparation) for arguments

[2] There are of course some discrepancies between production and comprehension that need to be accounted for in a full theory of language use. For example, most people can understand some expressions that they never use, including such things as dialect-specific words or accents. But these discrepancies are on the margins of speakers' knowledge of their languages. The vast majority of the words and structures that speakers know are used in both production and comprehension. Further, it seems to be generally true that what speakers can produce is a proper subset of what they can comprehend. Hence, the discrepancies can plausibly be attributed to performance factors such as memory or motor habits

[3] In the Bresnan et al. example, I know of no experimental evidence that clauses with third-person subjects and first- or second-person objects are difficult to process. But a plausible case can be made that the high salience of speaker and addressee makes the pronouns referring to them more accessible in both production and comprehension than expressions referring to other entities. In any event, clauses with first- or second-person subjects and third -person objects are far more frequent than clauses with the reverse pattern in languages where this has been checked. Thus, the Bresnan et al. example falls under the PGCH, at least with respect to "frequency of use" .

supporting this option, and thoughful discussion of literature on how to differentiate processing effects from grammar. For example, some HPSG feature structures might allow multiple values for the same feature, but with probabilities (adding up to 1) attached to each value.[4] I hasten to add that fleshing out this idea into a full-fledged probabilistic version of HPSG would be a large undertaking, well beyond the scope of this chapter; see Linadarki (2006), and Miyao & Tsujii (2008) for work along these lines. But the idea is fairly straightforward, and would allow, for example, English to have **in its grammar** a non-categorical constraint against clauses with third-person subjects and first- or second-person objects or obliques.

The second way for a theory adopting the competence hypothesis to represent Hawkins's PGCH would be to allow certain generalizations to be stated either as grammatical constraints (when they are categorical) or as probabilistic performance constraints. This requires a fit between the grammar and the other components of the performance model that is close enough to permit what is essentially the same generalization to be expressed in the grammar or elsewhere. In the case discussed by Bresnan et al. for example, treating the constraint in question as part of the grammar of Lummi but a matter of performance in English would require that both the theory of grammar and models of production would include, minimally, the distinction between third-person and other persons, and the distinction between subjects and non-subjects. Since virtually all theories of grammar make these distinctions, this observation is not very useful in choosing among theories of grammar. I will return later to phenomena that bear on the choice among grammatical theories, at least if one accepts the competence hypothesis.

Since its earliest days, HPSG research has been motivated in part by considerations of computational tractability (see Flickinger, Pollard & Wasow (2018), Chapter 2 of this volume, for discussion). Some of the design features of the theory can be traced back to the need to build a system that could run on the computers of the 1980s. Despite the obvious differences between human and machine information processing, some aspects of HPSG's architecture that were initially motivated on computational grounds have turned out to fit well with what is known about human language processing. A prime example of that is the computational analogue to the competence hypothesis, namely, the fact that the same grammar is used for parsing and generation. In section 3, I will discuss a

[4]I discussed this idea many times with the late Ivan Sag. He made it clear that he believed grammatical generalizations should be categorical. In part for that reason, this idea was not included in our joint publications on processing and HPSG

number of other high-level design properties of HPSG, arguing that they fit well with what is known about human language processing. In section 4, I will briefly discuss two phenomena that have been the locus of much discussion about the relationship between grammar and processing, namely, island constraints and differences between subject and object relative clauses.

2 Key Facts about Human Language Processing

In this section I review a number of well-known general properties of human language processing. Most of them seem evident from subjective experience of language use, but there is supporting experimental evidence for all of them.

2.1 Incrementality

Both language production and comprehension proceed incrementally, from the beginning to the end of an utterance. In the case of production, this is evident from the fact that utterances unfold over time. Moreover, speakers typically begin their utterances without having fully planned them out, as is evident from the prevalence of disfluencies. On the comprehension side, there is considerable evidence that listeners (and readers) begin analyzing input right away, without waiting for utterances to be complete. A grammatical theory that assigns structure and meaning to initial substrings of sentences will fit more naturally than one that doesn't into a processing model that exhibits this incrementality we see in human language use.

I hasten to add that there is also good evidence that both production and comprehension involve anticipation of later parts of sentences. While speakers may not have their sentences fully planned before they begin speaking, some planning of downstream words must take place. This is perhaps most evident from instances of nouns exhibiting quirky cases determined by verbs that occur later in the clause. For example, objects of German *helfen*, 'help', take the dative case, rather than the default accusative for direct objects. But in a sentence like (1), the speaker must know that the verb will be one taking a dative object at the time *dem* is uttered.

(1) Wir werden dem Kind bald helfen. (German)
we will the.DAT child soon help
'We will help the child soon.'

Likewise, in comprehension there is ample evidence that listeners and readers anticipate what is to come. This has been demonstrated using a variety experimental paradigms. Eye-tracking studies (see Tannenhaus, et al 1995, Altmann & Kamide 1999, Arnold et al. 2007, among many others) have shown that listeners use semantic information and world knowledge to predict what speakers will refer to next.

Thus, a theory of grammar that fits comfortably into a model of language use should provide representations of initial substrings of utterances that can be assigned (partial) meanings and be used in predicting later parts of those utterances.

2.2 Non-modularity

Psycholinguistic research over the past four decades has established that language processing involves integrating a wide range of types of information on an as-needed basis. That is, the various components of the language faculty interact throughout their operation. A model of language use should therefore **not** be modular, in the sense of Jerry Fodor's influential (1983) book, *The Modularity of Mind*.[5]

Some casual observations argue against modular language processing. For example, the famously ambiguous sentences (2) and (3) are disambiguated in speech by the stress patterns.

(2) a. I forgot how good beer tastes.
    b. Dogs must be carried.

The two meanings of (2) correspond to two different parses (one with *good* as part of the noun phrase *good beer* and the other with *how good* as a verb phrase modifier). The two meanings of (3) have the same syntactic structure, but differ in whether the requirement is that all dogs be carried, or that everyone carry a dog. This interaction of prosody with syntax (in the case of (2)) and with semantics (in the case of (3)) is produced and perceived before the end of the utterance,

[5] Much of the psycholinguistic research of the 1980s was devoted to exploring modularity – that is, the idea that the human linguistic faculty consists of a number of distinct "informationally encapsulated" modules. While Fodor's book was mostly devoted to arguing for modularity at a higher level, where the linguistic faculty was one module, many researchers at the time extended the idea to the internal organization of the linguistic faculty, positing largely autonomous mechanisms for phonology, syntax, semantics, and pragmatics, with the operations of each of these sub-modules unaffected by the operations of the others. The outcome of years of experimental studies on the linguistic modularity idea was that it was abandoned by most psycholinguists. For an early direct response to Fodor, see Marslen-Wilson & Tyler (1987).

suggesting that phonological information is available in the course of syntactic and semantic processing.

Moreover, non-linguistic knowledge influences the disambiguation in both of these cases. If (2) is preceded by "I just finished three weeks without alcohol," the natural interpretation of *good* is as a modifier of *tastes*; but following "I just finished three weeks drinking only Bud Light," *good* is more naturally interpreted as a modifier of *beer*. In the case of (3), only one interpretation (that anyone with a dog must carry it) is plausible, given our knowledge of the world. Indeed, most non-linguists fail to see the ambiguity of (3) without a lengthy explanation.

More rigorous evidence of the non-modular character of language processing has been provided by a variety of types of experiments. The work of Michael Tanenhaus and his associates, using eye-tracking to investigate the time-course of sentence comprehension, played an important role in convincing most psycholinguists that human language understanding is non-modular. See, for example, Eberhard et al. (1995), McMurray et al. (2008), Tanenhaus et al. (1995), Tanenhaus et al. (1996), and Tanenhaus & Trueswell (1995). A recent survey of work arguing against modularity in language processing is provided by Spevack et al. (2018).

2.3 Importance of Words

The individual properties of words play a central role in how people process phrases and sentences. Consider, for example, what is probably the most famous sentence in psycholinguistics, (4).

(3) The horse raced past the barn fell.

The extreme difficulty that people who have not previously been exposed to (4) have comprehending it depends heavily on the choice of words. A sentence like (5), with the same syntactic structure, is far easier to parse.

(4) The applicant interviewed in the morning left.

Numerous studies (e.g. Trueswell et al. (1993), MacDonald et al. (1994), Bresnan et al. (2007), Wasow et al. (2011)) have shown that such properties of individual words as subcategorization preferences, semantic categories (e.g. animacy), and frequency of use can influence the processing of utterances.

2.4 Influence of Context

Much of the evidence against modularity of the language faculty is based on the influences of non-linguistic context and world knowledge on language processing. The well-known McGurk effect (McGurk & MacDonald (1976)) and the Stroop effect (Stroop (1935)) demonstrate that, even at the word level, visual context can influence linguistic comprehension and production.

Linguistic context also clearly influences processing, as the discussion of examples (2) and (3) above illustrates. The same conclusion is supported by numerous controlled studies, including, among many others, those described by Crain & Steedman (1985), Altmann & Steedman (1988), Branigan (2007), **Tooley2007a** and Spevack et al. (2018). The last of these references concludes, "when humans and their brains are processing language with each other, there is no format of linguistic information (e.g., lexical, syntactic, semantic, and pragmatic) that cannot be rapidly influenced by context."

2.5 Speed and Accuracy of Processing

A good deal of psycholinguistic literature is devoted to exploring situations in which language processing encounters difficulties, notably work on garden paths (in comprehension) and disfluencies (in production). Much more striking than the existence of these phenomena, however, is how little they matter in everyday language use. While ambiguities abound in normal sentences (see Wasow (2015)), comprehenders very rarely experience noticeable garden paths. Similarly, disfluencies in spontaneous speech occur in nearly every sentence but rarely disrupt communication.

People are able to use speech to exchange information remarkably efficiently. A successful account of human language processing must explain why it works as well as it does.

3 Features of HPSG that Fit Well with Processing Facts

In this section, I review some basic design features of HPSG, pointing out ways in which they comport well with the properties of language processing listed in the previous section.

3.1 Constraint-based

Well-formedness of HPSG representations is defined by the simultaneous satisfaction of a set of constraints that constitutes the grammar. This lack of directionality allows the same grammar to be used in modeling production and comprehension.

Consider, for instance, the example of quirky case assignment illustrated in (1) above. A speaker uttering (1) would need to have planned to use the verb *helfen* before beginning to utter the object NP. But a listener hearing (1) would encounter the dative case on the article *dem* before hearing the verb and could infer only that a verb taking a dative object was likely to occur at the end of the clause. Hence, the partial mental representations built up by the two interlocutors during the course of the utterance would be quite different. But the grammatical mechanism licensing the combination of a dative object with this particular verb is the same for speaker and hearer.

In contrast, theories of grammar that utilize sequential operations to derive sentences impose a directionality on their grammars. If such a grammar is then to be employed as a component in a model of language use (as the competence hypothesis stipulates), its inherent directionality becomes part of the models of both production and comprehension. But production involves mapping meaning onto sound, whereas comprehension involves the reverse mapping. Hence, a directional grammar cannot fit the direction of processing for both production and comprehension.[6]

Branigan & Pickering (2017) argue at length that "structural priming provides an implicit method of investigating linguistic representations." They go on to conclude that the evidence from priming supports "frameworks that … assume nondirectional and constraint-based generative capacities (i.e., specifying well-formed structures) that do not involve movement." HPSG is one of the frameworks they mention that fit this description.

3.2 Surface-oriented

The features and values in HPSG representations are motivated by straightforwardly observable linguistic phenomena. HPSG does not posit derivations of observable properties from abstract underlying structures. In this sense it is surface-oriented.

The evidence linguists use in formulating grammars consists of certain types of performance data, primarily judgments of acceptability and meaning. Accounts of the data necessarily involve some combination of grammatical and processing mechanisms. The closer the grammatical descriptions are to the observable

[6]This was an issue for early work in computational linguistics that built parsers based on the transformational grammars of the time, which generated sentences using derivations whose direction went from an underlying structure largely motivated by semantic considerations to the observable surface structure. See, for example, Hobbs & Grishman (1975).

phenomena, the less complex the processing component of the account needs to be.

For example, the grammatical theory of Kayne (1994), which posits a universal underlying order of specifier-head-complement, requires elaborate (and directional) transformational derivations to relate these underlying structures to the observable data in languages whose surface order is different (a majority of the language of the world). In the absence of experimental evidence that the production and comprehension of sentences with different constituent orders involves mental operations corresponding to the grammatical derivations Kayne posits, his theory of grammar seems to be incompatible with the competence hypothesis.

Experimental evidence supports this reasoning. As Branigan & Pickering (2017) conclude, "[P]riming evidence supports the existence of abstract syntactic representations. It also suggests that these are shallow and monostratal in a way that corresponds at least roughly to the assumptions of … Pollard & Sag (1994) …. It does not support a second, underlying level of syntactic structure or the syntactic representation of empty categories associated with the movement of constituents in some transformational analyses."

3.3 Informationally Rich Representations

The feature structures of HPSG include all types of linguistic information relevant to the well-formedness and interpretation of expressions. This includes phonological, morphological, syntactic, semantic, and contextual information. They can also incorporate non-linguistic contextual information (e.g. social information), though this has not been extensively explored.

The cooccurrence of these different types of information within a single representation facilitates modeling production and comprehension processes that make reference to more than one of them. The architecture of the grammar is thus well suited to the non-modularity and context-sensitivity of language processing.

It is interesting in this regard to consider the conclusions of two papers by psycholinguists who surveyed experimental evidence and inferred what types of grammatical information was essential for processing.

The following series of quotes captures the essence of what MacDonald et al. (1994) wrote regarding lexical representation:

- "[T]he lexical representation for a word includes a representation of the word's phonological form, orthographic form, semantics, grammatical features (including grammatical category), morphology (at least inflectional), argument structure, and X-bar structure."

- "[T]he connection structure of the lexicon encodes relationships among different types of lexical information."
- "In addition to constraints that hold between various aspects of lexical representations, sentence and discourse contexts also constrain lexical representations during processing…"

With the possible exception of "X-bar structure", this sounds very much like a description of the types of information included in HPSG feature structures.

Over twenty years later, Branigan & Pickering (2017) came to the following conclusions about linguistic representations:

- "The syntactic representations capture local relationships between a 'mother' and its constituent 'daughter(s)' (e.g., a VP comprising a verb and two NPs), independent of the larger context in which the phrase appears (e.g., that the VP occurs within a subordinate clause), or the internal structure of the subphrases that constitute it (e.g., that the first NP comprises a determiner, adjective, and noun)."
- "[S]ome elements that are not phonologically represented may be syntactically represented."
- "Other priming evidence similarly indicates that some semantically specified elements are not specified syntactically."
- "[T]he semantic level of representation contains at least specifications of quantificational information, information structure, and thematic roles."
- "Evidence from priming supports a range of mappings between information encoded in the semantic representation and information encoded in the syntactic representation: between thematic roles and grammatical functions, between thematic roles and word order, between animacy and syntactic structure, and between event structures and syntactic structures."

The two lists are quite different. This is in part because the focus of the earlier paper was on lexical representations, whereas the later paper was on linguistic representations more generally. It may also be attributable to the fact that McDonald, et al, framed their paper around the issue of ambiguity resolution, while Branigan and Pickering's paper concentrated on what could be learned from structural priming studies. Despite these differences, it is striking that the conclusions of both papers about the mental representations employed in language processing are very much like those arrived at by work in HPSG.

3.4 Lexicalism

A great deal of the information used in licensing sentences in HPSG is stored in the lexical entries for words. A hierarchy of lexical types permits commonalities to be factored out to minimize what has to be stipulated in individual entries, but the information in the types gets into the representations of phrases and sentences through the words that instantiate those types. Hence, it is largely the information coming from the words that determines the well-formedness of larger expressions. Any lexical decomposition would have to be strongly motivated by the morphology.

Branigan & Pickering (2017) note that grammatical structures (what some might call "constructions") such as V-NP-NP can prime the use of the same abstract structure, even in the absence of lexical overlap. But they also note that the priming is consistently significantly stronger when the two instances share the same verb. They call this "the lexical boost." They write, "To explain abstract priming, lexicalist theories must assume that the syntactic representations … are shared across lexical entries." The types in HPSG's lexicon provide just such representations. Branigan and Pickering go on to say that the lexical boost argues for "a representation that encodes a binding between constituent structure and the lemma … of the lexical entry for the head." In HPSG, this "binding" is simply the fact that the word providing the lexical boost (say, *give*) is an instantiation of a type specifying the structures it appears in (e.g. the ditransitive verb type).

Similarly, the fact, noted in section 2.3 above, that a given structure may be more or less difficult to process depending on word choice is unsurprising in HPSG, so long as the processor has access to information about individual words and not just their types.

3.5 Underspecification

HPSG allows a class of linguistic structures that share some feature values to be characterized by means of feature structures that specify only the features whose values are shared. Such underspecification is very useful for a model of processing (particularly a model of the comprehender) because it allows partial descriptions of the utterance to be built up, based on the information that has been encountered. This property of the grammar makes it easy to incorporate into an incremental processing model.

4 Two Phenomena of Interest

4.1 Island Constraints

Ever since Ross's seminal dissertation (1967) introduced the notion of "island constraints," linguists have sought explanations for their existence, often suggesting that they were motivated by processing considerations (notably Grosu (1972), Fodor (1983), Deane (1991)). The basic idea is that island constraints restrict the search space the parser needs to consider in looking for a gap to match a filler it has encountered, thereby facilitating processing. This then raises the question of whether island constraints need to be represented in grammar (language particular or universal), or can be attributed entirely to processing and/or other factors, such as pragmatics.

In principle, this question is orthogonal to the choice among theories of grammar. But in recent years, a controversy has arisen between some proponents of HPSG and certain transformational grammarians, with the former (e.g. Chaves (2012 and 2018, Chapter 15 of this volume), Hofmeister & Sag (2010), Hofmeister, Jaeger, Arnon, Sag & Snider (2013)) arguing that certain island phenomena should be attributed entirely to extra-grammatical factors, and the latter (e.g. Phillips (2013), Sprouse et al. (2012)) arguing that island constraints are part of grammar.

I will not try to settle this dispute here. Rather, my point in this subsection is to note that a theory in which there is a close fit between the grammar and processing mechanisms allows for the possibility that some island phenomena should be attributed to grammatical constraints, whereas others should be explained in terms of processing. Indeed, if the basic idea that islands facilitate processing is correct, it is possible that some languages, but not others, have grammaticalized some islands, but not others. That is, in a theory in which the grammar is a tightly integrated component of a processing model, the question of whether a particular island phenomenon is due to a grammatical constraint is an empirical one, whose answer might differ from language to language.

Early work on islands assumed that, in the absence of negative evidence, island constraints could not be learned and hence must be innate and therefore universal. But cross-linguistic variation in island constraints, even between closely related languages, has been noted since the early days of research on the topic (see, e.g. Erteschik-Shir (1973) and Engdahl & Ejerhed (1982)).

This situation is what one might expect if languages differ with respect to the extent to which the processing factors that motivate islandhood have been grammaticalized. In short, a theory with a tight fit between its grammatical machinery and its processing mechanisms allows for hybrid accounts of islands that are not available to theories without such a fit.

One example of such a hybrid is Chaves's (2012) account of Ross's Coordinate Structure Constraint. Following much earlier work, he distinguishes between the "conjunct constraint," which prohibits a gap from serving as a conjunct in a coordinate structure (as in **What did you eat a sandwich and?*) and the "element constraint," which prohibits a gap from serving as an element of a larger conjunct (as in **What did you eat a sandwich and a slice of?*). The conjunct constraint, he argues, follows from the architecture of HPSG and is therefore built into the grammar. The element constraint, on the other hand, has exceptions and, he claims, should be attributed to extra-grammatical factors. See Chaves's chapter on islands in this volume for more detailed discussion.

4.2 Subject vs. Object Relative Clauses

One of the most discussed phenomena in the literature on human sentence processing is the difference in processing complexity between relative clauses (RCs) in which the gap is the subject and those in which the gap is the object – or, as they are commonly called, "subject RCs" and "object RCs"; see, among many others, Wanner & Maratsos (1978), Gibson (1998), Traxler et al. (2002), Gennari & MacDonald (2008). Relative clause processing complexity has been shown to be influenced by a number of other factors than the grammatical function of the gap, including the animacy and pronominality of the overt NP in the RC[7], as well as the frequency, animacy, and discourse properties of the head of the RC. When these factors are controlled for, however, most psycholinguists accept that it has been established that subject RCs are generally easier to process than object RCs, at least in English.[8]

[7] The stimuli in the experimental studies on this topic always have RCs with one overt NP, either in subject or object position and a gap corresponding to the other grammatical function.

[8] This processing difference corresponds to the top end of the "accessibility hierarchy" that Keenan & Comrie (1977) proposed as a linguistic universal. Based on a diverse sample of 50 languages, they proposed the hierarchy below, and hypothesized that any language allowing RC gaps at any point in the hierarchy would allow RC gaps at all points higher (to the left) on the hierarchy.

Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparison

One approach to explaining this asymmetry has been based on the distance between the filler and the gap (see, among others, Wanner & Maratsos (1978), Gibson (1998), Hawkins (2004)). In languages like English, with basic SVO clause order and RCs that follow the nouns they modify, the distance between the filler (the relativizer or head noun) and the gap is greater for an object gap than for a subject gap. If holding a filler in memory until the gap is encountered puts an extra burden on the processor, this could explain why object RCs are harder to process than subject RCs. This distance-based account makes an interesting prediction for languages with different word orders. In languages like Japanese with SOV order and RCs that precede the nouns they modify, the distance relationships are reversed – that is, the gaps in object RCs are closer to their fillers than those in subject RCs. The same is true of Chinese, with basic SVO order and RCs that precede the nouns they modify. So the prediction of distance-based accounts of the subject/object RC processing asymmetry is that it should be reversed in these languages.

The experimental evidence on this prediction is somewhat equivocal. While Hsiao & Gibson (2003) found a processing preference for object RCs over subject RCs in Chinese, their findings were challenged by Lin & Bever (2006) and Vasishth et al. (2013), who claimed that Chinese has a processing preference for subject RCs. In Japanese, Miyamoto & Nakamura (2003) found that subject RCs were processed more easily than object RCs. The issue remains controversial, but, for the most part, the evidence has not supported the idea that the processing preference between subject RCs and object RCs varies across languages with different word orders.

The most comprehensive treatment of English RCs in HPSG is Sag (1997). Based entirely on distributional evidence, Sag's analysis treats (finite) subject RCs as fundamentally different from RCs whose gap does not function as the subject of the RC. The difference is that the SLASH feature, which encodes information about long-distance dependencies in HPSG, plays no role in the analysis of subject RCs. Non-subject RCs, on the other hand involve a non-empty SLASH value in the RC.[9]

Sag deals with a wide variety of kinds of RCs. From the perspective of the

Keenan & Comrie speculated that the generality of this hierarchy of relativizability lay in processing, specifically on the comprehension side. The extensive experimental evidence that has been adduced in support of this idea in the intervening decades has been concentrated on subject RCs vs. (direct) object RCs. The remainder of the hierarchy remains largely untested by psycholinguists.

