Chapter 4

The nature and role of the lexicon in HPSG

Anthony Davis
Southern Oregon University

Jean-Pierre Koenig
University at Buffalo

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

1 Introduction

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

Anthony Davis & Jean-Pierre Koenig. 2019. The nature and role of the lexicon in HPSG. in Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 67–114. Berlin: Language Science Press. DOI:??

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

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

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

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

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

2 Lexicalism

2.1 Lexicalism and the origins of HPSG

Lexicalism began as a reaction to Generative Semantics, which treated any regularity in the structure of words (derivational patterns, broadly speaking) as only epiphenomenally 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 relations between 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 has relied heavily on treating relations between stems and between words as lexical rules, rather than the kind of generative devices that one finds in syntax. But, as accounts of linguistic phenomena in LFG focused increasingly on the lexicon, the question of whether lexical rules retain the character of redundancy rules or turn into yet another kind of generative device arose. Consequently, the necessity of lexical rules has been questioned as well (see Koenig & Jurafsky 1994 and Koenig 1999: 29–49 for potential issues that arise once lexical rules are assumed to be involved in the creation of new lexical items).

Another stream of research relying on a richly structured lexicon is Generative Lexicon theory (GL). Pustejovsky (1991; 1995) and Pustejovsky & Jezek (1996) present the elements of this approach to lexical representation, which focuses on semantic phenomena such as coercion and systematic polysemy. Within GL, lexical entries include, in addition to argument structure, an "event structure" and a "qualia structure", both of which play essential roles in GL accounts of semantic composition. For example, the natural interpretation of *enjoy the sandwich* as enjoying eating the sandwich arises from information in the event structures of *enjoy* and *sandwich* and the qualia structure of *sandwich*, which unify to yield this interpretation.

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

(1) I expected that he would leave yesterday.

(2) I expected him to leave yesterday.

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

2.2 Recent challenges to lexicalism

As there have been several challenges to lexicalism (see Bruening 2018 and Haspelmath 2011 among others for some recent challenges), we now explore lexicalism and lexical integrity in HPSG in more detail. We first note that lexicalism does not imply that word and phrase formation are necessarily different "components" as is often claimed (see Marantz 1997, Bruening 2018). Some lexicalist approaches *do* assume that word formation and phrase building belong to

[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.

two different components of a language's grammar (this is certainly true of Jackendoff 1975), but they need not. Within HPSG, there are approaches that treat every sign-formation (be it word-internal or word-external) as resulting from typed mother-daughter configurations: this is the hypothesis pursued in Koenig 1999, and is also the approach frequently taken in implementations of large-scale grammars where lexical rules are modeled as unary-branching trees; see the English Resource Grammar (Copestake 2002) and the grammars developed in the CoreGram project (Müller 2015) (see Müller 2018: 58 for a similar point in his response to Bruening's paper). Furthermore, recent approaches to inflectional morphology within HPSG model realizational rules through the very same tools the rest of a language's grammar uses (see Crysmann & Bonami 2016 and Crysmann 2019, Chapter 22 of this volume). There are also approaches to phrases where the same analytical tools developed to model lexical knowledge (see Section 4) are employed to model phrase-structural constructions (see Sag's 1997 analysis of relative clauses, for example). So, both in terms of the formal devices and in terms of analytical tools used to model datasets, words and phrases can be treated the same way in HPSG (although they need not be). Somewhat ironically, and despite claims to the contrary, word formation in the syntactocentric approach Marantz or Bruening advocates *does* make use of distinct formal machinery to model word formation, namely realizational rules as well as various readjustment rules, as well as fusion and fission rules, to model inflectional morphology (see Halle & Marantz 1993; Embick 2015).

