3

Sign-Based Construction Grammar: An Informal Synopsis

IVAN A. SAG

1 Introduction

This chapter[1] is intended as an introduction to some of the central notions of Sign-Based Construction Grammar (SBCG).[2] For a more general discussion of SBCG, including an informal, high-level summary of the framework, its historical development, motivation, and relation to other approaches to grammar, the reader is referred to Sag et al. this volume.

1.1 Preliminaries

SBCG is a framework blending ideas developed over a quarter century of research in Head-Driven Phrase Structure Grammar (HPSG[3]) with those pre-

[1]Above all, I'd like to thank Paul Kay for hours of discussion and useful feedback at every stage in the seemingly endless process of writing this chapter. I'm also grateful to Chuck Fillmore and Laura Michaelis for many useful discussions of both content and exposition and for comments on an earlier draft, which was also been improved by the comments of Rui Chaves, Adele Goldberg, Bill Croft, Adam Przepiórkowski, Russell Lee-Goldman, and Hans Boas. In addition, discussions with the following people have been quite helpful: Farrell Ackerman, Emily Bender, Dan Flickinger, Mark Gawron, Andreas Kathol, Beth Levin, Bob Levine, Stefan Müller, Frank Van Eynde, Tom Wasow, Gert Webelhuth, and Steve Wechsler. None of the views expressed here should be attributed to any of these individuals; nor should they be held responsible for any remaining errors.

[2]SBCG owes a considerable debt to the implementation work carried out within CSLI's LinGO Lab and the DELPHIN consortium, whose grammar development efforts have proceeded in parallel with the evolution of SBCG. See Copestake 2001, Flickinger 2000, Uszkoreit et al. 2000, and the online resources available at http://lingo.stanford.edu/ and http://www.delph-in.net/.

[3]See Pollard and Sag 1987 Pollard and Sag 1994, Ginzburg and Sag 2000, and Richter 2004, inter alia.

Sign-Based Construction Grammar.
Editors: Hans C. Boas & Ivan A. Sag.

sented within the tradition of Berkeley Construction Grammar (BCG[4]) over roughly the same period. Its goal is to expand the empirical coverage of HPSG, while at the same time putting BCG on a firmer theoretical footing.

To readers steeped in HPSG theory, SBCG will no doubt seem like a minor variant of constructional HPSG (as developed in Sag 1997, Ginzburg and Sag 2000, and elsewhere), with the principal innovation being the introduction of the distinction between signs and constructs. There is a certain truth to this, but at the same time, it is my sincere hope that construction grammarians of all stripes will find that SBCG is recognizable as a formalized version of BCG, with a few straightforward (and only minimal) notational adjustments. What is gained from the increased analytic precision of SBCG is plain: clearer empirical prediction and falsifiability, enhanced comparability of analyses across languages, and a general theoretical clarity. Certain differences between SBCG and BCG or other versions of Construction Grammar (CxG) will be noted where relevant. (See also Sag et al. this volume.)

Like Pollard and Sag (1987, 1994), Fillmore et al. (1988), Fillmore and Kay (1996), and Kay and Fillmore (1999), this study will take English as its focus. However, the goal of our enterprise, like that of other researchers in HPSG and BCG, is to provide a basis for the description of all human languages. Construction-based grammar has an advantage in this endeavor, as it concerns itself directly with words, generalizations about lexical classes, and the patterns according to which complex expressions are constructed. Every human language has these components; hence there is no need to transform a language into a mold that ill suits it in order to provide a typologically realistic theory of grammar – one that also meets the objections to the Chomskyan conception of Universal Grammar raised, for example, by Evans and Levinson (2009). It is clear that construction-based grammar has deep roots in Structural Linguistics. My goal here is to convince the reader that a properly formalized theory of construction-based grammar can in addition satisfy the demands of modern linguistic theory, e.g. those outlined by Lees (1957: 376):

1. freedom from contradiction,
2. maximal cohesion with other branches of knowledge,
3. maximal validity in coverage of known data, and
4. maximal elegance of statement.

1.2 Preview

Let us take a language to be an infinite set of **signs** and assume that the job of a grammarian is to provide a systematic account of those signs and their

[4]See, for example, Fillmore et al. 1988, Goldberg 1995, Fillmore and Kay 1996, Kay and Fillmore 1999, P. Kay 2002a, and Michaelis and Lambrecht 1996.

properties, including how they function in language processing and language use. The notion of 'sign' of course comes from Saussure (1916). However, while the Saussaurian sign is an association of sound (*signifiant*) and meaning (*signifié*), the signs of SBCG embody more components. These include at least phonological structure, (morphological) form, syntactic category, semantics, and contextual factors, including information structure.

Signs, like all linguistic entities in SBCG, are modeled as **feature structures** (FSs), which are of two basic kinds:

- atoms[5] (e.g. *accusative*, $+$, *finite*, ...),
- functions (as explained below).

A **functional** FS maps each feature in its domain (some proper subset of the set of features) to an appropriate value (atom or function).[6] In general terms then, functional FSs map features to feature structures. The particular features and values utilized in this chapter are given in the appendix.

As in LFG (Lexical-Functional Grammar; see Bresnan 2001), HPSG, and BCG, SBCG makes a strict distinction between entities in the language model (**model objects** for short) and **descriptions** of those objects.[7] As in HPSG, the most important model objects are signs (the formal representations of actual words and phrases, including sentences). Each lexical sign or fixed phrasal expression is licensed by a **listeme** (a 'listed' description of a word or phrase).[8] Another, distinct kind of model object in SBCG is the **construct**. As

[5]The set of atoms includes, for analytic convenience, an infinite set of indices.

[6]A function F can be defined simply as a set of ordered pairs, where the first member of each pair is a member of the set that is F's 'domain' and the second is a member of the set that is F's 'range'. The only further condition that has to be met is uniqueness – the condition that for any distinct b and c in F's range, F cannot contain both $\langle x, b\rangle$ and $\langle x, c\rangle$ (for any x in F's domain).

A total function F is a set of such pairs that contains a pair $\langle a, \dots \rangle$ for each member a of F's domain. Finally note that all FSs must 'bottom out' in atoms. That is, the most deeply embedded functions within a FS must map to an atom. Otherwise, i.e if F maps to another function, F is by definition not the most deeply embedded function in the feature structure.

[7]Though the distinction between model and model description may seem unfamiliar to many linguists, drawing such a distinction reflects the standard practice of research in most scientific disciplines. Representational models allow the theorist to abstract away from irrelevant properties of the phenomena under investigation, while the descriptions themselves constitute the theory of the modeling objects. By placing the modeling objects in correspondence with the real-world phenomena, the descriptions become a theory of those phenomena – more precisely, a theory of the properties of those objects that are being modeled. In the present case, for example, we want to model the observable phonetic and semantic properties of signs. Other model objects, e.g. syntactic categories, case values, or contextual conditions, are introduced because we believe they are necessary in order to model other properties of signs deemed of interest.

[8]The term 'listeme' is first proposed by Di Sciullo and Williams 1987 as a generalization of the notion 'lexical entry' to include multiword expressions of various kinds. Multiword expressions are discussed in section 7.1 below.

in Generalized Phrase Structure Grammar (GPSG – see Gazdar et al. 1985), constructs are local trees that are licensed by a particular kind of construction: a **combinatoric construction**. As we will see, a construct can be naturally accommodated within a FS-system as a functional FS that specifies values for the MOTHER (MTR) feature and the DAUGHTERS (DTRS) feature. The value of MTR is a sign and the value of DTRS is a nonempty list of signs.

Signs and constructs, as already noted, are FSs – they are part of the language model. Listemes and constructions are descriptions that license classes of linguistic objects (signs or constructs) – they are part of the grammar (the description of the language model; the theory of the language).

The linguistic objects in SBCG here are classified in terms of a system of **types**, which are organized into a lattice-like structure that reflects a linguistically motivated classification. Thus, *polar-interrogative-clause* is a **maximal type** (a type without subtypes) that is instantiated by clausal constructs like (1):[9]

(1) {[*Will*] [*Sandy*] [*be there*]?}

But in order to be well-formed according to our theory, this construct must also satisfy the constraints the grammar imposes on all the supertypes of *polar-interrogative-clause*. These constraints take the form of further constructions – those which define the particular properties of the supertypes *auxiliary-initial-construct*, *headed-construct*, *interrogative-clause*, *core-clause*, and *clause*.

An SBCG grammar also contains a **signature**. Just as the time and key signature of a musical composition specify how musical descriptions (e.g. notes, rests, and measures) are to be interpreted, the grammar signature delineates the basic ontology of the grammar, and thus specifies how grammatical descriptions are to be interpreted. It is here that the details of the type hierarchy are laid out, along with a characterization of the general properties of each type of feature structure. The signature associates each type of functional feature structure with a domain (a set of features) and assigns an appropriate value type to each feature of the domain, as illustrated in (2):

(2) $$type_0 : \begin{bmatrix} \text{FEATURE}_1 & type_1 \\ \dots & \\ \text{FEATURE}_n & type_n \end{bmatrix}$$

[9]The informal notation in (1), used for abbreviating constructs, is due to Chuck Fillmore. The entire construct is enclosed in curly braces, with each daughter in square brackets. Annotations are also possible, as indicated by the Fs in (i):

(i) $\{^F[^F D_1], \dots, [^F D_n]\}$

See Fillmore et al. this volume.

This is to be interpreted as: The grammar recognizes FSs of $type_0$, which are functions whose domain includes $\text{FEATURE}_1 \ldots \text{FEATURE}_n$ and which map each FEATURE_i to a FS of type $type_i$, as indicated.[10]

Against this background of possible FSs, a particular SBCG of a given language specifies which particular family of FSs exists in that language – those that are licensed by a listeme or a construction.[11] The grammar signature assumed in this chapter is summarized in the appendix.

2 Feature Structures

As already noted, grammatical objects of all kinds (including signs, case values, parts of speech, and constructions) are modeled as FSs, either atoms or else functions from features to FSs.[12] This is a simple, but powerful way of modeling linguistic objects, one that is already familiar from early work in generative phonology, where speech segments are often analyzed in this way. For example the following function characterizes the phone [*t*] in the feature system of Chomsky and Halle (1968):

(3) $$\begin{bmatrix} \text{CONTINUANT} & - \\ \text{VOICED} & - \\ \text{ANTERIOR} & + \\ \text{CORONAL} & + \\ \text{SONORANT} & - \\ \text{CONSONANTAL} & + \\ \text{VOCALIC} & - \end{bmatrix}$$

Similarly, a fundamental tenet of '$\overline{\text{X}}$ Theory'[13] is that familiar atomic categories like NP or VP are to be reanalyzed as functions, e.g. as in (4):[14]

(4) $$\begin{bmatrix} \text{NOUN} & + \\ \text{VERB} & - \\ \text{BAR} & 2 \end{bmatrix}$$

[10]The range of a given FS type, i.e. the union of the possible values of its features, does not generally constitute a particularly useful or coherent set. A FS of this sort can be regarded as a particular kind of function called an 'ad-hoc polymorphism' (Strachey 1967).

[11]Though the distinction between the signature's type declarations and the grammar's constructions is sometimes more one of convenience, there are computer implementations of type systems where the type declarations (but not the type constraints) are used in a 'top-down' manner during processing.

[12]Carpenter 1992. See also Sag et al. 2003.

[13]See Harris 1946, Chomsky 1970, Jackendoff 1977, and Kornai and Pullum 1990.

[14]These are the distinctive features of the category NP in the analysis proposed in Gazdar et al. 1985.

Note that the functional nature of this kind of analysis can be obscured by linguists' tendency to write the value of a feature before the feature's name, e.g. [+ CORONAL] or [+V] or to use other notations, e.g. X^1 (Harris 1946) or $\overline{X}$ (Chomsky 1970). Yet it is clear that the analytic intent of such notions is accurately rendered by functions whose domain is a set of features and whose range is a set of feature values (e.g. the set $\{+,-\}$ in the system of Chomsky and Halle 1968 or that of Chomsky 1974). The use of functions to model linguistic objects is thus nothing out of the ordinary, though lack of formalization and idiosyncratic or abstruse notation (especially in the case of generative-transformational syntactic theories) often obscures this fact.

Building on the more explicit ideas pioneered by computational linguistic work of the late 1970s, e.g. Martin Kay's Functional Unification Grammar (M. Kay 1979), and the extensive subsequent work in GPSG, LFG, and HPSG,[15] every grammatical object used here is modeled by a function that maps each member of a set of features to one of its possible values, as specified in (2) above. Grammatical categories, for example, are analyzed as complexes of various properties represented as feature-value pairs: nouns include specifications for the features CASE, NUMBER, and GENDER; verbs are specified in terms of the feature VERB-FORM (VF) for their inflection class (as [VF *finite*], [VF *present-participle*], etc.) and will have a '+' or '−' value for the feature AUXILIARY (AUX). This approach also takes advantage of the power of functions to model complex linguistic entities. Unlike the phonological and $\overline{X}$ illustrations given above, where the values of the features are all atomic, the value of an SBCG feature may be an atom or it may be a function (another complex FS). This allows for recursive embedding of feature structures within FSs, analogous to the embedding of functions that is now standard practice in formalized approaches to semantic analysis.

Signs are no exception. Saussure (1916) regarded a sign as an 'associative bond' between a sound concept and a semantic concept. Adding in syntactic information, we arrive at representations like the ones in Figure 1, rendered here in the Saussurean style (CN stands for common noun; N for proper noun; V for verb).

As already noted, signs are modeled as functions that specify a phonological and morphological structure, a meaning, contextual connections, and relevant syntactic information (including traditional syntactic category and combinatoric potential). These functions are described in terms of attribute-value matrices, i.e. diagrams like the following:[16]

[15] See Gazdar et al. 1985, Bresnan 1982, Dalrymple et al. 1995, Pollard and Sag 1987, 1994, King 1989, 1994, and Carpenter 1992.

[16] The informal semantics in (5) and (6) is a temporary expedient, and will be replaced by a more precise representation in section 3.4 below.

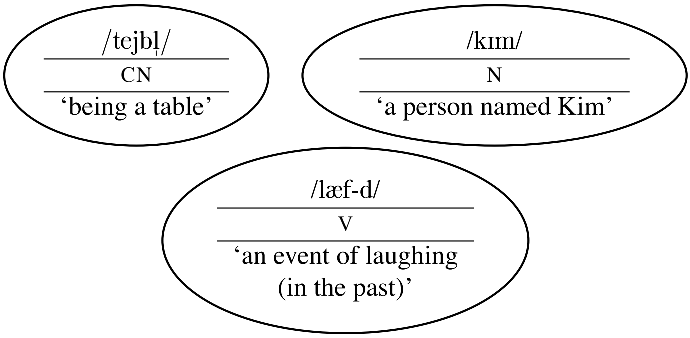


FIGURE 1 Saussurean Representation

(5) a.

PHONOLOGY	/kɪm/
SYNTAX	NP
SEMANTICS	'the intended person named Kim'

b.

PHONOLOGY	/læf-d/
SYNTAX	V[*fin*]
SEMANTICS	'a laughing event situated prior to the time of utterance'

And, following work in HPSG (Pollard and Sag 1987, 1994), the notion of 'sign' is extended to phrases, recognizing FSs like those in (6) for complex linguistic expressions:

(6) a.

PHONOLOGY	*/ɛvri#lɪŋwɪst/*
SYNTAX	NP
SEMANTICS	'the set of properties all linguists share'

b.

PHONOLOGY	*/pæt#læf-d/*
SYNTAX	S[*fin*]
SEMANTICS	'the proposition that there was a laughing event situated prior to the time of utterance where a certain person named Pat did the laughing'

The non-atomic FSs used to model linguistic objects are total functions. That is, once an appropriate feature domain is established for a particular type of feature structure, every FS of that type assigns an appropriate, fully determinate value to every feature in that domain. The value assigned to any feature must also be a feature structure, i.e. either an atom or a function that in turn assigns a value to every feature in *its* appropriate domain. A FS is thus always 'complete' in a simple, intuitive sense: every feature in a function's domain is assigned a value in the appropriate range.

It is important to note that whereas FSs are total functions, FS **descriptions** are in practice almost always partial. For this reason, most of the discussions below, indeed throughout the chapters of this book, will include partial descriptions of linguistic objects. Listemes are formulated as partial FS descriptions (typically being true of (or 'satisfied by') a large class of FSs); combinatoric constructions are also quite sparse, when compared with the feature structures that instantiate them. But underlying all our concerns will be the set of FSs that is licensed by the grammar we develop. If some aspect of our grammar goes awry, we should be able to learn why by isolating certain FSs that should satisfy the constraints of our theory, but do not. Alternatively, we should be able to find some FS that incorrectly satisfies our theory. In particular, an SBCG grammar must not license signs that fail to model something in the target language (they must not license an unacceptable sentence, for example). In addition, it must not fail to license a sign that is needed to serve as a model of a sentence that the grammarian decides is part of the target language.

FSs have one more property that is not part of the basic theory of functions (which I will assume only cursory familiarity with): FSs are organized in terms of linguistic types. A type is a classification associated with a set of FSs that have certain stated properties in common. One benefit derived from assigning FSs to types is that we can then better organize the properties that classes of grammatical objects have and simplify their description in the process. Intuitively, it makes no sense (in English, anyway) to ask what case a verb has or whether a noun is an auxiliary – certain grammatical featural distinctions are appropriate only for certain kinds of grammatical objects. This intuition is given formal expression in terms of the types that particular FSs instantiate. Each FS instantiates a particular maximal type and the feature appropriateness conditions together determine which subset of features is appropriate for FSs of that type, ruling out verbal categories that specify CASE values, nominal categories that specify VF values, and so forth.

The space of types is hierarchically structured. In fact, the types are interrelated in terms of a **multiple inheritance hierarchy**. If a type B **is a subtype of** another type A, then FSs of type B must satisfy all constraints that the grammar imposes on objects of type A, as well as the grammatical constraints imposed on type B.[17] In this case, the FSs of type B form a subset

[17]I am assuming that constraint inheritance is monotonic, i.e. that there is no 'overriding' of grammatical constraints. But this assumption is debatable, and the general framework of SBCG

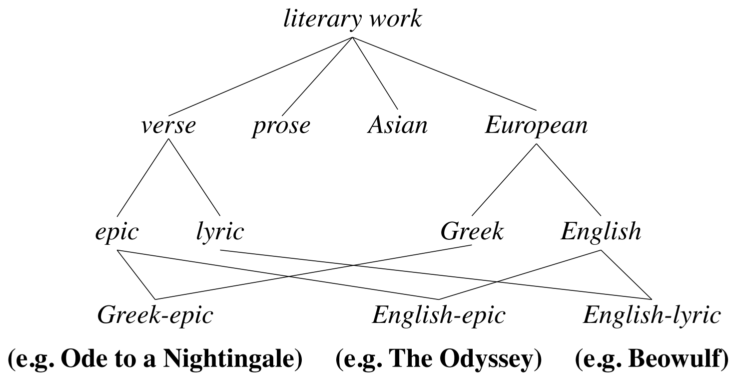


FIGURE 2 A Multiple-Inheritance Hierarchy

of the FSs of type A. This situation is informally characterized, following the terminology of BCG, by saying that 'type B **inherits from** type A'.[18] In a *multiple* inheritance hierarchy, a type can inherit from more than one immediate supertype. That is, type hierarchies behave like the nonlinguistic example in Figure 2, where an instance of each maximal type is given below it. Multiple-inheritance hierarchies are useful for analyzing cross-classifying properties of a set of objects, whether they are literary works, words, or constructs.

In SBCG, the more general notion of 'type hierarchy' takes over the inheritance functions that **constructional inheritance** performed in some earlier traditions of CxG.[19] For example, Fillmore (1999) treats the various kinds of auxiliary-initial clause in terms of constructional inheritance from the supertype he calls 'subject-auxiliary inversion' (SAI). Some of these SAI constructions are illustrated in (7):

(7) a. { [*Has*] [*he*] [*left?*] }
b. { [*Am*] [*I*] [*tired!*] }
c. Never { [*will*] [*I*] [*harm you.*] }
d. What { [*did*] [*Merle*] [*know?*] }
e. { [*May*] [*you*] [*live long and prosper!*] }
f. { [*Had*] [*he*] [*been on time*] }, he wouldn't have gone hungry.

could be outfitted to include default constraints and constraint overriding. Most linguistically interesting examples of default constraints, however, can be cast as 'nonpersistent' defaults in the sense of Lascarides and Copestake 1999, and hence can be straightforwardly reanalyzed in purely monotonic terms, as I have attempted to do here.