[9] The idea that at least some subject gaps differ in this fundamental way from non-subject gaps goes back to Gazdar (1981)

processing literature, the two crucial kinds are exemplified by (6) and (7), from Gibson (1998).

(5) a. The reporter who attacked the Senator admitted the error.
    b. The reporter who the Senator attacked admitted the error.

A well-controlled experiment on the processing complexity of subject and object RCs must have stimuli that are matched in every respect except the role of the gap in the RC. Thus, the conclusion that object RCs are harder to process than subject RCs is based on a wide variety of studies using stimuli like (6) and (7). Sag's analysis of (6) posits an empty SLASH value in the RC, whereas his analysis of (7) posits a non-empty SLASH value.

There is considerable experimental evidence supporting the idea that unbounded dependencies – that is, what HPSG encodes with the SLASH feature – add to processing complexity; see, for example, Wanner & Maratsos (1978), King & Just (1991), Kluender & Kutas (1993), Hawkins (1999). Combined with Sag's HPSG analysis of English RCs, this provides an explanation of the processing preference of subject RCs over object RCs. On such an account, the question of which other languages will exhibit the same preference boils down to the question of which other languages have the same difference in the grammar of subject and object RCs. At least for English, this is a particularly clear case in which the architecture of HPSG fits well with processing evidence.

5 Conclusion

This chapter opened with the observation that HPSG has not served as the theoretical framework for much psycholinguistic research. The observations in sections 2 through 4 argue for rectifying that situation. The fit between the architecture of HPSG and what is known about human sentence processing suggests that HPSG could be used to make processing predictions that could be tested in the lab.

To take one example, the explanation of the processing asymmetry between subject and object RCs offered above is based on a grammatical difference in the HPSG analysis: all else being equal, expressions with non-empty SLASH values are harder to process than those with empty SLASH values. Psycholinguists could test this idea by looking for other cases of phenomena that look superficially very similar but whose HPSG analyses differ with respect to whether SLASH is empty. One such case is occurs with pairs like Chomsky's famous minimal pair in (8) and (9).

(6) a. Chris is eager to please.
    b. Chris is easy to please.

Under the analysis of Pollard & Sag (1994), *to please* in (9) has a non-empty SLASH value but an empty SLASH value in (8). Processing (9) should therefore be easier. This prediction could be tested experimentally, and modern methods such as eye-tracking could pinpoint the locus of any difference in processing complexity to determine whether it corresponds to the region where the grammatical analysis involves a difference in SLASH values.

The disconnect between theoretical investigations of language structure and psycholinguistic studies is an unfortunate feature of our discipline. Because HPSG comports so well with what is known about processing, it could serve as the basis for a reconnection between these two areas of study.

References

Altmann, Gerry & Yuki Kamide. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent refrence. *Cognition* 73(7). 247–264.

Altmann, Gerry & Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30(3). 191–238.

Arnold, Jennifer E., Carla Hudson Kam & Michael K. Tanenhaus. 2007. If you say thee uh- you're describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33. 914–930.

Branigan, Holly. 2007. Syntactic priming. *Language and Linguistics Compass* 1(1–2). 1–16.

Branigan, Holly & Martin Pickering. 2017. An experimental approach to linguistic representation. *Behavioral and Brain Sciences* 40. 1–61.

Bresnan, Joan, Anna Cueni, Tatjana Nikitina & Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: KNAW.

Bresnan, Joan, Shipra Dingare & Christopher D. Manning. 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 01 Conference.* Stanford, CA: CSLI.

Chaves, Rui P. 2012. On the grammar of extraction and coordination. *Natural Language and Linguistic Theory* 30(2). 465–512.

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

Crain, Stephen & Mark J. Steedman. 1985. On not being led up the garden path: The use of context by the psychological syntax processor. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language processing*, 320–358. Cambridge, UK: Cambridge University Press.

Deane, Paul D. 1991. Limits to attention: A cognitive theory. *Cognitive Linguistics* 2(1). 1–63.

Eberhard, Kathleen M., Michael J. Spivey-Knowlton, Julie C. Sedivy & Michael K. Tannenhaus. 1995. Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguist Research* 24. 409–436.

Engdahl, Elisabeth & Eva Ejerhed. 1982. *Readings on unbounded dependencies in Scandinavian languages*. Stockholm: Almqvist & Wiksell International.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT dissertation.

Fodor, Janet Dean. 1983. Phrase structure parsing and the island constraints. *Linguisitcs and Philosophy* 6. 163–223.

Fodor, Jerry A., Thomas G. Bever & Merrill F. Garrett. 1974. *The psychology of language: An introduction to psycholinguistics and Generative Grammar*. New York: McGraw-Hill Book Co.

Francis, Elaine. In preparation. *Gradient acceptability and linguistic theory*.

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

Gennari, Silvia P. & Maryellen C. MacDonald. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 8. 161–187.

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68(1). 1–76.

Grosu, A. 1972. *The strategic content of island constraints* (Working Papers in Linguistics 13). Ohio Sate University.

Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75(2). 244–285.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hawkins, John A. 2014. *Cross-linguistic variation and efficiency*. Oxford: Oxford University Press.

Hobbs, Jerry R. & Ralph Grishman. 1975. The automatic transformational analysis of English sentences: An implementation. *International Journal of Computer Mathematics* 5(1–4). 267–283.

Hofmeister, Philip, T. Florian Jaeger, Inbal Arnon, Ivan A. Sag & Neal Snider. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28(1–2). 48–87.

Hofmeister, Philip & Ivan A. Sag. 2010. Cognitive constraints and island effects. *Language* 86(2).

Hsiao, Franny & Edward Gibson. 2003. Processing relative clauses in chinese. *Cognition* 90. 3–27.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

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

King, Jonathan & Marcel Just. 1991. Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30. 580–602.

Kluender, Robert & Marta Kutas. 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience* 5(2). 196–214.

Konieczny, Lars. 1996. *Human sentence processing: A semantics-oriented parsing approach*. Universität Freiburg Dissertation. IIG-Berichte 3/96.

Lin, Chien-Jer Charles & Thomas Bever. 2006. Subject preference in the processing of relative clauses in Chinese. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 254–260. Somerville, MA: Cascadilla Proceedings Project.

Linadarki, Evita. 2006. *Linguistic and statistical extensions of data oriented parsing*. University of Essex dissertation. Unpublished.

MacDonald, Maryellen C., Neal J. Pearlmutter & Mark. S. Seidenberg. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4). 676–703.

Marslen-Wilson, William & Lorraine Tyler. 1987. Against modularity. In Jay L. Garfield (ed.), *Modularity in knowledge representation and Natural Language Processing*, 37–62. Cambridge, MA: MIT Press.

McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264(5588). 746–748.

McMurray, Bob, Meghan A. Clayards, Michael K. Tanenhaus & Richard Aslin. 2008. Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic Bulletin* 15(6). 1064–1071.

Miyamoto, Edson T & Michiko Nakamura. 2003. *Subject/object asymmetries in the processing of relative clauses in Japanese*. Gina Garding & Mimu Tsujimura (eds.). Somerville, MA.

Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.

Phillips, Colin. 2013. On the nature of island constraints I: Language processing ad reductionist accounts. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and island effects*, 64–108. Cambridge: Cambridge University Press.

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

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT dissertation. http://files.eric.ed.gov/fulltext/ED016965.pdf, accessed 2018-2-25. Reproduced by the Indiana University Linguistics Club and later published as Ross (1986).

Ross, John Robert. 1986. *Infinite syntax!* Norwood, New Jersey: Ablex Publishing Corporation.

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

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/Cambridge, MA: Blackwell Publishing Ltd.

Sag, Ivan A. & Thomas Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguistic Research* 44(1). 47–63.

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

Spevack, Samuel C., J. Benjamin Falandays & Michael J. Spivey. 2018. Interactivity of language. *Language and Linguistics Compass* 12(7). e12282.

Sprouse, Jon, Matt Wagers & Colin Phillips. 2012. A test of the relation between working memory capacity and syntactic island effects. *Language* 88(1). 82–123.

Stroop, John Ridley. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18(6). 643–662.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1996. Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In Toshio Inui & James L. McClelland (eds.), *Information integration in perception and communication* (Attention and Performance XVI), 457–478. Cambridge, MA: MIT Press.

Tanenhaus, Michael K. & John C. Trueswell. 1995. Sentence comprehension. In Joanne L. Miller & Peter D. Eimas (eds.), *Handbook of cognition and perception*. San Diego, CA: Academic Press.

Traxler, Matthew J., Robin K. Morris & Rachel E. Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47(1). 69–90.

Trueswell, John C., Micahel K. Tanenhaus & Christopher Kello. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19(3). 528–553.

Vasishth, Shravan, Chen Zhong, Qiang Li & Gueilan Guo. 2013. Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS ONE* 8(10). e77006.

Wanner, Eric & Michael Maratsos. 1978. An ATN approach to comprehension. In Morris Halle, Joan Bresnan & George A. Miller (eds.), *Linguistic theory and psychological reality*, 119–161. Cambridge, MA: MIT Press.

Wasow, Thomas. 2015. Ambiguity avoidance is overrated. In Susanne Winkler (ed.), *Ambiguity: Language and communication*, 29–47. Berlin: De Gruyter.

Wasow, Thomas, Florian T. Jaeger & David Orr. 2011. Lexical variation in relativizer frequency. In Horst J. Simon & Heike Wiese (eds.), *Expecting the unexpected: Exceptions in grammar*, 175–195. Berlin: De Gruyter.

Chapter 29

Computational linguistics and grammar engineering

Emily M. Bender
University of Washington

Guy Emerson
University of Cambridge

We discuss the relevance of HPSG for computational linguistics, and the relevance of computational linguistics for HPSG.

1 Introduction

From the inception of HPSG in the 1980s, there has been a close integration between theoretical and computational work (for an overview, see Flickinger, Pollard & Wasow 2018, Chapter 2 of this volume). In this chapter, we discuss computational work in HPSG, starting with the infrastructure that supports it (both theoretical and practical) in Section 2. Next we describe several existing large-scale projects which build HPSG or HPSG-inspired grammars (see Section 3) and the deployment of such grammars in applications including both those within linguistic research and otherwise (see Section 4). Finally, we turn to linguistic insights gleaned from broad-coverage grammar development (see Section 5).

2 Infrastructure

2.1 Theoretical considerations

There are several properties of HPSG as a theory that make it well-suited to computational implementation. First, the theory is kept separate from the formalism:

Emily M. Bender & Guy Emerson. 2018. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 271–280. Berlin: Language Science Press. DOI:??

the formalism is expressive enough to encode a wide variety of possible theories. While some theoretical work does argue for or against the necessity of particular formal devices (e.g. the shuffle operator (**FIXME-Reape**)), much of it proceeds within shared assumptions about the formalism. This is in contrast to work in the context of the Minimalist Program (Chomsky 1993), where theoretical results are typically couched in terms of modifications to the formalism itself. From a computational point of view, the benefit of differentiating between theory and formalism is that it means that the formalism is relatively stable. That in turns enables the development and maintenance of software systems that target the formalism, for parsing, generation, and grammar exploration (see Section 2.3 below for some examples).[1]

A second important property of HPSG that supports a strong connection between theoretical and computational work is an interest in both so-called 'core' and so-called 'peripheral' phenomena. Most implemented grammars are built with the goal of handling naturally occurring text.[2] This means that they will need to handle a wide variety of linguistic phenomena not always treated in theoretical syntactic work (**FIXME-Baldwin-et-al-Beauty**). A syntactic framework that excludes research on 'peripheral' phenomena as uninteresting provides less support for implementational work than does one, like HPSG or Construction Grammar, that values such topics (for a comparison of HPSG and Construction Grammar, see Müller 2018, Chapter 36 of this volume).

Finally, the type hierarchy characteristic of HPSG lends itself well to developing broad-coverage grammars which are maintainable over time (**FIXME-find-cite?**) The use of the type hierarchy to manage complexity at scale comes out of the work of Flickinger (1987) and others at HP labs in the project where HPSG was originally developed. The core idea is that any given constraint is (ideally) expressed only once on types which serve as supertypes to all entities that bear that constraint.[3] Such constraints might represent broad generalizations that apply to many entities or relatively narrow, indiosyncratic properties. By isolating any given constraint on one type (as opposed to repeating it in mutiple places), we build grammars that are easier to update and adapt in light of new data that

[1]There are implementations of Minimalism, notably **FIXME-Stabler** and **FIXME-Indianadiss** However, writing an implementation requires fixing the formalism, and so these are unlikely to be useful for testing theoretical ideas as the theory moves on.

[2]Though it is possible to do implementation work strictly against testsuites of sentences constructed specifically to focus on phenomena of interest.

[3]Originally this only applied to lexical entries in Flickinger's work. Now it also applies phrase structure rules, lexical rules, and types below the level of the sign which are used in the definition of all of these.

require refinements to constraints. Having a single locus for each constraint also makes the types a very useful target for documentation (**FIXME:LTDB**) and grammar exploration (**FIXME:typediff**).

2.2 Practical considerations

The formalism of HPSG allows practical implementations, since feature structures are well-defined data structures. Furthermore, because HPSG is defined to be bi-directional, an implemented grammar can be used for both parsing and generation. In this section, we discuss how HPSG allows tractable algorithms, which enables linguists to empirically test hypotheses, and which also enables HPSG grammars to be used in a range of applications, as we will see in Sections 4.1 and 4.2, respectively.

2.2.1 Computational complexity

One way to measure how easy or difficult it is to use a syntactic theory is to consider the *computational complexity* of parsing and generation algorithms. For example, we can consider how much computational time a parsing algorithm needs to process a particular sentence. For longer sentences, we would expect the amount of time to increase, but the more complex the algorithm is, the more quickly the amount of time increases. If we consider sentences containing *n* tokens, we can find the average amount of time taken, or the longest amount of time taken. We can then increase *n*, and see how the amount of time changes, both in the average case, and in the worst case.

At first sight, analysing computational complexity would seem to paint HPSG in a bad light, because the formalism allows us to write grammars which can be arbitrarily complex; this means that the formalism can be called *Turing-complete* (Johnson 1988: Section 3.4). However, as discussed in the previous section, there is a clear distinction between theory and formalism. Although the feature-structure formalism rules out the possibility of efficient algorithms that could cope with any possible feature-structure grammar, a particular theory (or a particular grammar) might well allow efficient algorithms.

The difference between theory and formalism becomes clear when comparing HPSG to other computationally-friendly frameworks, such as Combinatory Categorial Grammar (CCG),[4] or Tree Adjoining Grammar (TAG; Joshi 1987; Schabes et al. 1988)). The formalisms of CCG and TAG inherently limit computational

[4]For an introduction, see Steedman & Baldridge (2011). For a comparison with HPSG, see Kubota (2018), Chapter 33 of this volume.

complexity: for both of them, as the sentence length n increases, worst-case parsing time is proportional to n^6 (Kasami et al. 1989). This is a deliberate feature of these formalisms, which aim to be just expressive enough to capture human language, and not any more expressive. Building this kind of constraint into the formalism itself highlights a different school of thought from HPSG. Indeed, Müller (2015) explicitly argues in favor of developing linguistic analyses first, and improving processing efficiency second. As discussed above in Section 2.1, separating the formalism from the theory means that the formalism is stable, even as the theory develops.

It would beyond the scope of this chapter to give a full review of parsing algorithms, but it is instructive to give an example. For grammars that have a context-free backbone (we can express every derivation as a phrase-structure tree plus constraints between mother and daughter nodes), it is possible to adapt the standard chart-parsing algorithm for context-free grammars. The basic idea is to parse "bottom-up" through the tree, starting by finding analyses for each token in the input, and then finding analyses for increasingly longer sequences of tokens, until we reach the entire sentence. The resulting algorithm is more computationally complex than for a context-free grammar, because we are dealing with feature structures, rather than nonterminal symbols. While a context-free grammar allows a finite number of nonterminals, a feature-structure grammar may allow a infinite number of possible feature structures. For an HPSG grammar without recursive unary rules, this algorithm has a worst-case complexity of exponential time. This is less complex than for an arbitrary grammar (which means that this class of grammars is *not* Turing-complete), but more complex than for CCG or TAG. However, when parsing real corpora, it turns out that the average-case complexity is much better than we might expect. On the one hand, grammatical constructions do not generally combine in the worst-case way, and on the other hand, when a grammar writer is confronted with multiple possible analyses for a particular construction, they may opt for the analysis that is more efficient for a particular parsing algorithm.

2.2.2 Parse ranking

For an ambiguous sentence, a grammar gives multiple valid parses. This is widely known, with attachment ambiguities and coordination ambiguities being particularly well-known examples. However, people are naturally very good at resolving ambiguity, which means most ambiguity is not apparent, even to linguists. It is only with the development of large-scale grammars that the sheer scale of ambiguity has become clear. For example, (1) might seem unambiguous, but there is a second reading, where *my favorite* is the topicalized object of *speak*, which would mean that town criers generally speak the speaker's favorite thing (perhaps a language) clearly. There is also a third, even more implausible reading, where *my favorite town* is the topicalized object. Such implausible readings don't easily come to mind, but with increasingly long sentences, such ambiguities stack up very quickly. For (2), the first line of a newspaper article,[5] the 1214 version of the English Resource Grammar (Flickinger 2000; 2011) gives TODO readings.

(1) My favorite town criers speak clearly.

(2) A small piece of bone found in a cave in Siberia has been identified as the remnant of a child whose mother was a Neanderthal and father was a Denisovan, a mysterious human ancestor that lived in the region.

In practical applications, considering all possible parses can be infeasible, and we may want to automatically disambiguate each sentence. This can be done by *ranking* the parses, so that the application only uses the most highly-ranked parse, or the top *N* parses. Parse ranking is not usually determined by the grammar itself, because of the difficulty of manually writing disambiguation rules. Typically, a statistical system is used (Oepen et al. 2004). First, a corpus is *treebanked*: for each sentence in the corpus, an annotator (often the grammar writer) chooses the best parse, out of all parses produced by the grammar. The set of all parses for a sentence is often referred to as the *parse forest*, and the selected best parse is often referred to as the *gold standard*. Given the gold parses for the whole corpus, a statistical system is trained to predict the gold parse from a parse forest, based on many features[6] of the parse. From the example in (1), we can see how a number of a different features all influence the preferred interpretation: the likelihood of a construction (such as topicalization), the likelihood of a valence frame (such as transitive *speak*), the likelihood of a collocation (such as *town crier*), the likelihood of a semantic relation (such as speaking a town), and so on.

Because of the large number of possible parses, it can be helpful to *prune* the search space: rather than ranking the full set of parses, we can restrict attention to a smaller set of parses, which hopefully includes the correct parse. By carefully choosing how to restrict our attention, we can drastically reduce processing time without hurting parsing accuracy. One method, called *supertagging*,[7]

[5] https://www.theguardian.com/science/2018/aug/22/offspring-of-neanderthal-and-denisovan-identified-for-first-time

[6] In the machine learning sense of "feature", not the feature-structure sense.

[7] The name refers to *part-of-speech tagging*, which predicts a part-of-speech for each input token, from a relatively small set of part-of-speech tags. Supertagging is "super", in that it predicts detailed lexical entries, rather than simple tags.

exploits the fact that HPSG is a lexicalized theory. Rather than considering all possible lexical entries for each token (which leads to high levels of ambiguity), supertagging aims to choose the correct lexical entry for each token, based on the surrounding context. Although there is a chance that the supertagger will predict the wrong lexical entry, using a supertagger can often improve parsing accuracy, by ruling out parses that the parse-ranking model might incorrectly rank too high. Supertagging was first applied to HPSG by Matsuzaki et al. (2007), building on previous work for TAG (Bangalore & Joshi 1999) and CCG (Clark & Curran 2004). To allow multi-word expressions (such as *by and large*), where the grammar assigns a single lexical entry to multiple tokens, Dridan (2013) has proposed an extension of supertagging, called *ubertagging*, which jointly predicts both a segmentation of the input and supertags for those segments. Dridan manages to increase parsing speed by a factor of four, while also improving parsing accuracy.

2.2.3 Semantic dependencies

…

Figure 1: Comparison of MRS, DMRS, and DM.

In practical applications of HPSG grammars, the full derivation trees and the full feature structures are often unwieldy, containing far more information than necessary for the task at hand. It is therefore often desirable to extract a concise semantic representation.

In computational linguistics, a popular approach to semantics is to represent the meaning of a sentence as a *dependency graph*, as this enables the use of graph-based algorithms.[8] Several types of dependency graph have been proposed based on Minimal Recursion Semantics (MRS; Copestake et al. 2005), with varying levels of simplification. The most expressive is Dependency Minimal Recursion Semantics (DMRS; Copestake 2009), which is fully interconvertible with MRS.[9] In contrast, Elementary Dependency Structures (EDS; Oepen & Lønning 2006) lose some scope information, which, for many applications, is less important than predicate-argument structure. Finally, Delph-in MRS Dependencies (DM;

[8]In this section, we are concerned with *semantic* dependencies. For *syntactic* dependencies, see Hudson (2018), Chapter 35 of this volume.

[9]More precisely, there is a one-to-one correspondence between DMRS and MRS structures, if every predicate as a unique *intrinsic variable*. As observed by Oepen & Lønning (2006), this allows a variable-free semantic representation, by replacing each reference to a variable with a reference to the corresponding predicate.

Ivanova et al. 2012) express predicate-argument structure purely in terms of the surface tokens, without introducing any abstract predicates. A comparison of MRS, DMRS, and DM is given in Figure 1. The existence of such dependency graphs has made it easier to use HPSG grammars in a number of practical tasks, as we will see in Section 4.2.

2.3 A brief history of HPSG grammar engineering

History: PAGE, VerbMobil, ??

Current platforms:

- LKB/ACE/PET/Agree
- Trale
- Other

3 Development of HPSG resources

- CoreGram
- DELPH-IN consortium
  - ERG
  - Other large-ish grammars
  - Grammar Matrix
- Systems inspired by HPSG:
  - Alpino
  - Enju

4 Deployment of HPSG resources

4.1 Language documentation and linguistic hypothesis testing

Deployment for linguistic goals.

4.1.1 CoreGram

4.1.2 Grammar Matrix

4.1.3 AGGREGATION

4.1.4 Derived resources: Redwoods-style treebanks

4.2 Downstream applications

Deployment for other tasks. Large number of applications. Focus on several important applications below. See also http://moin.delph-in.net/DelphinApplications

4.2.1 Language teaching

Redbird (McGraw-Hill)

4.2.2 NLP tasks

Information extraction

Summarisation

Machine translation

4.2.3 Data for machine learning

Unlike the previous sections, not providing analyses, but providing data. Not used at test time, only at training time.