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

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

We can provisionally accept for the sake of argument Bruening's contention that *don't-you-dare!* is a word in (3), despite its reliance on the (unjustified) assumption that the secondary object in (3) involves N-N compounding rather than an AP N structure (we refer readers to Bresnan & Mchombo 1995 or Müller 2018 for counter-arguments to Bruening's claim). Crucially, though, examples such as (3) have no bearing on HPSG's model of lexical knowledge, as HPSG-style lexicalism does not preclude constructions that form words from phrases. Nothing, as far as we know, rules out constructions of the form *stem/word* → *phrase* in HPSG. The two assumptions underlying the HPSG brand of lexicalism we mentioned above do not preclude a *lexical-sign* having a *phrasal-sign* as sole daughter (although we do not know of any HPSG work that exploits this possibility) and examples such as (3) are simply irrelevant to whether HPSG's lexicalist stance is empirically correct.

The second challenge to lexicalism presented in Bruening (2018) bears more directly on HPSG's assumption that words are syntactic atoms. Word-internal conjunction/ellipsis examples, illustrated in (4) (adapted from Bruening's (31a), p. 14), seem to violate the assumption that syntactic constraints cannot "see" the internal structure of words, 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 (p. 304) makes clear, such an analysis is fully compatible with lexical integrity, as there is no access to the internal structure of composite words, only to the (enriched) morphophonology of the entire word.

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

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

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

Endoclitics are clitics that at least appear to be situated within a word, rather than immediately preceding or following it, as clitics often do (Crysmann & Penn 2019, Chapter 11 of this volume). 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 extend to these as well. But some trickier cases have also come to light, in which the clitic appears within a morpheme, not at a boundary. Two of the best documented cases come from the Northeast Caucasian language Udi (Harris 2000) (see (6)) and from Pashto (Tegey 1977; Roberts 2000; Dost 2007) (see (7)).

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

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

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

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

If clitics are viewed as a syntactic phenomenon ("phrasal affixes", as Anderson 2005 puts it), these endoclitics must "see" into the internal structure of words (be it morphological, prosodic, or something else), thereby seemingly violating lexical integrity. Anderson's brief account invokes a reranking of optimality theoretic constraints from their typical ordering, whereby the clitic's positional requirements outrank lexical integrity requirements. Crysmann (2000) proposes an analysis, paralleling in many respects his account of European Portuguese clitics in Crysmann (2001), using Reape's constituent order domains (Reape 1994) and, in particular, Kathol's topological fields (Kathol 1999) (see Müller (2019a), Chapter 10 of this volume). The "morphosyntactic paradox" in Udi, to borrow a phrase from Crysmann (2002), is effectively "resolved on the basis of discontinuous lexical items"; this account then "parallels HPSG's representation of syntactic discontinuity" (Crysmann 2000), (Müller 2019a, Chapter 10 of this volume).

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

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

3 Lexical entries in HPSG

3.1 What are lexical entries?

Because lexical entries or items (the terms are used interchangeably, it seems, within the HPSG literature) play a central role in accounting for the syntax of natural languages, lexical entries are informationally rich in HPSG. An additional consequence of HPSG's lexicalist stance is that there will be many lexical entries where one might at first glance expect a single entry. We will see below how HPSG handles multiple entries and classes of entries while avoiding redundancy, but it is important at the outset to clarify what a lexical entry is in HPSG. One misunderstanding about lexical entries conflates descriptions and the entities they describe, or, in other words, fails to distinguish between constructions in the abstract and a particular word or phrase (i.e., a lexical entry vs. a fully instantiated lexeme). As Richter (2019), Chapter 3 of this volume makes explicit, grammars in HPSG consist of *descriptions* of structures, and the lexicon thus consists of descriptions of what a fully specified lexeme or word can be. To see the importance of the distinction between descriptions (stored entries) and the fully instantiated objects that are being described, consider HPSG's model of subcategorization with reference to the relevant portion of the tree for sentence (1). HPSG's model of the dependency between heads and complements stipulates identity between the syntactic and semantic information of each complement (the value of the SYNSEM attribute) and a member of the list of complements the head subcategorizes for. Since there are indefinitely many SYNSEM values, on the assumption that there are indefinitely many clausal meanings (a point Jackendoff 1990: 8–9 emphasizes), there are, in principle, indefinitely many fully instantiated entries for the verb *see* subcategorizing for a clausal complement (as in (1)). But each of these fully instantiated entries for *expect* – one for each clausal sentence that corresponds to the tree in Figure 1 – corresponds to a single abstract description, and it is this description that the lexicon contains.