[18]This terminology is in fact somewhat misleading, as it is the feature structures instantiating a given type that inherit the properties associated with the type's supertypes.

[19]See, for example, Fillmore et al. 1988, Goldberg 1995, Fillmore and Kay 1996, Fillmore 1999, Kay and Fillmore 1999, P. Kay 2002a, 2002b, Michaelis and Lambrecht 1996.

In the SBCG analog of Fillmore's analysis, each of the bracketed sequences in (7) instantiates a type of construct that is a subtype of the more general type *auxiliary-initial-construct* (*aux-initial-cxt*). The Aux-Initial Construction (a combinatoric construction) places general constraints on instances of the type *aux-initial-cxt*, as sketched in (8):

(8) Aux-Initial Construction (Preliminary Sketch)

An *aux-initial-cxt* must satisfy: $\begin{bmatrix} \text{MTR} & \text{S}[\ldots] \\ \text{DTRS} & \langle \text{V}[\text{AUX} +], \ldots \rangle \end{bmatrix}$

A more specific construction, i.e. a construction that characterizes a subtype of *aux-initial-cxt*, needs to specify only the properties that are specific to that subtype. In the case of aux-initial clauses, the subconstructions in question specify primarily semantic information, but also syntactic constraints about independent clause status, modifier status, etc. For example, the Inverted Wish Construction (of which (7e) is an instance) defines the characteristic properties of feature structures instantiating the maximal type *inverted-wish-construct* (*inv-wish-cxt*). It specifies the appropriate semantics and the constraint that a FS of this type must be an independent clause. Inverted wish constructs will of course also exhibit the general properties of (obey the general constraints on) aux-initial constructs. This is accomplished simply by specifying that *inv-wish-cxt* is a subtype of *aux-initial-cxt*.

3 Signs

The following sections introduce the specific features whose values serve to distinguish the signs of a language from one another.

3.1 PHONOLOGY and FORM

Little will be said here about morphology, and nothing at all about phonology, but the intention is that phonological and morphological entities be part of linguistic signs. I assume that a largely autonomous set of constraints characterize the relation between the phonological and morphological aspects of signs.[20] There are thus two distinct sign-level features: PHONOLOGY (PHON) and FORM:

(9) a. The value of the feature PHON is a phonological structure i.e. a FS of type *phonological-object*.

[20] A number of researchers have developed related approaches to the analysis of phonological and morphological structure in terms of typed feature structures of the sort assumed here. See Bird and Klein 1994, Orgun 1996, 2000, Klein 2000, Asudeh and Klein 2002, Bonami and Boyé 2002, 2006, Haji-Abdolhosseini 2003, Tseng 2003, 2008, and Bonami and Delais-Roussarie 2006, among others.

b. The value of the feature FORM is of type *morphological-object* (*morph-obj*); these are the elements that will be phonologically realized within the sign's PHON value.

The precise characterization of both these structures is left open here, though for convenience PHON values are treated as lists of segments (represented in phonemic brackets) and FORM values as lists whose members include **stems** and more complex entities built up from stems.

In order to deal with inflectional realization, morphological functions must make reference to stem identity. And because stems can exhibit idiosyncrasy like that shown in (10), stems must be individuated so as to allow homophonous elements to exhibit distinct inflectional patterns:

(10) a. *lie/lay/lain* 'rest, recline' vs. *lie/lied/lied* 'tell falsehoods'
b. *can/could* 'be able to' vs. *can/canned* 'put into cans'
c. *fly/flew* (basic sense) vs. *fly/flied* (various derived senses)
d. *sell/sold* vs. *cell/celled*
e. *write/wrote/written* vs. *right/righted/righted*

To this end, the FORM value of a sign will be represented as a list of conventional orthographic representations augmented by indices to distinguish homophonous stems with divergent inflectional realizations (e.g. lie_1 vs. lie_2). Morphological functions of the sort assumed here provide a basis for dealing with data like (10), as well as a precise way of expressing 'elsewhere' conditions in morphological realization and a means for avoiding unwarranted analysis via morpheme sequences.[21] The FORM feature provides a convenient shorthand for illustrating the combinatoric structure of both words and phrases.

3.2 ARGUMENT-STRUCTURE

The basic purpose of the ARGUMENT-STRUCTURE (ARG-ST) feature is to encode the combinatoric potential of a lexical sign by listing its potential syntactico-semantic **arguments**. The order of elements on the ARG-ST list corresponds in the main to that of the 'Accessibility Hierarchy' of Keenan and Comrie (1977). For example, the first NP of a verb's ARG-ST is the subject, the second NP (of a transitive verb's ARG-ST) is its direct object, and so forth. This 'rank-based' encoding of grammatical relations, as shown by Keenan and Comrie and other researchers in relation-based syntactic theory, is independently motivated by the cross-linguistic patterns that have been observed for such phenomena as relative clause accessibility (which NP in a

[21] That is, they provide a way of avoiding what Hockett (1987) has termed 'The Great Agglutinative Fraud' (for discussion, see Blevins 2008). For a different approach to elsewhere conditions, one more in line with Stump's (2001) adaptation of 'Pāṇini's Principle', see Bonami and Samvelian submitted.

clause can be 'relativized'), reflexive binding (which NP in a clause can bind a reflexive), and agreement (which NP in a clause the verb can mark agreement with). In a language like English, a verb's subject is identified as the first member of its ARG-ST list, which is also its 'external argument' (see the discussion of the feature XARG in section 3.3 below).

Variable polyadicity of a given lexeme, e.g. active vs. passive vs. middle, causative vs. inchoative, or oblique-recipient vs. ditransitive, involves differences in the ARG-ST list. These differences can arise in two distinct ways in SBCG: by derivational construction (e.g. familiar analyses of passivization or causativization) or by lexical underspecification (as in certain analyses of locative ('*spray/load*') alternations). The ARG-ST list, which includes 'extracted' and unexpressed arguments, works in tandem with the VALENCE list, which does not. The relation between these two lists is explained in section 3.3 below.

Some examples of lexical classes associated with particular ARG-ST lists are the following:[22]

(11)	$\langle$NP$\rangle$	*die*, *laugh*, . . .	*sintrans-v-lxm*
	$\langle$NP, NP$\rangle$	*hit*, *like*, . . .	*strans-v-lxm*
	$\langle$NP, NP, NP$\rangle$	*give*, *tell*, . . .	*ditrans-v-lxm*
	$\langle$NP, NP, PP$\rangle$	*put*, *place*, . . .	*loc-trans-v-lxm*

Lexemes, especially verbal lexemes (see below), fall into diverse classes, as determined in part by the length of their ARG-ST list and the constraints imposed on particular arguments. Only lexical signs (lexemes or words) specify a value for ARG-ST.

ARG-ST lists are also the locus of constraints on coindexation ('binding theory').[23] For example, a reflexive or reciprocal that is a member of an ARG-ST list must be coindexed with a sign preceding it on that list, if there is one

[22]Some abbreviations:
sintrans-v-lxm = *strict-intransitive-verb-lexeme*
strans-v-lxm = *strict-transitive-verb-lexeme*
ditrans-v-lxm = *ditransitive-verb-lexeme*
loc-trans-v-lxm = *locational-transitive-verb-lexeme*

$$\text{NP} = \begin{bmatrix} \textit{sign} & \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \textit{noun} \\ \text{VAL} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \qquad \text{PP} = \begin{bmatrix} \textit{sign} & \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \textit{prep} \end{bmatrix} \end{bmatrix}$$

$$\text{CP} = \begin{bmatrix} \textit{sign} & \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \textit{comp} \\ \text{VAL} & \langle\,\rangle \end{bmatrix} \end{bmatrix} \qquad \langle\,\rangle = \textit{elist} = \text{empty list.}$$

[23]This follows a tradition that begins with the Relational Grammar proposals of Johnson (1977). See also Pollard and Sag 1992, 1994 and Manning and Sag 1998. Although this tradition improves upon theories of binding based on constituent structure (including all theories based on the notion of 'c-command'), it may still be overly reliant on grammatical structure, as argued by Runner and Kaiser (2005).

(Principle A); personal pronominals must not be coindexed with any preceding element (Principle B).

3.3 SYNTAX

The value of the feature SYNTAX is a FS of type *syntax-object* (*syn-obj*). Functions of this type specify values for the features CATEGORY, VALENCE, and MARKING, which I will discuss in turn.[24]

CATEGORY

The values of the feature CATEGORY are complex grammatical categories, treated here as FSs of type *category* (*cat*).[25] The various subtypes of *cat* specify values for appropriate features. For example, the signature of the grammar of English assumed here includes the following information:

(12) a. The immediate subtypes of the type *category* are: *verbal* and *nonverbal*. The subtypes of *verbal* are *verb* and *complementizer* and those of *nonverbal* are *adverb* (*adv*), *adjective* (*adj*), and *nominal*. And the subtypes of *nominal* are *noun* and *preposition* (*prep*), yielding the hierarchy of CAT values shown in Figure 3.

b. CASE is appropriate only for FSs of type *noun* (in English). The possible values of CASE (in English) are *nominative* (*nom*) and *accusative* (*acc*).[26]

c. VERB-FORM (VF), appropriate only for FSs of type *verbal*, is used to specify the appropriate inflectional category of a verb or complementizer. The possible values of VF are *finite* (*fin*), *infinitive* (*inf*), *base*, *present-particple* (*prp*), *past-particple* (*psp*), and *passive-particple* (*pas*).

d. AUXILIARY (AUX) is used to specify whether a verb appears in one of the syntactic environments restricted to auxiliary verbs (e.g. sentential negation, inversion, contraction, or VP-Ellipsis; see section 9). The value of AUX is an atom of type *boolean*, the name often used to refer to a truth value, i.e. + (true) or − (false).

---

[24]There are three further features to be included in the domain of feature structures of this type: GAP, WH, and REL. These are discussed in section 10.

[25]Note that 'CATEGORY' denotes a feature and '*category*' denotes a type. Features are represented in small capitals and types in *lower case italics*. The FSs of type *category* play a role similar to that of $\overline{\text{X}}$ categories, as extended by work in GPSG. For discussion, see Sag 2010b.

[26]Note that genitive nominal expressions are not distinguished in terms of CASE. This is because case is a property of head nouns and the Modern English *'s* is a phrasal clitic that appears in final position of a genitive NP, rather than as an inflection on the head noun:

(i) [[The man on the radio's] voice] . . .

(ii)*[[The man's on the radio] voice] . . .

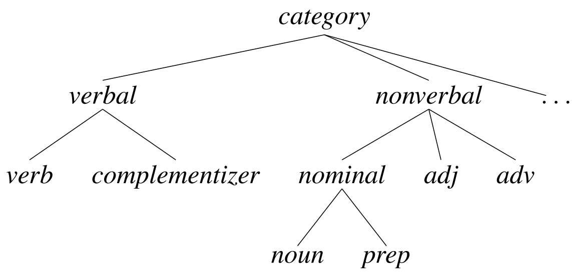


FIGURE 3 Hierarchy of CATEGORY Values

e. INVERTED (INV) is used to specify whether a verb is in clause-initial position; the values of INV are also of type *boolean*.

f. INDEPENDENT-CLAUSE (IC) is used to distinguish independent clauses (and the signs that project them) from their dependent clause counterparts. The values of IC, which is appropriate only for FSs of type *verbal*, are again of type *boolean*.

This partial signature countenances complex grammatical categories like those shown in (13), but none like the ones pictured in (14):

(13)

a. $\boxed{\begin{bmatrix} \textit{noun} & \\ \text{CASE} & \textit{nom} \end{bmatrix}}$ b. $\boxed{\begin{bmatrix} \textit{verb} & \\ \text{VF} & \textit{fin} \\ \text{INV} & + \\ \text{IC} & - \\ \text{AUX} & + \end{bmatrix}}$ c. $\boxed{\begin{bmatrix} \textit{complementizer} & \\ \text{VF} & \textit{inf} \\ \text{IC} & - \end{bmatrix}}$

(14)

a. * $\boxed{\begin{bmatrix} \textit{noun} & \\ \text{VF} & \textit{fin} \\ \text{IC} & - \\ \ldots & \end{bmatrix}}$ b. * $\boxed{\begin{bmatrix} \textit{verb} & \\ \text{AUX} & - \\ \text{CASE} & \textit{nom} \\ \ldots & \end{bmatrix}}$ c. * $\boxed{\begin{bmatrix} \textit{adv} & \\ \text{AUX} & - \\ \text{INV} & + \\ \text{VF} & \textit{prp} \\ \ldots & \end{bmatrix}}$

It is worth repeating that attribute-value matrices (AVMs) are being used here to formulate FS **descriptions**. This contrasts with the objects that are being described, which are models of particular linguistic entities, i.e. **total functions** of the appropriate type. When a particular FS, i.e. a linguistic model, is being depicted, rather than a description of a family of FSs, the AVM is displayed inside a box, as in (13)–(14).

As already noted, listemes are quite minimal, typically specifying just a form, a lexeme type and a meaning. But the set of possible FSs licensed by any given listeme is potentially vast, circumscribed only by the constraints of the grammar signature, which require that each appropriate feature have a value of an appropriate type. To take a simple example, the listeme licensing the proper noun *Dale* says nothing about the value of the feature CASE. But any given FS licensed by this listeme has a determinate value for CASE – one that is contextually resolved – in *Dale likes you*, it is resolved as *nom*; in *You like Dale*, it is resolved as *acc*. Similarly, a verbal listeme does not normally specify the FORM value of any of the signs on its ARG-ST list. However, since there is no upper bound on the length of a subject, object, or complement, there are infinitely many signs that could play each of these roles for any given verb, and hence infinitely many distinct feature structures licensed by the verbal listeme, each with a distinctive PHON, FORM, SEM, or CONTEXT value.

Although the functional entities in our linguistic models are always total functions,[27] essentially all of the business of grammar is conducted in terms of partial (or underspecified) descriptions of classes of these entities. Indeed, if a given grammar allows two distinct signs to have the same PHON specification, then it is predicting that this phonological structure exhibits a linguistic ambiguity. For example, the descriptions in (15) characterize more than one FS, and hence underspecify the indicated ambiguities:

(15) a. [PHON /aj#fɚgat#haw#gʊd#bir#ˈtejst-s/ ; FORM ⟨*I, forgot, how, good, beer, tastes*⟩]

([[*how good*] [*beer tastes*]] vs. [*how* [[*good beer*] *tastes*]])

b. [PHON /vɪzət-ɪŋ#rɛlətɪv-z#kɛn#bi#ˈbɔrɪŋ/ ; FORM ⟨*visiting, relatives, can, be, boring*⟩]

(*visiting* is an adjective modifying *relatives* or else a gerund whose direct object is *relatives*)

c. [PHON /aj#sɒ#hɚ#ˈdʌk/]

I saw her duck ('I saw her perform a rapid evasive downward movement' vs. 'I saw her barnyard fowl')

We will of course evaluate grammars in terms of their ability to model such ambiguities successfully. In addition, psycholinguistically plausible grammars must support the incremental computation of partial sign descriptions in such a way that supports partial semantic interpretation, the representation and resolution of local ambiguities, and appropriate underspecification (Sag and Wasow 2011).

[27] A caveat: a boxed AVM diagram will typically not mention all of the feature specifications included in a given feature structure.

The hierarachical organization of CAT values just presented defines a pattern of natural classes allowing, for instance, a concise account of the constructional variation in the category of the filler phrase in the diversity of English filler-gap constructions (Sag 2010a). Similarly, this inventory of CAT features has been integrated into a system that provides a reasonably well-worked out account of 'main clause phenomena', including construction-specific variation regarding the possibility of (or requirement for) auxiliary 'inversion' (Ginzburg and Sag 2000). In addition, this same feature inventory plays a critical role in the treatment of the English auxiliary system presented in Sag to appear (see section 9 below).

There are three other CAT features that must now be introduced:[28]

(16) a. SELECT is used to let an expression select what it can modify or combine with as a 'marker'. The value of SELECT is either the distinguished atom *none* (in the case of expressions that are neither modifiers nor specifiers) or else a sign. If an expression's SELECT value is a sign, then it is either a modifier (e.g. an adjective or adverb) or else a marker (e.g. a determiner) and its SELECT value imposes constraints on the element that it modifies or marks.

b. EXTERNAL-ARGUMENT is used to specify the argument of an argument-taking expression that is visible from outside its local domain (i.e. from outside the maximal phrase that expression projects). The value of XARG is either a sign or *none*.[29] The external argument of a clause is its subject; an NP's external argument is its prenominal genitive NP, if there is one (the XARG value of the NP is *none*, otherwise).

c. LEXICAL-IDENTIFIER (LID) is used to individuate lexical items semantically; the value of LID is a list of semantic frames that canonically specify the (fine-grained) meaning of a lexeme. In this paper, this list is always singleton or empty.

The features SELECT and XARG are discussed in more detail in section 8 below.

[28] Here I follow Van Eynde (1998), who builds directly on Allegranza 1998b, in eliminating Pollard and Sag's (1994) features MOD, SPEC, and SPR in favor of the single feature SELECT. The values of SELECT indicate properties of the phrasal head that are selected by a given modifier or specifier. See also Van Eynde 2006, 2007 and Allegranza 2007. The fundamental insights of the SELECT analysis presented here are indeed those of Van Eynde and Allegranza, despite minor differences of execution. For example, Van Eynde follows the basic feature inventory and more complex feature geometry of Pollard and Sag, which has been streamlined here, e.g. by eliminating the features HEAD and LOCAL.

[29] Sag and Pollard (1991), who first introduced this feature, assumed its value to be an index, an assumption preserved in ongoing computational grammar development work using the English Resource Grammar and the Grammar Matrix. See also Copestake et al. 2005.

VALENCE

The basic function of the feature VALENCE (VAL) is to specify which of an expression's syntactic-semantic arguments it has yet to combine with syntactically. VAL is thus closely related to the feature ARG-ST. While the ARG-ST list specifies all of a word's potential arguments, including those that could be 'extracted' in a filler-gap construction, those that could remain unexpressed ('null instantiated' in the sense of Fillmore (1986)), and those that could in some languages be realized morphologically instead of syntactically, the VAL list includes just the subset of these that are relevant to that word's local syntactic combinatorics. I will refer to the members of a lexical expression's VAL list as its **valents**.

In the simplest case, where no covert or nonlocal argument realization takes place, the ARG-ST value of a word is identical to its VAL list. That is, the grammar requires that a word's VAL list is the same as its ARG-ST list, except that all covert expressions (see below) are removed.[30] Although phrases have no ARG-ST in SBCG, a verb phrase like *persuaded me to go*, which is constructed by combining the verb with all but the first of its valents (i.e. its subject), is specified as follows:

(17) [SYN [VAL ⟨NP⟩]]

Similarly, the clause *My dad persuaded me to go*, a sentential sign whose construction has introduced all the verb's valents, is specified as in (18):

(18) [SYN [VAL ⟨ ⟩]]

The lexical head of the clause is the verb, and the phrases that it projects gradually 'saturate' the verb's valence by 'removing elements' from the valence list.[31] Clauses, NPs, pronouns, and proper nouns have an empty VAL list because they are already saturated, i.e. they need not – indeed they cannot – combine with subjects or complements. A VP or a predicative phrase of some other category has a singleton VAL list, reflecting the fact that it already contains (i.e. its head daughter has 'already combined with') all relevant complements, but it still has the potential to combine with a subject.

Discrepancies between a word's ARG-ST list and its VAL list can arise

[30]In Sag et al. 2003, this is accomplished by a single grammatical constraint – the Argument Realization Principle.