Training deep learning systems – semantic parsing – skip over HPSG, go straight to semantic representations

Evaluation of deep learning systems – ShapeWorld – use a grammar to produce annotations

4.3 Other?

Alpino?

Enju?

5 Linguistic Insights

- Ambiguity
- Long-tail phenomena (raising and control?)

- Scaling up (thematic roles)
- CLIMB methodology

6 Summary

Abbreviations

Acknowledgements

References

Bangalore, Srinivas & Aravind K Joshi. 1999. Supertagging: an approach to almost parsing. *Computational Linguistics* 25(2). 237–265. http://aclweb.org/anthology/J99-2004.

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

Clark, Stephen & James R Curran. 2004. The importance of supertagging for wide-coverage CCG parsing. In *Proceedings of the 20th international conference on computational linguistics (COLING)*. http://aclweb.org/anthology/C04-1041.

Copestake, Ann. 2009. Slacker semantics: why superficiality, dependency and avoidance of commitment can be the right way to go. In *Proceedings of the 12th conference of the European chapter of the Association for Computational Linguistics (EACL)*, 1–9. http://aclweb.org/anthology/E09-1001.

Copestake, Ann, Daniel P. Flickinger, Carl J. Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An introduction. *Research on Language and Computation* 4(3). 281–332.

Dridan, Rebecca. 2013. Ubertagging: joint segmentation and supertagging for English. In *Proceedings of the 2013 conference on empirical methods in natural language processing (EMNLP)*, 1201–1212. http://aclweb.org/anthology/D13-1120.

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

Flickinger, Daniel P. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering* 6(1). 15–28.

Flickinger, Daniel P. 2011. Accuracy vs. robustness in grammar engineering. In Emily M. Bender & Jennifer E. Arnold (eds.), *Language from a cognitive perspective: Grammar, usage, and processing*, 31–50. Stanford, CA: CSLI Publications.

Ivanova, Angelina, Stephan Oepen, Lilja Øvrelid & Dan Flickinger. 2012. Who did what to whom?: A contrastive study of syntacto-semantic dependencies. In *Proceedings of the 6th linguistic annotation workshop*, 2–11. http://aclweb.org/anthology/W12-3602.

Johnson, Mark. 1988. *Attribute-value logic and the theory of grammar* (CSLI Lecture Notes 16). Stanford, CA: CSLI Publications.

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

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

Matsuzaki, Takuya, Yusuke Miyao & Jun'ichi Tsujii. 2007. Efficient HPSG parsing with supertagging and CFG-filtering. In *Proceedings of the 20th international joint conference on artificial intelligence*, 1671–1676. https://www.aaai.org/Papers/IJCAI/2007/IJCAI07-270.pdf.

Müller, Stefan. 2015. The CoreGram project: Theoretical linguistics, theory development and verification. *Journal of Language Modelling* 3(1). 21–86. DOI:10.15398/jlm.v3i1.9

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

Oepen, Stephan & Jan Tore Lønning. 2006. Discriminant-based MRS banking. In *Proceedings of the 5th international conference on language resources and evaluation (LREC)*, 1250–1255. http://www.lrec-conf.org/proceedings/lrec2006/pdf/364_pdf.pdf.

Schabes, Yves, Anne Abeillé & Aravind K. Joshi. 1988. *Parsing strategies with 'lexicalized' grammars: Application to Tree Adjoining Grammars.* Technical Report MS-CIS-88-65. University of Pennsylvania Department of Computer & Information Science.

Steedman, Mark & Jason Baldridge. 2011. Combinatory Categorial Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: formal and explicit models of grammar*, chap. 5, 181–224. Blackwell. https://www.research.ed.ac.uk/portal/files/24976508/SteedmanBaldridgeNTSyntax.pdf.

Chapter 30

Sign languages

Markus Steinbach
Georg-August-Universität Göttingen

Anke Holler
Georg-August-Universität Göttingen

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

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

1 Introduction

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

Markus Steinbach & Anke Holler. 2018. Sign languages. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 281–282. Berlin: Language Science Press. DOI:??

orci. Donec faucibus metus dui, nec iaculis purus pellentesque sit amet. Sed fermentum lorem non augue cursus, eu accumsan risus ullamcorper. Suspendisse rhoncus magna vitae enim pellentesque, eget porttitor quam finibus. Nunc ultricies turpis at quam vehicula, at tempus justo molestie. Proin convallis augue ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

$$\begin{bmatrix} \textit{some-type} & \\ \text{FEAT-A} & \boxed{10} \begin{bmatrix} \textit{type-a} & \\ \text{FEAT-AA} & \textit{type-aa} \\ \text{FEAT-AB} & \left\langle \begin{bmatrix} \text{SYNSEM|LOC|CAT|HEAD} & \textit{type-aba} \\ \text{FEAT-ABC} & \textit{type-abc} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix} \\ \text{FEAT-B} & \boxed{10}\textit{type-b} \end{bmatrix}$$

$$\begin{bmatrix} \textit{some-type} & \\ \text{FEAT-A} & \boxed{10} \begin{bmatrix} \textit{type-a} & \\ \text{FEAT-AA} & \textit{type-aa} \\ \text{FEAT-AB} & \left\langle \begin{bmatrix} \text{SYNSEM|LOC|CAT|HEAD} & \textit{type-aba} \\ \text{FEAT-ABC} & \textit{type-abc} \end{bmatrix}, \text{NP} \right\rangle \end{bmatrix} \\ \text{FEAT-B} & \boxed{10}\textit{type-b} \end{bmatrix}$$

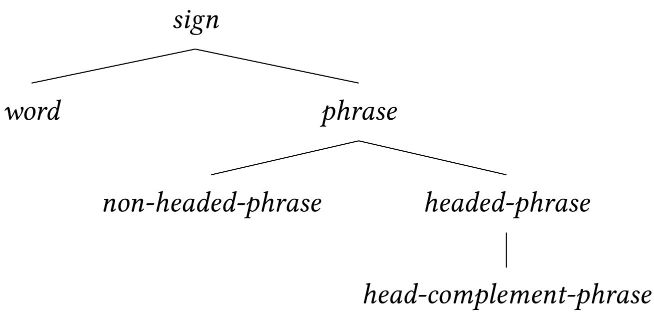


Figure 1: Type hierarchy for *sign*

Abbreviations

Acknowledgements

Chapter 31

Gesture

Andy Lücking
Goethe-Universität Frankfurt

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

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

1 Introduction

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

Andy Lücking. 2018. Gesture. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 283–284. Berlin: Language Science Press. DOI:??

ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Part V

The broader picture

Chapter 32

HPSG and Minimalism

Bob Borsley
University of Essex

Stefan Müller
Humboldt-Universität zu Berlin

Put abstract here with \abstract.

1 Introduction

The Minimalist framework, which was first outlined by Chomsky in the early 1990s (Chomsky 1993; 1995b), still seems to be the dominant approach to syntax. It is important, therefore, to consider how HPSG compares with this framework. The issues are clouded by the rhetoric that surrounds the framework. At one time "virtual conceptual necessity" was said to be its guiding principle. A little later, it was said to be concerned with the "perfection of language", with "how closely human language approaches an optimal solution to design conditions that the system must meet to be usable at all" Chomsky (2002: 58). Much of this rhetoric seems designed to suggest that Minimalism is quite different from other approaches and should not be assessed in the same way. In the words of Postal Postal (2003: 19), it looks like "an attempt to provide certain views with a sort of privileged status, with the goal of placing them at least rhetorically beyond the demands of serious argument or evidence". However, the two frameworks have enough in common to allow meaningful comparisons.

Both frameworks seek to provide an account of what is and is not possible both in specific languages and in language in general. Moreover, both are concerned not just with local relations such as that between a head and its complement or complements but also with non-local relations such as those in the following:

Bob Borsley & Stefan Müller. 2018. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 287–319. Berlin: Language Science Press. DOI:??

(1) The student knows the answer.

(2) It seems to be raining,

(3) Which student do you think knows the answer?

In (1), *the student* is subject of *thinks* and is responsible for the fact that *thinks* is a third person singular form, but they are not sisters if *knows* and *the answer* form a VP. In (2) the subject is it because the complement of *be* is *raining*, but it and raining are obviously not sisters. Finally, in (3), *which student* is understood as the subject of *thinks* and is responsible for the fact that it is third person singular, but again the two elements are structurally quite far apart. Both frameworks provide analyses for these and other central syntactic phenomena, and it is quite reasonable to compare them and ask which is the more satisfactory.[1]

Although HPSG and Minimalism have enough in common to permit comparisons, there are obviously many differences. Some are more important than others, and some relate to the basic approach and outlook, while others concern the nature of grammatical systems and syntactic structures. In this chapter we will explore the full range of differences.

The chapter is organized as follows. In Section 2, we look at differences of approach between the two frameworks. Then in Section 3, we consider the quite different views of grammar that the two frameworks espouse, and in Section 4, we look at the very different syntactic structures which result. Finally, in Section 5.2, we will look at a further issue which deserves some attention.

2 Differences of approach and outlook

As many of the chapters in this volume have emphasized, HPSG is a framework which places considerable emphasis on detailed formal analyses of the kind that one might expect within generative grammar. Thus, it is not uncommon to find lengthy appendices setting out formal analyses. See, for example, Sag's (1997) paper on English relative clauses and especially Ginzburg & Sag (2000), which has a 50 page appendix. One consequence of this, discussed by Bender & Emerson (2018), Chapter 29 of this volume, is that HPSG has had considerable influence in computational linguistics.

In Minimalism things are very different. Detailed formal analyses are virtually non-existent. There appear to be no appendices like those in Sag (1997) and

[1]As noted below, comparison is complicated somewhat by the fact that Minimalists typically provides only sketches of analyses in which various details are left quite vague.

Ginzburg & Sag (2000). In fact the importance of formalization has long been downplayed in Chomskyan work. Thus, in a 1980 conversation, Chomsky remarked that "I do not see any point in formalizing for the sake of formalizing" (see Huybregts & Riemsdijk 1982: 73), and this view seems fairly standard within Minimalism (see also Chomsky (1990: 146) and the discussion in Müller (2016: Section 3.6.2)). Chomsky & Lasnik (1995: 28) attempt to justify the absence of detailed analyses when they suggest that providing a rule system from which some set of phenomena can be derived is not "a real result" since "it is often possible to devise one that will more or less work". Instead, they say, 'the task is now to show how the phenomena … can be deduced from the invariant principles of UG with parameters set in one of the permissible ways'. In other words, providing detailed analyses is a job for unambitious drudges, and real linguists pursue a more ambitious agenda. Postal (2004: 5) comments that what we see here is "the fantastic and unsupported notion that descriptive success is not really that hard and so not of much importance". He points out that if this were true, one would expect successful descriptions to be abundant within transformational frameworks. However, he suggests that "the actual descriptions in these frameworks so far are not only not successful but so bad as to hardly merit being taken seriously". Postal does much to justify this assessment with detailed discussions of Chomskyan work on strong crossover phenomena and passives in Chapters 7 and 8 of his book.

There has also been a strong tendency to focus on just a subset of the facts in whatever domain is being investigated. As Culicover & Jackendoff (2005: 535) note, "much of the fine detail of traditional constructions has ceased to garner attention". This tendency has sometimes been buttressed by a distinction between core grammar, which is supposedly a fairly straightforward reflection of the language faculty, and a periphery of marked constructions, which are of no great importance and which can reasonably be ignored. However, as Culicover (1999) and others have argued, there is no evidence for a clear cut distinction between core and periphery. It follows that a satisfactory approach to grammar needs to account both for such core phenomena as *wh*-interrogatives, relative clauses, and passives but also with more peripheral phenomena such as the following:

(4) a. It's amazing the people you see here.
    b. The more I read, the more I understand.
    c. Chris lied his way into the meeting.

These exemplify the nominal extraposition construction (Michaelis & Lambrecht 1996), the comparative correlative construction Borsley (2011), and the *X's Way* construction (Sag 2012). As has been emphasized in other chapters, the HPSG system of types and constraints is able to accommodate broad linguistic generalizations and highly idiosyncratic facts and everything in between.

The general absence in Minimalism of detailed formal analyses is quite important. It means that Minimalists may not be fully aware of the complexity of the structures they are committed to and allows them to sidestep the question whether it is really justified. It also allows them to avoid the question of whether the very simple conception of grammar that they favour is really satisfactory. Finally, it may be that they are unaware of how many phenomena remain unaccounted for. These are all important matters.

The general absence of detailed formal analyses has also led to Minimalism having little impact on computational linguistics. There has been some work that has sought to implement Minimalist ideas (Stabler 2001; Fong & Ginsburg 2012; Fong 2014), but Minimalism has not had anything like the productive relation with computational work that HPSG has enjoyed. Existing implementations are rather toy grammars analyzing very simple sentences and some do not even this and require pre-segmented input. See Müller (2016: Section 4.7.2) for discussion.

There are, then, issues about the quantity of data that is considered in Minimalist work. There are also issues about its quality. Research in HPSG is typically quite careful about data and often makes use of corpus and experimental data. Research in Minimalism is often rather less careful. In a review of a collection of Minimalist papers, Bender (2002: 434) comments that: "In these papers, the data appears to be collected in an off-hand, unsystematic way, with unconfirmed questionable judgments often used at crucial points in the argumentation". She goes on to suggest that the framework encourages "lack of concern for the data, above and beyond what is unfortunately already the norm in formal syntax, because the connection between analysis and data is allowed to be remote." Similar things could be said about a variety of Minimalist work. Consider, for example, Aoun & Li (2003), who argue for quite different analyses of *that*-relatives and *wh*-relatives on the basis of the following (supposed) contrasts, which appear to represent nothing more than their own judgements:

(5) a. The headway that Mel made was impressive.
    b. ?? The headway which Mel made was impressive.

(6) a. We admired the picture of himself that John painted in art class.
    b. * We admired the picture of himself which John painted in art class.

(7) a. The picture of himself that John painted in art class is impressive.

b. *? The picture of himself which John painted in art class is impressive.

None of the native speakers we have consulted find significant contrasts here which could support different analyses.

This material comes from my Lexington paper. Maybe there are other examples it would be good to use. You commented that 'One has to be careful here, since there are people who do work with corpora and experiments' Maybe this should be noted. This needs to be replaced by something better.

There are also differences in the kind of arguments that the two frameworks find acceptable. It is common within Minimalism to assume that some phenomenon which cannot be readily observed in some languages must be part of their grammatical system because it is clearly present in other languages. Notable examples would be case or agreement. This stems from the longstanding Chomskyan assumption that language is the realization of a complex innate language faculty. From this perspective, there is much in any grammatical system that is a reflection of the language faculty and not in any simple way of the observable phenomena of the language in question. If some phenomenon plays an important role in many languages it is viewed as a reflection of the language faculty, and hence it must be a feature of all grammatical systems even those in which it is hard to see any evidence for it. An example – taken from a textbook on Minimalism (Hornstein, Nunes & Grohmann 2005: 124) – is an analysis of prepositional phrases in English. Figure 1 shows the analysis. Due to theory internal assumptions the case requirement of the preposition cannot be checked in the P-DP combination. According to the version of the theory explained by the authors, case has to be checked in specifier positions. Therefore it was assumed that the preposition moves to an Agr head and the DP moves to the specifier position of this Agr head. The problem is of course that DP and P are in the wrong order now. However, the authors argue that this is the order that is manifested in Hungarian and that Hungarian is a language which has postpositions and these are agreeing with their nominal dependent. It is claimed that the movement exists both in Hungarian and in English but that the movement is covert (that is, invisible) in the latter language.

@Bob, please check what I inserted here

This line of argument would be reasonable if a complex innate language faculty was an established fact, but it isn't, and since Hauser, Chomsky & Fitch (2002), it seems to have been rejected within Minimalism. It follows that ideas about an innate language faculty should not be used to guide research on individual languages. Rather, as Müller (2015: 25) puts it, "grammars should be motivated

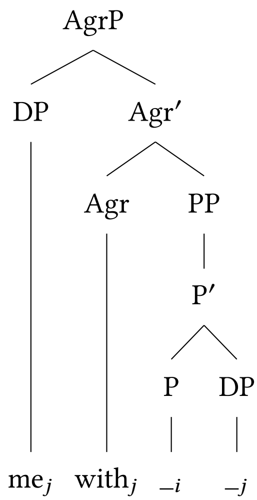


Figure 1: Minimalist analysis of a PP according to (Hornstein, Nunes & Grohmann 2005: 124)

on a language-specific basis". Does this mean that other languages are irrelevant when one investigating a specific language? Clearly not. As Müller also puts it, "In situations where more than one analysis would be compatible with a given dataset for language X, the evidence from language Y with similar constructs is most welcome and can be used as evidence in favor of one of the two analyses for language X." (2015: 43). In practice, any linguist working on a new language will use apparently similar phenomena in other languages as a starting point. It is important, however, to recognize that apparently similar phenomena may turn out on careful investigation to be significantly different.[2]

3 Different views of grammar

We turn now to more substantive differences between HPSG and Minimalism, differences in their conceptions of grammar, especially syntax, and differences in their views of syntactic structure. As we will see, these differences are related. In this section we consider the former, and in the next we will look at the latter.

As has been emphasized throughout this volume, HPSG assumes a declarative

[2]Equally, of course, apparently rather different phenomena may turn out on careful investigation to be quite similar. For further discussion of HPSG and comparative syntax, see Borsley (forthcoming).

or constraint-based view of grammar. It also assumes that the grammar involves a complex systems of types and constraints. Finally, it assumes that syntactic analyses complemented by separate semantic and morphological analyses. In each of these areas, Minimalism is different. It assumes a procedural view of grammar. It assumes that grammar involves just a few general operations. Finally, it assumes that semantics and morphology are simple reflections of syntax. We comment on each of these matters in the following paragraphs.

Whereas HPSG is a declarative or constraint-based approach, Minimalism seems to be firmly committed to a procedural approach. Chomsky (1995b: 219) remarks that: "We take L [a particular language] to be a generative procedure that constructs pairs (π, λ) that are interpreted at the articulatory-perceptual (A-P) and conceptual-intentional (C-I) interfaces, respectively, as 'instructions' to the performance systems.". Various arguments have been presented within HPSG for a declarative view, but no argument seems to be offered within Minimalism for a procedural view. Obviously, speakers and hearers do construct representations and must have procedures that enable them to do so, but this is a matter of performance, and there is no reason to think that the knowledge that is used in performance has a procedural character. Rather, the fact that it used in both production and comprehension suggests that it should be neutral between the two and hence declarative. For further discussion of the issues, see e. g. Pullum & Scholz (2001), Postal (2003) and Sag & Wasow (2011; 2015).

Bob: Maybe we should say more here.

The declarative-procedural contrast is an important one, but the contrast between the complex systems of types and constraints that are assumed within HPSG and the few general operations that form a Minimalist grammar is arguably more important.[3] Much work in Minimalism has three main operations Merge, Agree, and Move or Internal Merge. Merge combines two expressions, either words or phrases, to form a larger expression with the same label as one of the expressions (Chomsky 1995b: 244). Its operation can be presented as shown in Figure 2:
In the case of English, the first alternative is represented by situations where a lexical head combines with a complement while the second is represented by situations where a specifier combines with a phrasal head.

Agree, as one might suppose, offers an approach to various kinds of agreement phenomena. It involves a probe, which is a feature or features of some kind on head, and a goal, which the head c-commands. At least normally, the probe is

[3] A procedural approach doesn't necessarily involve a very simple grammatical system. The Standard Theory of transformational grammar (Chomsky 1965) is procedural but has many different rules, both phrase structure rules and transformations.

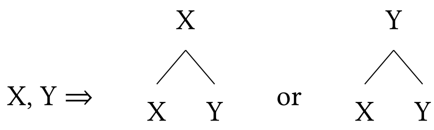


Figure 2: Merge

an uninterpretable feature or features with no value and the goal has a matching interpretable feature or features with appropriate values. Agree values the uninterpretable feature or features and they are ultimately deleted, commonly after they have triggered some morphological effect. Agree can be represented as in Figure 3 (where the "*u*" prefix identifies a feature as uninterpretable).[4]

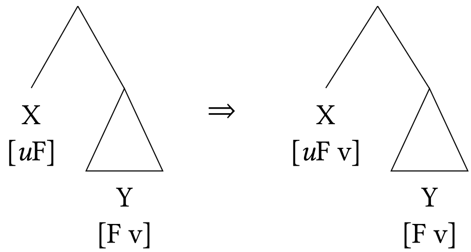


Figure 3: Agree

Unsurprisingly subject-verb agreement is one manifestation of Agree, where X is a T(ense) and Y is a nominal phase, for Minimalism a DP, inside the complement of T. Here, and elsewhere, Agree is a non-local relation involving elements which are not sisters. This contrasts with the situation in HPSG, in which subject-verb agreement is a consequence of a relation between the subject and its VP sister and a relation between the VP and the V that heads it.

Finally, Move or Internal Merge is an operation which makes a copy of a constituent of some expression and merges it with the expression. The original element that is copied normally undergoes deletion. The process can be presented as in Figure 4.

This covers both the A'-movement process assumed for unbounded dependency constructions such as *wh*-interrogatives and the A-movement process assumed for raising sentences and passives. A question arises about so-called head-

add references

[4]On standard assumptions, the goal also has some uninterpretable feature, which renders it "active", i. e. capable of undergoing Agree. This is ultimately deleted, possibly after they have triggered some morphological effect.

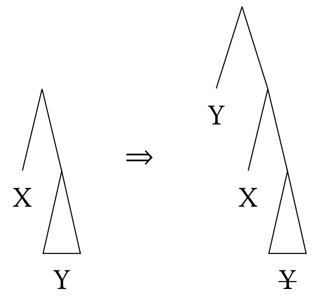


Figure 4: Move

movement, where a head moves to a higher head position. This appears to mean that it must be possible for the copy to be merged with the head of the expression that contains it. However, this is incompatible with the widely assumed extension condition, which requires Merge to produce a larger structure. One response is the idea espoused in Chomsky (1995a: 368; 2001: 37) that head-movement takes place not in the syntax but in the PF component, which maps syntactic representations to phonetic representations. It seems that its status is currently rather unclear.

The three operations just outlined interact with lexical items to provide syntactic analyses. It follows that the properties of constructions must largely derive from the lexical items that they contain. Hence, the properties of lexical items are absolutely central to Minimalism. Oddly, the obvious implication – that the lexicon should be a major focus of research – seems to be ignored. As Newmeyer (2005: 95, fn. 9) comments:

> […] in no framework ever proposed by Chomsky has the lexicon been as important as it is in the MP [Minimalist Program]. Yet in no framework proposed by Chomsky have the properties of the lexicon been as poorly investigated. (Newmeyer 2005: 95, fn. 9)

Sometimes it is difficult to derive the properties of constructions from the properties of visible lexical elements. But there is a simple solution: postulate an invisible element. The result is a large set of invisible functional heads. As we will see in the next section, these heads do the work in Minimalism that is done by phrase types and the constraints on them in HPSG.