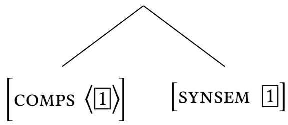

Figure 1: A head-complement phrase

The formal status of lexical entries has engendered a fair amount of theoretical work and some debate. We will touch on some aspects of this further below, in connection with online type construction. For further discussion of these kinds of issues, see Richter (2019), Chapter 3 of this volume and Borsley & Abeillé (2019), Chapter 1 of this volume.

3.2 What information is in lexical entries?

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

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

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

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

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

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

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

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

One other question arises with regard to the information in lexical entries. Are there attributes or values that occur solely within lexical signs, and not in phrasal ones? If so, they would provide a diagnostic for distinguishing lexical signs from others. Specific phonological information, for instance, is something we would expect to be introduced by lexical entries, and not elsewhere. Such information would thus be specific to lexical signs. Likewise, the ARG-ST list, which we included in the categorial information of signs in (10) and (11), might be regarded as a feature confined to lexical signs (see, among others, Ginzburg & Sag 2000), on the premise that lexical items alone specify combinatorial requirements (but see Przepiórkowski 2001 for a contrary view, and see Müller 2019b, Chapter 36 of this volume for other views questioning this assumption). But HPSG researchers have generally not explored this question in depth, and we will leave this issue here.

3.3 The role of the lexicon in HPSG

As we hope is evident by now, the lexicon plays a critical role in HPSG's explanatory mechanisms, as lexical entries encode not merely their idiosyncratic phonological and semantic characteristics, but their distributional and combinatorial potential as well. Much of the information contained in lexical entries is geared towards modeling how words interact with one another, as we have already seen. As detailed in Wechsler, Koenig & Davis (2019), Chapter 9 of this volume, their combinatorial potential is recorded using two kinds of information, a list of syntactic arguments or syntactic requirements to be satisfied, and distinct lists that indicate how these requirements are to be satisfied (as local dependents, as non-local dependents, or as clitics/affixes). Not only are syntactic arguments recorded; so is their relative obliqueness (in terms of grammatical function), as per the partial hierarchy in (12) from Pollard & Sag (1992).

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

We illustrate this explanatory role by noting the role of the lexicon in HPSG's approach to binding, as described in Pollard & Sag (1992) (see Branco (2019), Chapter 21 of this volume for details). As lexical entries of heads record both syntactic and semantic properties of their dependents, constraints between properties of heads and properties of dependents, e.g., subject-verb agreement, or between dependents, e.g., binding constraints, illustrated in (14), can be stated, at least partially, as constraints on classes of lexical entries. The principle in (14) is such a constraint.

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

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

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

(15) a. $\left[\text{ARG-ST}\ \langle \text{NP}_{i,+ana}, ... \rangle\right]$

b. $\left[\text{ARG-ST}\ \langle \text{NP}_{i,-ana}, ..., \text{NP}_{i,+ana}, ... \rangle\right]$

c. $^{*}\left[\text{ARG-ST}\ \langle \text{XP}_{j}, ..., \text{NP}_{i,+ana}, ... \rangle\right]$

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

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

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

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

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

3.4 Lexical vs. constructional explanations

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

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

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

4 The hierarchical lexicon

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

HPSG is, of course, not the only linguistic framework to exploit inheritance, although HPSG researchers, perhaps more than others, have emphasized its central role in expressing lexical generalizations. Appeals to similar mechanisms feature prominently in Generative Lexicon accounts of lexical semantics, for example. Both the lexical typing structure and qualia structures within GL, in particular the formal quale, have values situated in type hierarchies Pustejovsky & Jezek (1996) and GL accounts of coercion and metonymy rely crucially on multiple inheritance within qualia values.