[31]This way of looking at things, which has its origin in the argument cancellation of Categorial Grammar (Ajdukiewicz 1935; see also http://en.wikipedia.org/wiki/Categorial_grammar), involves a 'bottom-up' procedural metaphor where one starts with a predicator and builds successively larger phrases of which that predicator is the lexical head. It is important to recognize, however, that the constructions of an SBCG, like the rules of a Context-Free Phrase Structure (CFG) Grammar or a Categorial Grammar, are static constraints defining well-formed local tree configurations. The importance of this fundamental design property has been discussed by many over the years, including Kaplan and Bresnan 1982, Sag et al. 1986, Fenstad et al. 1987, Jackendoff 1997, 2002, Pullum and Scholz 2005, and Sag and Wasow 2011.

in several ways. One of these is **nonlocal realization**, discussed in section 10 below, where one of the word's valents appears in a dislocated syntactic position. Another is the phenomenon of **null instantiation** (Fillmore 1986), which arises when a lexical sign undergoes a derivational construction whose syntactic consequence is that an argument (and element of the ARG-ST list) fails to appear on its VAL list. Finally, there is **morphological realization**. For example, in many varieties of the Romance languages, so-called 'clitic' pronouns have been shown to require reanalysis as inflectional affixes (see Miller and Monachesi 2003). This is a third kind of noncanonical argument expression, where a verb bearing pronominal affixes has a VAL list that omits specific elements of its ARG-ST list.[32]

MARKING

The feature MARKING (MRKG), introduced by Pollard and Sag (1994) and refined in crucial ways by Van Eynde and Allegranza,[33] is used to distinguish expressions like *than Kim read* and *the book* from their respective 'unmarked' counterparts *Kim read* and *book*. The MRKG value is *unmarked* (*unmk*) in the case of all unmarked signs, but we will assume various other MRKG values, such as those in (19). Some prepositions also lead a life as markers, as in the case of *than* and *as* (Hankamer 1973) and certain uses of *of*.[34]

(19)	*than*	compared phrases, e.g. *than we read*
	as	equated phrases, e.g. *as I could*
	of	some *of*-phrases, e.g. *of mine*
	det	'determined' signs of category *noun* (see below)
	a	a subtype of *det*, e.g. *a book*
	def	definite signs, e.g. *the table*, *Prince*, *we*

An element that specifies a MRKG value other than *unmk* is informally called a 'marker'; all such elements also specify a nonempty value for the feature SELECT. Not all marked phrases, however, contain such an element, for example genitive NPs, proper nouns, and pronouns are all specified as [MRKG *def*]. The MRKG value of a marker is passed up to its mother via a constraint imposed by the Head-Functor Construction, which is introduced in section 8 below. The MRKG value of a head is passed up to its mother via constraints on certain other types of construct, e.g. by the head-complement constructions.

[32] For an analysis of this phenomenon in French, broadly compatible with the framework developed here, see Miller and Sag 1997. A detailed, compatible treatment of this and related matters in Italian and Romanian can be found in Monachesi 2005.

[33] See footnote 28. Van Eynde's MRKG values are more complex than those assumed here, in large part because of his need to analyze complex morphological and agreement patterns absent in English. I leave open the possibility of modifying the theory of MRKG to incorporate further of Van Eynde's insights.

[34] See Abeillé et al. 2006 for an analogous distinction in French.

3.4 SEMANTICS

A central thesis of CxG is that constructions can bear meaning. But there has been a disagreement in the CxG literature about whether or not 'constructions must have meaning'. This debate has centered around the Aux-Initial (or 'Subject-Auxiliary Inversion') Construction discussed earlier. Goldberg (2006: Ch. 8) has argued that there are defeasible semantic properties associated with the general SAI construction, including "non-positive polarity, non-predicate focus information structure, non-assertive, and non-declarative speech act function". By contrast, Fillmore (1999) argued (1) that there were no such grammatically encoded properties, (2) that individual SAI-subconstructions specify their own semantic properties, and hence (3) that constructions need not involve semantics.

Fillmore's third conclusion seems inevitable, even if Goldberg's position turns out to be correct about aux-initial constructs. For example, it is commonplace, within both HPSG and BCG, to assume a construction that applies just to headed constructs. This is a superordinate construction that expresses properties common to all headed constructs, just as the Aux-Initial Construction generalizes over the various aux-initial construct types discussed by Fillmore and Goldberg. Yet it seems quite unlikely that there is some general meaning common to aux-initial, subject-predicate, head-modifier, head-functor, head-complement, and other kinds of headed constructs. Note further that such a meaning would have to be restricted to the headed constructs, i.e. absent from nonheaded constructs (otherwise, it would be formulated at some level in the type hierarchy higher than *headed-construct*). Surely there are some generalizations in grammar not grounded in meaning and some of these seem plausibly viewed as constructions.[35] For further discussion of these issues, see Boas 2008a, Borsley and Newmeyer 2009, Goldberg 2009, and the references cited therein.

In sum, it is not necessary for a construction to bear meaning in SBCG. All that is at issue is whether or not a given class of signs or constructs is individuated in terms of semantic information. Although it is typical for both lexical class constructions and maximal combinatoric constructions (see section 5 below) to make reference to semantic properties, there is no reason to expect that this is always the case. Section 5 also includes a discussion of compositionality in SBCG, which arises from a single principle governing semantically canonical constructs.

SBCG is committed to the inclusion of a precise theory of meaning that can be used to describe the semantics of linguistic expressions in general, and the

[35]Note that Goldberg characterizes her Aux-Initial Construction as 'semantically polysemous'. This notion can be eliminated entirely, however, if each subtype specifies an appropriate semantics, as seems independently necesary, leaving the Aux-Initial Construction with no semantic work to do.

semantics of constructions in particular. In this regard, it should be noted that SBCG is in principle compatible with almost any explicit approach to semantic analysis. Most work in BCG has assumed some version of 'Frame Semantics' (Fillmore 1982, 1985; Fillmore and Baker 2010), while work in construction-based HPSG (e.g. Ginzburg and Sag 2000) has usually embraced some version of Situation Semantics.[36] In addition, Sag (2010a) presents a comprehensive discussion of English filler-gap constructions couched in terms of an SBCG that embraces a conventional, Montague-style possible-worlds semantics.

Here we will utilize a version of Frame Semantics, blended together with the basic approach provided by Minimal Recursion Semantics.[37] Following the insights of recent and ongoing work in computational semantics, the representations we use are not hierarchically structured. Rather, the value of the feature SEM will include a list of predications (here taken to be frames), accompanied by a set of constraints that limits the way these labels can be connected. This 'flat' conception of semantics simulates embedding by identifying the value of a feature in one frame with the label that identifies another frame. Thus, to represent the meaning of a sentence like *Lee says that Garfield ate Mighty Mouse*, we will write (20a), instead of (20b) or the like:

(20) a.
$$\left\langle \begin{bmatrix} \textit{saying-fr} & \\ \text{LABEL} & l \\ \text{SPEAKER} & \textit{Lee} \\ \text{MESSAGE} & l' \end{bmatrix}, \begin{bmatrix} \textit{eating-fr} & \\ \text{LABEL} & l' \\ \text{INGESTOR} & \textit{Garfield} \\ \text{INGESTIBLE} & \textit{Mighty-Mouse} \end{bmatrix} \right\rangle$$

b.
$$\begin{bmatrix} \textit{saying-fr} & \\ \text{SPEAKER} & \textit{Lee} \\ \text{MESSAGE} & \begin{bmatrix} \textit{eating-fr} & \\ \text{INGESTOR} & \textit{Garfield} \\ \text{INGESTIBLE} & \textit{Mighty-Mouse} \end{bmatrix} \end{bmatrix}$$

Each semantic frame is assigned to a FS type, with the types organized in a

[36] I am referring here to the original framework of Situation Semantics. See Barwise and Perry 1983, 1985; Gawron and Peters 1990; Devlin 1991; Cooper and Ginzburg 1996; Seligman and Moss 1997.

[37] Wherever possible, I will borrow frames and features from FrameNet's website, found at http://framenet.icsi.berkeley.edu/. The FrameNet lexicon, the practical implementation of Fillmore's (1982) Frame Semantics, is a lexicographic database employing semantic frames as its main structuring device (Fillmore et al. 2003, Fillmore and Baker 2010).

For a general introduction to Minimal Recursion Semantics (MRS), see Copestake et al. 2005. Blackburn and Bos 2005 provides a useful introduction to the related framework of 'Hole Semantics', as well as a general discussion of issues in computational semantics.

multiple inheritance hierarchy so that shared common properties can be assigned to common supertypes: *saying-fr* is a subtype of *statement-fr*; *eating-fr* is a subtype of *ingestion-fr*. And since instances of a given subtype must specify values for all the features declared for that type and for those features inherited from its supertypes, each type of frame exhibits both idiosyncratic properties and the properties of more general frame classes. The frame hierarchy is used to account for many lexical entailments. For example, the fact that *glimpse-fr* is a subtype of *see-fr*, which in turn is a subtype of *perceive-fr* plays a crucial role in explaining the inference from *She glimpsed it* to *She perceived it*.

Our semantic discussion will be couched in terms of the following three features of the semantic objects serving as values of the feature SEM:

(21) a. INDEX is used to identify the referent of an expression. Its value is an index, functioning essentially as a variable assigned to an individual in the case of an NP. (Situational indices, corresponding to VPs or Ss, are discussed below.)

b. LTOP (LOCAL-TOP) takes a label (of a frame) as its argument. This label is a pointer to the sign's fully resolved semantics. It indicates the 'top' frame in the semantics of a sentence viewed as a tree of frames in which the hierarchical relation is embedding (see below).

c. The feature FRAMES is used to specify the list of predications that together determine the meaning of a sign. The value of FRAMES is a (possibly empty) list of frames.

In order to treat modification, as well as quantification over situations or events, it is useful for frames to be able to make reference to the situations they are used to describe. To this end, we will make use of a further feature, SITUATION (SIT), whose value is a situational index. This corresponds to a Davidsonian event variable, which is, roughly, a referential index denoting the situation described by an elementary predication. Thus we have semantic objects like (22):

(22)
$$\begin{bmatrix} \textit{sem-obj} & \\ \text{INDEX} & s \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{eating-fr} & \\ \text{LABEL} & l \\ \text{SIT} & s \\ \text{INGESTOR} & i \\ \text{INGESTIBLE} & j \end{bmatrix} \right\rangle \end{bmatrix}$$

The world of frames must encompass elements appropriate for the analysis of all parts of speech. I will assume that the type of the frame on the FRAMES list of a common noun like *book*, for example, can resolve either to *book*^{po}*-fr*, corresponding to its 'physical object' sense, or to *book*^{to}*-fr*, which represents its 'textual object' sense. The *book* listeme, however, will be specified as in (23), where *book-fr* is an immediate supertype of both these types (*cn-lxm* stands for *common-noun-lexeme*):

(23)
$$\begin{bmatrix} \textit{cn-lxm} & \\ \text{FORM} & \langle \textit{book} \rangle \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & i \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{book-fr} & \\ \text{LABEL} & l_0 \\ \text{ENTITY} & i \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Polysemy is systematic in English and is appropriately analyzed via lexical underspecification of this sort. In any model of the noun *book*, the frame's type must be either *book*^{po}*-fr* or *book*^{to}*-fr*, since the type of all feature structures must be maximally resolved. In addition, note that *book*^{po}*-fr* will also be a subtype of *physical-object-fr* and *book*^{to}*-fr* of *textual-object-fr*, allowing more general linguistic properties of physical objects and textual objects to be expressed in terms of constraints operating at a higher level (see also Pustejovsky 1995).

Note further that the lexical underspecification proposed here is independently motivated on psycholinguistic grounds. Frazier and Raynor (1990) found that examples like (24a) and (24b) are both read without any 'garden path' effect:

(24) a. Apparently, the book didn't sell, after having so many pages torn.
b. Apparently, the book didn't sell, after taking so long to write.

The following context in (24a) resolves the interpretation of *book* to its physical object sense, while in (24b), *book* is resolved to its textual object interpretation. The fact that subjects effortlessly resolved the interpretation either way suggests, as Frazier and Raynor argue, that the initial processing of polysemous words like *book* involves an underspecified semantic representation – like the one sketched in (23). In particular contexts, of course, semantic and pragmatic factors may force an early resolution of such ambiguities. The point here is rather that such resolution is not forced.

This result contrasts with Frazier and Raynor's findings with respect to words like *band, fan, bug, date, pitcher*, and *club* – words which exhibit two unrelated meanings. Here subjects committed to one interpretation initially, which could cause a garden-path effect given inconsistent subsequent material. This contrast can be explained by assuming that each of these words corresponds to multiple listemes, rather than to a single, underspecified listeme. Hence the grammar provides no basis for an underspecified representation constructed at an intermediate stage of processing.

Frame Semantics must also accommodate determiners and NPs, whose semantic status as generalized quantifiers is now firmly established.[38] This can be accomplished by positioning a general type in the frame space that has all generalized quantifier frames as a subtype. Generalized quantifier frames specify values for the features RESTRICTION (RESTR) and SCOPE, as well as BOUND-VARIABLE (BV), and of course LABEL. The frames *most-fr* and *few-fr* are maximal in this part of the frame hierarchy, while intermediate types may be assumed to provide a locus for stating constraints that define such classes as monotone-increasing or anti-additive quantifiers, which have been discussed at length in the considerable literature on generalized quantifiers.

Thus a model of the NP *every book* (a sign) will include all the information shown in (25):

(25)

$$\left[\begin{array}{ll} \text{FORM} & \langle \textit{every, book} \rangle \\ \text{SEM} & \left[\begin{array}{ll} \textit{sem-obj} & \\ \text{INDEX} & i \\ \text{LTOP} & l_0 \\ \text{FRAMES} & \left\langle \left[\begin{array}{ll} \textit{every-fr} & \\ \text{LABEL} & l_1 \\ \text{BV} & i \\ \text{RESTR} & l_2 \\ \text{SCOPE} & l_3 \end{array}\right], \left[\begin{array}{ll} \textit{book}^{po}\textit{-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \end{array}\right] \right\rangle \end{array}\right] \end{array}\right]$$

Note that in (25) the label of *book*^{po}*-fr* (i.e. l_2) has been identified with the RESTR argument of *every-fr*, which intuitively conveys the information that the range of the quantification is restricted to books (in the physical object sense). Neither the SCOPE value of the generalized quantifier (l_3) nor its LABEL value (l_1) are identified with the label of any other frame. This indicates that the quantifier's scope is not determined in (25).

To further illustrate how the semantics works, consider example (26):

(26) Some student knows every answer.

This sentence is ambiguous – either 'one student knows all the answers', or

[38] See, for example, Keenan and Westerståhl 1997, Keenan 2002, and Peters and Westerståhl 2006 (but also Hobbs 1983, 1996, Kamp and Reyle 1993, and Steedman 2012, for interesting alternatives).

else the weaker 'there is no answer unknown to every student'. This is usually represented as a scope ambiguity, where quantifiers are here represented as '*(quantifier variable, restriction)*' (*restriction* is a sentence), as illustrated in (27):

(27) a. *(some i, student(i))(every j, answer(j))(know(i,j))*
b. *(every j, answer(j))(some i, student(i))(know(i,j))*

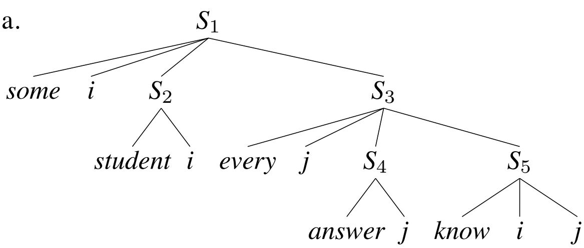
A slightly different, but equivalent, way of representing these meanings, in terms of *(quantifier, variable, restriction, scope)* formulas (*restriction* and *scope* are both sentences), is shown in (28):

(28) a. *(some, i, student(i), (every, j, answer(j), know(i,j)))*
b. *(every, j, answer(j), (some, i, student(i), know(i,j)))*

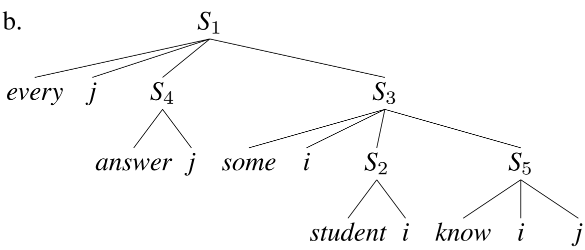
It is quantificational representations of this kind that are provided by our feature structures of type *quantifier-frame*.

Notice that the formulas in (28), indeed all standard logic formulas, may equivalently be expressed as trees. For example, (28a) and (28b) could be represented as (29a) and (29b), respectively:

(29) a.



b.



Trees of this kind correspond to resolved MRS structures, where all scope is explicitly indicated. For example, (30) is equivalent to (29a) and (31) to (29b):[39]

[39]Resolved MRS structures must meet a further condition requiring that every occurrence of a variable be within the restriction or scope of the quantifier that binds it. For convenience, I am ignoring the feature SIT in this discussion.

(30)

$$\begin{bmatrix} \textit{sem-obj} & \\ \text{LTOP} & l_1 \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LABEL} & l_1 \\ \text{BV} & i \\ \text{RESTR} & l_2 \\ \text{SCOPE} & l_3 \end{bmatrix}, \begin{bmatrix} \textit{student-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \end{bmatrix}, \begin{bmatrix} \textit{every-fr} & \\ \text{LABEL} & l_3 \\ \text{BV} & j \\ \text{RESTR} & l_4 \\ \text{SCOPE} & l_5 \end{bmatrix}, \begin{bmatrix} \textit{answer-fr} & \\ \text{LABEL} & l_4 \\ \text{ENTITY} & j \end{bmatrix}, \begin{bmatrix} \textit{knowing-fr} & \\ \text{LABEL} & l_5 \\ \text{COGNIZER} & i \\ \text{COGNIZED} & j \end{bmatrix} \right\rangle \end{bmatrix}$$

(31)

$$\begin{bmatrix} \textit{sem-obj} & \\ \text{LTOP} & l_3 \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LABEL} & l_1 \\ \text{BV} & i \\ \text{RESTR} & l_2 \\ \text{SCOPE} & l_5 \end{bmatrix}, \begin{bmatrix} \textit{student-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \end{bmatrix}, \begin{bmatrix} \textit{every-fr} & \\ \text{LABEL} & l_3 \\ \text{BV} & j \\ \text{RESTR} & l_4 \\ \text{SCOPE} & l_1 \end{bmatrix}, \begin{bmatrix} \textit{answer-fr} & \\ \text{LABEL} & l_4 \\ \text{ENTITY} & j \end{bmatrix}, \begin{bmatrix} \textit{knowing-fr} & \\ \text{LABEL} & l_5 \\ \text{COGNIZER} & i \\ \text{COGNIZED} & j \end{bmatrix} \right\rangle \end{bmatrix}$$

Note further that the local top (the value of LTOP) is the label of the node at the top of the semantic tree, which corresponds to the quantifier with the widest scope in the formula. This is a 'local' top because the entire clause (*some student knows each answer*) might be embedded within a larger semantic structure, whose semantics would have its own local top. The LTOP is thus a useful bookkeeping device that lets us specify a sign's semantics (the top of its semantic tree) without committing to a particular interpretation in the case of scope ambiguity.

In a simple clause (i.e. one without verbal modifiers), the LTOP of the verb will be identified with that of the VP and the S that it projects. Since the clause may contain quantifiers, e.g. in examples like those just considered, the verb's listeme cannot identify its LTOP value with the label of the frame on its FRAMES list. However, that listeme should impose the restriction that in the resolved MRS, the LTOP is either identical to the label of the verb's frame or else 'higher' in the structure, where some quantifier embeds the verb's frame in its SCOPE (and some other quantifier might embed that quantifier in its SCOPE, and so forth). Such a restriction is the same as requiring that the verb's LTOP be identical to the simple clause's LTOP 'modulo quantification'.[40] Whenever we need to impose such a condition, i.e. that l_0 is either equal to l_1 or higher than l_1, separated from it only by quantifiers, we will informally write '$l_{0=_q 1}$'.

Scope underspecification is the heart and soul of MRS, which was originally developed for computational purposes.[41] Scope resolution – the problem of how (in a given context) to completely resolve the scope of however many quantifiers appear in a given sentence – is a difficult and currently unsolved research problem. Hence, having semantic representations that do not require full scope resolution is a significant advantage in much computational work, e.g. machine translation. However, as pointed out by Sanford and Sturt (2002), it may be useful in psycholinguistics, as well. Certain experimental results, for example those of Tunstall (1998) about the processing of two-quantifier sentences, can be explained by assuming that sentence processing, at least sometimes, initially involves a scope-neutral representation. Tunstall showed that subjects found dialogues like (32) and (33) to be equally (and fully) natural:

(32) Kelly showed every photo to a critic last month.
The critic was from a major gallery.

(33) Kelly showed every photo to a critic last month.
The critics were from a major gallery.

The continuation in (32) imposes an 'A $>$ Every' ('A outscopes Every') interpretation on the preceding two-quantifier sentence, while the one in (33) resolves to an 'Every $>$ A' interpretation. The fact that both continuations are equally natural, Sanford and Sturt argue, can be explained simply by assum-

[40] This is not true in more complex sentences. For example, scopal modifiers in sentences like (i) introduce their own LTOP value:

(i) Some student frequently knows every answer.

Here the clause's LTOP is inherited from the adverb, which embeds the verb's LTOP as an argument. For further discussion, see Copestake et al. 2005.