Although Minimalism is a procedural approach and HPSG a declarative approach, there are some similarities between Minimalism and early HPSG, the approach presented in Pollard & Sag (1987; 1994). In much the same way as Minimalism has just a few general mechanisms, early HPSG had just a few general phrase types. Research in HPSG in the 1990s led to the conclusion that this is too simple and that a more complex system of phrase types is needed to accommodate the full complexity of natural language syntax. Nothing like this happened within Minimalism, almost certainly because there was little attempt within this approach to deal with the full complexity of natural language syntax. As noted above, the approach has rarely been applied in detailed formal analyses. It looks too simple and it appears problematic in various ways. It is also a major source of the complexity that is characteristic of Minimalist syntactic structures, as we will see in the next section.

The Minimalist machinery has various implications for syntactic structure which we will explore in the next section. Here we will just highlight two points. First, the fact that Merge combines two expressions entails that syntactic structures are confined to binary branching and excludes various analyses that have been assumed within HPSG and other frameworks. Second, the assumption that expressions produced by Merge have the same label as one of the expressions that they consist of is essentially the assumption that all complex expressions are headed. For HPSG, as for many other approaches, there are headed expressions and non-headed expressions, e.g., coordinate structures. We will discuss both of these matters in the next section.

As emphasized above, a further important feature of Minimalism is the view that semantics and morphology are simple reflections of syntax. The idea that semantics is a simple reflection of syntax goes back to the early years of transformational grammar. One aspect of this idea was formalized as the Uniform Theta Assignment Hypothesis (UTAH) by Baker (1988: 46).

(8) Uniform Theta Assignment Hypothesis
Identical thematic relationships between items are represented by identical structural relationships between those items at the level of D-structure.

Minimalism abandoned the notion of D-structure, but within Minimalism the Hypothesis can be reformulated as follows:

(9) Uniform Theta Assignment Hypothesis (revised)
Identical thematic relationships between items are represented by identical structural relationships between those items when introduced into the structure.

We will look at some of the implications of this in the next section.

The idea that morphology is a simple reflections of syntax is also important. As we will discuss in the next section, it leads to abstract underlying structures and complex derivations and to functional heads corresponding to various suffixes. Again, we will say more about this in the next section.

4 Different views of syntactic structure

The very different views of grammar that are assumed in Minimalism and HPSG naturally lead to very different views of syntactic structure. The syntactic structures of Minimalism are both very complex and very simple. This sounds paradoxical but it isn't. They are very complex in that is that they involve much more structure than those assumed in HPSG and other approaches. But they are very simple in that they have just a single ingredient – they consist entirely of local trees in which there is a head and a single non-head. From the standpoint of HPSG, they are both too complex and too simple. We will consider the complexity in 4.1 and then turn to the simplicity in 4.2.

4.1 The complexity of Minimalist structures

For HPSG, as the chapters in this volume have illustrated, linguistic expressions have a single relatively simple constituent structure with a minimum of phonologically empty elements.[5] For Minimalism, they have a complex structure containing a variety of empty elements and with various constituents occupying more than one position in the course of the derivation. Thus the structures assumed within Minimalism are not at all minimalist. But this complexity is a more or less inevitable consequence of the Minimalist view of grammar outlined above.

There are a variety of sources of complexity, and some predate Minimalism.[6] This is true especially of the idea that semantics and morphology are simple reflections of syntax. For the syntax-semantics relation, UTAH, which we introduced above, is particularly important. It leads to a variety of abstract representations and movement processes. Consider, for example, the following:

(10) a. Who did Lee see?
b. Lee saw who

[5] The relatively simple structures of HPSG are not an automatic consequence of its declarative nature. Postal's Metagraph Grammar framework (formerly known Arc Pair Grammar) is a declarative framework with structures that are similar in complexity to those of Minimalism (see Postal 2010).

[6] For interesting discussion of the historical development of the ideas that characterize Minimalism, see Culicover & Jackendoff (2005: Chapters 2 and 3).

Who bears the same thematic relation to the verb see in (10a) as in (10b). Assuming UTAH, it follows that *who* in (10a) should be introduced in the object position which it occupies in (10b) and then be moved to its superficial position. Consider next the following:

(11) a. Lee was seen by Kim.
     b. Kim saw Lee.

Here, *Lee* bears the same thematic relation to the verb *see* in (11a) as in (11b). Hence, it follows that *Lee* in (11a) should be introduced in the object position which it occupies in (11b) and then be moved to its superficial subject position. Finally, consider these examples:

(12) a. Lee seems to be ill.
     b. It seems that Lee is ill.

Here, *Lee* bears the same thematic relation to *ill* in (12a) as in (12b). Thus, it follows that *Lee* in (12a) should be introduced in the same position as *Lee* in (12a). The standard Minimalist approach assumes that *Lee* in both examples originates in a position adjacent to *ill* and is moved a short distance in (12a) but a longer distance in (12a).

These analyses are more or less inevitable if one accepts UTAH. But how sound is UTAH? Work in HPSG shows that it is quite possible to capture both the syntactic and the semantic properties of these sentence types without the assumption that the crucial constituents occupy more than one position. Thus, there is no reason to accept UTAH.

The idea that semantics is a simple reflection of syntax has led to other kinds of complexity. For example, it has led to revival of the idea once characteristic of Generative Semantics that lexical items may derive from complex expressions which in some sense represent their meanings. Thus, Hale & Keyser (1993) argue that (4a) derives from a structure like that of (4b).

(13) a. Kim shelved the books.
     b. Kim put the books on the shelf.

One problem with this proposal is that *shelve X* means more than just *put X on the shelf.* Thus, (14a) is not equivalent to (14b).

(14) a. Kim put his elbow on the shelf.
     b. Kim shelved his elbow.

Moreover, denominal verbs can have many different interpretations.

(15) a. Kim saddled the horse.
(Kim put the saddle on the horse.)

b. Lee chaired the meeting.
(Lee was the chairperson of the meeting.)

c. Sandy skinned the rabbit.
(Sandy removed the skin from the rabbit.)

d. Kim pictured the scene.
(Kim constructed a mental picture of the scene.)

e. They stoned the criminal.
(They threw stones at the criminal.)

f. He fathered three children.
(He was the biological father of three children.)

g. He mothers his students.
(He treats his students the way a mother would.)

Denominal verbs need to be associated with the correct meanings, but there is no reason to think that syntax has a role in this.[7]

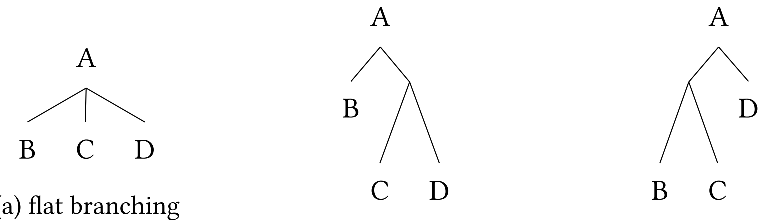
The idea that morphology is a simple reflection of syntax also leads to syntactic complexity. The fact that verbs in English and many other languages are marked for tense leads to the assumption that there is a T(ense) head at the heart of clause structure. Similarly the fact that nouns in English and other languages are marked for number leads to the assumption that there is a Num(ber) head at the heart of noun phrase structure. These elements are not solely motivated by morphology. The assumption that verbs move to T and nouns to Num in some languages but not others provides a way of accounting for cross-linguistic word order differences. However, assumptions about morphology are an important part of the motivation.

Bob: Probably there should be some reference to HPSG work on morphology.

Another source of complexity which also predates Minimalism is the assumption that all structures are binary branching. As Culicover & Jackendoff (2005: 112–116) note, this idea goes back to the 1980s. It entails that there can be no structures of the form in Figure 5a. Rather all structure must take the form in Figure 5b or Figure 5c.

As Culicover & Jackendoff discuss, the arguments for the binary branching restriction have never been very persuasive. Moreover, it is incompatible with various analyses which have been widely accepted in HPSG and other frameworks. We will return to this topic in Section 4.2.

[7] See Culicover & Jackendoff (2005: 53–56) for further discussion.



(a) flat branching (b) binary branching (c) binary branching

Figure 5: Flat and binary branching

As noted in Section 3, the simplicity of the Minimalist grammatical system means the properties of constructions must largely derive from the lexical items that they contain. Hence, the properties of lexical items are absolutely central to Minimalism and often this means the properties of phonologically empty items, especially empty functional heads. Thus, a large set of such elements is a central feature of Minimalism. These elements do much the same work as phrase types and the associated constraints in HPSG.

The contrast between the two frameworks can be illustrated with unbounded dependency constructions. Detailed HPSG analyses of various unbounded dependency constructions are set out in Sag (1997; 2010) and Ginzburg & Sag (2000), involving a complex system of phrase types. For Minimalism, unbounded dependency constructions are headed by a phonologically empty complementizer (C) and have either an overt filler constituent or an invisible filler (an empty operator) in their specifier position. Essentially, then, they have the structure in Figure 6. All the properties of the construction must stem from the properties of

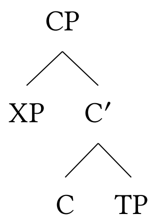


Figure 6: CP structures in Minimalism

the C that heads it.

An important unbounded dependency construction is relative clauses. In English there are *wh*-relatives and non-*wh*-relatives and finite and non-finite relatives. *Wh*-relatives are illustrated by the following:

(16) a. someone [who you can rely on]
 b. someone [on whom you can rely]

(17) a. * someone [who to rely on]
 b. someone [on whom to rely]

These show that whereas finite *wh*-relatives allow either an NP or a PP as the filler, non-finite *wh*-relatives only allow a PP. In the HPSG analysis of Sag (1997), the facts are a consequence of constraints on two phrase types. A constraint on the type *fin-wh-fill-rel-cl* allows the first daughter to be an NP or a PP while a constraint on *inf-wh-fill-rel-cl* requires the first daughter to be a PP. For Minimalism, the facts must be attributed to the properties of the complementizer. There must be a complementizer which takes a finite TP complement and allows either an NP or a PP as its specifier and another complementizer which takes a non-finite TP complement (with an unexpressed subject) and only allows a PP as its specifier.

Non-*wh*-relatives require further phrase types within HPSG and further complementizers in Minimalism. However, rather than consider this, we will look at another unbounded dependency construction: *wh*-interrogatives. The basic data that needs to be accounted for is illustrated by the following:

(18) a. Who knows?
 b. I wonder [who knows].
 c. Who did Kim talk to?
 d. I wonder [who Kim talked to].
 e. I wonder [who to talk to].

Like *wh*-relatives, *wh*-interrogatives can be finite and non-finite. When they are finite their form depends on whether the *wh*-phrase is subject of the highest verb or something else. When it is subject of the highest verb, it is followed by what looks like a VP although it may be a clause with a gap in subject position. When the *wh*-phrase is something else, the following clause shows auxiliary-initial order if it is a main clause and subject-initial order if it is not. Non-finite *wh*-interrogatives are a simple matter, especially as the filler does not have to be restricted in the way that it does in non-finite *wh*-relatives. Ginzburg & Sag (2000) present an analysis which has two types for finite *wh*-interrogatives, one for subject-*wh*-interrogatives such as those in (18a) and (18b), and another for non-subject-*wh*-interrogatives such as those in (18c) and (18d). The latter is subject to a constraint requiring it to have the same value for the features IC (INDEPENDENT-CLAUSE) and INV (INVERTED). Main clauses are [IC +] and auxiliary-initial clauses are [INV +]. Hence the constraint ensures that a non-subject-*wh*-interrogative shows auxiliary-initial order just in case it is a main clause.

How can the facts be handled within Minimalism? As noted above, Minimalism analyses auxiliary-initial order as a result of movement of the auxiliary to C. It is triggered by some feature of C. Thus C must have this feature just in case (18a) it heads a main clause and (18b) the *wh*-phrase in its specifier position is not the subject of the highest verb. There are no doubt various ways in which this might be achieved, but the key point is the properties of a phonologically empty complementizer are crucial.

Bob: We could probably say more here, e.g. referring to Pesetsky and Torrego's account of the contrast between subject and non-subject-wh-interrogatives, but I'm not sure if it is worth it.

Borsley (2006b; 2017) discusses Minimalist analyses of relative clauses and *wh*-interrogatives and suggests that at least eight complementizers are necessary. One is optionally realized as *that*, and another is obligatorily realized as *for*. The other six are always phonologically empty. But it has been clear since Ross (1967) and Chomsky (1977) that relative clauses and *wh*-interrogatives are not the only unbounded dependency constructions. Here are some others:

(19) a. What a fool he is! (*Wh*-exclamative clause)
b. The bagels, I like. (Topicalized clause)
c. Kim is more intelligent [than Lee is]. (Comparative-clause)
d. Kim is hard [to talk to]. (*Tough*-complement-clause)
e. Lee is too important [to talk to]. (*Too*-complement-clause)
f. [The more people I met], [the happier I became]. (*The*-clauses)

Each of these constructions will require at least one empty complementizer. Thus, a comprehensive account of unbounded dependency constructions will require a large number of such elements. But a large set of complementizers makes no distinction between properties shared by some or all elements and properties restricted to a single element. There are a variety of shared properties. Many of the complementizers will take a finite complement, many others will take a non-finite complement, and some will take both. There will also be complementizers which take the same set of specifiers. Most will not attract an auxiliary, but some will, not only the compementizer in an example like (18c) but also the complementizers in the following, where the auxiliary is in bold:

(20) a. Only in Colchester *could* such a thing happen.

b. Kim is in Colchester, and so *is* Lee.

c. Such *is* life.

d. The more Bill smokes, the more *does* Susan hate him.

Thus, there are generalizations to be captured here. The obvious way to capture them is with the approach developed in the 1980s in HPSG work on the hierarchical lexicon, i.e. a detailed classification of complementizers which allows properties to be associated not just with individual complementizers but also with classes of complementizers. With this it should be possible for Minimalism not just to get the facts right but to capture the full set of generalizations. In many ways such an analysis would be mimicking the HPSG approach with its hierarchy of phrase types.[8] But in the present context the main point is that the Minimalist approach to unbounded dependency constructions which leads to considerable complexity.

Thus, a variety of features of Minimalism lead to structures that are much more complex than those of HPSG. HPSG shows that this complexity is unnecessary given a somewhat richer conception of grammar.

4.2 The simplicity of Minimalist structures

As we emphasized above, while minimalist structures are very complex, they are also simple in the sense that they have just a single ingredient, local trees consisting a head and a single non-head. To most outsiders this looks too simple.

We look first at binary branching. As we noted above, the assumption that all branching is binary is incompatible with various analyses which have been widely accepted in HPSG and other frameworks. For example, it means that the bracketed VP in (21), which contains two complements, cannot have the ternary branching structure in Figure 7.

(21) Kim gave a book to Lee.

Instead it has been assumed since Larson (1988) that it has something like the structure in Figure 8. It is assumed that the verb originates in the lower VP and is moved into the higher VP.[9] The main argument for such an analysis appears to involve anaphora, especially contrasts like the following:

(22) a. John showed Mary herself in the picture.

[8] For a fuller discussion of the issues see Borsley (2006b; 2017)

[9] The higher V position to which the verb moves is commonly labelled v ("little v") and the higher phrase vP.

Bob: It might make sense to say something about Rizzi's 'split CP' analysis which adds a lot more complexity, and Cinque's approach to adverb order, which I believe postulates an empty head for every type of adverb (but I don't know this work very well and I don't

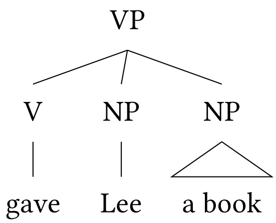


Figure 7: Flat structure for the VP *gave Lee a book*

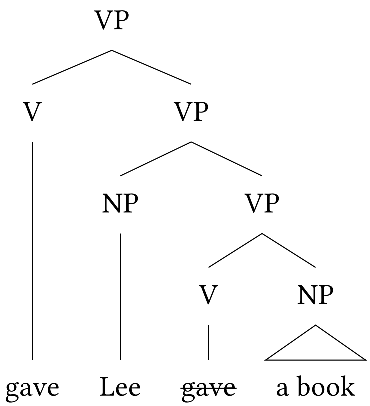


Figure 8: Larson-type analysis of VPs

b. * John showed herself Mary in the picture.

The first complement can be the antecedent of a reflexive which is the second complement, but the reverse is not possible. If constraints on anaphora refer to constituent structure, the contrast suggests that the second NP should be lower in the structure than the first NP. But, as discussed in Branco (2018), Chapter 20 of this volume, it is assumed in HPSG that constraints on anaphora refer not to constituent structure but to ARG-ST lists. On this view, anaphora can provide no argument for the complex structure in (24).

Bob: We should probably note that some HPSG work, e.g. yours involves a different binary branching structure.

The fact that Merge combines two expressions also means that the auxiliary-initial clause in (23) cannot have a flat structure with both subjects and complement(s) as sisters of the verb, as in (9).

(23) Will Kim be here?

It is standardly assumed that it has a structure of the following form:

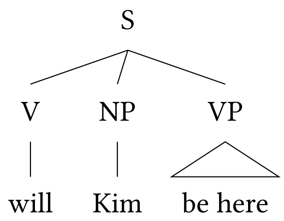


Figure 9: Flat structure for *Will Kim be there?*

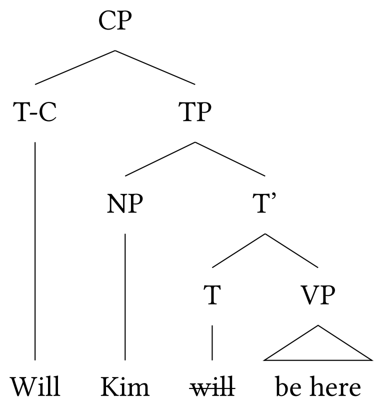


Figure 10: CP/TP structure for *Will Kim be there?*

Will is analysed as a T(ense) element which moves to the C(omplementizer) position. An analysis like this is essentially the only possibility within Minimalism.

It is not just English auxiliary-initial clauses that cannot have a ternary branching analysis within Minimalism but verb-initial clauses in any language. A notable example is Welsh, which has verb-initial order in all types of finite clause. Here are some relevant examples:[10]

(24) a. Mi/Fe gerddith Emrys i 'r dre.
PRT walk.FUT.3SG Emrys to the town
'Emrys will walk to the town.'

[10]Positive main clause verbs are optionally preceded by a particle (*mi* or *fe*). We have included this in (24a) but not in (24b). When it appears it triggers so-called soft mutation. Hence (24a) has *gerddith* rather than the basic form *cerddith*, which is seen in (24b).

b. Dywedodd Megan [cerddith Emrys i 'r dre].
say.PAST.3SG Megan walk.FUT.3SG Emrys to the town
'Megan said Emrys will walk to the town.'

A variety of transformational work, including work in Minimalism, has argued for an analysis like (23) for Welsh finite clauses (see e.g., Jones & Thomas 1977, Sproat 1985, Sadler 1988, Rouveret 1994, and Roberts 2005). But Borsley (2006a) argues that there is no theory-neutral evidence for a structure of this kind. Hence, at least for Welsh, it seems that a simpler flat structure like (22) is preferable.[11]

Bob: I assume you favour a structure rather like (30) for German. I suppose we need to make it clear that (29) seems right for some verb-initial clauses, not all.

Bob: Insert NPN-construction here

We turn now to the idea that all structures are headed. For HPSG, and many other approaches, there are headed structures and non-headed structure. Probably the most important example of the latter are coordinate structures such as those in (25) (see Sag 2003 for an HPSG analysis).

(25) [Kim and Lee] [sang and danced].

Much work in Minimalism assumes that coordinate structures are headed by the conjunction. This suggests that both coordinate structures in (25) are conjunction phrases. If they are phrases of the same kind, one would expect them to be interchangeable, but of course they are not.[12]

(26) * [Sang and danced] [Kim and Lee].

It is fairly clear that conjunctions cannot be ordinary heads. One might suggest that they are heads which have the properties of their specifier and complement, and are thus nominal if they are nominal, verbal if they are verbal, etc. This would make them a unique type of a head with a unique relation to their specifier and complement. A problem for this approach is coordinate structures in which the conjuncts belong to different categories, e.g., the following:

(27) a. Hobbs is [a linguist and proud of it].
b. Hobbs is [angry and in pain].

Such examples have led to HPSG analyses in which coordinate structures have whatever properties are common to the two conjuncts (Sag 2003). Within Minimalism, one might try to mimic such analyses by proposing that conjunctions have whatever properties are common to their specifier and complement. But a problem arises with an example like (12), where the conjuncts are not words but phrases.

[11] Borsley (2016) argues for a similar flat structure for the Caucasian ergative SOV language Archi.
[12] For a more detailed critique of this approach see Borsley (2005).

(28) Kim [criticized and insulted] his boss.

To accommodate such examples, conjunctions would have to acquire not only part of speech information from the conjuncts but also selectional information. They would be heads which combine with a specifier and a complement to form an expression which, like a typical head, combines with a specifier and a complement. This would be a very strange situation.[13] Perhaps recognizing the weaknesses of the ConjP analysis, Chomsky (2013) sketches a different approach to coordinate structures, in which the first conjunct is the head. This approach has a problem with a simple example like (36).

(29) [Kim and Lee] were late.

Since the first conjunct *Kim* is singular, Chomsky's approach will identify the coordinate structure as singular and one would expect the singular form *was* and not the plural form *were*. Further problems arise with the following examples:

(30) a. [You and he] know yourselves well.
b. [You and I] know ourselves well.

In both examples the first conjunct is the second person, and in (30a), the form of the reflexive suggests that the coordinate structure is too. However, in (30b), the form of the reflexive suggests that the coordinate structure is first person. Clearly, this is because the second conjunct is first person. It is clear, then, that the properties of a coordinate structure reflects both conjuncts in a way that makes them very different from ordinary headed structures. This suggests rather strongly that the idea that all structures are headed is untenable.

Finally, we want to consider the Move/Internal Merge approach to unbounded and other non-local dependencies. It is important to emphasize that this mechanism is involved not only in unbounded dependency constructions but also in passives, unaccusatives, and raising sentences, such as the examples in (31).

(31) a. Kim has been hit.
b. Kim has disappeared.
c. Kim seems to be clever.

The two types of construction have rather different. For example, whereas the gaps in unbounded dependency constructions are positions in which overt NPs can appear, this is not true of the supposed gap positions.

[13] There have been attempts to argue that conjuncts are always phrases (Kayne 1994, Bruening 2018). But this position seems untenable (Abeillé 2006, Müller 2018).