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

4.1 Inheritance

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

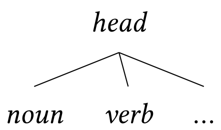

Figure 2: A hierarchy of types of *head*

Each of these types is a partial specification of a lexical entry's head properties. Typing of HEAD information allows the ascription of appropriate properties to different classes of lexical entries. For example, case information is only relevant to nouns and adjectives, and whether a verb is an auxiliary or not is only relevant to verbs. The subtypes of *head* in (2) allow us to define additional specifications of the properties appropriate for different parts of speech. For example, lexical entries with a HEAD value of *noun* contain an attribute for CASE, while those for *verb* contain the attributes AUX, TENSE, and ASPECT, as shown in (17) (we use implicational statements in (17) to indicate feature appropriateness conditions for the types *noun* and *verb* for perspicuity only; such conditions would be part of the grammar's signature, see Richter (2019), Chapter 3 of this volume). In other words, the grammar's signature will specify that only for nouns (those lexical entries whose HEAD value is of type *noun*) is the attribute CASE appropriate. Similarly, only for verbs are the attributes AUX, TENSE, and ASPECT appropriate.

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

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

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

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

4.2 Representing lexical generalizations

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

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

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

However, a hierarchy of subtypes of *word* is, while formally feasible, potentially rather inelegant. Note first that types like *verb* and *noun* are already defined as subtypes of the type *head*. There is an obvious danger of redundancy if we additionally posit parallel subtypes of *word* such as *verb-word* and *noun-word*, serving no other function than as types with the corresponding HEAD values. Furthermore, signs in HPSG are structured objects, with their various kinds of information deliberately arranged in a way that associates pieces of information that "travel together." The information within HEAD, for example, is grouped there because it is all subject to the Head Feature Principle. Both part-of-speech and control information are found within SYNSEM, as phonological information has no bearing on these things. So rather than creating subtypes of *word* to capture regularities in the lexicon, we would prefer to express those regularities as constraints on subtypes that encompass only the information that's pertinent. These are the smallest, "narrowest" portions of *word* objects that include all that information; the remaining portions of a *word* can be ignored in this context. In other words, we take advantage of HPSG's feature geometry and of the hierarchies of types appropriate for a particular substructure within signs to express generalizations as "locally" as possible (see Richter (2019), Chapter 3 of this volume).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

4.3 Cross-cutting types in the lexicon

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

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

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

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

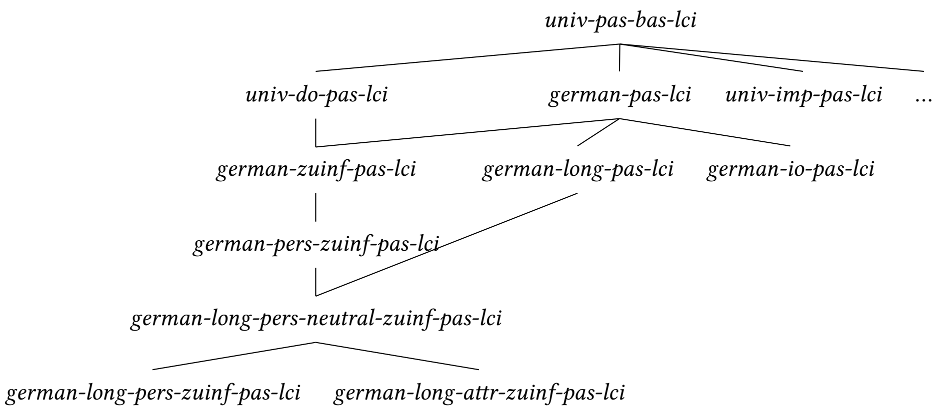

Figure 3: A portion of the hierarchy of passive lexical types according to Ackerman & Webelhuth (1998: 244)

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

4.4 Default inheritance in the lexicon

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

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

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

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

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

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

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

$$|$$

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

$$|$$

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

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

morphology and phonology here). In the more specific type *pst-t-verb*, however, the default *-ed* is overriden by (default) information that the suffix is *-t*.