[41] That research was conducted in the context of the LINGO Lab at Stanford's Center for the Study of Language and Information (CSLI), which has participated in a number of computational projects involving SBCG-style grammars, including the VERBMOBIL speech-to-speech translation project (see Uszkoreit et al. 2000).

ing that the initial processing of the two-quantifier sentence involves a scope-neutral representation that is unproblematically (monotonically) refined into either interpretation. Thus, MRS, which allows partially resolved, as well as fully scope-neutral representations, may prove useful in terms of the broader goal of embedding SBCG within a realistic model of human language use.

Our grammar associates *some student knows each answer* with an unresolved MRS description like the one shown in (34):[42]

(34)
$$\begin{bmatrix} \textit{sem-obj} & \\ \text{LTOP} & l_{0=_q 5} \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LABEL} & l_1 \\ \text{BV} & i \\ \text{RESTR} & l_2 \\ \text{SCOPE} & l_6 \end{bmatrix}, \begin{bmatrix} \textit{student-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \end{bmatrix}, \begin{bmatrix} \textit{every-fr} & \\ \text{LABEL} & l_3 \\ \text{BV} & j \\ \text{RESTR} & l_4 \\ \text{SCOPE} & l_7 \end{bmatrix}, \begin{bmatrix} \textit{answer-fr} & \\ \text{LABEL} & l_4 \\ \text{ENTITY} & j \end{bmatrix}, \begin{bmatrix} \textit{knowing-fr} & \\ \text{LABEL} & l_5 \\ \text{COGNIZER} & i \\ \text{COGNIZED} & j \end{bmatrix} \right\rangle \end{bmatrix}$$

This description leaves open which scoping is intended by not identifying the value of LTOP with the any frame's LABEL value. However, by imposing the requirement that a top-level MRS list must be scope-resolved (and hence representable as a tree), we ensure that a description like (34) has two distinct models, depending on whether the resolution proceeds by identifying the labels as in (35a) or as in (35b):

(35) a. $l_0 = l_1, l_3 = l_6$, and $l_5 = l_7$
b. $l_0 = l_3, l_1 = l_7$, and $l_5 = l_6$

The former gives us (30) and the latter, (31), thus accounting for the ambiguity of our two-quantifier sentence. Note finally that these are the only scope resolutions of (34), i.e. the only ways of identifying the labels so that all of the frames shown in (34) are included within a single connected tree.

[42]Note that the treatment of quantifier restrictions presented here is in fact a simplification. In MRS, a quantifier that appears inside an NP may be scoped inside the restriction of that NP's quantifier, as in examples like (i):

(i) Each solution to a problem was posted on the internet.

3.5 CONTEXT

I will have relatively little to say about CONTEXT (CNTXT) values here. A basic approach to the structure of CNTXT values (FSs of type *context*) is developed by Pollard and Sag (1994).[43] This is based on such features as BACKGROUND (BCKGRND) and CONTEXTUAL-INDICES (C-INDS), where the latter specifies values for such features as SPEAKER (SPKR), ADDRESSEE (ADDR), and UTTERANCE-LOCATION (UTT-LOC). The context-objects in such a theory look like (36):

(36)
$$\begin{bmatrix} \textit{context} & \\ \text{C-INDS} & \begin{bmatrix} \text{SPKR} & \textit{index} \\ \text{ADDR} & \textit{index} \\ \text{UTT-LOC} & \textit{index} \\ \ldots & \end{bmatrix} \\ \text{BCKGRND} & \textit{list(proposition)} \end{bmatrix}$$

The various contextual indices specify contextual elements that underly an account of indexical and deictic expressions formulated in the style of D. Kaplan's (1989) seminal work. The propositions specified in the BCKGRND value correspond to the set of utterance felicity conditions, which may in principle be contributed by any expression within the sentence being uttered. These conditions are recursively inherited by a general principle requiring that the BCKGRND of a construct's mother must include the elements in the BCKGRND lists of all its daughter signs.

This brings us to the analysis of proper nouns, which will be treated here, building on Pollard and Sag (1994), in terms of a background condition (a 'presupposition') that the individual being referred to by a particular proper noun in fact exists (or exists in some more specific sense involving relevance to the current conversation) and is 'so named'.[44] A word corresponding to the proper noun *Pat* will thus include information like that sketched in (37):

[43] See also Green 1996, Green 2000; Ginzburg and Sag 2000, Ginzburg and Cooper 2004, Ginzburg 2012, P. Kay 1997, Fillmore 1985, and Lee-Goldman 2011.

[44] I am of course skirting around myriad subtle issues about the meaning and multiple uses of names. This essentially Fregean analysis is also close to those suggested by Recanati (1997), Geurts (1997), and the variant adapted to 'naming' uses by Matushansky (2008). Clearly, this is not the only analysis available within SBCG. However, by making context ever-present in the derivation of signs, SBCG is compatible with a wide range of context-dependent analyses of interpretational phenomena.

(37)
$$\begin{bmatrix} \text{FORM} & \langle \textit{Pat} \rangle \\ \text{SEM} & \begin{bmatrix} \textit{sem-obj} & \\ \text{INDEX} & i \\ \text{FRAMES} & \langle\,\rangle \end{bmatrix} \\ \text{CONTEXT} & \begin{bmatrix} \text{BCKGRND} & \left\langle \begin{bmatrix} \textit{naming-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \\ \text{NAME} & \langle \textit{Pat} \rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

It is also possible to augment these assumptions about contextual features, incorporating, for example, the features TOPIC and/or FOCUS, as in Michaelis and Lambrecht 1996. In a similar vein, Engdahl and Vallduvi (1994, 1996) analyze ‘information packaging’ in terms of CONTEXT values that are structured as shown in (38):

(38)
$$\begin{bmatrix} \textit{context} & \\ \text{INFO-STRUCTURE} & \begin{bmatrix} \text{FOCUS} & \ldots \\ \text{GROUND} & \begin{bmatrix} \text{LINK} & \ldots \\ \text{TAIL} & \ldots \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

I will not explore such elaborations here.[45]

Finally, it is worth observing that there has been considerable interesting work on embedding contextual structures of the sort illustrated in this section within larger models of dialogue. Ginzburg (2012)’s theory of dialogue, which crucially employs such context-objects (associated with diverse sentence types in Ginzburg and Sag 2000), expands the inventory of relevant features to include DIALOGUEGAMEBOARD (DGB), which allows him to draw a distinction between ‘public’ and ‘private’ information relevant to dialogue. Other features, e.g. QUESTIONSUNDERDISCUSSION (QUD), FACTS, and MOVES provide part of a bridge between the contextual requirements of particular sentences and the general requirements of the theory of dialogue;

[45]Nor will I explore the ways in which a set of background propositions may be structured by relations such as unilateral entailment, as in a scalar model (Fillmore et al. 1988, P. Kay 1990, Israel 1996, Schwenter 1999, Schwenter and Vasishth 2000). Additionally, some contextual propositions may be mentioned in a sign type or construction, not for acceptance of their content’s being a felicity condition on the utterance, but as ‘context propositions’, whose content is acknowledged as being ‘on the floor’, although not necessarily accepted – perhaps specifically denied – by a conversational participant. See, for example P. Kay 1997.

this gives an account of such notions as utterance coherence, conversational moves, updating and downdating, and conversational repair.[46]

3.6 Signs: A Synthesis

Signs are analyzed as FSs that specify values for the five features PHON, FORM, SYN, SEM, and CNTXT, whose values have now all been introduced. Accordingly, the grammar signature contains the following type declaration:

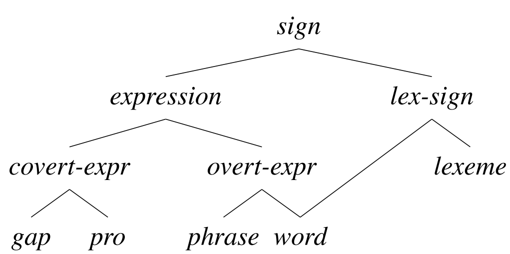
(39)
$$sign : \begin{bmatrix} \text{PHON} & \textit{phon-obj} \\ \text{FORM} & \textit{morph-obj} \\ \text{SYN} & \textit{syn-obj} \\ \text{SEM} & \textit{sem-obj} \\ \text{CNTXT} & \textit{context-obj} \end{bmatrix}$$

The immediate subtypes of *sign* are *lexical-sign* (*lex-sign*) and *expression*. Lexical signs include ARG-ST specifications in virtue of the type declaration in (40):

(40) *lex-sign* : [ARG-ST *list*(*expression*)]

The immediate subtypes of *expression* are *covert-expr*(*ession*) and *overt-expr*(*ession*), the latter of which has the two immediate subtypes *word* and *phrase*. There are two immediate subtypes of *lexical-sign* – *word* and *lexeme*. The supertype relations of *word* thus reflect the fact that words share properties with phrases that lexemes lack (e.g. the ability to be a daughter in a phrasal construct, discussed in the next section) and that words share a property with lexemes that phrases lack (that is, having an ARG-ST list). These cross-cutting properties of words are analyzed by treating FSs of type *word* as both overt expressions and lexical signs, as indicated in the multiple-inheritance hierarchy shown in (41):

(41)



[46]Building on some of Ginzburg's ideas, but introducing modifications of the theory of DGB, Marandin and colleagues develop interesting approaches to the problems of focus and intonational meaning (Beyssade and Marandin 2007), evidential meaning (Marandin 2008), and the meaning/force of negation (Godard and Marandin 2006).

$$\begin{bmatrix} \textit{word} & \\ \text{PHON} & /\textit{pæt}/ \\ \text{FORM} & \langle \textit{Pat} \rangle \\ \text{ARG-ST} & \langle\,\rangle \\ \text{SYN} & \begin{bmatrix} \textit{syn-obj} & \\ \text{CAT} & \begin{bmatrix} \textit{noun} & \\ \text{CASE} & \textit{nom} \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \textit{none} \end{bmatrix} \\ \text{VAL} & \langle\,\rangle \\ \text{MRKG} & \textit{det} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \textit{sem-obj} & \\ \text{IND} & i \\ \text{FRAMES} & \langle\,\rangle \end{bmatrix} \\ \text{CNTXT} & \begin{bmatrix} \textit{context-obj} & \\ \text{BCKGRND} & \left\langle \begin{bmatrix} \textit{naming-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \\ \text{NAME} & \langle \textit{Pat} \rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

FIGURE 4 A Model of the Word *Pat*

We are now in a position to illustrate in more detail what the various signs introduced earlier will look like in SBCG. Recall that boxed AVMs, like those in Figures 4, 5,[47] and 6, indicate a function – i.e. a model of a particular sign, rather than a function description (a constraint satisfied by a family of FSs). Hence the FS in Figure 6 must have a determinate value for each of its features, including (at the appropriate internal level) VF. So the value *psp* illustrated here (like the choice of *nom* as the CASE value in Figure 4) represents an arbitrary expositional choice – any value VF would satisfy the requirements imposed by the *laugh* listeme. And each such choice gives rise to a family of well-formed FSs licensed by that listeme. Finally, observe that each of these FSs instantiates a maximal type.

[47] Further abbreviations:

NP_i = NP & [SEM|INDEX i] NP[*nom*] = NP & [SYN|CAT|CASE *nom*]

S-SRCE = SOUND-SOURCE

Recall that a feature structure is a total function on the appropriate domain but that a diagram illustrating such a function may not include every feature-value pair that it contains.

$$
\begin{bmatrix}
\textit{word} & \\
\text{PHON} & \textit{/læf-d/} \\
\text{FORM} & \langle \textit{laughed} \rangle \\
\text{ARG-ST} & \langle \text{NP}[\textit{nom}]_i \rangle \\
\text{SYN} & \begin{bmatrix}
\textit{syn-obj} & \\
\text{CAT} & \begin{bmatrix}
\textit{verb} & \\
\text{VF} & \textit{fin} \\
\text{SELECT} & \textit{none} \\
\text{XARG} & \text{NP}[\textit{nom}]_i \\
\text{LID} & \left\langle \begin{bmatrix}
\textit{laughing-fr} & \\
\text{LABEL} & l_3 \\
\text{SIT} & s \\
\text{S-SRCE} & i
\end{bmatrix} \right\rangle
\end{bmatrix} \\
\text{MRKG} & \textit{unmk} \\
\text{VAL} & \langle \text{NP}[\textit{nom}]_i \rangle
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix}
\textit{sem-obj} & \\
\text{IND} & s \\
\text{LTOP} & l_0 \\
\text{FRAMES} & \left\langle \begin{bmatrix}
\textit{some-fr} & \\
\text{LABEL} & l_1 \\
\text{BV} & s \\
\text{RESTR} & l_2 \\
\text{SCOPE} & l_3
\end{bmatrix}, \begin{bmatrix}
\textit{laughing-fr} & \\
\text{LABEL} & l_3 \\
\text{SIT} & s \\
\text{S-SRCE} & i
\end{bmatrix}, \begin{bmatrix}
\textit{past-fr} & \\
\text{LABEL} & l_2 \\
\text{ARG} & s
\end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

FIGURE 5 A Model of the Word *laughed*

$$
\begin{bmatrix}
\textit{sintrans-v-lxm} & \\
\text{PHON} & /\textit{læf}/ \\
\text{FORM} & \langle \textit{laugh} \rangle \\
\text{ARG-ST} & \langle \text{NP}_i \rangle \\
\text{SYN} & \begin{bmatrix} \textit{syn-obj} & \\ \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{SELECT} & \textit{none} \\ \text{VF} & \textit{psp} \\ \text{XARG} & \text{NP}_i \\ \text{LID} & \left\langle \begin{bmatrix} \textit{laughing-fr} & \\ \text{LABEL} & l_1 \\ \text{SIT} & s \\ \text{S-SRCE} & i \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{MRKG} & \textit{unmk} \\ \text{VAL} & \langle \text{NP}_i \rangle \end{bmatrix} \\
\text{SEM} & \begin{bmatrix} \textit{sem-obj} & \\ \text{IND} & s \\ \text{LTOP} & l_1 \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{laughing-fr} & \\ \text{LABEL} & l_1 \\ \text{SIT} & s \\ \text{S-SRCE} & i \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

FIGURE 6 A Model of the Lexeme *laugh*

The sign corresponding to the (sentential) phrase *Pat laughed* is shown in Figure 7. It should be noted that the diagram does not represent a construct, but rather a single *sign*. Indeed, the model that SBCG provides for each expression of a language is a sign, even though the analysis (or the 'construction') of that sign if it is not 'listemically licensed' (see below) must involve a construct with that sign as mother and one or more signs as daughters.[48]

[48]Note also that a resolved semantics for this sign must have the quantifier as its 'top' operator. In this case, the quantifier's label is identified with the LTOP value.

$$
\begin{bmatrix}
\textit{phrase} & \\
\text{PHON} & \textit{/pæt\#'læf-d/} \\
\text{FORM} & \langle \textit{Pat, laughed} \rangle \\
\text{SYN} & \begin{bmatrix} \textit{syn-obj} & \\ \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{VF} & \textit{fin} \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \text{NP}[\textit{nom}]_i \\ \text{LID} & \left\langle \begin{bmatrix} \textit{laughing-fr} & \\ \text{LABEL} & l_3 \\ \text{SIT} & s \\ \text{S-SRCE} & i \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{VAL} & \langle\ \rangle \\ \text{MRKG} & \textit{unmk} \end{bmatrix} \\
\text{SEM} & \begin{bmatrix} \textit{sem-obj} & \\ \text{IND} & s \\ \text{LTOP} & l_1 \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LABEL} & l_1 \\ \text{BV} & s \\ \text{RESTR} & l_2 \\ \text{SCOPE} & l_3 \end{bmatrix}, \begin{bmatrix} \textit{laughing-fr} & \\ \text{LABEL} & l_3 \\ \text{SIT} & s \\ \text{S-SRCE} & i \end{bmatrix}, \begin{bmatrix} \textit{past-fr} & \\ \text{LABEL} & l_2 \\ \text{ARG} & s \end{bmatrix} \right\rangle \end{bmatrix} \\
\text{CNTXT} & \begin{bmatrix} \textit{context-obj} & \\ \text{BCKGRND} & \left\langle \begin{bmatrix} \textit{naming-fr} & \\ \text{LABEL} & l_5 \\ \text{ENTITY} & i \\ \text{NAME} & \langle \textit{Pat} \rangle \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

FIGURE 7 The Phrase *Pat laughed*

4 Well-Formed Feature Structures

As explained earlier, the grammar signature of an SBCG defines a space of well-formed FSs by placing general constraints on the domain and range of each type of (nonatomic) FS. The minimal conditions that must be satisfied in order for a FS to be well formed are shown in (42):

(42) A feature structure F is well-formed just in case:

1. F instantiates a maximal type τ,
2. F is either an atom or a total function whose domain and range are specified for τ by the grammar signature, and
3. F satisfies all constraints that the grammar imposes on FSs of type τ and τ's supertypes.

But if the only constraints imposed on feature structures were the general domain/range specifications of the grammar signature, then the grammar would license many FSs that are not well-formed. To see this, consider the FS in Figure 8, which is not a sign of English. Given what has been said so far, Figure 8 illustrates a well-formed FS of type *sign*. Each feature (PHON, FORM, SYN, SEM, and CNTXT) has a value of an appropriate type and each of these values is a feature structure conforming to the grammar signature. However, even though this is the case, there are numerous problems with the FS in Figure 8, such as:

(43) a. This is a finite clause whose FORM value is ⟨Kim, the⟩, yet *Kim the* is not a well-formed English expression.

b. The FORM value ⟨Kim, the⟩ cannot be phonologically realized as /pæt#tuː/.

c. The meaning of the sentence is (roughly) that a person named Bo sneezed at some time in the past, yet that meaning cannot be expressed in English by uttering /pæt#tu/

Clearly, the inventory of types and the feature declarations that are specified in the grammar's signature are insufficient for doing the business of grammar.

Unwanted signs like the one in Figure 8 are ruled out by the inventory of listemes and constructions that are part of an SBCG grammar of English. Each construction is a constraint defining the properties that are common to all instances of a given FS type. That is, a construction is a conditional constraint of the form shown in (44), where τ (the antecedent) is a type name and D (the consequent) is a FS description:

(44) $\tau \Rightarrow D$

(All FSs of type τ must satisfy D.)

Once the grammar is appropriately populated with its 'lexicon' (a set of listemes) and its 'constructicon' (a set of lexical class and combinatoric con-

$$
\begin{bmatrix}
\textit{phrase} & \\
\text{PHON} & \textit{/pæt\#tu/} \\
\text{FORM} & \langle \textit{Kim, the} \rangle \\
\text{SYN} & \begin{bmatrix}
\textit{syn-obj} & \\
\text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{VF} & \textit{fin} \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \text{NP}[\textit{nom}]_j \end{bmatrix} \\
\text{VAL} & \langle\,\rangle \\
\text{MRKG} & \textit{that}
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix}
\textit{sem-obj} & \\
\text{IND} & s \\
\text{FRAMES} & \left\langle \begin{bmatrix} \textit{sneezing-fr} & \\ \text{ACTOR} & i \\ \text{SIT} & s \end{bmatrix}, \begin{bmatrix} \textit{past-fr} & \\ \text{ARG} & s \end{bmatrix} \right\rangle
\end{bmatrix} \\
\text{CNTXT} & \begin{bmatrix}
\textit{context-obj} & \\
\text{BCKGRND} & \left\langle \begin{bmatrix} \textit{naming-fr} & \\ \text{ENTITY} & i \\ \text{NAME} & \textit{Bo} \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

FIGURE 8 An Ungrammatical Sign

structions), then the definition of well-formed FS can be augmented to include the following grammatical principle:[49]

[49] Note that my terminology here differs somewhat from that commonly assumed in BCG, where a constructicon is said to include both lexical entries and constructions. But the listemes and constructions of SBCG all state conventional constraints on the sound-meaning correspondences of a language; hence this discrepancy is one of terminology, rather than analytic content. Similarly, what Croft (2003) refers to as 'verb-specific constructions' correspond to SBCG listemes and what he calls 'verb-class specific constructions' correspond to lexical class constructions.

In SBCG, constructions are constraints on classes of signs or constructs and listemes are likewise constraints on signs. That is, because of the Sign Principle, the listemes in the lexicon together provide a disjunctive constraint on the set of 'basic' lexical signs (those not licensed by combinatoric constructions). Note that we could eliminate clause a. of the Sign Principle in order to emphasize the functional similarlity between listemes and constructions, for example by positing a schematic 'Ur Construction' formulated along the lines of (i):

(i) *ur-construct* $\Rightarrow \begin{bmatrix} \text{MTR} & \lambda \\ \text{DTRS} & \langle\,\rangle \end{bmatrix}$, where λ is any listeme.