(32) a. * It has been hit Kim.
b. * It has disappeared Kim.
c. * It seems Kim to be clever.

This is a complication if they involve the same mechanism, but is unsurprising if they involve different mechanisms, as in HPSG and most other frameworks.

The Move/Internal Merge approach seems quite plausible for typical examples of an unbounded dependency, but issues arise with less typical examples. Within this approach one expects to see a constituent in the tree and a matching gap somewhere inside its sister. This is what we commonly find, but there are unbounded dependency constructions in which there is a gap but no visible higher constituent matching it. Consider e.g., the following:

(33) a. the book [Kim bought _]
b. Lee is too important [for you to talk to _].
c. Lee is important enough [for you to talk to _].
d. Kim is easy [for anyone to talk to _].

Within Minimalist assumptions, it is more or less necessary to assume that such examples contain an invisible filler (a so-called empty operator). Unless there is some independent evidence for such invisible fillers, they are little more than an ad hoc device to maintain the Move/Internal Merge approach. Within the HPSG SLASH-based approach to unbounded dependencies, there is no assumption that there should always be a filler at the top of an unbounded dependency. Hence, the examples in (33) are completely unproblematic.

There are also unbounded dependency constructions which seem to have not a gap but a resumptive pronoun (RPs). Among many languages that are relevant here is Welsh, which has RPs in both *wh*-interrogatives and relative clauses, as the following illustrate:

(34) a. Pa ddyn werthodd Ieuan y ceffyl iddo *fo*?
which man sell.PAST.3SG Ieuan the horse to.3SGM he
'Which man did Ieuan sell the horse to?'

b. y dyn werthodd Ieuan y ceffyl iddo *fo*
the man sell.PAST.3SG Ieuan the horse to he
'the man that Ieuan sold the horse to'

Willis (2011) and Borsley (2010; 2013)) present evidence that Welsh RPs involve the same mechanism as gaps. Within Minimalism, this means that they must involve Move/Internal Merge. But one expects to see a gap where Move/Internal Merge has applied. One Minimalist response suggests that instead of being deleted, the copy left behind by Move/Internal Merge is somehow turned into a pronoun (see McCloskey 2006). Another approach exploits the complexity of Minimalist structures and proposes that there is a gap in the structure somewhere near the RP. See Willis (2011), Aoun et al. (2001), and Boeckx (2003). For detailed objections to both approaches, see Borsley (2013: Section 3). Within the SLASH-based approach of HPSG, there is no reason to think that there will always be a gap at the bottom of a dependency, and it is not difficult to accommodate RPs. See Borsley (2013) and Crysmann (2012; 2016) for slightly different approaches.[14]

Bob: Could obviously say more here, but maybe it's not necessary.

Thus, there are variety of phenomena which suggest that the Minimalist view of constituent structure is too simple. The restriction to binary branching, the assumption that all structures are headed, and Move/Internal Merge all seem problematic. It looks, then, as if the Minimalist view is both too complex and too simple.

5 Psycholinguistic issues

Although they differ in a variety of ways, HPSG and Minimalism agree that grammatical theory is concerned with linguistic knowledge. They focus first and foremost on the question: what form does linguistic knowledge take? But there are other questions that arise here, notably the following:

- How is linguistic knowledge put to use?
- How is linguistic knowledge acquired?

Both questions are central concerns for psycholinguistics. Thus, in considering the answers that HPSG and Minimalism can give we are considering their relevance to psycholinguistics. Chomskyan approaches, including Minimalism, have focused mainly on the second question and have paid little attention to the first. HPSG has had more to say about the first and has shown less interest in the

[14] Also relevant here are examples with more than one gap such as the following:

(i) a. Who does Kim like _ and Lee hate _?
    b. Which book did you criticize _ without reading _?

There have been various attempts to accommodate such examples within the Move/Internal Merge approach, but it is not clear that any of them is satisfactory. In contrast such examples are expect within the SLASH-based approach Levine & Sag (2003).

second. Clearly an adequate grammatical theory should be able to give satisfactory answers to both questions. In this section we will look briefly at the relation of the two theories to processing and then consider more fully their relation to acquisition.

5.1 Processing

We noted in section 3 that whereas HPSG is a declarative or constraint-based approach to grammar, Minimalism has a procedural view of grammar. This contrast means that HPSG is much more suitable then Minimalism for incorporation into an account of the processes that are involved in linguistic performance.

The most obvious fact about linguistic performance is that it involves both production and comprehension. As noted in section 3, this suggests that the knowledge that is used in production and comprehension should have a declarative character as in HPSG and not a procedural character as in Minimalism.

A second important feature of linguistic performance is that it involves different kinds of information utilized in any order that is necessary. Sag & Wasow (2011) illustrate with the following examples:

(35) a. The sheep that was sleeping in the pen stood up.

b. The sheep in the pen had been sleeping and were about to wake up.

In (35a), morphological information determines the number of sheep before non-linguistic information determines that pen means 'fenced enclosure' and not 'writing implement'. In (35b), on the other hand, non-linguistic information determines that pen means 'fenced enclosure' before morphological information determines the number of sheep. This is unproblematic for an approach like HPSG in which linguistic and non-linguistic knowledge takes the form of constraints which are not ordered in any way. It is quite unclear how the facts can be accommodated within Minimalism given that linguistic knowledge with its procedural form is quite different from non-linguistic knowledge.

Other features of HPSG also make it attractive from a processing point of view. Firstly, there is the fact emphasized earlier that linguistic expressions have a single relatively simple constituent structure with a minimum of phonologically empty elements. Secondly there is the fact that all constraints are purely local and never affect anything larger than a local tree consisting of an expression and its daughters. Both these properties make processing easier than it would otherwise be. Minimalism has neither property and hence again seems less satisfactory than HPSG in this area.

5.2 Acquisition

Acquisition has long been a central concern for Chomskyans and it has long been argued that acquisition is made possible by the existence of a complex innate language faculty. Since the early 1980s the dominant view has been that the language faculty consists of a set of principles responsible for the properties which they share and a set of parameters responsible for the ways in which they may differ. On this view acquiring a grammatical system is a matter of parameter-setting. Proponents of HPSG have always been sceptical about these ideas (see e.g., the remarks about parameters in Pollard & Sag (1994: 31) and have favoured accounts with 'an extremely minimal initial ontology of abstract linguistic elements and relations' (Green 2011: 378). Thus, the two frameworks appear to be very different in this area. It is not clear, however, that this is really the case.

The idea that acquiring a grammatical system is a matter of parameter-setting is only as plausible as the idea of a language faculty with a set of parameters. It seems fair to say that this idea has not been as successful as was hoped when it was first introduced in the early 1980s. Outsiders have always been sceptical, but they have been joined in recent times by researchers sympathetic to many Chomskyan ideas. Thus, Newmeyer (2005: 75) writes as follows:

> [...] empirical reality, as I see it, dictates that the hopeful vision of UG as providing a small number of principles each admitting of a small number of parameter settings is simply not workable. The variation that one finds among grammars is far too complex for such a vision to be realized.

At least some Minimalists have come to similar conclusions. Thus, Boeckx (2011: 206) suggests that:

> some of the most deeply-embedded tenets of the Principles-and-Parameters approach, and in particular the idea of Parameter, have outlived their usefulness. (Boeckx 2011: 206)

Much the same view is expressed in Hornstein (2009: 164–168).

A major reason for scepticism about parameters is that estimates of how many there are seem to have steadily increased. Fodor (2003)considers that there might be just twenty parameters, so that acquiring a grammatical system is a matter of answering twenty questions. Newmeyer (2005: 44) remarks that 'I have never seen any estimate of the number of binary-valued parameters needed to capture all of the possibilities of core grammar that exceeded a few dozen'. However,

page

Fodor (1998: 346–347) assumes that there are 20 to 30 parameters, Gibson &

Roberts & Holmberg (2005) comment that '[n]early all estimates of the number of parameters in the literature judge the correct figure to be in the region of 50-100'. Clearly, a hundred is a lot more than twenty. This is worrying. As Newmeyer (2006: 6) observes, 'it is an ABC of scientific investigation that if a theory is on the right track, then its overall complexity decreases with time as more and more problematic data fall within its scope. Just the opposite has happened with parametric theory. Year after year more new parameters are proposed, with no compensatory decrease in the number of previously proposed ones. Just the opposite has happened with parametric theory. Year after year more new parameters are proposed, with no compensatory decrease in the number of previously proposed ones'.

The growing scepticism appears to tie in with the proposal by Hauser, Chomsky & Fitch (2002: 1573)) that "FLN [the "Narrow Language Faculty"] comprises only the core computational mechanisms of recursion as they appear in narrow syntax and the mappings to the interfaces". On this view there seems to be no place for parameters within FLN. This conclusion is also suggested by Chomsky's remarks (2005) that 'There is no longer a conceptual barrier to the hope that the UG [Universal Grammar] might be reduced to a much simpler form' (p. 8) and that 'we need no longer assume that the means of generation of structured expressions are highly articulated and specific to language' (p. 9). It's hard to see how such remarks are compatible with the assumption that UG includes 50–100 parameters. But if parameters are not part of UG, it is not at all clear what their status might be.

It looks, then, as Chomskyans are gradually abandoning the idea of parameters. But if it is abandoned, grammar acquisition is not a matter of parameter-setting. Hence, it is not clear that Chomskyans can invoke any mechanisms that are not available to HPSG.

This might suggest that HPSG and Minimalism are essentially in the same boat where acquisition is concerned. However, this is not the case given the very different nature of grammatical systems in the two frameworks. The complex and abstract structures that are the hallmark of Minimalism and earlier Chomskyan frameworks pose major problems for acquisition. It is this that has led to the assumption that acquisition must be assisted by a complex language faculty. In contrast, HPSG structures are quite closely related to the observable data and so pose less of a problem for acquisition and hence create less need for some innate apparatus. Thus, HPSG probably has an advantage over Minimalism in this area too.

There is one further issue that we should discuss here. It appears to be quite widely assumed that one advantage that Minimalism has over alternatives like HPSG is that it is more 'restrictive', in other words that it makes more claims about what is and is not possible in language. It is true that HPSG makes few claims about what is and is not possible in language. It is also true that it is good other things being equal for a theory to make such claims. It looks then as if there might be an argument for Minimalism here. It is not clear, however, that this is really the case.

Minimalism would be a restrictive theory making interesting claims about language assumed a relatively small number of parameters. However, it seems that the idea that there is just small number of parameters seems to have been abandoned, and at least some minimalists have abandoned the idea of parameters altogether. If there is either a large number of parameters or no parameters at all, Minimalism is not restrictive in the way that it once was. However, it does still embody some restrictions on grammatical systems. The assumption that syntactic structures are confined to binary branching is an important restriction, as is the assumption that expressions produced by Merge have the same label as one of the expressions that they consist of. But we have argued that both assumptions are quite dubious. It also seems to be assumed that case and agreement are features of all grammatical systems. This would be another important restriction, but this also seems dubious given that many languages show no clear evidence for one or both of these features. It looks to us, then, as if the restrictiveness of Minimalism is largely a matter of imposing certain dubious restrictions on grammatical systems.

These remarks should not be understood as a suggestion that vary without limit, as **Joos58a-u** suggested. No doubt there are language universals and variation is limited. However, most HPSG linguists would think that we don't have enough detailed formal analyses of enough phenomena in enough languages to have any firm conclusions about these matters. If this is right, a restrictive theory if not a realistic expectation at the present time.

Acknowledgements

6 To do

Tom Wasow: Mention Generative Semantics.

References

Abeillé, Anne. 2006. In defense of lexical coordination. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in formal syntax and semantics*, vol. 6, 7–36. Paris: CNRS. http://www.cssp.cnrs.fr/eiss6/, accessed 2018-2-25.

Aoun, Joseph E. & Yen-hui Audrey Li. 2003. *Essays on the representational and derivational nature of grammar*. Cambridge, MA: MIT Press.

Aoun, Joseph, Lina Choueiri & Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32(3). 371–404.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function change*. Chicago: The University of Chicago Press.

Baker, Mark C. 2003. Linguistic differences and language design. *Trends in Cognitive Sciences* 7(8). 349–353.

Bender, Emily M. 2002. Number names in Japanese: A head-medial construction in a head-final language. In *Linguistic society of america*. USA.

Boeckx, Cedric. 2003. *Islands and chains*. Amsterdam: Benjamins.

Boeckx, Cedric. 2011. Approaching parameters from below. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 205–221. Oxford: Oxford University Press.

Borsley, Robert D. 2005. Against ConjP. *Lingua* 115(4). 461–482.

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

Borsley, Robert D. 2006b. *Syntactic and lexical approaches to unbounded dependencies*. Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.

Borsley, Robert D. 2010. An HPSG approach to Welsh unbounded dependencies. In Stefan Müller (ed.), *Proceedings of the hpsg10 conference*, 47–67. Stanford: CSLI Publications.

Borsley, Robert D. 2011. Constructions, functional heads, and comparative correlatives. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics*, vol. 8, 7–20. Paris: CNRS. http://www.cssp.cnrs.fr/eiss8/borsley-eiss8.pdf, accessed 2018-8-23.

Borsley, Robert D. 2013. On the nature of Welsh unbounded dependencies. *Lingua* 133. 1–29. DOI:10.1016/j.lingua.2013.03.005

Borsley, Robert D. 2016. Hpsg and the nature of agreement in Archi. In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi:*

Complexities of agreement in cross-theoretical perspective, 118–149. Oxford: Oxford University Press.

Borsley, Robert D. 2017. Explanations and 'engineering solutions'? Aspects of the relation between Minimalism and HPSG? In Stefan Müller (ed.), *Proceedings of the hpsg17 conference*, 82–102. Stanford: CSLI Publications.

Bruening, Benjamin. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language*. To appear.

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

Chomsky, Noam. 1977. *On wh-movement*. Peter Culicover, Tom Wasow & Adrian Akmajian (eds.). New York.

Chomsky, Noam. 1990. On formalization and formal linguistics. *Natural Language and Linguistic Theory* 8(1). 143–147.

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

Chomsky, Noam. 1995a. Bare phrase structure. In Hector Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero*, 51–109. Washington, DC: Georgetown U Press.

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

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2002. *On nature and language*. Cambridge, UK: Cambridge University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.

Chomsky, Noam & Howard Lasnik. 1995. The theory of principles and parameters. In Noam Chomsky (ed.), *The minimalist program*, chap. 1, 13–127. cambridgeusa: MIT Press.

Crysmann, Berthold. 2012. Resumption and island-hood in Hausa. In Philippe de Groote & Mark-Jan Nederhof (eds.), *Formal Grammar. 15th and 16th International Conference on Formal Grammar, FG 2010 Copenhagen, Denmark, August 2010, FG 2011 Lubljana, Slovenia, August 2011*, vol. 7395 (Lecture Notes in Computer Science). Berlin: Springer Verlag.

Crysmann, Berthold. 2016. An underspecification approach to Hausa resumption. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on head-driven phrase structure grammar and lexical functional grammar*, 194–214. Stanford: CSLI Publications.

Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition*. Vol. 1 (Foundations of Syntax). Oxford: Oxford University Press.

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

Fodor, Janet Dean. 1998. Parsing to learn. *Journal of Psycholinguistic Research* 27(3). 339–374.

Fodor, Janet Dean. 2003. Setting syntactic parameters. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 730–767. Oxford: Blackwell Publishing Ltd.

Fong, Sandiway. 2014. Unification and efficient computation in the Minimalist Program. In L. Francis & L. Laurent (eds.), *Language and recursion*, 129–138. Berlin: Springer Verlag.

Fong, Sandiway & Jason Ginsburg. 2012. Computation with doubling constituents: Pronouns and antecedents in Phase Theory. In Anna Maria Di Sciullo (ed.), *Towards a Biolinguistic understanding of grammar: Essays on interfaces* (Linguistik Aktuell/Linguistics Today 194), 303–338. Amsterdam: John Benjamins Publishing Co.

Gibson, Edward & Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25(3). 407–454.

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

Green, Georgia M. 2011. Elementary principles of Head-driven Phrase Structure Grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 9–53. Oxford, UK/Cambridge, MA: Blackwell Publishing Ltd.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (Current Studies in Linguistics 24), 53–109. Cambridge, MA/London: MIT Press.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298. 1569–1579. DOI:10.1126/science.298.5598.1569

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge University Press. DOI:10.1017/CBO9780511575129

Hornstein, Norbert, Jairo Nunes & Kleantes K. Grohmann. 2005. *Understanding Minimalism* (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.

Huybregts, Riny & Henk van Riemsdijk. 1982. *Noam chomsky on the Generative Enterprise*. Dordrecht: Foris.

Jones, Morris & Alun R. Thomas. 1977. *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press.

Kayne, Richard S. 1994. *The antisymmetry of syntax* (Linguistic Inquiry Monographs 25). Cambridge, MA: MIT Press.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391.

Levine, Robert D. & Ivan A. Sag. 2003. Some empirical issues in the grammar of extraction. In Stefan Müller (ed.), *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, 236–256. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-6-7.

McCloskey, James. 2006. Resumption. In Martin Everaert & Henk van Riemsdijk (eds.), *The blackwell companion to syntax*, 94–117. Oxford: Blackwell.

Michaelis, Laura A. & Knud Lambrecht. 1996. Toward a Construction-based theory of language function: The case of nominal extraposition. *Language* 72(2). 215–247.

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

Müller, Stefan. 2016. *Grammatical theory: From Transformational Grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2018. The end of lexicalism as we know it? *Language* 94(1). e54–e66. DOI:10.1353/lan.2018.0014

Newmeyer, Frederick J. 2005. *Possible and probable languages: A Generative perspective on linguistic typology*. Oxford: Oxford University Press.

Newmeyer, Frederick J. 2006. *A rejoinder to 'On the role of parameters in Universal Grammar: A reply to Newmeyer' by Ian Roberts and Anders Holmberg*. http:

//people.ds.cam.ac.uk/mtb23/NSP/Rejoinder%20to%20R&H.pdf, accessed 2018-6-7.

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

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

Postal, Paul M. 2003. (virtually) conceptually necessary. *Journal of Linguistics* 39. 599–620.

Postal, Paul M. 2004. *Skeptical linguistic essays*. Oxford: Oxford University Press.

Postal, Paul M. 2010. *Edge-based clausal syntax: A study of (mostly) English object structure*. Cambridge, MA: MIT Press.

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between generative-enumerative and model-theoretic syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retor (eds.), *Logical Aspects of Computational Linguistics: 4th International Conference* (Lecture Notes in Computer Science 2099), 17–43. Berlin: Springer Verlag.

Roberts, Ian. 2005. *Principles and Parameters in a VSO language: A case study in Welsh*. Oxford: Oxford University Press.

Roberts, Ian F. & Anders Holmberg. 2005. On the role of parameters in Universal Grammar: A reply to Newmeyer. In Hans Broekhuis, N. Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk* (Studies in Generative Grammar 86), 538–553. Berlin: Mouton de Gruyter.

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT dissertation. http://files.eric.ed.gov/fulltext/ED016965.pdf, accessed 2018-2-25. Reproduced by the Indiana University Linguistics Club and later published as Ross (1986).

Ross, John Robert. 1986. *Infinite syntax!* Norwood, New Jersey: Ablex Publishing Corporation.

Rouveret, Alain. 1994. *Syntaxe du gallois: principes généraux et typologie*. Paris: Editions du CNRS.

Sadler, Louisa. 1988. *Welsh syntax: a government binding approach*. London: Croom Helm.

Sag, Ivan. 2003. Coordination and underspecification. In Jong-Bok Kim & Stephen Wechsler (eds.), *Proceedings of the ninth international conference on head-driven phrase structure grammar, kyung-hee university, seoul*, 267–291. Stanford: CSLI Publications.

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

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

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

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/Cambridge, MA: Blackwell Publishing Ltd.

Sag, Ivan A. & Thomas Wasow. 2015. Flexible processing and the design of grammar. *Journal of Psycholinguistic Research* 44(1). 47–63.

Sproat, Richard. 1985. Welsh syntax and vso structure. *Natural Language and Linguistic Theory* 3(2). 173–216.

Stabler, Edward P. 2001. Minimalist Grammars and recognition. In Christian Rohrer, Antje Rossdeutscher & Hans Kamp (eds.), *Linguistic form and its computation* (Studies in Computational Linguistics 1), 327–352. Stanford, CA: CSLI Publications.

Willis, David. 2011. The limits of resumption in welsh *wh*-dependencies. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces*, 189–222. Amsterdam: Benjamins.

Chapter 33

HPSG and Categorial Grammar

Yusuke Kubota
University of Tsukuba

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

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

1 Introduction

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

Yusuke Kubota. 2018. HPSG and Categorial Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 321–322. Berlin: Language Science Press. DOI:??

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

Abbreviations

Acknowledgements

Chapter 34

HPSG and Lexical Functional Grammar

Doug Arnold
University of Essex

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

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

1 Introduction

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

Doug Arnold. 2018. HPSG and Lexical Functional Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 323–324. Berlin: Language Science Press. DOI:??

ut turpis cursus rhoncus. Donec sed convallis justo. Sed sed massa pharetra ex aliquet eleifend. finality

Abbreviations

Acknowledgements

Chapter 35

HPSG and Dependency Grammar

Dick Hudson
London

Put abstract here with \abstract.

1 Two centuries of syntactic theory

In the early 19th century, European grammar was still dominated by the Latin grammar of Priscian which focused on individual words, their morphosyntactic properties and their relations (controlled especially by government and agreement); grammars and grammatical theory were mainly focused on school pedagogy, where the dominant model was the parsing of individual words. But these ideas, and especially government, defined 'dependency' relations holding most words together. The exception was the relation between the verb and its subject, which was still described in terms of the dominant classical logic based on the subject-predicate split. Putting these two traditions together, grammarians produced a mixed theory of sentence structure and a number of diagramming systems to represent such structures – most famously, the diagramming system invented in the USA by Reed & Kellogg (1877) (and still taught in the 21st century in some American schools). This is also the theory that Bloomfield brought back to the USA from Germany, and which he developed into Immediate Constituent analysis (which later turned into phrase-structure analysis); as in the earlier theory, the subject and predicate were equal, in contrast with other 'endocentric' constructions. Bloomfield combined this mixed theory with Wundt's theory of cognition, with the sentence as the 'whole' which defines its parts (and the word no longer in prime position), which allowed a consistent geometry, but phrase-structure trees did not appear till the middle of the 20th century. Meanwhile, however, both Humboldt and Grimm had suggested that the verb was

Dick Hudson. 2018. HPSG and Dependency Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 325–328. Berlin: Language Science Press. DOI:??

the sole head of the sentence, with the subject as one of its dependents, and by the 1860s and 1870s, grammarians in Hungary, Russia and Germany (apparently working independently) were arguing for this view, half a century before it was formalised by Tesniere and named ‘dependency analysis’. The first ‘stemma’ diagram appeared (in Hungary) in 1873. Another 19th-century reaction against classical logic was the logical tradition started (in Germany) by Frege, who may have learned to draw stemmas at school; this tradition gave rise (in Poland) to categorial grammar, which some (including Chomsky) see as a version of dependency analysis. One outcome of this history was the present-day geographical split between American phrase structure (PS) and European dependency structure (DS). Variations on the dependency theme Unsurprisingly, therefore, dependency theory has had more impact on Europeans than on Americans. The general idea of word-word dependencies was built into a number of different theoretical packages which combined it with other ideas, notably multiple levels (the Russian Mel’cuk) and information structure (the Czechs Sgall and Hajicova). However, dependency structure has also been popular internationally in natural-language processing (represented perhaps most notably by the Stanford Parser). ‘Plain-vanilla’ versions of DS and PS are very similar and are weakly equivalent, but as with phrase structure, such theories need to be supplemented, giving rise to theories in which structures are much richer. One such theory is Word Grammar (WG), which is probably closer to HPSG than any of the other DS theories. In WG, a word is allowed to depend on more than one other word (like re-entrance in HPSG) and dependencies are combined with extra mechanisms for coordination and for word order. This theory will be the main point of comparison with HPSG in the rest of the chapter.