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

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

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

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

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

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

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

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

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

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

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

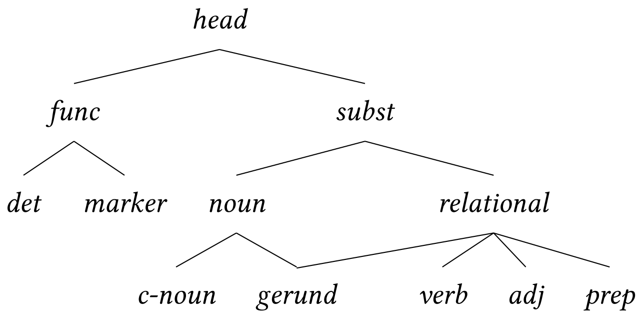

Figure 5: A cross-cutting hierarchy of types of *head* according to Malouf (2000: 65)

(30) $\begin{bmatrix} \textit{v} & \\ \textsc{head} & \textit{relational} \,/\textit{verb} \\ \textsc{cont} & \textit{psoa} \end{bmatrix}$

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

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

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

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

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

5 Lexical rules

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

5.1 What is the nature of lexical rules in HPSG?

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

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

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

(32) a. $\begin{bmatrix} \text{IN} & \begin{bmatrix} \text{SYNSEM} & a \\ \text{PHON} & b \end{bmatrix} \\ \text{OUT} & \begin{bmatrix} \text{SYNSEM} & c \\ \text{PHON} & d \end{bmatrix} \end{bmatrix}$

b. $\begin{bmatrix} \text{SYNSEM} & c \\ \text{PHON} & d \\ \text{IN} & \begin{bmatrix} \text{SYNSEM} & a \\ \text{PHON} & b \end{bmatrix} \end{bmatrix}$

In the variant in (32b), lexical rules are treated as subtypes of a *derived-word* type, which can combine with other types in the lexical hierarchy, merely adding the derivational source via the IN value. Formulated in either fashion, lexical rules are essentially equivalent to unary syntactic rules, with the IN attribute corresponding to the daughter and the OUT attribute to the mother (or the rest of the information in the rule, if the OUT attribute is done away with). This is the way lexical rules are implemented in the English Resource Grammar (see http://www.delph-in.net/erg/ for demos and details about this large-scale implemented grammar of English) as well in the CoreGram Project and the Grammix grammar development environment (see Müller 2007 and https://hpsg.hu-berlin.de/Software/Grammix/ for details on the Grammix software) (Bender & Emerson 2019, Chapter 28 of this volume).

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

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

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

5.2 Phenomena accounted for by lexical rules

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

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

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

A similar use of lexical rules to model alternative realizations of arguments can be found in Monachesi (1993), who analyzes alternations between complements and pronominal object affixes (traditionally called object clitics) in Italian in a way that parallels the French examples in (9). $\bigcirc$ in the rule, shown in (36), a.k.a. the "shuffle" operation, stands for the unordered concatenation of two lists, since any member of the input's COMPS list can be realized as a clitic and therefore not included in the output's COMPS list (see Müller (2019a), Chapter 10 of this volume for a more formal explanation of $\bigcirc$). In the output of the lexical rule in (36), a subset of the list of complements in the input ($\boxed{2}$) corresponds to a list of clitic SYNSEMS, realized as prefixes through inflectional rules not shown here.