(45) **The Sign Principle:**

Every sign must be listemically or constructionally licensed, where:

a. a sign is listemically licensed only if it satisfies some listeme, and

b. a sign is constructionally licensed only if it is the mother of some well-formed construct.

The Sign Principle works in tandem with the lexicon, the constructicon, and the well-formedness conditions in (42) above. That is, it specifies a further condition that must be satisfied by FSs that are of type *sign*: lexemes, words, and phrases. The goal of the next two sections is to lay out some of the particulars of this relatively simple formulation of grammatical theory.

5 Constructs and Constructions

The Construction Grammar community usually defines **construction** informally as 'any conventionalized pairing of form and meaning'.[50] This essence of this conception of construction is preserved, but it is slightly refined and made more precise in SBCG. A listeme is a possibly complex constraint on a particular form-meaning correspondence.[51] The lexemic listeme *love*, for example, enforces the basic form-meaning correspondence that permeates nominal and verbal words based on this lexeme. Lexical class constructions, which define the more general properties of classes of lexemes and words (see below), typically associate formal properties (e.g. grammatical category, argument structure list) with more abstract semantic information (e.g. a common frame supertype). But since an SBCG grammar is a recursive system, the constructions that define the ways that signs can combine to build more complex signs (combinatoric constructions) specify form-meaning correspondences indirectly – by constraining the relation between the form and meaning of a construct's mother sign and those of its daughters.

Constructs are modeled as FSs of the form specified in (46):[52]

But such modifications seem to introduce an otherwise unnecessary level of analytic complexity and hence will be avoided here.

[50]Goldberg (2006: 5) offers a slightly different statement of this definition:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency.

[51]The term 'form' here is construed broadly, so as to include syntactic, lexical, and morphological form.

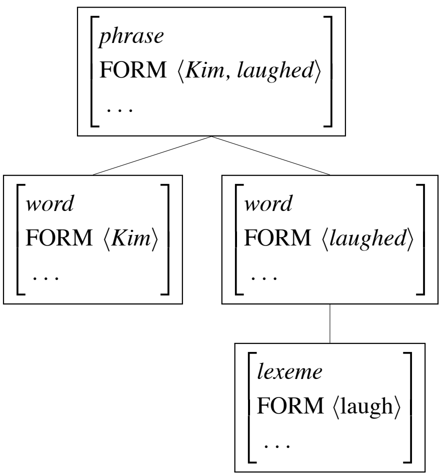
[52]For any type τ, *nelist*(τ) stands for a nonempty list, each of whose members is a feature structure of type τ. (Note that if τ is nonmaximal, then each feature structure will belong to some maximal subtype of τ.)

(46)
$$construct : \begin{bmatrix} \text{MTR} & \textit{sign} \\ \text{DTRS} & \textit{nelist(sign)} \end{bmatrix}$$

The MOTHER (MTR) feature is used to place constraints on the set of signs that are licensed by a given construct. The feature DAUGHTERS (DTRS) specifies information about the one or more signs that contribute to the analysis of a construct's mother; the value of DTRS is a nonempty list of signs. The inventory of constructions determines which constructs are well-formed and this inventory of constructs in turn licenses a set of signs, as per the Sign Principle in (45).

The term 'construct' thus has a meaning here that is somewhat different from the way it has been used in the tradition of BCG. In BCG, a construct was sometimes regarded as a fleshed-out 'FS tree' of any degree of configurational complexity (See P. Kay 2002b). Even a single node FS tree was a construct in this conception of BCG, as well as a non-local tree like (47), referred to here as an **analysis tree**:

(47)

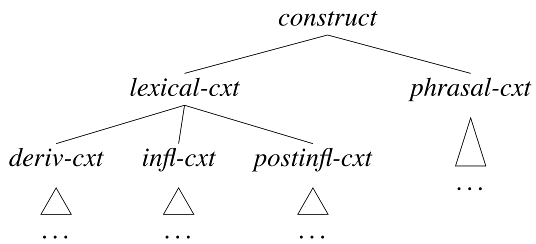


In SBCG, an expression – a word, phrase or sentence – is modeled by a sign (or family of signs). Although such signs are usually licensed by reference to a well-formed construct (i.e. a local tree), an analysis tree like (47) (which is the rough equivalent in the present architecture of an FS tree in BCG) has no 'official' status in SBCG.

Such diagrams are of course useful for illustrating the recursive effect of grammatical constructions or for demonstrating that a given sign is licensed by the grammar, but they are not part of the language model, nor are they part of the grammar.[53] From the BCG perspective then, the major theoretical changes embodied in SBCG are: (1) the use of signs and constructs as language models, rather than FS trees and (2) the imposition of locality on constructions, limiting their purview to local trees.

The immediate subtypes of *construct* are *lexical-construct* (*lexical-cxt*) and *phrasal-construct* (*phrasal-cxt*). Lexical constructs (following Sag et al. (2003, Ch.16)) are further classified in terms of the subtypes *derivational-construct* (*deriv-cxt*), *inflectional-construct* (*infl-cxt*), and *postinflectional-construct* (*postinfl-cxt*). This constructional hierarchy is sketched in (48):

(48)



For each type of construct, the constructicon will contain a combinatoric construction – a conditional constraint – with that type as its antecedent. In this way, there is a one-to-one relation between constructions and the types that name the class of FSs that they characterize.

6 Licensing Words

The properties of a word – a verb, say – are partly determined by a listeme (a lexeme-description in the lexicon), partly by lexical class constructions, and partly by derivational and/or inflectional constructions. Following traditional terminology, derivational constructions define ways that lexemes can be formed from other lexemes and inflectional constructions define the patterns by which words can be constructed from lexemes.[54] The morphological

[53] It should be noted, however, that a minor modification of SBCG would redefine the licensed objects as analysis trees, rather than local trees, as was done, for instance, in GPSG (Gazdar et al. 1985).

[54] My approach to morphology here is realizational (Matthews 1991, Anderson 1992), perhaps closest in spirit to the approach developed by Stump (2001) and related work. Morphological affixes are not signs in the analyses presented here. Rather, affixation (as well as more complex morphological processes) is effected by the morphological functions associated with specific lexical constructions. I have not addressed the issues that distinguish this kind of approach from others that have been developed. In particular, there are constructional approaches to morphology that have challenged a number of the premises that I have simply assumed here, e.g. that of Booij 2010. I believe that insights of Booij's approach and others like it can also be preserved within SBCG. However, the relevant issues, some of which involve the matter of constructional locality (see section 8.2 below), are beyond the scope of the present chapter.

stem, semantic frame, and ARG-ST list of *laughed*, for example, are specified in the listeme licensing the lexeme *laugh* (see below), but its inflected form and the added constraints – that (i) its VF value must be *fin* and (ii) the CASE value of its subject (its first, and only, ARG-ST member) must be *nom* – are determined by the preterite construction, one of a handful of inflectional constructions that remain in the grammar of Present-Day English.

Derivational and inflectional constructions fit uniformly into a two-level mode, one that is articulated in terms of a mother and its daughter(s).[55] For example, verbal words whose form is *laughed* are constructed from verbal lexemes whose form is *laugh*, in accordance with the Preterite Construction, which licenses constructs whose mother is a word and whose daughter is a lexeme.

6.1 Listemes and Lexical Class Constructions

Let us begin with listemes. As in most feature-based theories of grammar, a lexical entry is specified as a constraint relating form, syntactic category, and meaning. In SBCG, listemes contain varying degrees of information about all aspects of lexemes. A listeme is thus usually a description of a set of FSs of type *lexeme* or *word*. Since the value of the FORM feature is a list, the generalization from lexical entry to listeme is quite natural in SBCG.

Listemes are typically quite spartan. This is because an SBCG grammar contains type constraints that play a significant role in FS well-formedness. These are the **lexical class constructions** mentioned above. The lexical class construction in Figure 9, for example, defines once and for all the properties that distinguish proper nouns from other lexical classes (*pn-lxm* stands for *proper-noun-lexeme*):[56]

Once constraints like this are in place, a proper noun's listeme can be reduced to a specification of (1) a FORM value, (2) the relevant lexical type and (3) any exceptional properties. For example, a listeme like (49) licenses a FS like the one in Figure 10:

(49) $\begin{bmatrix} \textit{pn-lxm} & \\ \text{FORM} & \langle \textit{Kim} \rangle \end{bmatrix}$

[55] The theory of lexical constructions presented here builds directly on a long tradition of work in HPSG and BCG. See Flickinger et al. (1985); Pollard and Sag (1987); Flickinger (1987); Copestake (1992); Orgun (1996, 2000); Koenig (1999); Meurers (1999, 2001) and Sag et al. (2003: Ch.16), among others. The notion of lexical construction developed here is analogous to Meurers' conception of 'description-level' lexical rule.

[56] *L* variables range over lists of feature structures. These seem necessary in (49) to allow for arbitrarily long names, e.g. *John Jacob Jingleheimer Smith*, or definite descriptions that function as names, e.g. *The Lord God Almighty*. The notation '↑*type*' is used to indicate the immediate supertype(s) of the type being characterized by a given construction. This is provided purely for the reader's convenience, as the type hierarchy of the grammar signature provides a complete specification of the hierarchical relations among the types of the language.

$$
\textit{pn-lxm} \Rightarrow \begin{bmatrix} \text{FORM} & L \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{noun} & \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \textit{none} \end{bmatrix} \\ \text{VAL} & \langle\ \rangle \\ \text{MRKG} & \textit{def} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{IND} & i \\ \text{FRAMES} & \langle\ \rangle \end{bmatrix} \\ \text{CNTXT} & \begin{bmatrix} \text{BCKGRND} & \left\langle \begin{bmatrix} \textit{naming-fr} & \\ \text{ENTITY} & i \\ \text{NAME} & L \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}
$$

FIGURE 9 Proper Noun Construction (↑*invariant-lxm*)

The FS in Figure 10 must obey the type constraint sketched in Figure 9 because it is of type *pn-lxm*.

But this lexeme is not yet ready to combine with a VP to form a subject-predicate clause. Only overt expressions, signs of type *overt-expr*, may participate in phrasal constructions, as we will see. Hence a lexeme like the one in Figure 10 must give rise to a corresponding word, which is accomplished via an inflectional construction. Of course, in many other languages, this constructional 'pumping' of a lexeme to a phonologically identical word is replaced by a block of constructions that add inflectional morphology to nouns, adjectives, or other types of lexeme.

The basic intuition behind the theoretical and terminological innovations presented here (which distinguish SBCG from earlier work in Construction Grammar), is that constructions define the patterns that organize lexical classes and the patterns for building words and phrases from other signs or sequences of signs. The constructions thus form a recursive system for generating signs. Crucially, a combinatoric construction – like a rule of a Context-Free Grammar – is a static constraint that licenses a particular kind of mother-daughter configuration (i.e. a construct). An SBCG grammar, since it contains no other structure-licensing devices, provides a declarative and order-independent characterization of sign well-formedness.

For various reasons, the class of lexical items that satisfy any listeme is infinite. Strictly speaking, this is true even in the case of the listeme in (49) above, because there are infinitely many indices that could serve as the value of the feature INDEX in Figure 10, corresponding to the fact that there is no principled limit to the number of people who might be named *Kim*. How-

$$
\begin{bmatrix}
\textit{pn-lxm} & \\
\text{FORM} & \boxed{L}\langle \textit{Kim} \rangle \\
\text{ARG-ST} & \langle\,\rangle \\
\text{SYN} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \textit{noun} & \\ \text{SELECT} & \textit{none} \\ \text{CASE} & \textit{nom} \\ \text{XARG} & \textit{none} \end{bmatrix} \\
\text{VAL} & \langle\,\rangle \\
\text{MRKG} & \textit{def}
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix} \text{IND} & i \\ \text{FRAMES} & \langle\,\rangle \end{bmatrix} \\
\text{CNTXT} & \begin{bmatrix} \text{BCKGRND} & \left\langle \begin{bmatrix} \textit{naming-fr} & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \\ \text{NAME} & \boxed{L} \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

FIGURE 10 A Model of the Proper Noun *Kim*

ever, all feature structures that differ only in this way are equivalent for grammatical purposes; the only grammatically significant distinction among these functions is the value for the feature CASE.

In other circumstances, a given construction or listeme will license infinitely many FSs that are substantively different from one another. This arises whenever nonempty ARG-ST lists are involved. For example, consider the listeme in (50):

(50)
$$
\begin{bmatrix}
\textit{sintrans-v-lxm} & \\
\text{FORM} & \langle \textit{laugh} \rangle \\
\text{SEM} & [\text{FRAMES} \ \langle [\textit{laughing-fr}] \rangle\ ]
\end{bmatrix}
$$

This interacts with the lexical class constructions associated with *sintrans-v-lxm* and its supertypes to license infinitely many FSs like the one in Figure 11. It is important to see here that the NP_i included in the ARG-ST list (and the identical FS which is the XARG value; see below) must be fully specified in the FS depicted in Figure 11, even though neither the lexical entry in (50) nor any of the constraints that affect Figure 11 places any further restrictions on this FS.[57] This is as it should be, since this NP will be the

[57] Unlike some other approaches to construction-based grammar, SBCG does not use lexical

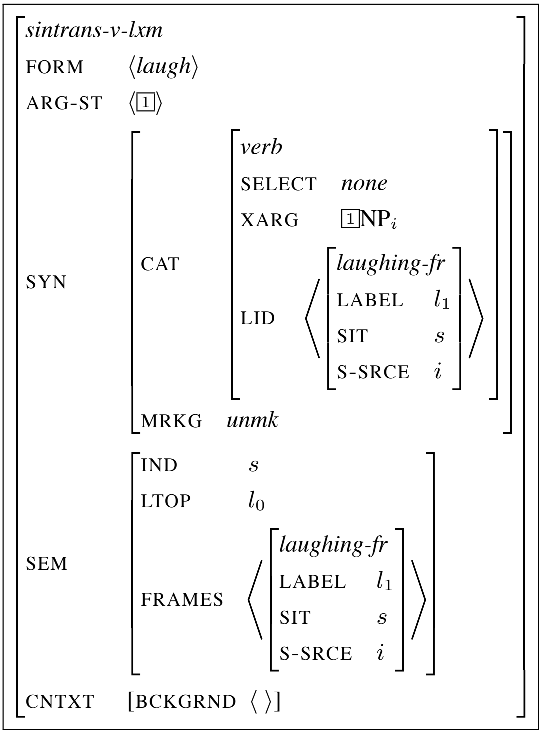


FIGURE 11 A Lexeme Licensed by Listeme (50)

subject of *laugh* and there are infinitely many NP signs that could perform that function, corresponding to infinitely many sentences of the form: 'NP *laugh/laughed/laughs*'.[58]

As noted earlier, the semantic and ARG-ST properties of lexeme classes are organized by the hierarchy of lexeme types, i.e. the subtypes of the type *lexeme*. This method is illustrated by the partial lexeme hierarchy in (51):[59]

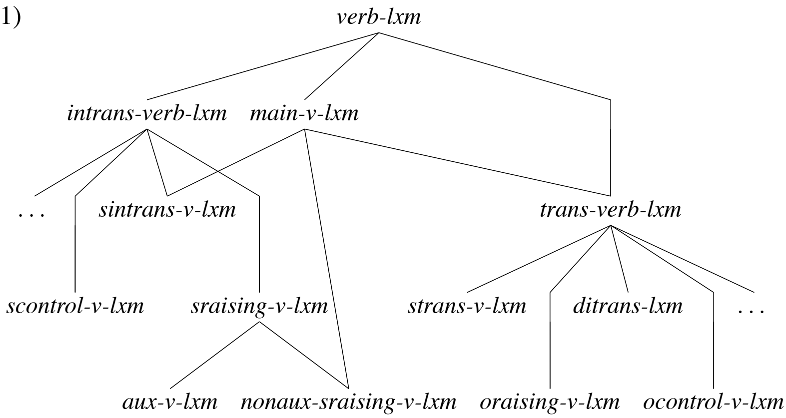
---

entries to impose 'selectional restrictions' (in the sense of Chomsky 1965). Most selectional phenomena, following McCawley's (1971, 219) sage advice, are assumed to be nonlinguistic in nature.

[58] Following the practice of the HPSG community, I use two occurrences of the same 'tag' to call attention to the fact that some grammatical constraint has identified two pieces of a feature structure. This is normally a boxed integer, e.g. [1], or (in the case of lists) a boxed capital letter, e.g. [L].

[59] Some further abbreviations:
intrans-v-lxm = intransitive-verb-lexeme
scontrol-v-lxm = subject-control-verb-lexeme
ocontrol-v-lxm = object-control-verb-lexeme
aux-v-lxm=auxiliary-verb-lexeme
sraising-v-lxm = subject-raising-verb-lexeme
oraising-v-lxm = object-raising-verb-lexeme
main-v-lxm = main-verb-lexeme

(51)



A given lexeme must obey the constraints specified in the listeme that licenses it; in addition, it must obey all the relevant lexical class constructions (those characterizing every type that the lexeme instantiates). For example, all verbal lexemes must obey the lexical class construction in (52), which ensures that a verbal lexeme is unmarked, that the first member of its ARG-ST list is its external argument, and that its LID value matches its FRAMES list, which is usually a singleton list (whose member is scope-bounded by the LTOP), but sometimes the empty list (in the case of light verbs, e.g. auxiliary *do* or *be*):

(52) **Verb Lexeme Construction (↑*lexeme*):**

$$
\textit{verb-lxm} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \langle X, \ldots \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{LID} & L \\ \text{SELECT} & \textit{none} \\ \text{XARG} & X \end{bmatrix} \\ \text{MRKG} & \textit{unmk} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{LTOP} & l_{0 =_q 1} \\ \text{FRAMES} & L : \langle ([\text{LABEL} \;\; l_1]) \rangle \end{bmatrix} \end{bmatrix}
$$

But lexical class constructions are stated at diverse levels, so as to affect, for example, all lexemes, all verb lexemes (as in (52)), all main-verb lexemes (as in (53)), all intransitive verb lexemes, or all instances of a particular maximal type of verb lexeme.

(53) **Main Verb Lexeme Construction** (↑*verb-lxm*)**:**

$$
\textit{main-v-lxm} \Rightarrow \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{AUX} & - \\ \text{INV} & - \\ \ldots & \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{IND} & s \\ \text{FRAMES} & \langle [\text{SIT} \;\; s] \rangle \end{bmatrix} \end{bmatrix}
$$

A ditransitive lexeme, for example, must simultaneously satisfy the lexical class constructions that characterize the types *ditrans-verb-lxm*, *trans-verb-lxm*, *main-v-lxm*, *verb-lxm*, and *lexeme*. A given verbal listeme can thus be streamlined, leaving it to the theory of lexical classes, as embodied in the lexical class constructions, to determine which lexical properties are required of a given type of lexeme and which properties are compatible with it.

In this approach, an underspecified listeme may sometimes be compatible with more than one maximal subtype of *lexeme*. This provides one way of analyzing argument structure variations, as we will see in the discussion of dative and locative verb alternations in section 7 below.

6.2 Morphological Functions

The next two sections discuss inflectional and derivational constructions. A key part of such word-building constructions are the **morphological functions**, which determine the morphological shape of a given lexeme or word. In a well-worked out theory of morphology like that of Stump 2001, there are both morpholexical rules and paradigm realization functions. The former map lexemes to lexemes and the latter map lexemes to the shape they exhibit in a particular paradigm slot, associated with an appropriate feature bundle. These entities in Stump's theory play roughly the same role as derivational and inflectional constructions in SBCG.[60] Thus when I speak here of 'morphological functions', I am talking more narrowly about the relation between the forms of two lexemes or the relation between the form of a lexeme and the form of a word that realizes that lexeme. The constructions in which the morphological functions are embedded will do some of the work that the corresponding entities in Stump's theory are intended to do.

Let us first consider the preterite forms of verbal lexemes. Constructs licensed by the Preterite Construction – FSs of type *preterite-construct* (*preterite-cxt*) – have a DTRS list containing exactly one FS of type *lexeme* and a mother of type *word*. The mother must include an appropriate FORM value

[60] For a more detailed consideration of how to integrate Paradigm Function Morphology and constraint-based frameworks like HPSG/SBCG, see Bonami and Samvelian submitted.

and also the additional semantic bits corresponding to the meaning of the preterite word. The form of the mother is the image of the daughter's form under the morphological function $\mathbf{F}_{pret}$; the semantics of the mother situates the lexeme's situation argument in past time.