2 Signs, constructions and levels

The contrast between PS and DS is orthogonal to choices about the number of levels (syntax, morphology, etc) and how they are related, but of course these choices are essential for any theoretical package. As in PS theories, different DS theories assume different answers, but Word Grammar takes a rather conservative position in which syntax is distinct both from morphology and from semantics. This view is hard to reconcile with the claim that language consists of ‘constructions’ or ‘signs’, both of which assume a direct link between ‘form’ and ‘meaning’. In this view, units of phonological ‘form’ are only indirectly linked to units of meaning. Approaches which evoke ‘signs’ or ‘constructions’ can also be challenged for their conservative assumptions about plain-vanilla surface PS.

Arguably, DS is a better basis for capturing the fine detail of idiosyncratic constructions since these always involve individual lexical items linked by dependencies, and typically focus on just one dependent of a given lexeme rather than on entire multi-dependent phrases. Networks WG takes the whole of language (not just the lexicon) to be a gigantic network, which is a step further than HPSG (where PS rules are outside the network); the network is also not assumed to be a DAG because mutual dependency is allowed. One of the characteristics of network analyses is the central role of relation types (i.e. HPSG attributes). According to WG, but not HPSG, these types form a typed hierarchy which parallels the typed hierarchy of non-relational 'entities' such as words, phonemes and so on; and in both hierarchies, properties are inherited by (a special formalisation of) default inheritance. One of the consequences of this treatment of relations is that, just like entities, they can freely be created and learned as required, so there is no need to assume a universal hard-wired reservoir of relations. This is particularly helpful in DS, where dependencies are typed but different languages require different classifications and distinctions. Word order Another similarity between WG and HPSG is in the treatment of word order. In both theories, dominance (i.e. daughterhood in HPSG and dependency in WG) is separated from linear precedence. In WG, a word's position is treated as one of the word's property's linked to a second property ('landmark'), the word from which it takes its position; the word's landmark is normally the word on which it depends, but exceptions are allowed in cases such as extraction and pied piping. The landmark relation allows a treatment of pied piping which avoids the feature-percolation of HPSG.

3 Words, nodes and semantic phrases

The final topic is the Achille's heel of DS: the completely flat structures where a word has two or more dependents. This is problematic in DS (but not, of course, in HPSG) in examples such as typical French house, meaning 'typical for a French house', because there is no syntactic node that could carry the meaning 'French house'. Current WG provides a solution which moves WG in the direction of PS by distinguishing types from tokens, and then distinguishing 'sub-tokens' of tokens. In this analysis, the token house is distinct not only from the type HOUSE, but also from the sub-token house' which is modified by the dependent French, which in turn is distinct from house'' modified by typical. Sub-tokens are very similar in function to the phrases of HPSG but arguably not quite equivalent.

Abbreviations

Acknowledgements

References

Reed, Alonzo & Brainerd Kellogg. 1877. *Higher lessons in English: A work on English grammar and composition*. Projekt Gutenberg. http://www.gutenberg.org/ebooks/7188, accessed 2018-3-23.

Chapter 36

HPSG and Construction Grammar

Stefan Müller
Humboldt-Universität zu Berlin

This chapter discusses the main tenets of Construction Grammar (CxG) and shows that HPSG adheres to them. This discussion includes surface orientation, language acquisition without UG, inheritance networks and shows how HPSG (and other frameworks) are positioned along these dimensions. Formal variants of CxG will be briefly discussed and their relation to HPSG will be pointed out. It is argued that lexical representations of valence are more appropriate than phrasal approaches, which are assumed in most variants of CxG. Other areas of grammar seem to require headless phrasal constructions (e.g., the NPN construction and certain extraction constructions) and it is shown how HPSG handle these. Derivational morphology is discussed as a further example of an early constructionist analysis in HPSG.

This chapter deals with Construction Grammar (CxG) and its relation to HPSG. The short version of the message is: HPSG is a Construction Grammar. It was one right from the beginning and over the years certain aspects were adapted allowing to capture generalizations over phrasal patterns. In what follows I will first say what Construction Grammars are (Section 1), I will explain why HPSG as developed in Pollard & Sag (1987; 1994) was a Construction Grammar and how it was changed to become even more Constructive (Section 1.2.3). Section 2 deals with so-called argument structure constructions, which are usually dealt with by assuming phrasal constructions in CxG and explains why this is problematic and why lexical approaches are more appropriate. Section 3 explains Construction Morphology, Section 4 shows how cases that should be treated phrasally can be handled in HPSG. Section 5 sums up the paper.

Stefan Müller. 2018. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 329–361. Berlin: Language Science Press. DOI:??

1 What is Construction Grammar?

The first question to answer in a chapter like this is: what is Construction Grammar? While it is relatively clear what a Construction is, the answer to the question regarding Construction Grammar is less straight-forward. Section 1.1 provides the definition for the term *Construction* and Section 1.2 states the tenets of CxG and discusses to what extent the main frameworks currently on the market adhere to them.

1.1 The notion Construction

Goldberg (2006: 5) defines Construction as follows:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (Goldberg 2006: 5)

She provides Table 1 with examples for Constructions.

BB: I understand her view is that constructions must have a distinctive meaning. I believe Jackendoff argues against this, and Newmeyer and I did here: Borsley & Newmeyer (2009)

Table 1: Examples of constructions, varying in size and complexity according to Goldberg (2009)

Word	e.g., *tentacle*, *gangster*, *the*
Word (partially filled)	e.g., *post*-N, V-*ing*
Complex word	e.g., *textbook*, *drive-in*
Idiom (filled)	e.g., *like a bat out of hell*
Idiom (partially filled)	e.g., *believe* <one’s> *ears/eyes*
Covariational Conditional	The Xer the Yer (e.g., *The more you watch the less you know*)
Ditransitive	Subj V Obj1 Obj2 (e.g., *She gave him a kiss*; *He fixed her some fish tacos*.)
Passive	Subj aux VPpp (PPby) (e.g., *The cell phone tower was struck by lightening*.)

If one just looks at the definition of Construction, all theories currently on the market could be regarded as Construction Grammars. As Peter Staudacher pointed out in the discussion after a talk by Knud Lambrecht in May 2006 in Potsdam, lexical items are form-meaning pairs and the rules of phrase structure grammars come with specific semantic components as well, even if it is just functional application. So, Categorial Grammar, GB, GPSG, TAG, LFG, HPSG and even Minimalism would be Construction Grammars. If one looks at the examples of Constructions in Table 1 things change a bit. Idioms are generally not the focus of work in Mainstream Generative Grammar (MGG)[1]. MGG is usually concerned with explorations of the so-called Core Grammar as opposed to the Periphery, to which the idioms are assigned. The Core Grammar is the part of the grammar that is supposed to be acquired with help of innate domain specific knowledge, something the existence of which Construction Grammar denies. But if one takes Hauser, Chomsky & Fitch (2002) seriously and assumes that only the ability to form complex linguistic objects out of less complex linguistic objects (Merge) is part of this innate knowledge then the core/periphery distinction does not have much content and after all Minimalists could adopt a version of Sag's local, selection-based analysis of idioms (Sag 2007; Kay et al. 2015; Kay & Michaelis 2017). However, as is discussed in the next subsection, there are other aspects that really set Construction Grammar apart from MGG.

1.2 Basic tenets of Construction Grammar

Goldberg (2003) names the following tenets as core assumptions standardly made in CxG. form-meaning pairs):

Tenet 1 All levels of description are understood to involve pairings of form with semantic or discourse function, including morphemes or words, idioms, partially lexically filled and fully abstract phrasal patterns. (See Table 1)

Tenet 2 An emphasis is placed on subtle aspects of the way we conceive of events and states of affairs.

Tenet 3 A 'what you see is what you get' approach to syntactic form is adopted: no underlying levels of syntax or any phonologically empty elements are posited.

[1]The term *Mainstream Generative Grammar* is used to refer to work in Transformational Grammar, for example Government & Binding (Chomsky 1981) and Minimalism (Chomsky 1995). Some authors working in Construction Grammar see themselves in the tradition of Generative Grammar in a wider sense, see for example Fillmore, Kay & O'Connor (1988: 501).

Tenet 4 Constructions are understood to be learned on the basis of the input and general cognitive mechanisms (they are constructed), and are expected to vary cross-linguistically.

Tenet 5 Cross-linguistic generalizations are explained by appeal to general cognitive constraints together with the functions of the constructions involved.

Tenet 6 Language-specific generalizations across constructions are captured via inheritance networks much like those that have long been posited to capture our non-linguistic knowledge.

Tenet 7 The totality of our knowledge of language is captured by a network of constructions: a 'construct-i-con.'

I already commented on Tenet 1 above. Tenet 2 concerns semantics and the syntax-semantics interface, which are part of most HPSG analyses. In what follows I want to look in more detail at the other tenets.

1.2.1 Surface orientation and empty elements

Tenet 3 requires a surface-oriented approach. Underlying levels and phonologically empty elements are ruled out. This excludes derivational models of transformational syntax assuming a D-structure and some derived structure or more recent derivational variants of Minimalism. There was a time where representational models of GB that did not assume a D-structure but just one structure with traces (Koster 1978: 1987: 235; Kolb & Thiersch 1991; Haider 1993: Section 1.4; Frey 1993: 14; Lohnstein 1993: 87–88, 177–178; Fordham & Crocker 1994: 38; Veenstra 1998: 58). Some of these analyses are rather similar to HPSG analyses as they are assumed today (Kiss 1995; Bouma & van Noord 1998; Meurers 2000; Müller 2005; 2017a; 2018b). Chomsky's Minimalist work (Chomsky 1995) assumes a derivational model and comes with a rhetoric of building structure in a bottom-up way and sending complete phases to the interfaces for pronunciation and interpretation. This is incompatible with Tenet 3, but in principle Minimalist approaches are very similar to Categorial Grammar, so there could be representational approaches adhering to Tenet 3.

Rui: In Top-down Phase-based Minimalist Grammar (TPMG) as developed by Chesi (2012; 2007), and Bianchi & Chesi (2006; 2012), there is no movement. Rather, wh-phrases are linked to their "in situ" positions with the aid of a short-term memory buffer that functions like a stack. See also Hunter (2010) and Hunter (2018) for a related account where the information about the presence of a wh-phrase is percolated in the syntax tree, like in GPSG/HPSG. These references may be of use here.

A comment on empty elements is in order: all articles introducing Construction Grammar state that CxG does not assume empty elements. Most of the alternative theories do use empty elements: see König (1999) on Categorial Grammar, Gazdar, Klein, Pullum & Sag (1985: 143) on GPSG, Bresnan (2001: 67) on LFG, Bender (2000) and Sag, Wasow & Bender (2003: 464) on HPSG/Sign-Based Construction Grammar. There are results from the 60s that show that phrase structure grammars containing empty elements can be translated into grammars that do not contain empty elements (Bar-Hillel, Perles & Shamir 1961: 153, Lemma 4.1). Grammars with empty elements often are more compact than those without empty elements and express generalizations more directly. See for example Bender (2000) for copulaless sentences in African American Venacular English and Müller (2014) on nounless NPs in German. The argument against empty elements usually refers to language acquisition: it is argued that empty elements cannot be learned since they are not present in the input. However, if the empty elements alternate with visible material it can be argued that what is learned is the fact that a certain element can be left out. What is true though is that things like empty expletives cannot be learned since these empty elements are neither visible nor do they contribute to meaning. Their only purpose in grammars is to keep uniformity. For example, Grewendorf (1993) suggests an analysis of the passive in German that is parallel to the movement-based analysis of English passives. In order to account for the fact that the subject does not move in German, he suggests an empty expletive pronoun that takes the subject position and that is connected to the original non-moved subject. Such elements cannot be acquired without innate knowledge about the IP/VP system and constraints about the obligatory presence of subjects. The CxG criticism is justified here.

A frequent argumentation for empty elements in MGG is based on the fact that there are overt realizations of an element in other languages (e.g., object agreement in Basque and focus markers in Gungbe). But since there is no language internal evidence for these empty elements they cannot be learned and one would have to assume that they are innate. This kind of empty elements is rightly rejected.

1.2.2 Language acquisition without the assumption of UG

Tenet 4 and 5 are basically what everybody should assume in MGG if Hauser, Chomsky & Fitch (2002) are taken seriously. Of course this is not what is done in large parts of the filed. The most extreme variant being Cinque & Rizzi (2010), who assume at least 400 functional heads being part of Universal Grammar (UG) and being present in all grammars of all languages although sometimes invisibly. Such assumptions beg the question why the genera of Bantu languages should be part of our genome and how they got there. Researchers working on language acquisition realized that the Principles & Parameters approach (Meisel 1995) makes wrong predictions. They now talk about Micro-Cues instead of parameters (Westergaard 2014) and these Micro-Cues are just features that can be learned. However, Westergaard still assumes that the features are determined by UG, a dubious assumption seen from a CxG perspective (and from the perspective of Hauser, Chomsky, Fitch and genetics in general (Bishop 2002)).

Note that even those versions of Minimalism that do not follow the Rizzi-style Cartographic approaches are far from being minimalist in their assumptions. Some distinguish between strong and weak features, some assume enumerations of lexical items from which a particular derivation draws its input from, some assume that all movement has to be feature driven. Some assume that derivations work in so-called phases and that a phase once completed is "shipped to the interfaces". Construction of phases is bottom up, which is incompatible with psycholinguistic results (see also **chapter/minimalism** Chapter 32 in this volume). None of these assumptions is a natural assumption to make from a language acquisition point. Most of these assumptions do not have any real motivation in data, the only motivation usually given is that they result in "restrictive theories". But if there is no motivation for them, this means that the respective architectural assumptions have to be part of our innate domain-specific knowledge, which is implausible according to Hauser, Chomsky & Fitch (2002).

As research in computational linguistics shows, our input is rich enough to form classes, to determine the part of speech of lexical items and even to infer syntactic structure thought to be underdetermined by the input. For instance, Bod (2009) shows that the classical auxiliary inversion examples that Chomsky still uses in his Poverty of the Stimulus arguments (Chomsky 1971: 29–33; Berwick, Pietroski, Yankama & Chomsky 2011) can also be learned from language input available to children. See also Freudenthal et al. (2006; 2007) on input-based language acquisition.

HPSG does not make any assumptions about complicated mechanisms like feature driven movement and so on. HPSG states properties of linguistic objects like part of speech, case, gender and so on and states relations between such features like agreement and government. In this respect it is like other Construction Grammars and hence experimental results regarding and theories of language acquisition can be carried over to HPSG. See also **chapters/aquisition** Chapter **??** of this volume on language acquisition.

1.2.3 Inhertiance networks

This leaves us with Tenet 6 and 7, that is *inheritance networks* and the constructicon. MGG does not make reference to inheritance hierarchies. HPSG did this right from the beginning in 1985 (Flickinger et al. 1985) for lexical items and since 1995 also for phrasal constructions (Sag 1997). LFG rejected the use of types but used macros in computer implementations. The macros were abbreviatory devices and did not play any role in theoretical work. This changed in 2004 where macros were suggested in theoretical work (Dalrymple, Kaplan & King 2004). And although any connection to constructionist work is vehemently denied by some of the authors, recent work in LFG has a decidedly constructional flavor (Asudeh, Dalrymple & Toivonen 2008; 2014).[2] LFG differs from frameworks like HPSG though in assuming a separate level of c-structure. c-structure rules are basically context free phrase structure rules and they are not modeled by feature value pairs (although they could be (Kaplan 1995)). This means that it is not possible to capture generalizations regarding lexical items, lexical rules and phrasal schemata. While HPSG describes all of these elements with the same inventory and hence can use common supertypes in the description of all three, this is not possible in LFG. TAG is also using inheritance in the Meta Grammar (Lichte & Kallmeyer 2017).

citep* yields wrong results here

add reference

Since HPSG's lexical entries, lexical rules and phrasal schemata are all described by typed feature descriptions one could call the set of these descriptions the constructicon. Therefore, tenet 7 is also adhered to.

1.2.4 Summary

If all these points are taken together, it is clear that most variants of MGG are not Construction Grammars. However, CxG had considerable influence on other frameworks so that there are constructionist variants of LFG and TAG. HPSG in the version of Sag (1997) (also called Constructional HPSG) and the HPSG dialect

[2] See Toivonen (2013: 516) for an explicit reference to construction-specific phrase structure rule in the senese of Construction Grammar.

Sign-Based Construction Grammar are Construction Grammars that follow all the tenets mentioned above.

1.3 Variants of Construction Grammar

The previous section discussed the tenets of CxG and to what degree other frameworks adhere to them. This section deals with frameworks that have Construction Grammar explicitly in their name. The following variants are usually named:

- Berkeley Construction Grammar (Fillmore 1988; Kay & Fillmore 1999; Fried 2015)
- Cognitive Construction Grammar (Lakoff 1987; Goldberg 1995; 2006)
- Cognitive Grammar (Langacker 1987; 2000; 2008; Dąbrowska 2004)
- Radical Construction Grammar (Croft 2001)
- Embodied Construction Grammar (Bergen & Chang 2005)
- Fluid Construction Grammar (Steels & De Beule 2006; Steels 2011)
- Sign-Based Construction Grammar (Sag 2010; 2012)

Berkely Construction Grammar, Embodied Construction Grammar, Fluid Construction Grammar, and Sign-Based Construction Grammar are the ones that are more formal. All of these variants use feature value pairs and are constraint-based. They are sometimes also referred to as unification-based approaches. Berkeley Construction Grammar never had a consistent formalization. The variant of unification assumed by Kay & Fillmore (1999) was formally inconsistent (Müller 2006a: Section 2.4) and the computation of construction-like objects (CLOs) suggested by Kay (2002) did not work either (Müller 2006a: Section 3). Berkeley Construction Grammar was dropped by the authors, who joined forces with Ivan Sag, Tom Wasow, and Laura Michaelis and eventually came up with an HPSG variant named Sign-Based Construction Grammar (Sag 2012). The differences between Constructional HPSG (Sag 1997) and SBCG are to some extent cosmetic: semantic relations got the suffix *-fr* for *frame* (*like-rel* became *like-fr*), phrases were called constructions (*hd-subj-ph* became *subj-head-cxt*) and lexical rules were called *derivational constructions*.[3] While this renaming would not

[3]This renaming trick was so successful that it even confused some of the co-editors of the volume about SBCG (Boas & Sag 2012). See for example Boas (2014) and the reply in Müller & Wechsler (2014b).

have changed anything in terms of expressiveness of theories, there was another change that was not motivated by any of the tenets of Construction Grammar but rather by the wish to get a more restrictive theory: Sag, Wasow & Bender (2003) and Sag (2007) changed the feature geometry of phrasal signs in a way that signs do not contain daughters. The information about mother-daughter relations is contained in lexical rules and phrasal schemata (Constructions) only. The phrasal schemata are more like GPSG phrase structure rules in licensing a mother node when certain daughters are present but without the daughters being part of the mother as it was common in HPSG from 1985 till Sag, Wasow & Bender (2003). This differs quite dramatically from what was done in Berkeley Construction Grammar, since BCxG explicitly favored a non-local approach. Arguments were not cancelled but passed up to the mother node. Adjuncts were passed up as well so that the complete internal structure of an expression is available at the top-most node. The advantage of BCxG (Fillmore, Kay & O'Connor 1988) and Constructional HPSG (Sag 1997) is that complex expressions (e.g., idioms and other more transparent expressions with high frequency) can be stored as chunks containing the internal structure. This is not possible with SBCG, since phrasal signs never contain internal structures. For a detailed discussion of Sign-Based Construction Grammar see Müller (2016: Section 10.6.2).

Rui: Head feature principle, LID from Berkeley CxG

Embodied Construction Grammar (Bergen & Chang 2005) uses typed feature descriptions for the description of linguistic objects and allows for discontinuous constituents. As argued by Müller (2016: Section 10.6.3), it is a notational variant of Reape-style HPSG (Reape 1994) (see Müller 2018: Section 5, Chapter 9 of this volume for discontinuous constituents in HPSG).

Fluid Construction Grammar is also rather similar to HPSG. An important difference is that FCG attaches weights to constraints, something that is usually not done in HPSG. But in principle there is nothing that forbids to add weights to HPSG as well and in fact it has been done (Brew 1995; Briscoe & Copestake 1999; Miyao & Tsujii 2008) and it should be done to a larger extent (Miller 2013). Van Trijp (2013) tried to show that Fluid Construction Grammar is fundamentally different from SBCG but I think he failed in every single respect. See Müller (2017b) for a detailed discussion, which cannot be repeated here for space reasons.

What makes SBCG different from other Construction Grammars is that SBCG assumes a strongly lexicalist stance (Sag & Wasow 2011): argument structure is encoded lexically. A ditransitive verb is a ditransitive verb since it selects for three NP arguments. This selection is encoded in valence features of lexical items. It is not assumed that phrasal configurations can license additional arguments as it is in Radical Construction Grammar, Embodied Construction Grammar and in Fluid Construction grammar. The next section discusses phrasal CxG approaches in more detail. Section 4 then discusses patterns that should be analyzed phrasally and which are problematic for entirely head-driven (or rather functor-driven) theories like Categorial Grammar, Dependency Grammar and Minimalism.

2 Valence vs. phrasal patterns

Much work in Construction Grammar starts from the observation that children acquire patterns and get more abstract leaving slots to be filled in in later acquisition stages (Tomasello 2003). The conclusion that is drawn from this is that language should be described with reference to phrasal patterns. Most Construction Grammar variants assume a phrasal approach to argument structure constructions (Goldberg 1996; 2006; Goldberg & Jackendoff 2004), Constructional
Boas HPSG and SBCG being the two exceptions.