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

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

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

5.3 Alternatives to lexical rules

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

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

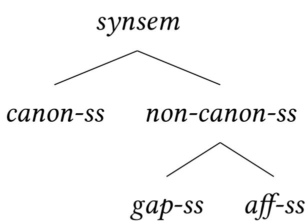

Figure 6: Types of *synsem*

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

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

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

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

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

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

Figure 7 is a simplified hierarchy of verb lexemes we use for strictly expository purposes, where the boxed labels in small caps VFORM and ARG-ST are mnemonic names of orthogonal dimensions of classification of subcategories of verbs (and

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

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).

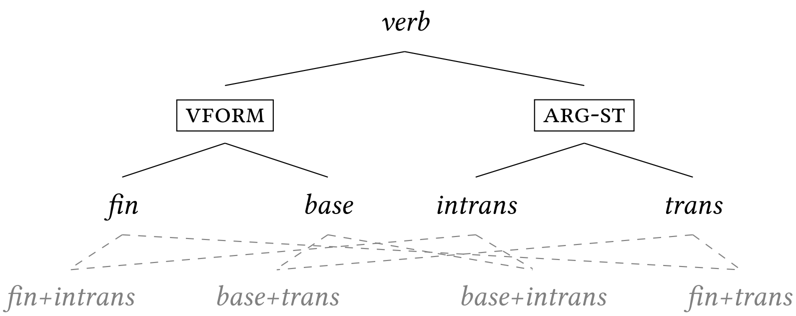

Figure 7: An example of on-line type construction

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 (Stump 1993; Bonami & Crysmann 2016).

For reasons of space we illustrate the usefulness of this dynamic approach to type creation, the Type Underspecified Hierarchical Lexicon (TUHL), with an example from Koenig (1999): the cross-cutting classification of syntactic/semantic information and stem form in the entry for the French verb *aller* (see Bonami & Boyé 2001 for a much more thorough discussion of French stem allomorphy along similar lines; Crysmann & Bonami's much more developed approach to stem allomorphy would model the same phenomena differently and we use Koenig's simplified presentation for expository purposes only). The forms of *aller* are based on four different suppletive stems: *all-* (1st and 2nd person plural of the indicative and imperative present, infinitive, past participle, and imperfective past), *i-*

(future and conditional), *v-* (1st, 2nd, or 3rd person singular and 3rd person plural of the indicative present), and *aill-* (subjunctive present). These four suppletive stems are shared by all entries (i.e., senses) of the lexeme *aller*: the one which means 'to fit' as well as the one which means 'to leave', as shown in (39) (see Koenig 1999: 40–41). The cross-cutting generalizations over lexemes and stems are represented in Figure 8. Any *aller* stem combines one entry and one stem form. In a traditional HPSG type hierarchy, each combination of types (grayed out in Figure 8), would have to be stipulated. In a TUHL, these combinations can be dynamically created when an instance of *aller* needs to be produced or comprehended.

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

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

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.'

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

Lexical rules have played a crucial role in the rise of lexicalist approaches to syntax. But the two alternative analytical tools we discussed in this section (which, of course, can be combined in an analysis) have chipped away at their use in HPSG. Inflectional morphology is now dealt with through lexical types associating morphosyntactic features with forms/positions and constraints on words (ensuring that all morphosyntactic features are realized, see Crysmann (2019), Chapter 22 of this volume). Non-canonical realization of syntactic arguments as