Morphological functions allow us to model 'elsewhere' phenomena in SBCG morphology. In addition, they provide a way of dealing with other problems posed by various kinds of irregularity. The FORM value of a given lexeme is a singleton list containing the lexical formative associated with that lexeme, which we will assume is a stem.[61]

The domain of an inflectional realization function is the set of stems and its range is the set of inflected forms, including those constructed via affixation. Note that the stems must be distinguished in some fashion in order to license inflectional variation among homophonous stems. For example, while the preterite form of *have* is *had* for all of the semantically distinct *have* lexemes, the preterite form of *lie* is *lay* if the lexeme's semantics is 'recline', and *lied* if it is 'prevaricate'. We will therefore follow Stump (2001) in positing distinct stems lie_1 and lie_2, each specified as the FORM value of the appropriate listeme.

$\mathbf{F}_{pret}$ can be defined along the following lines:

(54)

Stem	$\mathbf{F}_{pret}$(Stem)
be	**undefined**
have	*had*
lie_1	*lay*
swim	*swam*
buy	*bought*
keep	*kept*
...	
otherwise	
x	x-*ed*

We will need special specifications for the preterite forms of *be* (*was* for non-second person singular and *were* for 2nd person or plural). Note that this is properly analyzed in terms of constructions, as there are multiple *be*-listemes (in order to accommodate different uses of *be*, including a large number of multiword expressions) and each of these has all the same preterite forms. Similarly, all *have*-listemes show the same irregular preterite realizations, i.e. *had*, irrespective of meaning.

[61]For some languages, morphological functions must effect stem alternations as necessary (e.g. French *va-/all-* 'go'; Sanskrit *gam-/gacch-* 'go'). However, it is arguably the case that no English lexeme requires a muliple stem analysis. For an interesting discussion and a precise, compatible proposal for treating multiple stems in French morphology, see Bonami and Boyé 2006.

The lexeme *lie*$_1$ 'recline' is realized as *lay* by the third line of the function definition in (54) and *lie*$_2$ 'prevaricate' is realized as *lied* by the 'otherwise' clause. *Swam* and *bought* are unproblematic irregular forms without doublets.[62]

6.3 Inflectional Constructions

Lexical constructs are constrained by the following type declaration, specified in the grammar's signature:

(55) *lex-cxt*: [DTRS *list*(*lex-sign*)]
(The daughters of a lexical construct are all of type *lex-sign*, i.e. they are words or lexemes.)

In addition to the general constraints on lexical constructs, inflectional constructs have more specific properties that are also specified as part of the grammar signature:

(56) $\textit{infl-cxt}: \begin{bmatrix} \text{MTR} & \textit{word} \\ \text{DTRS} & \textit{list(lexeme)} \end{bmatrix}$
(The mother of an inflectional construct is of type *word*; the daughters must be lexemes.[63])

This treatment embodies the traditional intuition that inflectional constructions are resources for building words from lexemes. (Recall that affixes are not signs in the analyses presented here.)

An inflected word like *laughed* is modeled as a FS of type *word*, built in accordance with the Preterite Construction, sketched in (57):

[62] Non-past-tense uses of the preterite morphological form, such as the antecedent clause of a counterfactual conditional (*If I had my way,...*) could in principle be licensed by a separate inflectional construction that also avails itself of $\mathbf{F}_{pret}$. Alternatively, one might explore a semantics-changing, unary post-inflectional construction whose mother and daughter do not differ in FORM. In either case, special arrangements must be made to license and distinguish these two entities (counterfactual *had* and its past time homophone). I will not resolve these matters here.

[63] In an English inflectional construct, there is always a single daughter, but this restriction may not apply to all languages. The requirement that the daughters must be lexemes may not be universal, either. Languages with more 'layered' morphology may warrant analysis in terms of constructions that add inflections to already inflected words. Alternatively such languages might be treated in terms of a distinction among morphological types that is more fine-grained than the standard derivational/inflectional dichotomy that I have assumed here. See Orgun 1996, Miller and Sag 1997, Ackerman and Webelhuth 1998, and Malouf 2000 for some relevant discussion.

(57) **Preterite Construction** (↑*infl-cxt*):

$$
\textit{preterite-cxt} \Rightarrow
\left[\begin{array}{ll}
\text{MTR} & \left[\begin{array}{ll}
\text{FORM} & \langle \mathbf{F}_{pret}(X) \rangle \\
\text{SYN} & Y : [\text{CAT } [\text{VF } \textit{fin}]] \\
\text{SEM} & \left[\begin{array}{ll}
\text{IND} & s \\
\text{LTOP} & l_0 \\
\text{FRAMES} & \left\langle \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{some-fr}} \\ \text{LBL} & l_2 \\ \text{BV} & s \\ \text{RESTR} & l_1 \end{array}\right], \left[\begin{array}{l} \textit{past-fr} \\ \text{LBL } l_1 \\ \text{ARG } s \end{array}\right] \right\rangle \oplus L
\end{array}\right]
\end{array}\right] \\
\text{DTRS} & \left\langle \left[\begin{array}{ll}
\text{FORM} & \langle X \rangle \\
\text{ARG-ST} & \langle \text{NP}[\textit{nom}] , \ldots \rangle \\
\text{SYN} & Y \\
\text{SEM} & \left[\begin{array}{ll} \text{IND} & s \\ \text{LTOP} & l_0 \\ \text{FRAMES} & L \end{array}\right]
\end{array}\right] \right\rangle
\end{array}\right]
$$

This construction requires explanation. First, the variables *X*, *Y*, and *Z* range over feature structures, while *L*-variables range over lists of feature structures, as before. The notation '[FEATURE $Y : D$]' indicates that the value of FEATURE must satisfy the description *D* and that it is being tagged as Y for the purposes of identifying the value of FEATURE with some feature structure elsewhere in the FS being described. Thus the mother's SYN value in (57) must satisfy the constraint specified after the colon – its VF value must be *finite*. The daughter's SYN value in (57) (since it is also specified as Y) must be just like the SYN value of the mother. In addition, the sole member of the mother's FORM list must be the image under $\mathbf{F}_{pret}$ of the daughter's FORM value X. Also, the mother's SEM value must differ from that of the daughter in the way indicated in (57). Note finally that the LTOP value of the mother is identified with that of the daughter but that this construction imposes no constraint on the scope of the existential quantifier introduced on the mother's FRAMES list.

One way of paraphrasing (57) – albeit more procedurally – is as follows: Given a verbal lexeme, one can construct a verbal word meeting the following four conditions:

(58) a. the word's VF value is *finite*, as is the VF value of the lexeme,

b. the word's FORM value is related to that of the lexeme via the morphological function $\mathbf{F}_{pret}$,

c. the word's SYN and ARG-ST values are identified with those of the lexeme daughter, thus requiring that everything in the listeme that licensed the lexeme be consistent with the constraints introduced by this construction, e.g. requiring the subject's CASE value be nominative, and

d. the word's FRAMES list includes that of the lexeme, but it also includes an existential quantifier binding a situation index (the 'bound variable' (BV)) restricted to past time. This index is also identified with the situation index specified in the lexeme's frame, and thus functions semantically to spatio-temporally situate the event described by the verb.[64]

This construction therefore licenses constructs like the one shown in Figure 12; and because this is a well-formed construct, the FS in (59) (the mother of the construct in Figure 12) is constructionally licensed.

(59)
$$\begin{bmatrix} \textit{word} & \\ \text{FORM} & \langle \textit{laughed} \rangle \\ \text{ARG-ST} & \langle \boxed{1} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{VF} & \textit{fin} \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \boxed{1} \\ \text{LID} & \langle \boxed{2} \rangle \end{bmatrix} \\ \text{VAL} & \langle \boxed{1}\text{NP}_i[\textit{nom}] \rangle \\ \text{MRKG} & \textit{unmk} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{IND} & s \\ \text{LTOP} & l_0 \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LBL} & l_0 \\ \text{BV} & s \\ \text{RESTR} & l_1 \\ \text{SCOPE} & l_2 \end{bmatrix}, \begin{bmatrix} \textit{past-fr} & \\ \text{LBL} & l_1 \\ \text{ARG} & s \end{bmatrix}, \boxed{2}\begin{bmatrix} \textit{laughing-fr} & \\ \text{LBL} & l_2 \\ \text{S-SRCE} & i \\ \text{SIT} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

[64] The lack of further constraints on the situation-binding quantifier correctly allows for scope ambiguities in examples like *Every man laughed*: 'Every man laughed at some time in the past' or 'at some time in the past, every man laughed':

(i) $l_0{:}\textit{every}(i, l_1, l_2), l_1{:}\textit{man}(i), l_2{:}\textit{some}(s, l_3, l_4), l_3{:}\textit{past}(s), l_4{:}\textit{laugh}(s, i)$
'For every man i, there is a past situation where i was laughing'

(ii) $l_2{:}\textit{some}(s, l_3, l_0), l_3{:}\textit{past}(s), l_0{:}\textit{every}(i, l_1, l_4), l_1{:}\textit{man}(i), l_4{:}\textit{laugh}(s, i)$
'There is a past situation where every man was laughing'

For details, see Copestake et al. 2005.

$$
\begin{bmatrix}
\textit{preterite-cxt} & \\
\text{MTR} & \begin{bmatrix}
\textit{word} & \\
\text{FORM} & \langle \textit{laughed} \rangle \\
\text{ARG-ST} & \boxed{L} \\
\text{SYN} & \boxed{1} \\
\text{SEM} & \begin{bmatrix}
\text{IND} & s \\
\text{LTOP} & l_0 \\
\text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LBL} & l_0 \\ \text{BV} & s \\ \text{RESTR} & l_1 \\ \text{SCOPE} & l_2 \end{bmatrix}, \begin{bmatrix} \textit{past-fr} & \\ \text{LBL} & l_1 \\ \text{ARG} & s \end{bmatrix}, \boxed{2} \right\rangle
\end{bmatrix}
\end{bmatrix} \\
\text{DTRS} & \left\langle \begin{bmatrix}
\textit{sintrans-v-lxm} & \\
\text{FORM} & \langle \textit{laugh} \rangle \\
\text{ARG-ST} & \boxed{L} \langle \text{NP}_i[\textit{nom}] \rangle \\
\text{SYN} & \boxed{1} \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{VF} & \textit{fin} \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \text{NP}_i[\textit{nom}] \\ \text{LID} & \langle \boxed{2} \rangle \end{bmatrix} \\
\text{MRKG} & \textit{unmk} \\
\text{VAL} & \boxed{L}
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix}
\text{IND} & s \\
\text{LTOP} & l_0 \\
\text{FRAMES} & \left\langle \boxed{2} \begin{bmatrix} \textit{laughing-fr} & \\ \text{LBL} & l_2 \\ \text{S-SRCE} & i \\ \text{SIT} & s \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix} \right\rangle
\end{bmatrix}
$$

FIGURE 12 A FS of Type *preterite-construct*

Notice that it would be redundant for the construction in (57) to explicitly require that the MTR value be of type *word* or that the daughter be of type *lexeme*. Because the constructs licensed by the Preterite Construction are all instances of the type *preterite-cxt*, they must obey all constraints the grammar imposes on FSs of that type and its supertypes, including the supertype *infl-cxt*. As we have already seen, (56) requires that the MTR value of all inflectional constructs be of type *word*, hence the MTR value of any preterite construct is already guaranteed to be of this type. It also follows that the 'output' of this construction (a word) can never 'undergo' the construction again (since the daughter of any inflectional construction must be of type *lexeme*). I make the further assumption that in all English inflectional constructions, the mother and daughter share ARG-ST and CNTXT values (see the appendix); hence the construct in Figure 12 obeys these constraints, as well. Finally, the information encoded in Figure 12 is identical to what is represented in a more familiar diagram – the unary local tree in Figure 13. Because of their familiarity, I will use tree diagrams whenever possible to illustrate constructs.

Finally, as a trivial example of an inflectional construction, let us consider the Zero Inflection Construction, which pumps noninflecting lexemes to word status. This construction takes a lexeme as daughter, licensing mothers that are of type *word*, but otherwise identical to their daughter:

(60) **Zero Inflection Construction** ($\uparrow$*infl-cxt*)**:**

$$\textit{zero-infl-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & X \,!\, \textit{word} \\ \text{DTRS} & \langle X : \textit{invariant-lxm} \rangle \end{bmatrix}$$

This analysis assumes that *invariant-lxm* is a supertype of *pn-lxm* and the other lexeme types we posit for the analysis of noninflecting words, e.g. prepositions, adverbs, and conjunctions.

6.4 Derivational Constructions

Derivational constructions are structured as shown in (61):

(61)

$$\textit{deriv-cxt}: \begin{bmatrix} \text{MTR} & \textit{lexeme} \\ \text{DTRS} & \textit{list(lex-sign)} \end{bmatrix}$$

(The mother of a derivational construct is of type *lexeme*; the daughters of a derivational construct are lexical signs (words or lexemes).)

Constructions of this type allow new lexemes to be built from one or more lexical signs. For example, there is an *un*-prefixation construction, sketched in (62), which allows *un*-verb lexemes to be derived from a specifiable class of verb lexemes:[65]

[65] The morphological function $\mathbf{F}_{un}$ is utilized by more than one construction (see below).

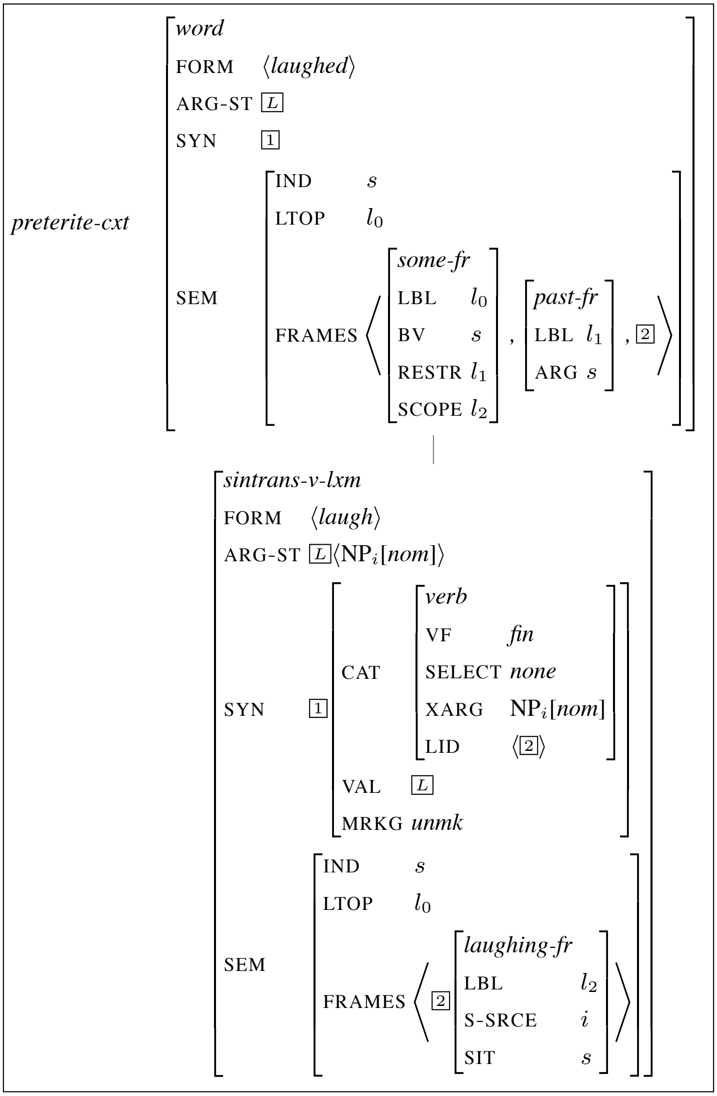


FIGURE 13 A FS of Type *preterite-construct* in Tree Notation

(62) ***Un*-Verb Construction** (↑*deriv-cxt*)**:**

$$
\textit{un-verb-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \textit{derived-trans-v-lxm} & \\ \text{FORM} & \langle \mathbf{F}_{un}(X) \rangle \\ \text{ARG-ST} & L_1 \\ \text{SYN} & Y \\ \text{SEM} & [\text{FRAMES } L_2 \oplus \ldots] \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \textit{strans-v-lxm} & \\ \text{FORM} & \langle X \rangle \\ \text{ARG-ST} & L_1 \\ \text{SYN} & Y \\ \text{SEM} & [\text{FRAMES } L_2] \end{bmatrix} \right\rangle \end{bmatrix}
$$

The formulation in (62) presupposes that only strict-transitive verbal lexemes (lexemes of type *strans-v-lxm*) can give rise to *un*-verb lexemes. However, the restrictions on this construction are partly semantic. For example, the impossibility of *uncrush* and *unlift* may be attributed to the fact that the events denoted by *crush* and *lift* are 'irreversible'. Moreover, there are various intransitive verbs that give rise to *un*-verbs, e.g. *roll* ↦ *unroll*, *congeal* ↦ *uncongeal*, via what may be the same derivational construction. Likewise, there are nonce occurrences of other intransitive verbs with *un-* (e.g. *unstink*, *unwaffle*, *unburn*, *unshine*). Hence, there is room for disagreement about what constraints govern *un*-verbs and whether their nature is syntactic, semantic, or some combination of the two. There is corresponding room for minor modifications of (62) that will accommodate such ideas.

Since inflectional constructs are required to have a daughter of type *lexeme*, a natural relation exists between the two types of construction: derivational constructions feed inflectional constructions. That is, a derived lexeme, one that is the mother of a construct licensed by some derivational construction, can then serve as the daughter of a construct licensed by an inflectional construction. Derivational constructions can also feed other derivational constructions (as in Figure 14 below) and inflectional constructions can sometimes feed derivational constructions, e.g. in noun-noun compounds such as *grants secretary* or *weapons specialist*, where the modifying noun bears plural inflection.[66]

[66] Thus on empirical grounds, it seems prudent not to build any strong theory of 'level ordering' (Kiparsky 1982) into grammatical theory, though it would be easy enough to incorporate this idea into SBCG, simply by tightening the condition on the value of DTRS in (61) above. See note 70.

Sag et al. (this volume) discuss a number of issues that have been raised in the CxG literature about the analysis of lexical constructions. The debate that is of relevance here concerns lexical phenomena that appear to allow a 'one-level' analysis, such as resultatives and passives. The one-level analysis, advanced by Goldberg (1995, 2003, 2006) and others, involves an underspecifie listeme that then can simply 'unify' with one or more constructions to produce a well-formed sentence. These constructions determine the number of dependent elements a verb combines with and how they (both the verb and the dependents, in the case of passive) are realized. Thus the one-level analysis of 'directed-motion' uses of an intransitive verb like *sneeze*, shown in (63), involves unifying the basic listeme for *sneeze* with a construction that specifies that it combines with a direct object NP and a directional phrase:

(63) Sandy sneezed the tissue off the table.

The alternative analysis advocated by Müller and discussed briefly in section 7.4 below, involves a two-step analysis. First the intransitive verb lexeme *sneeze* is licensed with the singleton ARG-ST list characteristic of all strictly intransitive verbs. And from this lexeme, another can be built – via a lexical construction – and this constructed lexeme will have a longer ARG-ST list that includes an object NP and a directional phrase.[67] Since the result of this two-level analysis is a lexeme, it can then be inflected either as an active verbal word or as a passive word, with differing ARG-ST lists that will determine the verbs' differing combinatoric potentials. Both active and passive verbs project VPs via the Predicational Head-Complement Construction discussed in section 8.3 below.

Further evidence that multiple levels of structure may be required in order to make sense of constructional interaction (as also noted by Müller (2006)) comes from the ambiguity of expressions like *unlockable*. In addition to the analysis sketched in Figure 14, there is another, shown in Figure 15, where the *Able*-Adjective Construction creates a derived lexeme *lockable* from *lock*. This adjective may then serve as the daughter of a construct licensed by the *Un*-Adjective Construction, resulting in another adjective, as shown. The meanings of the two words are structured in accordance with this difference in analysis as well, as indicated by the scopal abbreviations summarized in (64):

(64) a. **able**(**un**(**lock**)) [Figure 14]
b. **un**(**able**(**lock**)) [Figure 15]

Thus, simply allowing the *Able*-Adjective Construction to unify with either the *Un*-Adjective Construction or the *Un*-Verb Construction fails to account

[67] For a similar analysis, see Croft (2003), who also offers a useful discussion of the role of lexical idiosyncrasy in the one-level/two-level debate.