I argued in several publications that the language acquisition facts can be explained in lexical models as well (Müller 2010: Section 6.3; Müller & Wechsler 2014a: Section 9). While a pattern-based approach claims that (1) is analyzed by inserting *Kim*, *loves*, and *Sandy* into a phrasal schema stating that NP[nom] verb NP[acc] or subject verb object are possible sequences in English, a lexical approach would state that there is a verb *loves* selecting for an NP[nom] and an NP[acc] (or for a subject and an object).

(1) Kim loves Sandy.

Since objects follow the verb in English (modulo extraction) and subjects precede the verb, the same sequence is licensed in the lexical approach. The lexical approach does not have any problems with accounting for patterns in which the sequence of subject, verb and object is discontinuous. For example, an adverb may intervene between subject and verb:

(2) Kim really loves Sandy.

In a lexical approach it is assumed that verb and object may form a unit (a VP). The adverb attaches to this VP and the resulting VP is combined with the subject. The phrasal approach has to assume that either adverbs are part of phrasal schemata licensing cases like (2) (see Uszkoreit (1987: Section 6.3.2) for such a proposal in a GPSG for German) or that the phrasal construction may license discontinuous patterns. Bergen & Chang (2005: 170) follow the latter approach and assume that subject and verb may be discontinuous but verb and object(s) have to be adjacent. While this accounts for adverbs like the one in (2), it does not solve the general problem since there are other examples showing that verb and object(s) may appear discontinuously as well:

(3) Mary tossed me a juice and Peter a water.

Even though *tossed* and *Peter a water* are discontinuous in (3), they are an instance of the ditransitive construction. The conclusion is that what has to be acquired is not a phrasal pattern but rather the fact that there are dependencies between certain elements in phrases. I return to ditransitive constructions in Section 2.3.

I discussed several phrasal approaches to argument structure and showed where they fail (Müller 2006a,b; 2007; 2010; Müller & Wechsler 2014a,b; Müller 2018a). Of course the discussion cannot be repeated here but I want to repeat two points showing that lexical valence representation is necessary. The first two are problems that were around at GPSG times and basically were solved by abandoning the framework and adopting a new framework which was a fusion of GPSG and Categorial Grammar: HPSG.

fix me, add point

2.1 Derivational morphology and valence

The first argument (Müller 2016: Section 5.5.1) is that certain patterns in derivational morphology refer to valence. For example, the -*bar* 'able' derivation productively applies to transitive verbs only, that is to verbs that govern an accusative.

(4) a. unterstützbar
supportable

b. * helfbar
helpable

c. * schlafbar
sleepable

Note that -*bar* 'able' derivation is like passive: it surpresses the subject and promotes the accusative object: the accusative object is the element adjectives derived with -*bar* 'able' derivation predicate over. There is no argument realized with the adjective *unterstützbaren* 'supportable' attaching to *Arbeitsprozessen* 'work.processes' in *unterstützbaren Arbeitsprozessen*.[4] Hence one could not claim

[4] Adjectives realize their arguments preverbally in German:

that the stem enters a phrasal construction with arguments and *-bar* attaches to this phrase. It follows that information about valency has to be present at the stem.

Note also that the resultative construction interacts with *-bar* 'able' derivation. (5) shows an example of this construction in which the accusative object is introduced by the construction: it is the subject of *leer* 'empty' but not a semantic argument of the verb *fischt* 'fishes'.

(5) Sie fischt den Teich leer.
she fishes the pond empty

So even though the accusative object is not a semantic argument of the verb, the *-bar* 'able' derivation is possible and an adjective like *leerfischbar* can be derived. This is explained by lexical analyses of the *-bar* 'able' derivation and the resultative construction since if one assumes that there is a lexical item for *fisch-* selecting the accusative object and the result predicate then this item may function as the input for the *-bar* 'able' derivation. See Section 3 for further discussion of *-bar* 'able' derivation and Verspoor (1997), Wechsler (1997), Wechsler & Noh (2001), Müller (2002: Chapter 5) for lexical analyses of the resultative construction in the framework of HPSG.

2.2 Partial verb phrase fronting

The second argument concerns partial verb phrase fronting (Müller 2016: Section 5.5.2). (6) gives some examples: in (6a) the bare verb is fronted and its arguments are realized in the middle field, in (6b) one of the objects is fronted together with the verb and in (6c) both objects are fronted with the verb.

(6) a. Erzählen wird er seiner Tochter ein Märchen können.
tell will he his daughter a fairy.tale can

b. Ein Märchen erzählen wird er seiner Tochter können.
a fairy.tale tell will he his daughter can

c. Seiner Tochter ein Märchen erzählen wird er können.
his daughter a fairy.tale tell will he can

'He will be able to tell his daughter a fairy tale.'

(i) der [seiner Frau treue] Mann
the his wife faithful man

'the man who is faithful to his wife'

unterstützbaren 'supportable' does not take an argument it is a complete adjectival projection like *seiner Frau treue*.

The problem with sentences such as those in (6) is that the valence requirements of the verb *erzählen* 'to tell' are realized in various positions in the sentence. For fronted constituents, one requires a rule which allows a ditransitive to be realized without its arguments or with one or two objects. This basically destroys the idea of a fixed phrasal configuration for the ditransitive construction and points again into the direction of dependencies.

Furthermore, it has to be ensured that the arguments that are missing in the prefield are realized in the remainder of the clause. It is not legitimate to omit obligatory arguments or realize arguments with other properties like a different case, as the examples in (7) show:

(7) a. Verschlungen hat er es nicht.
devoured has he.NOM it.ACC not
'He did not devour it.'

b. * Verschlungen hat er nicht.
devoured has he.NOM not

c. * Verschlungen hat er ihm nicht.
devoured has he.NOM him.DAT not

The obvious generalization is that the fronted and unfronted arguments must add up to the total set belonging to the verb. This is scarcely possible with the rule-based valence representation in GPSG. In theories such as Categorial Grammar, it is possible to formulate elegant analyses of (7) (Geach 1970). Nerbonne (1986) and Johnson (1986) both suggest analyses for sentences such as (7) in the framework of GPSG which ultimately amount to changing the representation of valence information in the direction of Categorial Grammar. With a switch to CG-like valence representations in HPSG the phenomenon of partial verb phrase fronting found elegant solutions (Höhle 2018: Section 4; Müller 1996; Meurers 1999). reference to UDC chapter?

2.3 Coercion

An important observation in constructionist work is that in certain cases verbs can be used in constructions that differ from the constructions they are normally used in. For example, verbs that are usually used with one or two arguments may be used in the ditransitive construction:

(8) a. She smiled.
b. She smiled herself an upgrade.[5]

[5] Douglas Adams. 1979. *The Hitchhiker's Guide to the Galaxy*, Harmony Books. Quoted from Goldberg (2003: 220).

c. He baked a cake.
d. He baked her a cake.

The usual explanation for sentences like (8b) and (8d) is that there is a phrasal pattern with three arguments into which intransitive and strictly transitive verbs may be entered. It is assumed that the phrasal patterns are associated with a certain meaning (Goldberg 1996; Goldberg & Jackendoff 2004). For example, the benefactive meaning of (8d) is contributed by the phrasal pattern (Goldberg 1996; Asudeh, Giorgolo & Toivonen 2014: 81).

page number

The insight that a verb is used in the ditransitive pattern and thereby contributes a certain meaning is of course also captured in lexical approaches. Briscoe & Copestake (1999) suggested a lexical rule-based analysis mapping a transitive version of verbs like *bake* onto a ditransitive one and adding the benefactive semantics. This is parallel to the phrasal approach in that it says: three-place *bake* behaves like other three-place verbs (e.g., *give*) in taking three arguments and by doing so it comes with a certain meaning (see Müller 2018a for a lexical rule-based analysis of the benefactive constructions working for both English and German despite the surface-differences of the respective languages). The lexical rule is a form-meaning pair and hence a construction. As Croft put it 15 years ago: Lexical rule vs. phrasal schema is a false dichotomy (Croft 2003).

BB: You quote Croft as saying that "lexical rule vs. phrasal schema is a false dichotomy". This might suggest that you think there is no real issue here, but of course you have argued that there is.

Briscoe & Copestake (1999) paired their lexical rules with probabilities to be able to explain differences in productivity. This corresponds to the association strength that van Trijp (2011: 141) used to relate lexical items to phrasal constructions of various kinds.

2.4 Non-predictability of valence

The last subsection discussed phrasal proposals of coercion that assume that verbs can be inserted into constructions that are compatible with the semantic contribution of the verb. Müller & Wechsler (2014a: Section 7.4) pointed out that this is not sufficiently constrained. Müller & Wechsler discussed the examples in (9), among others:

(9) a. John depends on Mary. (*counts, relies,* etc.)
b. John trusts (*on) Mary.

While *depends* can be combined with a *on*-PP, this is impossible for *trusts.* Also the form of the preposition of prepositional objects is not always predictable from semantic properties of the verb. So there has to be a way to state that certain verbs go together with certain kinds of arguments and others do not. A lexical specification of valence information is the most direct way to do this. Phrasal approaches sometimes assume other means to establish connections between lexical items and phrasal constructions. For instance, Goldberg (1995: 50) assumes that verbs are "conventionally associated with constructions". The more technical work in Fluid CxG assumes that every lexical item is connected to various phrasal constructions via coapplication links (van Trijp 2011: 141). This is very similar to Lexicalized Tree Adjoining Grammar (LTAG, Schabes, Abeillé & Joshi 1988), where a rich syntactic structure is associated to a lexical anchor. So, the phrasal approaches that link syntactic structure to lexical items are actually lexical approaches as well. As in GPSG some means makes sure that the lexical items enter into correct constructions. In GPSG this was taken care of by a number. I already discussed the GPSG shortcomings in previous subsections.

Concluding this section, it can be said that there has to be a connection between lexical items and their arguments and that a lexical representation of argument structure is the best way to establish such a relation.

3 Construction Morphology

The first publication in Construction Morphology was the masters thesis of Riehemann (1993), which later appeared as Riehemann (1998). Riehemann called her framework *Type-Based Derivational Morphology* since it was written before influential work like Goldberg (1995) appeared and before the term *Construction Morphology* (Booij 2005) was used. Riehemann did a careful corpus study on adjective derivations with the suffix *-bar* '-able'. She noticed that there is a productive pattern that can be analyzed by a lexical rule relating a verbal stem to the adjective suffixed with *-bar*.[6] The productive pattern applies to verbs governing an accusative as in (10a) but is incompatible with verbs taking a dative as in (10b):

(10) a. unterstützbar
   supportable

 b. * helfbar
   helpable

[6] She did not call her rule lexical rule but the difference between her template and the formalization of lexical rules by Meurers (2001) is the naming of the feature MORPH-B vs. DTR.

c. * schlafbar
sleepable

Intransitive verbs are also excluded as (10c) shows. Riehemann suggests the schema in (11):

fix page number

(11) Schema for productive adjective derivations with the suffix *-bar* in German according to Riehemann (1998: 17):

$$\begin{bmatrix} \textit{reg-bar-adj} \\ \text{PHONOLOGY } \boxed{1} + \textit{bar} \\ \text{MORPH-B } \left\langle \begin{bmatrix} \textit{trans-verb} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM|LOC } \begin{bmatrix} \text{CAT|VAL|COMPS } \langle \text{NP}[\textit{acc}]{:}\boxed{2} \rangle \oplus \boxed{3} \\ \text{CONT|NUC } \boxed{4} \begin{bmatrix} \text{ACT} \\ \text{UND } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{SYNSEM|LOCAL } \begin{bmatrix} \text{CATEGORY } \begin{bmatrix} \text{HEAD } \textit{adj} \\ \text{VALENCE } \begin{bmatrix} \text{SUBJ } \langle \text{NP:}\boxed{2} \rangle \\ \text{COMPS } \boxed{3} \end{bmatrix} \end{bmatrix} \\ \text{CONTENT|NUCLEUS } \begin{bmatrix} \text{RELN } \Diamond \\ \text{ARG1 } \boxed{2} \\ \text{ARG2 } \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

MORPH-B is a list that contains a description of a transitive verb (something that governs an accusative object which is linked to the undergoer role ($\boxed{2}$) and has an actor.[7] The phonology of this element ($\boxed{1}$) is combined with the suffix *-bar* and forms the phonology of the complete lexical item. The resulting object is of category *adj* and the semantics of the accusative object of the input verb ($\boxed{2}$) is identified with the one of the subject of the resulting adjective. The semantics of

[7]Note that the specification of the type *trans-verb* in the list under MORPH-B is redundant since it is stated that there has to be an accusative object and that there is an actor and an undergoer in the semantics. Depending on further properties of the grammar the specification of the type is actually wrong: productively derived particle verbs may be input to the *-bar* ‘able’ derivation and these are not a subtype of *trans-verb* since the respective particle verb rule derives both transitive (*anlachen* ‘laught at somebody’) and intransitive verbs (*loslachen* ‘start to laugh’) (Müller 2003: 296). *anlachen* does not have an undergoer in the semantic representation suggested by Stiebels (1996). See Müller (2003: 308) for a version of the *-bar* ‘able’ derivation schema that is compatible with particle verb formations as input.

the input verb ($\boxed{4}$) is embedded under a modal operator in the semantics of the adjective.

While the description of *-bar* 'able' derivation given so far captures the situation quite well, there are niches and isolated items that are exceptions. According to Riehemann (1998: 5), this was the case for 7 % of the adjectives she looked at in her corpus study. Examples are verbs ending in *-ig* like *entschuldigen* 'to excuse'. The *-ig* is dropped in the derivation:

(12) entschuldbar
excuseable

Other cases are lexicalized forms like *essbar* 'safely edible', which have a special lexicalized meaning. Exceptions of the accusative requirement are verbs selecting a dative (13a), a prepositional object (13b), reflexive verbs (13c), and even intransitive, mono-valent verbs (13d):

(13) a. unentrinnbar
inescapable

b. verfügbar
available

c. regenerierbar
regenerable

d. brennbar
inflammable

To capture generalizations about productive, semi-productive and fixed patterns/items Riehemann suggests a type hierarchy, parts of which are provided in Figure 1. The type *bar-adj* stands for all *-bar* adjectives and comes with the

why doesn't the figure start at the left?

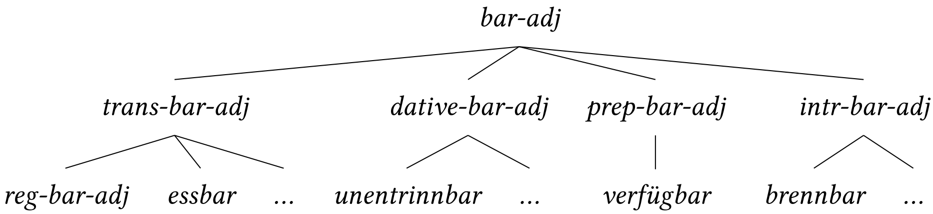


Figure 1: Parts of the type hierarchy for *-bar* 'able' derivation adapted from Riehemann (1998: 15)

constraints that apply to all of them. One subtype of this general type is *trans-bar-adj*, which subsumes all adjectives that are derived from transitive verbs. This includes all regularly derived -*bar*-adjectives, which are of the type *reg-bar-adj* but also *essbar* 'edible' and *sichtbar* 'visible'.

As this recapitulation of Riehemann's proposal shows, the analysis is a typical CxG analysis: V-*bar* is a partially filled word (see Goldberg's examples in Table 1). The schema in (11) is a form-meaning pair. Exceptions and subregularities are represented in an inheritance network.

4 Phrasal patterns

Section 2 discussed the claim that Constructions in the sense of CxG have to be phrasal. I showed that this is not true and that in fact lexical approaches to valence have to be preferred under the assumptions usually made in non-transformational theories. However, there are other areas of grammar that give exclusively head-driven approaches like Categorial Grammar, Minimalism, and Dependency Grammar a hard time. In what follows I discuss the NPN construction and various forms of filler gap constructions.

4.1 The N-P-N Construction

Matsuyama (2004) and Jackendoff (2008) discuss the NPN Construction, examples of which are provided in (14):

(14) a. Student after student left the room.
  b. Day after day after day went by, but I never found the courage to talk to her. (Bargmann 2015)

The properties of the NPN construction (with *after*) are summarized by Bargmann (2015) in a concise way and I will repeat his examples and summarization below to motivate his analysis in (22).

The examples in (14) show that the N-after-N Construction has *NP distribution*.

As (15) shows, the construction is *partially lexically fixed*: *after* cannot be replaced by any other word (Matsuyama 2004: 73).

(15) Alex asked me question { after / * following / * succeeding } question.

The construction is *partially lexically flexible*: The choice of Ns is free, except for that the Ns must be identical (16a), the Ns must be count nouns (16b), Ns must be in the singular (16c), and the Ns must be bare (16d).

(16) a. * bus after car (N1 ≠ N2)

b. * water after water (Ns = mass nouns)
c. * books after books (Ns = plurals)
d. * a day after a day (Ns have determiners)

The construction is *syntactically fixed*: N-after-N cannot be split by syntactic operations as the contrast in (17) shows (Matsuyama 2004):

(17) a. Man after man passed by.
b. * Man passed by after man.

If extraposition of the *after*-N constituent were possible, (17b) with an extraposed *after man* should be fine but it is not, so NPN seems to be a fixed configuration.

There is a syntax-semantics mismatch: while N-after-N is singular, syntactically as (18) shows, it is plural semantically as (19) shows:

(18) Study after study { reveals / *reveal } the dangers of lightly trafficked streets.

(19) a. John ate { apple after apple / apples / *an apple } for an hour.
b. John ate { *apple after apple / *apples / an apple } in an hour.

Furthermore there is an aspect of semantic sequentiality: N-after-N conveys a temporal or spatial sequence: as Bargmann (2015) states the meaning of (20a) is something like (20b).

(20) a. Man after man passed by.
b. First one man passed by, then another(, then another(, then another(, then …))).

The Ns in the construction do not refer to one individual each, rather they contribute to a holistic meaning.

The NPN construction allows adjectives to be combined with the nouns but this is restricted. N1 can only be preceded by an adjective if N2 is preceded by the same adjective:

(21) a. bad day after bad day (N1 and N2 are preceded by the same adjective.)
b. * bad day after awful day (N1 and N2 are preceded by different adjectives.)
c. * bad day after day (Only N1 is preceded by an adjective.)
d. day after bad day (Only N is preceded by an adjective.)

Finally, *after* N may be iterated to emphasize the fact that there are several referents of N as the example in (14b) shows.

This empirical description is covered by the following phrasal construction, which is adapted from Bargmann (2015):[8]

(22)
$$\left[\begin{array}{l}
\text{PHON}\ \langle \ldots N \ldots, \textit{after}, \ldots N \ldots \rangle \\
\text{SS|LOC|CAT}\ \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{noun}} \\ \text{COUNT} & - \\ \text{AGR} & \textit{3rdsing} \end{array}\right] \\
\text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{array}\right]
\end{array}\right] \\
\text{SR}\ \lambda P.\exists X.|X| > 1\ \&\ \forall x \in X\text{: } N'(x)\ \&\ \exists R^{order} \subseteq X^2\ \&\ P(x) \\
\text{DTRS}\ \left\langle \left[\begin{array}{l}
\text{PHON}\ \langle \ldots N \ldots \rangle \\
\text{SS|L|C}\ \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{l} \textit{noun} \\ \text{COUNT} + \\ \text{AGR}\ \textit{3rdsing} \end{array}\right] \\
\text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle \text{Det} \rangle \\ \text{COMPS} & \langle\rangle \end{array}\right]
\end{array}\right] \\
\text{SR} \ldots \lambda x.N'(x) \ldots
\end{array}\right],
\left(\left[\begin{array}{l}
\text{PHON}\ \langle \textit{after} \rangle \\
\ldots\ \text{HEAD}\ \textit{prep} \\
\text{SR}\ \exists R^{order} \subseteq X^2
\end{array}\right],
\left[\begin{array}{l}
\text{PHON}\ \langle \ldots N \ldots \rangle \\
\text{SS|L|C}\ \left[\begin{array}{ll}
\text{HEAD} & \left[\begin{array}{l} \textit{noun} \\ \text{COUNT} + \\ \text{AGR}\ \textit{3rdsing} \end{array}\right] \\
\text{VAL} & \left[\begin{array}{ll} \text{SPR} & \langle \text{Det} \rangle \\ \text{COMPS} & \langle\rangle \end{array}\right]
\end{array}\right] \\
\text{SR} \ldots \lambda x.N'(x) \ldots
\end{array}\right] \right)^{+} \right\rangle
\end{array}\right]$$

There is a list of daughters consisting of a first daughter and an arbitrarily long list of *after* N pairs. The '+' means that there has to be at least one *after* N pair. The nominal daughters select for a determiner via SPR, so they can be either bare nouns or nouns modified by adjectives. The semantic representation, non-standardly represented as the value of SR, says that there have to be several objects in a set X ($\exists X.|X| > 1$) and for all of them the meaning of the $\overline{\text{N}}$ has to hold ($\forall x \in X : N'(x)$). Furthermore there is an order between the elements of X as stated by $\exists R^{order} \subseteq X^2$.

[8] Jackendoff and Bargmann assume that the result of combining N, P, and N is an NP. However this is potentially problematic as Matsuyama's example in (22) shows (Matsuyama 2004: 71):

(i) All ranks joined in hearty cheer after cheer for every member of the royal family …

As Matsuyama points out the reading of such examples is like the reading of *old men and women* in which *old* scopes over both *men* and *women*. This is accounted for in structures like the one indicated in (ii):

(ii) hearty [cheer after cheer]

Since adjectives attach to $\overline{\text{N}}$s and not to NPs this means that NPN constructions should be $\overline{\text{N}}$s. Of course (ii) cannot be combined with determiners, so one would have to assume that NPN constructions select for a determiner that has to be dropped obligatorily. This is also the case for mass nouns with a certain reading.

From looking at this construction it is clear that it cannot be accounted for by standard $\overline{\text{X}}$ rules. Even without requiring $\overline{\text{X}}$ syntactic rules, there seems to be no way to capture these constructions in head-based approaches like Minimalism, Categorial Grammar or Dependency Grammar. For simple NPN constructions one could claim that *after* is the head. *after* would be categorized as 3rd singular mass noun and select for two $\overline{\text{N}}$s. It would (non-compositionally) contribute the semantics stated above. But it is unclear how the general schema with arbitrarily many repetitions of *after* N could be accounted for. If one assumes that *day after day* forms a constituent, then the first *after* in (23) would have to combine an N with an NPN sequence.