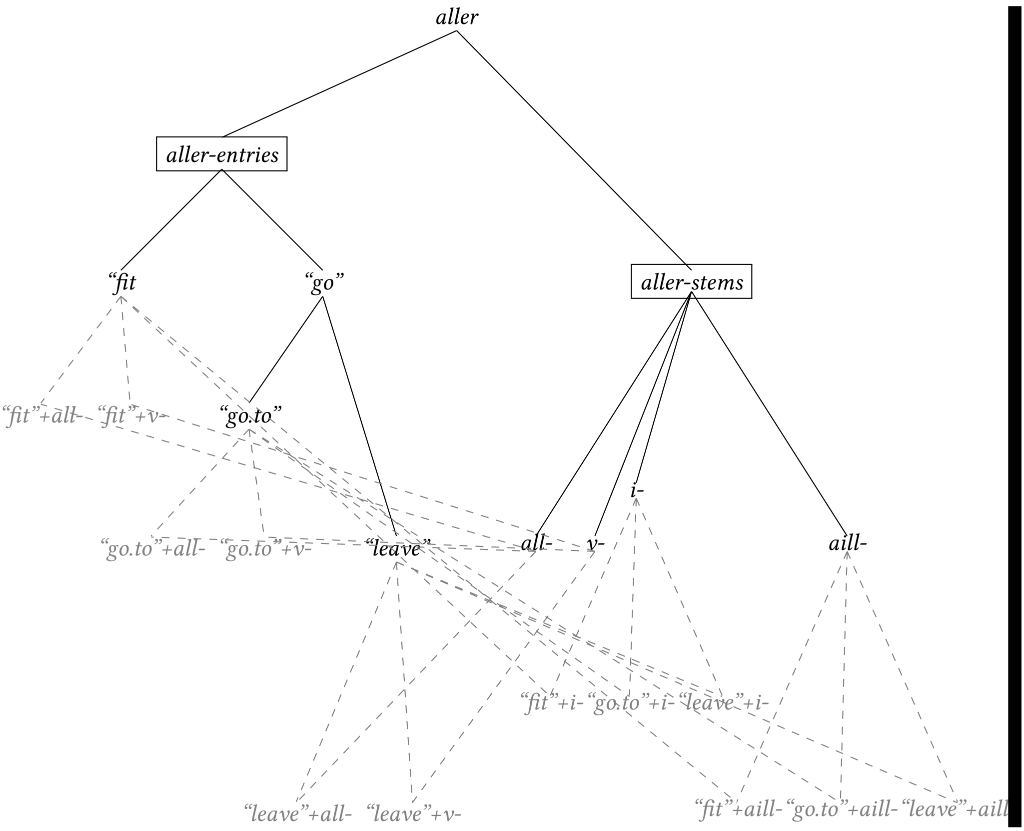

Figure 8: A hierarchy of lexical entries and stem-forms for the French verb *aller*, from Koenig (1999)

affixes or fillers in unbounded dependencies is now modeled by distinguishing kinds of members of the ARG-ST list and constraints on words that relate valence, argument structure, and dependents lists.

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

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

6 Conclusion

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

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

References

Ackerman, Farrell & Gert Webelhuth. 1998. *A theory of predicates* (CSLI Lecture Notes 76). Stanford, CA: CSLI Publications.

Anderson, Stephen. 1992. *A-morphous morphology* (Cambridge Studies in Linguistics 62). Cambridge: Cambridge University Press.

Anderson, Stephen. 2005. *Aspects of the theory of clitics* (Oxford Studies in Theoretical Linguistics 11). Oxford: Oxford University Press.

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

Bender, Emily M. & Guy Emerson. 2019. Computational linguistics and grammar engineering. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 863–907. Berlin: Language Science Press. DOI:??

Bochner, Harry. 1993. *Simplicity in Generative Morphology* (Publications in Language Sciences 37). Berlin: Mouton de Gruyter.

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

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

Booij, Gert. 2005. Compounding and derivation: evidence for Construction Morphology. In Wolfgang Dressler, Dieter Kastovsky, Oskar Pfeiffer & Franx Rainer (eds.), *Morphology and its demarcations* (Current Issues in Linguistic Theory 264), 109–132. Amsterdam: John Benjamins.

Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: syntactic projections and the lexicon. In John Moore & Maria Polinsky (eds.), *The nature of explanation in linguistic theory* (CSLI Lecture Notes 162), 31–67. Stanford: CSLI Publications.

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

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

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

Branco, António. 2019. Anaphoric Binding. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 677–729. Berlin: Language Science Press. DOI:??

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

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

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

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

Carpenter, Bob. 1992a. Categorial Grammars, lexical rules, and the English predicative. In Bob Levine (ed.), *Formal grammar: theory and implementation* (New Directions in Cognitive Science 2), 168–242. New York, NY: Oxford University Press.