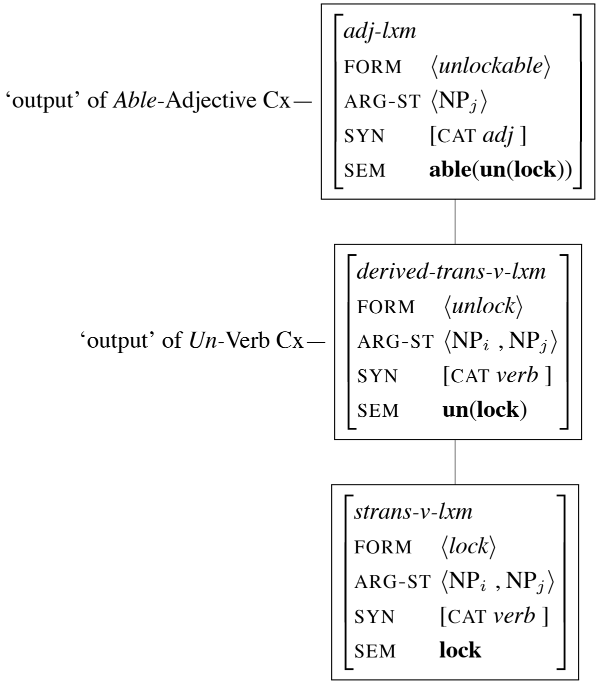


FIGURE 14 The *Un*-Verb Construction Feeding the *Able*-Adjective Construction

for the observed interpretational difference. In fact, it is hard to see how any variant of the one-level, 'construction unification' approach is going to provide an account of facts such as these.

In addition to the constructions just discussed, it is reasonable to assume that all of the following phenomena should be analyzed in terms of derivational constructions: passivization (which feeds overt inflectional constructions in many languages), word-formation processes like adjectivalization in English (see Bresnan 2001), denominal verb formation ($porch_{noun} \mapsto porch_{verb}$; see Clark and Clark 1979), agentive noun formation (*drive* $\mapsto$ *driver*), and various other kinds of nominalization. An example of a binary derivational construction is English noun-noun compounding. By specifying the DTRS value of a *deriv-cxt* to be a list of lexical signs, members of compounds are permitted to be inflected words, as well as lexemes, subject to the particular constraints of individual derivational constructions.

The general compounding construction, which appeals to a contextually salient (but otherwise arbitrary) property to relate the interpretations of two

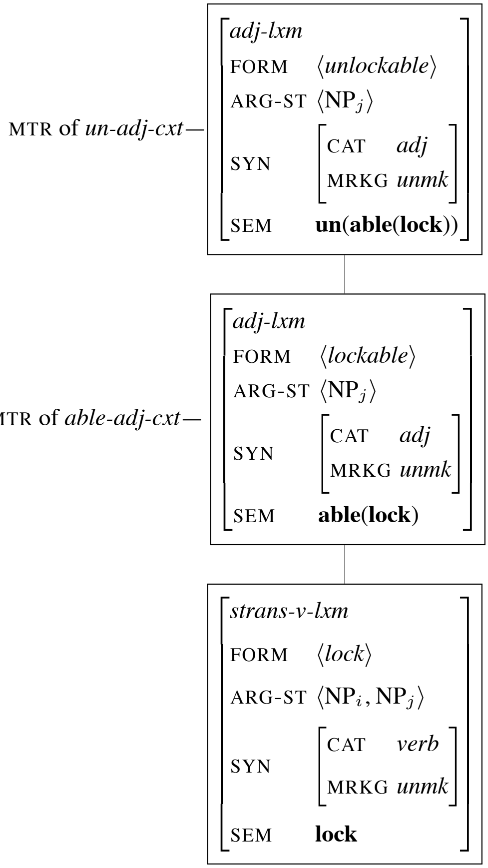


FIGURE 15 The *Able*-Adjective Construction Feeding the *Un*-Adjective Construction

nouns, accounts for compounds like the following:[68]

(65) a. *pumpkin bus*: ‘bus that was used in some previous excursion to a pumpkin patch familiar to the relevant interlocutors’ (Downing 1977)

b. *Jaeger potato*: ‘potato of the kind that the speaker once used for something when spending an evening with someone named Jaeger’

c. *Beatles fan*, *women friends*, *people smugglers*, *pubs inspector*, *munitions dump*

[68] See Kay and Zimmer 1978, Downing 1977, Levi 1978, and Copestake and Lascarides 1997.

Examples (65a) and (65b) illustrate attested innovative compounds. The examples in (65c) are also attested and exemplify some of the subtypes of noun-noun compounds exhibiting internal inflection.[69]

It is also possible to incorporate proposals like that of Copestake and Lascarides (1997) (extending the basic intuition of Levi 1978), who posit a number of more specific constructions that specify patterns fitting particular classes of nouns together in conventionalized ways. Such explorations are beyond the scope of this chapter.[70]

6.5 Postinflectional Constructions

Postinflectional constructs are structured as follows:

(66) *pinfl-cxt*: $\begin{bmatrix} \textsc{mtr} & \textit{word} \\ \textsc{dtrs} & \textit{list(word)} \end{bmatrix}$

(The mother and daughters of a postinflectional construct are of type *word*.)

Postinflectional constructions thus allow for words to be derived from other words. Sag et al. (2003) introduce this category as a way of incorporating a number of proposals that have been made (e.g. by Warner (1993b), Bouma and van Noord (1994), Kim (2000), and Kim and Sag (2002)) to use lexical rules for the purpose of creating adverb-selecting auxiliary verbs (e.g. a variant of *will* that must combine with *not*), as well as *not*-contracted words (*didn't*, *couldn't*, and the like).

The analysis of finite negation thus involves imparting to finite forms of auxiliary verbs the ability to take the adverb *not* as a complement. The general construction that accomplishes this, shown in Figure 16, involves the postlexical construct type *negative-aux-construct* (*neg-aux-cxt*). Here the notation '[$\textsc{f}_1$ X ! [$\textsc{f}_2$ D]]' means (1) that $\textsc{f}_1$'s value must be identical to the feature structure tagged as X elsewhere in the description, except with respect to the value of the feature $\textsc{f}_2$, and (2) that $\textsc{f}_2$'s value must satisfy the description D.[71] In (66), this means that the mother's SYN value must be identical to that of the daughter, except with respect to the features CAT and VAL. In addition,

[69] See Bauer and Reynaud 2001.

[70] As noted, the first member of most noun-noun compounds is a lexeme (*computer screen*, *pumpkin bus*, etc.), but in certain types of compounds, the first element is a word: *algorithms course*, *sales tax*, etc. Pinker (1999) and Kiparsky (1982) argue that the first member of a noun-noun compound cannot be a word. Bauer and Reynaud (2001), in a detailed corpus study, discuss the circumstances under which it is likely to be one. For further critical discussion and experimental evidence that there are multiple factors determining the well-formedness of plural nouns in noun-noun compounds (rather than a grammatical constraint blocking any such combination), see Harrison 2001 and Haskell et al. 2003.

[71] It is sometimes useful to restrict feature identity so as to simply exclude values of particular features. Thus 'X ! [F]' is equivalent to 'X ! [F [ ]]', where no constraint is imposed on the value of F. More generally, (i) means that the values of the features A and B are identical except for the

$$
\textit{neg-aux-cxt} \Rightarrow \left[\begin{array}{ll}
\text{MTR} & \left[\begin{array}{ll}
\text{FORM} & \langle W \rangle \\
\text{SYN} & X \,!\, \left[\begin{array}{ll}
\text{CAT} & Y \,!\, [\text{AUX} \ -] \\
\text{VAL} & \langle \text{Adv}[\textit{neg}] \rangle \oplus L
\end{array}\right] \\
\text{SEM} & Z
\end{array}\right] \\
\text{DTRS} & \left\langle \left[\begin{array}{ll}
\text{FORM} & \langle W \rangle \\
\text{SYN} & X : \left[\begin{array}{ll}
\text{CAT} & Y : \left[\begin{array}{ll}
\text{AUX} & + \\
\text{VF} & \textit{fin}
\end{array}\right] \\
\text{VAL} & L
\end{array}\right] \\
\text{SEM} & Z
\end{array}\right] \right\rangle
\end{array}\right]
$$

FIGURE 16 Negative Auxiliary Construction (↑*post-infl-cxt*)

the mother's CAT and VAL values must differ from those of the daughter in the ways indicated in (66) and illustrated in Figure 17. Nonauxiliary verb lexemes are instances of the type *main-v-lxm* (see (53) above) and are always lexically specified as [AUX −]. And because (66) requires the daughter to be [AUX +], words derived from nonauxiliary lexemes can never acquire the potential to select a negative adverbial as complement. Once an auxiliary verb acquires the adverb-augmented valence, it is specified as [AUX −], and hence is ineligible to be the daughter of another negative-auxiliary construct, thus ruling out double finite negation. The auxiliary verb illustrated as the mother in Figure 17 can combine with its two complements in accordance with the Predicative Head-Complement Construction (discussed in section 8.3 below) to license verb phrases like the one in (67):

(67) Kim { ^{VP} [*will*] [*not*] [*sign the letter*] }

For discussion of semantic aspects of this construction, see Sag to appear.[72]

values of the features F_1,. . . ,F_n, whose values are free to differ as they may. The difference may also be one of types, as in (ii):

$$
\text{(i)} \left[\begin{array}{ll}
\text{A} & X \,!\, \left[\begin{array}{l} \text{F}_1 \\ \ldots \\ \text{F}_n \end{array}\right] \\
\text{B} & X
\end{array}\right]
\qquad
\text{(ii)} \left[\begin{array}{ll}
\text{A} & X \,!\, \textit{type}_1 \\
\text{B} & X : \textit{type}_2
\end{array}\right]
$$

[72] It should be noted that it can be difficult to discern the differing consequences of a postinflectional analysis and a derivational one. Sometimes the issue is decided by the feeding relations between the construction in question and other derivational constructions. For example, a word licensed by a postinflectional construction cannot usually serve as the daughter of a derivational construct because most derivational constructions require a daughter of type *lexeme*. Hence, treating a given alternation via a postinflectional construction ensures that the result cannot feed most derivational constructions.

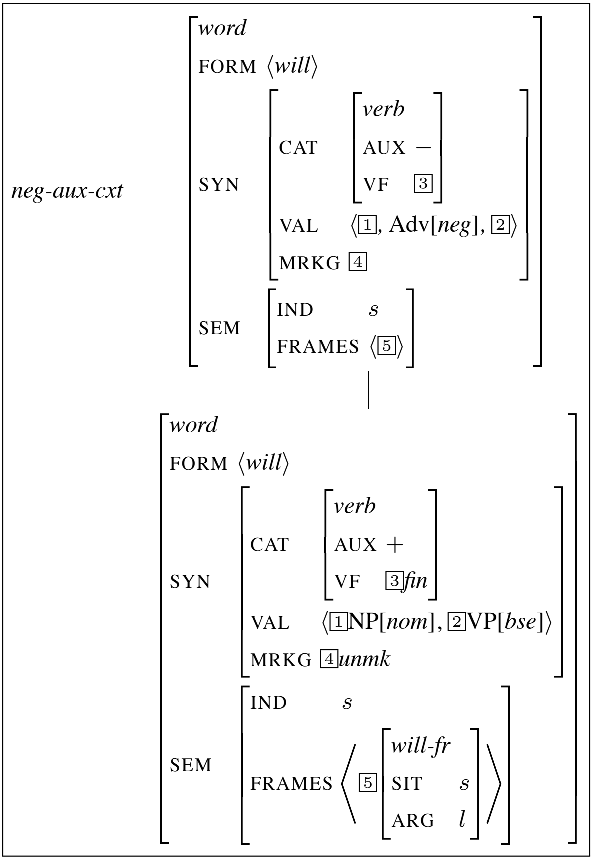


FIGURE 17 Constructing a *Not*-Selecting Auxiliary Verb

7 Some Expressions of Interest

7.1 Multiword Expressions

Multiword expressions (MWEs) are a diverse lot and do not have a unified analysis, contrary to what has often been assumed in the generative literature. Consider first the inert MWEs that Bauer (1983) refers to as 'fixed expressions', including: *by and large*, *in short*, *kingdom come*, *every which way*, *ad hoc*, *jack of all trades*, *Palo Alto*, *Newcastle Upon Tyne*, etc. These 'words with spaces' can be accommodated simply by positing listemes which specify the appropriate lexical type and whose FORM value is a non-singleton list. Since these expressions are fully lexicalized (**in shorter*, **in very short*), their treatment need be no more complicated than this.[73]

Semantically decomposable idioms (Nunberg et al. 1994, Sag et al. 2002) present a greater challenge. These are MWEs where the idiomatic meaning is distributed throughout the meaning of the subexpressions, including *spill the beans*, *keep tabs on*, *pull strings*, and many others. As Nunberg et al. observe, only semantically decomposable idioms are syntactically flexible (in English) and only semantically decomposable idioms allow internal quantification and modification:

(68) **Syntactic Flexibility**:

a. Strings had been pulled to get Sandy the job.

b. It was the close tabs they kept on our parents that upset us most.

Internal Quantifiability:

c. The FBI kept closer tabs on Kim than they kept on Sandy.

d. They took more advantage of the situation than they should have.

Internal Modifiability:

e. Many Californians jumped on the bandwagon that Perot had started.

f. She spilled the beans that cost them the contract.

These observations lead directly to the view that MWEs in this class have internal combinatorics – and this is inconsistent with the simple treatment of fixed expressions just discussed. Instead, in all the cases illustrated in (68), there must be a listeme for the verbal head and separate listemes for other appropriate parts of the dependents that the head combines with, in accordance with independently motivated grammatical constructions. For example, we might treat the MWE *pull strings* via two listemes: an idiomatic *pull* whose meaning is 'manipulate' and an idiomatic *strings* whose meaning is 'connections'. The frames required for such an analysis, presumably grounded in a metaphorical relation between situation types, will be indicated as $\textit{pulling}^{\textit{manipulating}}$*-fr* and $\textit{i-strings}^{\textit{connections}}$*-fr*, respectively. The listemes in question are sketched in (69):[74]

[73]For further discussion, see Sag et al. 2002 and Kay and Sag 2012.

[74]I will henceforth also use sign abbreviations like the following to suppress short pathnames involving 'SYN', 'SEM', 'SYN [CAT', and the like:

$$\begin{bmatrix}\text{NP}_j \\ \text{LID} \ \langle \textit{i-strings}^{\textit{connections}}\textit{-fr}\rangle\end{bmatrix} = \text{NP}_i \ \& \ [\text{SYN} \ [\text{CAT} \ [\text{LID} \ \langle \textit{i-strings}^{\textit{connections}}\textit{-fr}\rangle]]]$$

$$\begin{bmatrix}\text{NP} \\ \text{MRKG} \ \textit{unmk}\end{bmatrix} = \text{NP} \ \& \ [\text{SYN} \ [\text{MRKG} \ \textit{unmk}]]$$

(69)

$$\begin{bmatrix} \textit{strans-v-lxm} & \\ \text{FORM} & \langle \textit{pull} \rangle \\ \text{ARG-ST} & \left\langle X_i, \begin{bmatrix} \text{NP}_j \\ \text{LID} \ \langle \textit{i-strings}^{connections}\textit{-fr} \rangle \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{pulling}^{manipulating}\textit{-fr} & \\ \text{AGENT} & i \\ \text{ENTITY} & j \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

(70)

$$\begin{bmatrix} \textit{cn-lxm} & \\ \text{FORM} & \langle \textit{strings} \rangle \\ \text{SEM} & \begin{bmatrix} \text{IND} & i \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{i-strings}^{connections}\textit{-fr} & \\ \text{ENTITY} & i \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The listeme in (69) licenses a strict-transitive verb lexeme (*strans-v-lxm*; see (51) above) whose second ARG-ST member (the direct object NP) must be specified as [LID $\langle$*i-strings*^{connections}*-fr*$\rangle$].[75] The only NPs compatible with this specification (given that in general a common noun also uses its meaning as its LID value) are those headed by the idiomatic noun *strings*, whose listeme is shown in (70). In this way, the idiomatic interpretation is decomposed into an idiomatic pulling frame (*pulling*^{manipulating}*-fr*) involving two participants, an AGENT and an entity, where the latter has idiomatic strings (*i-strings*^{connections}*-fr*) predicated of it. Because the idiomatic meaning is distributed over the parts, it is possible to modify or quantify these components of the idioms using the very same constructions that are responsible for the modification and quantification of nonidiomatic expressions.

The lexemes licensed by the listemes in (69)–(70) will give rise to inflected words (*pulled*, *strings*) and phrases that contain those words, such as those in (71):

(71) a. They're pulling strings to get you the job.
b. We have pulled strings more than once.
c. We pulled strings to get invited.

---

[75]Note that nothing more need be specified in a listeme like (69), since any restrictions on the syntactic category of relevant ARG-ST members will be specified by the appropriate lexical class construction (e.g. the Strict-Transitive Verb Lexeme Construction).

If, following Pollard and Sag 1994, Ch. 2, English agreement features are part of referential indices, then the index i in (70) should be replaced by i : [NUM *pl*].

d. He pulls strings whenever he can.

Moreover, this treatment, which makes crucial use of independent constructions of the language, allows morphological and syntactic flexibility, as well as internal modification and quantification. For example, the Passive Construction (a derivational construction) will give rise to a passive lexeme licensing the passive analogue of (69), and this in turn will interact with the rest of the grammar to provide an analysis of examples like *Strings had been pulled to get Sandy the job*. Finally, notice that what emerges is an intuitively satisfying characterization of the correlation between flexibility and modifiability on the one hand and semantic decomposability and lexicalization on the other: the more 'decomposable' the meaning of a MWE is, the more likely it is for speakers to analyze it in terms of the general principles for composing complex expressions.

This analysis, further developed in Kay and Sag 2012, specifies that frames are classified as idiomatic frames (via subtypes of *i-frame*) or canonical frames (via subtypes of *c-frame*), with idiomatic predicators that project decomposable MWEs (e.g. $pulling^{manipulating}$*-fr*, $spilling^{revealing}$*-fr*) being classified as c-frames. The *i-frame* analysis is motivated by the basic fact that an idiomatic argument (e.g. *strings* in its idiomatic sense) can only appear together with the right governor (e.g. *pull* in its appropriate idiomatic sense). Hence, the reason why the examples in (72) only allow a nonidiomatic interpretation (and are therefore hard to contextualize) is that the listemes for the verbs in these sentences select arguments that are lexically identified in terms of a particular *c-frame*:

(72) a. Leslie found the strings that got Pat the job.
b. We resented their tabs.
c. The beans impressed us.

The motivation for classifying idiomatic predicators as *c-frame*s is this: in spite of their idiomatic meanings, they project phrases (typically VPs, Ss, or PPs) that freely appear in nonidiomatic environments. That is, their distribution shows none of the restrictions that idiomatic arguments must obey, as illustrated in (73):

(73) a. I think [Kim spilled the beans].
b. They tried to [pull strings to get Lee the job].
c. [With [my kids [keeping tabs on the stock market]]], I can finally think of retiring.
d. [Taking care of homeless animals] is rewarding.

Each bracketed expression in (73) appears in a syntactic environment where nonidiomatic expressions freely occur.

MWEs exhibit varying degrees of morphological and syntactic flexibility. For example, *kick the bucket* 'die', which Bauer classifies as a 'semi-fixed expression', allows inflectional variants of *kick* (*kicked*, *kicking*, *kicks*,...), but otherwise exhibits none of the flexibility and modifiability just illustrated for decomposable idioms. Unlike the sentences in (68), all of the following examples allow only nonidiomatic interpretations:

(74) a. The bucket had been kicked many times in that community.
b. It was the bucket(s) that they had kicked that upset us most.
c. Cancer will cause Europeans to kick fewer buckets this year than last.
d. Fewer Californians are kicking the cancer bucket each year.