(23) day after [day [after day]]

This means that we would have to assume two different items for *after*: one for the combination of $\overline{\text{N}}$s and another one for the combination of $\overline{\text{N}}$ with NPN combinations. Note that an analysis of the type in (23) would have to project information about the $\overline{\text{N}}$s contained in the NPN construction since this information has to be matched with the single $\overline{\text{N}}$ at the beginning. In any case a lexical analysis would require several highly idiosyncratic lexical items (prepositions projecting nominal information and selecting items they usually do not select). It is clear that a reduplication account of the NPN construction as suggested by G. Müller (2011) does not work since patterns with several repetitions of PN as in (23) cannot be accounted for as reduplication. G. Müller (p. 241) stated that reduplication works for word-size elements only (in German) and hence his account does not extend to the English examples given above. (24) shows an attested German example containing adjectives, which means that G. Müller's approach is not appropriate for German either.

(24) Die beiden tauchten nämlich geradewegs wieder aus dem heimischen
the two surfaced namely straightaway again from the home
Legoland auf, wo sie im Wohnzimmer, schwarzen Stein um
Legoland PART where they in.the living.room black brick after
schwarzen Stein, vermeintliche Schusswaffen nachgebaut hatten.[9]
black brick alledged firearms recreated had
'The two surfaced straightaway from their home Legoland where they had recreated alledged firearms brick after brick.'

This subsection showed how a special phrasal pattern can be analyzed within HPSG. The next section will discuss filler-gap constructions, which were ana-

[9] taz, 05.09.2018, p. 20

lyzed as instances of a single schema by Pollard & Sag (1994) but which were later reconsidered and analyzed as a family of subconstructions by Sag (2010).

4.2 Specialized sub-constructions

HPSG took over the treatment of nonlocal dependencies from GPSG (Gazdar 1981) (see also Flickinger, Pollard & Wasow (2018), Chapter 2 of this volume on the history of HPSG and Borsley & Crysmann (2018), Chapter 13 of this volume on unbounded dependencies). Pollard & Sag (1994: Chapters 4 and 5) had an analysis of topicalization constructions like (25) and an analysis of relative clauses. However, more careful examination revealed that more fine-grained distinctions have to be made. Sag (2010: 491) looked at the following examples:

(25) a. [My bagels,] she likes. (topicalized clause)
b. [*What* books] do they like? (*wh*-interrogative)
c. (the person) [*who* (*se* book)] they like (*wh*-relative)
d. [*What a* play] he wrote! (*wh*-exclamative)
e. [*the more* books] they read … (the-clause)

As Sag shows, the fronted element is specific to the construction at hand:

(26) a. * [*Which* bagels] / [*Who*], she likes. (topicalized clause)
b. * [*What a* book] do they like? (*wh*-interrogative)
c. % the thing [[*what*] they like] (*wh*-relative)
d. * [*Which* bagels] / [*What*] she likes! (*wh*-exclamative)
e. * [*which* books] they read, the more they learn. (the-clause)

A topicalized clause should not contain a *wh* item (26a), a *wh*-interrogative should not contain a *what a* sequence as known from *wh*-exclamatives (26b) and so on.

Furthermore, some of these constructions allow non-finite clauses and others do not:

(27) a. * Bagels, (for us) to like. (topicalized clause)
b. * It's amazing [what a dunce (for them) to talk to]. (*wh*-exclamative)
c. * The harder (for them) to come, the harder (for them) to fall. (the-clause)
d. I know how much time (* for them) to take. (*wh*-interrogative)
e. The time in which (*for them) to finish. (*wh*-relative)

So there are differences as far as fillers and as far as sentences from which something is extracted are concerned. Sag discussed further differences like invertion/non-invertion in the clauses out of which something is extracted. I do not repeat the full discussion here but refer the reader to the original paper.

In principle there are several ways to model the phenomena. One could assume empty heads as Pollard & Sag (1994: Chapter 5) suggested for the treatment of relative clauses. Or one could assume empty heads as they are assumed in Minimalism: certain so-called operators have features that have to be checked and cause items with the respective properties to move. Borsley (2006) discussed potential analyses of relative clauses involving empty heads and showed that one would need a large number of such empty heads and since there is no theory of the lexicon in Minimalism, generalizations are missed (see also Borsley & Müller (2018), Chapter 32 of this volume). The alternative suggested by Sag (2010) is to assume a general Filler-Head Schema of the kind assumed in Pollard & Sag (1994) and then define more specific sub-constructions. To take an example, the wh-exclamative is a filler-head structure, so it inherits everything from the more general construction, but in addition it specifies the filler daughter to contain a *what a* part and states the semantics that is contributed by the exclamative construction.

reference

5 Summary

This paper summarized the properties of Construction Grammar or rather Construction Grammars and showed that HPSG can be seen as a Construction Grammar. I showed why lexical analyses of argument structure should be preferred over phrasal ones and that there are other areas in grammar where phrasal analyses are superior to lexical ones. I showed that they can be covered in HPSG while they are problematic for proposals assuming that all structures have to have a head.

Acknowledgements

I thank Bob Borsley, Rui Chaves, and Jean-Pierre Koenig for comments on the outline for this chapter and for discussion in general.

I thank Bob Borsley and Rui Chaves for comments on an earlier version of this paper.

References

Asudeh, Ash, Mary Dalrymple & Ida Toivonen. 2008. Constructions with lexical integrity: Templates as the lexicon-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2008 conference*. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/13/, accessed 2018-2-25.

Asudeh, Ash, Gianluca Giorgolo & Ida Toivonen. 2014. Meaning and valency. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2014 conference*, 68–88. Stanford, CA: CSLI Publications.

Bargmann, Sascha. 2015. *Syntactically flexible VP-idioms and the N-after-N Construction*. Poster presentation at the 5th General Meeting of PARSEME, Iasi, 23–24 September 2015.

Bar-Hillel, Yehoshua, Micha A. Perles & Eliahu Shamir. 1961. On formal properties of simple phrase-structure grammars. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 14(2). 143–172.

Bender, Emily M. 2000. *Syntactic variation and linguistic competence: The case of AAVE copula absence*. Stanford University dissertation. http://faculty.washington.edu/ebender/dissertation/, accessed 2018-2-25.

Bergen, Benjamin K. & Nancy Chang. 2005. Embodied Construction Grammar in simulation-based language understanding. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction Grammars: Cognitive grounding and theoretical extensions*, 147–190. Amsterdam: John Benjamins Publishing Co.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the Stimulus revisited. *Cognitive Science* 35(7). 1207–1242.

Bianchi, Valentina & Cristiano Chesi. 2006. Phases, left branch islands, and computational nesting. In *Proceedings of the 29th Penn Linguistic Colloquium* (Upenn Working Papers in Linguistics 12.1), 15–28.

Bianchi, Valentina & Cristiano Chesi. 2012. Subject islands and the subject criterion. In Valentina Bianchi & Chesi Cristiano (eds.), *Enjoy linguistics! papers offered to luigi rizzi on the occasion of his 60th birthday*. Siena: CISCL Press.

Bishop, Dorothy V. M. 2002. Putting language genes in perspective. *TRENDS in Genetics* 18(2). 57–59.

Boas, Hans C. 2014. Lexical approaches to argument structure: Two sides of the same coin. *Theoretical Linguistics* 40(1–2). 89–112.

Boas, Hans C. & Ivan A. Sag (eds.). 2012. *Sign-based Construction Grammar* (CSLI Lecture Notes 193). Stanford, CA: CSLI Publications.

Bod, Rens. 2009. From exemplar to grammar: integrating analogy and probability in language learning. *Cognitive Science* 33(4). 752–793.

Booij, Geert E. 2005. Construction-Dependent Morphology. *Lingue e linguaggio* 4. 31–46.

Borsley, Robert D. 2006. *Syntactic and lexical approaches to unbounded dependencies*. Essex Research Reports in Linguistics 49. University of Essex. 31–57. http://core.ac.uk/download/pdf/4187949.pdf#page=35, accessed 2018-2-25.

Borsley, Robert D. & Frederick J. Newmeyer. 2009. On subject-auxiliary inversion and the notion "purely formal generalization". *Cognitive Linguistics* 20(1). 135–143. DOI:10.1515/COGL.2009.007

Bouma, Gosse & Gertjan van Noord. 1998. Word order constraints on verb clusters in German and Dutch. In Erhard W. Hinrichs, Andreas Kathol & Tsuneko Nakazawa (eds.), *Complex predicates in nonderivational syntax* (Syntax and Semantics 30), 43–72. San Diego: Academic Press. http://www.let.rug.nl/~vannoord/papers/, accessed 2018-2-25.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford, UK/Cambridge, USA: Blackwell.

Brew, Chris. 1995. Stochastic HPSG. In Steven P. Abney & Erhard W. Hinrichs (eds.), *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, 83–89. Dublin: Association for Computational Linguistics.

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

Chesi, Cristiano. 2007. An introduction to phase-based Minimalist Grammars: Why *move* is top-down from left to right. In.

Chesi, Cristiano. 2012. *Phases and Cartography in linguistic computation: Toward a cognitively motivated computational model of linguistic competence*. University of Siena dissertation.

Chomsky, Noam. 1971. *Problems of knowledge and freedom*. London: Fontana.

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

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

Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 51–65. Oxford: Oxford University Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William. 2003. Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in language: Studies in honour of Günter Radden*, 49–68. Amsterdam: John Benjamins Publishing Co.

Dąbrowska, Ewa. 2004. *Language, mind and brain: Some psychological and neurological constraints on theories of grammar*. Washington, D.C.: Georgetown University Press.

Dalrymple, Mary, Ronald M. Kaplan & Tracy Holloway King. 2004. Linguistic generalizations over descriptions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2004 conference*, 199–208. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/LFG/9/, accessed 2018-2-25.

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

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3). 501–538.

Flickinger, Daniel P., Carl J. Pollard & Thomas Wasow. 1985. Structure-sharing in lexical representation. In William C. Mann (ed.), *Proceedings of the Twenty-Third Annual Meeting of the Association for Computational Linguistics*, 262–267. Chicago, IL.

Fordham, Andrew & Matthew Walter Crocker. 1994. Parsing with principles and probabilities. In Judith L. Klavans Philip Resnik (ed.), *The balancing act: combining symbolic and statistical approaches to language*. Las Cruces, New Mexico, USA: Association for Computational Linguistics.

Freudenthal, Daniel, Julian M. Pine, Javier Aguado-Orea & Fernand Gobet. 2007. Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science* 31(2). 311–341.

Freudenthal, Daniel, Julian M. Pine & Fernand Gobet. 2006. Modeling the development of children's use of optional infinitives in Dutch and English using MOSAIC. *Cognitive Science* 30(2). 277–310.

Frey, Werner. 1993. *Syntaktische Bedingungen für die semantische Interpretation: Über Bindung, implizite Argumente und Skopus* (studia grammatica 35). Berlin: Akademie Verlag.

Fried, Mirjam. 2015. Construction Grammar. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, vol. 42 (Handbooks of Linguistics and Communication Science), 974–1003. Berlin: Mouton de Gruyter.

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

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

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22. 3–17.

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

Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. *Berkeley Linguistic Society* 22. 132–146. http://roa.rutgers.edu/files/415-0900/roa-415-goldberg-2.pdf, accessed 2018-3-20.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language* (Oxford Linguistics). Oxford, New York: Oxford University Press.

Goldberg, Adele E. 2009. The nature of generalization in language. *Cognitive Linguistics* 20(1). 93–127.

Goldberg, Adele E. & Ray S. Jackendoff. 2004. The English resultative as a family of Constructions. *Language* 80(3). 532–568.

Grewendorf, Günther. 1993. German: A grammatical sketch. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax – Ein internationales Handbuch zeitgenössischer Forschung*, vol. 9.2 (Handbücher zur Sprach- und Kommunikationswissenschaft), 1288–1319. Berlin: Walter de Gruyter Verlag. DOI:10.1515/9783110095869.1

Haider, Hubert. 1993. *Deutsche Syntax – generativ: Vorstudien zur Theorie einer projektiven Grammatik* (Tübinger Beiträge zur Linguistik 325). Tübingen: Gunter Narr Verlag.

Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298. 1569–1579. DOI:10.1126/science.298.5598.1569

Höhle, Tilman N. 2018. Spuren in HPSG. In Stefan Müller, Marga Reis & Frank Richter (eds.), *Beiträge zur Grammatik des Deutschen: Gesammelte Schriften von Tilman N. Höhle* (Classics in Linguistics 5), 491–498. Berlin: Language Science Press. DOI:10.5281/zenodo.1145680

Hunter, Tim. 2010. Unifying adjunct islands and freezing effects in Minimalist Grammars. In *Proceedings of the 10th international workshop on tree adjoining grammar and related frameworks (tag+ 10)*, 17–24.

Hunter, Tim. 2018. Left-corner parsing of Minimalist Grammars. In Bob Berwick & Ed Stabler (eds.). Oxford: Oxford University Press.

Jackendoff, Ray S. 2008. Construction after Construction and its theoretical challenges. *Language* 84(1). 8–28.

Johnson, Mark. 1986. A GPSG account of VP structure in German. *Linguistics* 24(5). 871–882.

Kaplan, Ronald M. 1995. The formal architecture of Lexical-Functional Grammar. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.), *Formal issues in Lexical-Functional Grammar* (CSLI Lecture Notes 47), 7–27. Stanford, CA: CSLI Publications.

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

Kay, Paul & Charles J. Fillmore. 1999. Grammatical Constructions and linguistic generalizations: The What's X Doing Y? Construction. *Language* 75(1). 1–33.

Kay, Paul & Laura A. Michaelis. 2017. A few words to do with multiword expressions. In Cleo Condoravdi & Tracy Holloway King (eds.), *Festschrift for lauri karttunen*. Stanford, CA: CSLI Publications. To appear.

Kay, Paul, Ivan A. Sag & Daniel P. Flickinger. 2015. *A lexical theory of phrasal idioms*. Ms. CSLI Stanford.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum* (Linguistische Arbeiten 333). Tübingen: Max Niemeyer Verlag.

Kolb, Hans-Peter & Craig L. Thiersch. 1991. Levels and empty categories in a Principles and Parameters based approach to parsing. In Hubert Haider & Klaus Netter (eds.), *Representation and derivation in the theory of grammar* (Studies in Natural Language and Linguistic Theory 22), 251–301. Dordrecht: Kluwer Academic Publishers.

König, Esther. 1999. LexGram: A practical Categorial Grammar formalism. *Journal of Language and Computation* 1(1). 33–52.

Koster, Jan. 1978. *Locality principles in syntax*. Dordrecht: Foris Publications.

Koster, Jan. 1987. *Domains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris Publications.

Lakoff, George. 1987. *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. Vol. 1. Stanford, CA: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford, CA: CSLI Publications.

Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

Lichte, Timm & Laura Kallmeyer. 2017. Tree-Adjoining Grammar: A tree-based constructionist grammar framework for natural language understanding. In *Proceedings of the AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding* (Technical Report SS-17-02). Association for the Advancement of Artificial Intelligence.

Lohnstein, Horst. 1993. *Projektion und Linking: Ein prinzipienbasierter Parser fürs Deutsche* (Linguistische Arbeiten 287). Tübingen: Max Niemeyer Verlag.

Matsuyama, Tetsuya. 2004. The N *after* n construction: A constructional idiom. *English Linguistics* 1(11). 55–84.

Meisel, Jürgen. 1995. Parameters in acquisition. In Paul Fletcher & Brian MacWhinny (eds.), *The handbook of child language*, 10–35. Oxford: Blackwell Publishing Ltd.

Meurers, Walt Detmar. 1999. German partial-VP fronting revisited. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and Constructional aspects of linguistic explanation* (Studies in Constraint-Based Lexicalism 1), 129–144. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar. 2000. *Lexical generalizations in the syntax of German non-finite constructions*. Arbeitspapiere des SFB 340 No. 145. Tübingen: Eberhard-Karls-Universität. http://www.sfs.uni-tuebingen.de/~dm/papers/diss.html, accessed 2018-2-25.

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

Miller, Philip. 2013. Usage preferences: The case of the English verbal anaphor *do so*. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 121–139. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2013/miller.pdf, accessed 2018-8-17.

Miyao, Yusuke & Jun'ichi Tsujii. 2008. Feature forest models for probabilistic HPSG parsing. *Computational Linguistics* 34(1). 35–80.

Müller, Gereon. 2011. Regeln oder Konstruktionen? Von verblosen Direktiven zur sequentiellen Nominalreduplikation. In Stefan Engelberg, Anke Holler & Kristel Proost (eds.), *Sprachliches Wissen zwischen Lexikon und Grammatik* (Institut für Deutsche Sprache, Jahrbuch 2010), 211–249. Berlin: de Gruyter.

Müller, Stefan. 1996. Yet another paper about partial verb phrase fronting in German. In Jun-ichi Tsuji (ed.), *Proceedings of COLING-96: 16th International Conference on Computational Linguistics (COLING96). Copenhagen, Denmark, August 5–9, 1996*, 800–805. Copenhagen, Denmark: Association for Computational Linguistics.

Müller, Stefan. 2002. *Complex predicates: Verbal complexes, resultative constructions, and particle verbs in German* (Studies in Constraint-Based Lexicalism 13). Stanford, CA: CSLI Publications.

Müller, Stefan. 2003. Solving the bracketing paradox: An analysis of the morphology of German particle verbs. *Journal of Linguistics* 39(2). 275–325. DOI:10.1017/S0022226703002032

Müller, Stefan. 2005. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201. 3–39.

Müller, Stefan. 2006a. Phrasal or lexical Constructions? *Language* 82(4). 850–883. DOI:10.1353/lan.2006.0213

Müller, Stefan. 2006b. Resultativkonstruktionen, Partikelverben und syntaktische vs. lexikonbasierte Konstruktionen. In Kerstin Fischer & Anatol Stefanowitsch (eds.), *Konstruktionsgrammatik: Von der Anwendung zur Theorie* (Stauffenburg Linguistik 40), 177–202. Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2007. Phrasal or lexical Constructions: Some comments on underspecification of constituent order, compositionality, and control. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 373–393. Stanford, CA: CSLI Publications.

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

Müller, Stefan. 2014. Elliptical constructions, multiple frontings, and surface-based syntax. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, 91–109. Stanford, CA: CSLI Publications.

Müller, Stefan. 2016. *Grammatical theory: From Transformational Grammar to constraint-based approaches* (Textbooks in Language Sciences 1). Berlin: Language Science Press. DOI:10.17169/langsci.b25.167

Müller, Stefan. 2017a. *German sentence structure: An analysis with special consideration of so-called multiple fronting* (Empirically Oriented Theoretical Morphology and Syntax). Berlin: Language Science Press. Revise and resubmit.

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

Müller, Stefan. 2018a. *A lexicalist account of argument structure: Template-based phrasal LFG approaches and a lexical HPSG alternative* (Conceptual Foundations of Language Science 3). Berlin: Language Science Press. http://hpsg.hu-berlin.de/~stefan/Pub/phrasal-lfg-langsci.html, accessed 2018-3-20. To appear.

Müller, Stefan. 2018b. *Germanic syntax*. Ms. Humboldt Universität zu Berlin, to be submitted to Language Science Press. Berlin. https://hpsg.hu-berlin.de/~stefan/Pub/germanic.html, accessed 2018-3-20.

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

Müller, Stefan & Stephen Mark Wechsler. 2014b. Two sides of the same slim Boojum: Further arguments for a lexical approach to argument structure. *Theoretical Linguistics* 40(1–2). 187–224.

Nerbonne, John. 1986. 'Phantoms' and German fronting: Poltergeist constituents? *Linguistics* 24(5). 857–870.

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

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

Reape, Mike. 1994. Domain union and word order variation in German. In John Nerbonne, Klaus Netter & Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar* (CSLI Lecture Notes 46), 151–198. Stanford, CA: CSLI Publications.

Riehemann, Susanne. 1993. *Word formation in lexical type hierarchies: A case study of* bar-*adjectives in German*. Eberhard-Karls-Universität Tübingen MA thesis. Also published as SfS-Report-02-93, Seminar für Sprachwissenschaft, University of Tübingen.

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

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

Sag, Ivan A. 2007. Remarks on locality. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, 394–414. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2007/, accessed 2018-2-25.

Sag, Ivan A. 2010. English filler-gap constructions. *Language* 86(3). 486–545.

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

Sag, Ivan A. & Thomas Wasow. 2011. Performance-compatible competence grammar. In Robert D. Borsley & Kersti Börjars (eds.), *Non-transformational syntax: Formal and explicit models of grammar: A guide to current models*, 359–377. Oxford, UK/Cambridge, MA: Blackwell Publishing Ltd.

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

Schabes, Yves, Anne Abeillé & Aravind K. Joshi. 1988. *Parsing strategies with 'lexicalized' grammars: Application to Tree Adjoining Grammars.* Technical Report MS-CIS-88-65. University of Pennsylvania Department of Computer & Information Science.

Steels, Luc (ed.). 2011. *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11). Amsterdam: John Benjamins Publishing Co.

Steels, Luc & Joachim De Beule. 2006. *A (very) brief introduction to Fluid Construction Grammar.* Paper presented at the Third International Workshop on Scalable Natural Language Understanding (ScaNaLU 2006) June 8, 2006, following HLT/NAACL, New York City.

Stiebels, Barbara. 1996. *Lexikalische Argumente und Adjunkte: Zum semantischen Beitrag verbaler Präfixe und Partikeln* (studia grammatica 39). Berlin: Akademie Verlag.

Toivonen, Ida. 2013. English benefactive NPs. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG 2013 conference*, 503–523. Stanford, CA: CSLI Publications.

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

Uszkoreit, Hans. 1987. *Word order and constituent structure in German* (CSLI Lecture Notes 8). Stanford, CA: CSLI Publications.

van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Luc Steels (ed.), *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11), 115–145. Amsterdam: John Benjamins Publishing Co.

van Trijp, Remi. 2013. A comparison between Fluid Construction Grammar and Sign-Based Construction Grammar. *Constructions and Frames* 5(1). 88–116.

Veenstra, Mettina Jolanda Arnoldina. 1998. *Formalizing the Minimalist Program*. Rijksuniversiteit Groningen Ph.D. thesis.

Verspoor, Cornelia Maria. 1997. *Contextually-dependent lexical semantics*. University of Edinburgh dissertation.

Wechsler, Stephen Mark. 1997. Resultative predicates and control. In Ralph C. Blight & Michelle J. Moosally (eds.), *Texas Linguistic Forum 38: The syntax and semantics of predication: Proceedings of the 1997 Texas Linguistics Society Conference*, 307–321. Austin, Texas: University of Texas, Department of Linguistics.

Wechsler, Stephen Mark & Bokyung Noh. 2001. On resultative predicates and clauses: parallels between Korean and English. *Language Sciences* 23(4). 391–423.

Westergaard, Marit. 2014. Linguistic variation and micro-cues in first language acquisition. *Linguistic Variation* 14(1). 26–45.

Name index

Language index

Subject index

Did you like this book?

This book was brought to you for free

Head-Driven Phrase Structure Grammar: The handbook

change blurb in localmetadata.tex