Carpenter, Bob. 1992b. *The logic of typed feature structures* (Cambridge Tracts in Theoretical Computer Science 32). Cambridge: Cambridge University Press.

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

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

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

Copestake, Ann. 2002. *Implementing typed feature structures* (CSLI Lecture Notes 110). Stanford: CSLI Publications.

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

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

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

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

Crysmann, Berthold. 2019. Morphology. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 733–734. Berlin: Language Science Press. DOI:??

Crysmann, Berthold & Olivier Bonami. 2016. Variable morphotactics in Information-based Morphology. *Journal of Linguistics* 52(2). 311–374. DOI:10.1017/S0022226715000018

Crysmann, Berthold & Gerald Penn. 2019. Clitics. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 335–336. Berlin: Language Science Press. DOI:??

Davis, Anthony. 2001. *Linking by types in the hierarchical lexicon* (Studies in Constraint-Based Lexicalism). Stanford: CSLI Publications.

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

Di Sciullo, Anna-Maria & Edwin Williams. 1987. *On the definition of word* (Linguistic Inquiry Monographs 14). 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*, 89–110. Stanford, CA: CSLI Publications.

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

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

Emerson, Guy & Ann Copestake. 2015. Lacking integrity: HPSG as a morphosyntactic theory. In Stefan Müller (ed.), *Proceedings of the 22nd international conference on head-driven phrase structure grammar*, 75–95. Stanford: CSLI.

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

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

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

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

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20* (Current Studies in Linguistics 24), 111–176. Cambridge, MA: MIT Press.

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

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

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

Jackendoff, Ray. 1990. *Semantic structures* (Current Studies in Linguistics 18). Cambridge, MA: MIT Press.

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

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

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

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

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

Koenig, Jean-Pierre & Anthony Davis. 2003. Semantically transparent linking in HPSG. In Stefan Müller (ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University, East Lansing*, 222–235. Stanford, CA: CSLI Publications. http://csli-publications.stanford.edu/HPSG/2003/, accessed 2018-2-25.

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

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

Lakoff, George. 1970. *Irregularity in syntax* (Transatlantic series in linguistics). 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* (Studies in Constraint-Based Lexicalism). Stanford, CA: CSLI Publications.

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

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

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

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

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

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

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

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

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

Müller, Stefan. 2018. The end of lexicalism as we know it? *Language* 94(1). e54–e66. DOI:10.1353/lan.2018.0014

Müller, Stefan. 2019a. Constituent order. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 289–334. Berlin: Language Science Press. DOI:??

Müller, Stefan. 2019b. HPSG and Construction Grammar. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 1191–1243. Berlin: Language Science Press. DOI:??

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.

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

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.

Pustejovsky, James. 1991. The generative lexicon. *Computational Linguistics*. 4th ser 17. 409–441.

Pustejovsky, James. 1995. *The generative lexicon* (Language, Speech, and Communication). Cambridge, MA: MIT Press.

Pustejovsky, James & Elisabetta Jezek. 1996. Integrating Generative Lexicon and lexical semantic resources. In *Proceedings of LREC 2016*. Portorož, Slovenia.

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* (CSLI Lecture Notes 46), 151–197. Stanford University: CSLI Publications.

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

Richter, Frank. 2019. Formal Background. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 45–65. Berlin: Language Science Press. DOI:??

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

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

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

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

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

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

van Trijp, Remi. 2011. A design pattern for argument structure constructions. In Luc Steels (ed.), *Design patterns in Fluid Construction Grammar* (Constructional Approaches to Language 11), 115–145. Amsterdam: John Benjamins Publishing.

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

Wechsler, Stephen, Jean-Pierre Koenig & Anthony Davis. 2019. Argument structure and linking. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-Driven Phrase Structure Grammar: The handbook*, 247–287. Berlin: Language Science Press. DOI:??