Kick the bucket may be analyzed in terms of the following two listemes:

(75)
$$\begin{bmatrix} \textit{pseudo-trans-v-lxm} & \\ \text{FORM} & \langle \textit{kick} \rangle \\ \text{ARG-ST} & \left\langle X_i, \begin{bmatrix} \text{NP} & \\ \text{LID} & \langle \textit{i-bucket-fr} \rangle \\ \text{MRKG} & \textit{i-the} \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{kicking}^{dying}\textit{-fr} & \\ \text{PROTAGONIST} & i \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

(76)
$$\begin{bmatrix} \textit{cn-lxm} & \\ \text{FORM} & \langle \textit{bucket} \rangle \\ \text{SYN} & [\text{CAT} \ [\text{LID} \ \langle \textit{i-bucket-fr} \rangle]] \\ \text{SEM} & \begin{bmatrix} \text{IND} & \textit{none} \\ \text{FRAMES} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$$

Although a lexeme licensed by (76) has ⟨*i-bucket-fr*⟩ as its LID value, it has no frame on its FRAMES list and an IND value of *none*.[76] Hence *the bucket* contributes nothing to the semantic composition of the sentence and provides nothing for a modifier to modify or for a quantifier to restrict, predicting the absence of idiomatic readings for (74b-d). The failure of idiomatic *kick* to passivize (cf. (74a)) is accounted for because (75) assigns all the lexemes it licenses to the type *pseudo-transitive-verb-lexeme*, which is not a subtype of *transitive-verb-lexeme*, the type that limits the domain of passivization; refinements are of course possible. Note that although the idiomatic *bucket*

[76]For a discussion of LID values in relation to a lexeme's meaning, see Kay and Sag 2012.

provides no semantic argument for an internal modifier, that expression may nonetheless be modified by metalinguistic elements, which do not make reference specifically to the common noun's meaning or index. Thus we find contrasts like the following:

(77) a. Kim kicked the purple bucket. (only literal interpretation)
b. They kicked the proverbial bucket. (only idiomatic interpretation)

LID is not the only feature relevant to the analysis of MWEs; XARG has a significant role to play, as well. There are many English idioms that require referential and agreement identity between a possessor within an NP and some other argument of the idiom, or which assign a semantic role to the embedded possessor. Some of these are illustrated in (78)–(79):

(78) a. He$_i$ lost [his$_i$/*her$_j$ marbles].
b. They$_i$ kept [their$_i$/*our$_j$ cool].

(79) a. That$_i$ made [her$_j$ hair] stand on end.
b. That$_i$ tickled [your$_j$ fancy].

As noted in section 3.3 above, the presence of a prenominal genitive within an NP is encoded via a nonempty value for the feature XARG. If an object NP includes information about its prenominal genitive in its XARG value, then the listeme of a verb like *lose* (in its 'lose your cool' sense) can be formulated as in (80):[77]

(80)
$$\begin{bmatrix} \textit{strans-v-lxm} & \\ \text{FORM} & \langle \textit{lose} \rangle \\ \text{ARG-ST} & \left\langle X_i\ , \begin{bmatrix} \text{NP}_j & \\ \text{XARG} & [\textit{pron}]_i \\ \text{LID} & \langle \textit{i-cool}^{\textit{composure}}\textit{-fr} \rangle \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{losing-fr} & \\ \text{AGENT} & i \\ \text{ENTITY} & j \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

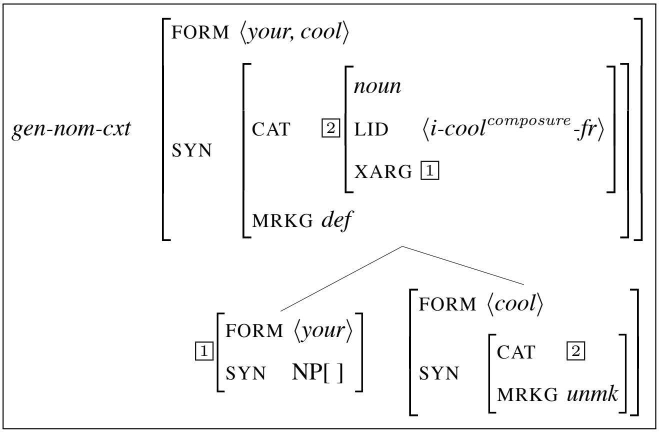
This specification requires both that the object NP contain a prenominal pronominal genitive NP and that that pronoun be coindexed with the subject of *lose* (blocking **He lost your cool* and the like).

I am assuming that NPs like *your cool* are built via the same Genitive Nominal Construction that is used for such NPs generally. This construction

[77] Note that I use '*pron*' as a shorthand for an overt pronominal sign. By contrast, '*pro*' designates a covert pronominal.

requires that the mother's XARG value be identified with the prenominal genitive NP, as shown in (81):

(81)



Thus, because only certain verbs, e.g. *keep*, *lose* and *blow* (in their relevant idiomatic senses) select a direct object whose LID value is ⟨*i-cool-fr*⟩, these are the only lexical elements that can govern NPs headed by *cool* in its relevant idiomatic sense. The possessor within the *cool* NP and the subject of governor are always coindexed. Various semantic treatments are possible. The lexical entry in (80) assumes that *lose* is dyadic, with the direct object NP forming a second semantic argument (the ENTITY argument, in FrameNet's terms).

The phenomena just discussed are outside the analytic scope of the version of HPSG developed by Pollard and Sag (1987, 1994). As argued in Sag (2007, 2010b), these data (and others discussed in section 8.1 below) provide considerable motivation for the analysis of verbal and nominal signs in terms of nonempty XARG specifications. Finally, note that XARG values, unlike VAL lists, do not 'shrink' in a bottom-up progression from head daughter to mother within an analysis tree. That is, no elements are 'cancelled off' an XARG list – the information about the external argument is locally visible at the top of the phrasal domain projected by the lexical head because XARG is a CAT feature and hence is generally 'passed up' from a head daughter to its mother.

7.2 Dative Alternations

As of June, 2010, FrameNet (see footnote 37), posits a single lexical entry to account for both of the following subcategorizations of *give* (in its 'caused possession' sense):

(82) a. Sandy gave Bo the beer. [Ditransitive]
b. Sandy gave the beer to Bo. [*To*-transitive]

This analysis is easily accommodated in SBCG by positing a single listeme like (83), which mentions the nonmaximal type *trans-verb-lxm*:

(83)
$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \text{FORM} & \langle \textit{give} \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \langle\, [\textit{giving-fr}]\, \rangle \end{bmatrix} \end{bmatrix}$$

This listeme licenses lexemes of two distinct types. That is, (83) is compatible with either of the following lexical class constructions:[78]

(84) a. *ditrans-v-lxm* ⇒
$$\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_x, \text{NP}_z, \text{NP}_y \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{giving}^{+}\textit{-fr} & \\ \text{DONOR} & x \\ \text{THEME} & y \\ \text{RECIPIENT} & z \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

b. *to-trans-v-lxm* ⇒
$$\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_x, \text{NP}_y, \text{PP}_z[\textit{to}] \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{giving}^{+}\textit{-fr} & \\ \text{DONOR} & x \\ \text{THEME} & y \\ \text{RECIPIENT} & z \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The result is that the two sentences in (82) are assigned the same basic semantics.

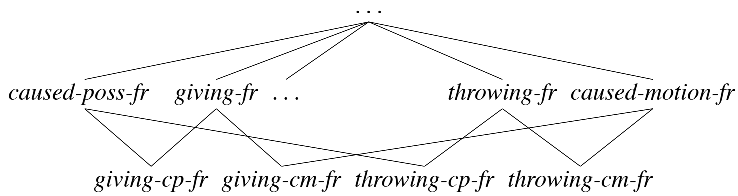
But meaning has played a critical role in the literature on dative alternations. In a tradition going back at least to Green 1974, Oehrle 1975, Jackendoff 1983 and Pinker 1989, researchers have attempted to associate semantic differences of various kinds with the ditransitive and *to*-transitive variants of dative alternations. A particularly influential proposal involves distinguishing the 'caused possession' semantics (associated with the ditransitive) from the 'caused motion' semantics (associated with the *to*-transitive). Such an analysis is also easily implemented within SBCG by organizing the type hierarchy so as to distinguish these two meaning types, perhaps as in (85):[79]

---

[78] '*giving*$^{+}$*-fr*' names the supertype classifying the meanings of all alternating *give*-type verbs and 'PP[*to*]' abbreviates a PP headed by the 'case-marking' (or 'bleached') preposition *to*.

[79] *giving-cp-fr* abbreviates *giving-caused-possession-frame*; *giving-cm-fr* abbreviates *giving-caused-motion-frame*; etc.

(85)



With this type hierarchy in place, the listeme for *give* can remain as formulated in (83) above, but the lexical class constructions will need to be revised along the lines of (86) in order to resolve the semantics of the ditransive construction to 'caused possession' and that of the *to*-transitive to 'caused motion' (AGENT replaces FrameNet's DONOR role):

(86) a. *ditrans-lxm* ⇒

$$\begin{bmatrix} \text{ARG-ST} & \langle NP_x, NP_z, NP_y \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{caused-poss-fr} & \\ \text{AGENT} & x \\ \text{THEME} & y \\ \text{RECIPIENT} & z \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

b. *to-trans-lxm* ⇒

$$\begin{bmatrix} \text{ARG-ST} & \langle NP_x, NP_y, PP_s[\textit{to}] \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{caused-motion-fr} & \\ \text{AGENT} & x \\ \text{THEME} & y \\ \text{PATH} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The constructions will thus interact with the listemes to associate the appropriate valence patterns with the appropriate meanings.

However, this analysis assumes that verbs like *give* are ambiguous between a 'caused possession' and a 'caused motion' interpretation, a pervasive claim that has been challenged by Rappaport Hovav and Levin (2008) (henceforth RH&L; see also Levin 2008). RH&L argue that *give*-type verbs have only a 'caused possession' interpretation, while other classes, e.g. the *throw*-type and *send*-type verbs, allow both 'caused motion' and 'caused possession' interpretations (though in different ways). Their arguments for this are based on certain contrasts involving the possibility of extraction with *where* (*Where did you kick/throw/*give/*hand the ball?*), the observation that *give*-type verbs lack a motion semantics in any of their uses, and further contrasts involving metaphorical interpretations of 'caused possession' verbs.

One way of implementing the RH&L analysis is based on the frame hierarchy sketched in Figure 18. On this approach, the familiar semantic conditions

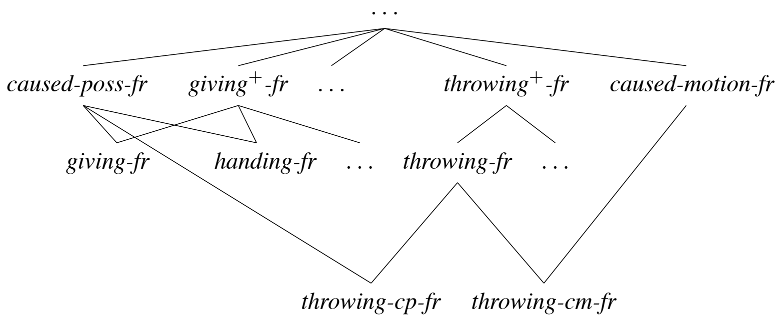


FIGURE 18 A Frame Hierarchy for the Dative Alternation

that RH&L associate with ‘caused possession’ and ‘caused motion’, namely those in (87), are recast as entailments (or meaning postulates) associated with the relevant frame types, in the manner outlined in Davis and Koenig 2000 and Koenig and Davis 2001:

(87) a. In any situation framed by a *caused-poss-fr*, where x is the AGENT, y is the RECIPIENT and z is the THEME, x’s action causes it to be the case that y has z.

b. In any situation framed by a *caused-motion-fr*, where x is the AGENT, y is the THEME, and z is the PATH, x’s action causes it to be the case that y goes along z.

The correspondences between argument structure patterns and meaning involves three lexical class constructions, as indicated in (88):[80]

(88) a. *ditrans-v-lxm* ⇒

$$\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_x, \text{NP}_z, \text{NP}_y \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{caused-poss-fr} & \\ \text{AGENT} & x \\ \text{THEME} & y \\ \text{RECIPIENT} & z \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

b. *to-trans-v-lxm* ⇒

$$\begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_x, \text{NP}_y, \text{PP}_z[\textit{to}] \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{caused-poss-fr} & \\ \text{AGENT} & x \\ \text{THEME} & y \\ \text{RECIPIENT} & z \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

[80]Here, ‘PP[*to*]’ continues to abbreviate a PP headed by the case-marking preposition *to* (RH&L refer to this as an ‘allative’ use), while ‘PP[*dir*]’ designates the class of PPs headed by a directional preposition, a class that includes the homophonous directional preposition *to*.

c. *trans-motion-v-lxm* ⇒

$$\begin{bmatrix} \textsc{arg-st} & \left\langle \text{NP}_x, \text{NP}_y, \begin{bmatrix} \text{PP}[\textit{dir}]_s \\ \textsc{val}\ \langle \textit{pro}_y \rangle \end{bmatrix} \right\rangle \\ \textsc{sem} & \begin{bmatrix} \textsc{frames} & \left\langle \begin{bmatrix} \textit{caused-motion-fr} & \\ \textsc{agent} & x \\ \textsc{theme} & y \\ \textsc{path} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

The PATH argument in (88c) is a situation and the unexpressed subject (*pro*) of the directional PP is identified with the verb's THEME argument, in effect treating caused-motion verbs as a kind of object-control.

The verb classes at play in this analysis, according to RH&L, are sketched in (89), with sample listemes shown in (90):

(89) a. **give-type verbs:** *give*, *hand*, *lend*, *loan*, *rent*, *sell*, ...; includes 'verbs of future having': *allocate*, *allow*, *bequeath*, *forward*, *grant*, *offer*, *promise*, ...

b. **send-type verbs:** *mail*, *send*, *ship*, ...

c. **throw-type verbs:** *fling*, *flip*, *kick*, *lob*, *slap*, *shoot*, *throw*, *toss*, ...

(90) a.

$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \textsc{form} & \langle \textit{give} \rangle \\ \textsc{sem} & \begin{bmatrix} \textsc{frames} & \langle [\textit{giving-fr}] \rangle \end{bmatrix} \end{bmatrix}$$

b.

$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \textsc{form} & \langle \textit{send} \rangle \\ \textsc{sem} & \begin{bmatrix} \textsc{frames} & \langle [\textit{sending-fr}] \rangle \end{bmatrix} \end{bmatrix}$$

c.

$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \textsc{form} & \langle \textit{throw} \rangle \\ \textsc{sem} & \begin{bmatrix} \textsc{frames} & \langle [\textit{throwing-fr}] \rangle \end{bmatrix} \end{bmatrix}$$

In sum, various proposals for the analysis of the dative alternation that have been sketched in the literature find a natural home in SBCG, where the consequences of particular analyses can be explored in detail and where analytic debates can take place with increased precision.

7.3 Locative Alternations

The range of analyses available for locative alternations like (91)–(92) are similar to those just examined for the dative alternation:

(91) a. They sprayed the wall with paint.
  b. They sprayed paint on the wall.

(92) a. We loaded the truck with hay.
  b. We loaded hay onto the truck.

And analogous to the treatment in (84) above for the FrameNet analysis of dative alternations, it is possible to specify a single listeme for each *spray*/*load* verb in such a way that it is compatible with two lexical class constructions, each of which assigns the same basic meaning, but links the ARG-ST members to semantic roles in distinct ways:

(93)
$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \text{FORM} & \langle \textit{spray} \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \langle [\textit{spray-fr}] \rangle \end{bmatrix} \end{bmatrix}$$

Under these assumptions, the alternations in (91)–(92) arise because both *spray-fr* and *load-fr* are instances of the type *loc-motion-fr*, and hence a listeme like (93) would be compatible with two lexical class constructions, resulting in identical semantics for pairs like (91a-b) or (92a-b).

But investigations into locative alternations, like studies of the dative alternation, have focused on semantic differences between the two relevant ARG-ST patterns, as well as differences in the semantic interactions between these patterns and particular verbs. Various analyses since Pinker 1989 have attempted to accommodate such semantic differences. For example, Iwata (2005, 2008) proposes to distinguish two broad classes of meanings that he associates with subcategorization frames in the manner of (94):

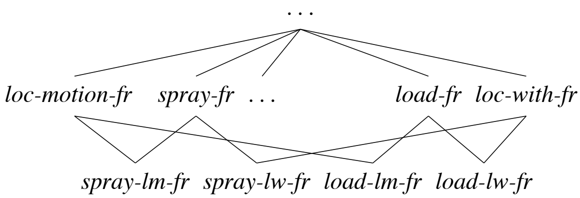
(94) a. ⟨NP, NP, PP[*dir*]⟩
  ‘X acts upon Y, thereby causing Y to go Z’
  b. ⟨NP, NP, PP[*with*]⟩
  ‘X acts upon Y, by exerting force over the surface[81] of Y with Z’

Iwata distinguishes a verb’s ‘lexical’ meaning from the ‘phrasal’ meaning that results from the conceptualization of that meaning that is compatible with the choice of subcategorization frame. Of course, the latter kind of meaning can be regarded as lexemic (rather than phrasal) in SBCG, as in the various analyses of dative variants we examined in the previous section. This is because a lexeme can make reference to the relevant phrasal information, e.g. the particular complements that appear on its ARG-ST list.

[81] Iwata imposes a further requirement that the force be exerted horizontally over the surface of Y. This is questionable, or at least in need of clarification, since it is clear that one can speak of spraying a wall, for example, even if all the movement involved is vertical and the event fails to be ‘over’ (horizontally overlapping; vertically superior to) the wall.

Despite this difference, the basic intuition of proposals like Iwata's – that there are two alternate conceptualizations of a given lexical meaning, either of which can be associated with a given listeme – can be formulated straightforwardly. For example, the hierarchy of frame types relevant to the locative alternation can be organized as shown in (95):

(95)



This partial frame hieararchy works in tandem with lexical class constructions like the following:

(96) a. **Transitive Locative Construction: (↑*trans-verb-lxm*):**

$$\textit{trans-loc-v-lxm} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \left\langle \text{NP}_x, \text{NP}_y, \begin{bmatrix} \text{PP}[\textit{dir}]_s \\ \text{VAL}\ \langle \textit{pro}_y \rangle \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{loc-motion-fr} & \\ \text{AGENT} & x \\ \text{THEME} & y \\ \text{PATH} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

b. **Applicative Construction: (↑*trans-verb-lxm*):**

$$\textit{trans-with-v-lxm} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_x, \text{NP}_z, \text{PP}_s[\textit{with}] \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{loc-with-fr} & \\ \text{AGENT} & x \\ \text{THEME} & z \\ \text{MEANS} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

In this proposal, like the analyses of the dative alternation reviewed earlier, the classifying types (in this case, *loc-motion-fr* and *loc-with-fr*) are associated with the appropriate entailments that Iwata posits to characterize his phrasal meanings. The listemes required (one per verb) are just as before, involving semantic and lexemic underspecification like that shown in (93) above. Nonalternating verbs of various kinds (see Levin 1993) are accommodated by positing listemes that specify more information, so as to be compatible with only one maximal type of verb lexeme.

Other analyses of locative alternations are of course possible. Note that there is ample room here for proposals that involve a more fine-grained semantic analysis (of both dative and locative alternations), as urged by numerous researchers (see, for example, Boas 2005, 2008b, 2010). These may be executed in terms of lexical subclasses, semantically tuned lexical class constructions, and exceptional lexical specifications. My aim here is not to choose among alternatives, but rather to show that SBCG provides a comfortable environment in which analyses of the kind that have been proposed in the literature can be precisely formulated and better evaluated.

7.4 Extended Valence Constructions

Construction Grammarians have spent considerable time examining extended valence patterns like those illustrated in (97):

(97) a. Pat sneezed the napkin off the table. [Caused Motion]
b. They drank themselves silly. ['Fake' Reflexive]
c. Chris lied his way into the meeting. [*X's Way*]

Researchers have been influenced by different aspects of these examples. For example, Goldberg (1995: 200) is struck by the enormous variety of verbs that can cooccur with *X's way*, while Boas (2003) argues that both the resultative and caused-motion constructions lack productivity.[82] For present purposes, I will follow Goldberg in assuming that all the phenomena in (97) are appropriately analyzed in terms of grammatical constructions.

Further examples containing the idiomatic phrase *X's way* are given in (98):

(98) a. The hunters$_i$ made their$_i$ way into the forest.
b. Bo$_i$ ate his$_i$ way across Europe.
c. But he$_i$ consummately ad-libbed his way through a largely secret press meeting. (OUP; cited by Goldberg 1995)
d. The players will maul their way up the middle of the field. (OUP; cited by Goldberg 1995)

The support verb *make* occurs with *X's way* more frequently than any other and involves a separate listeme whose instantiations give rise to sentences like (98a). This listeme remains separate from the construction I will propose.

[82] P. Kay (to appear) is similarly impressed by the lack of productivity shown by the caused motion phenomenon, proposing to treat it as a 'pattern of coinage', rather than a grammatical construction.

Sentences containing *X's way* typically situate a verb in an extended valence pattern of the form: verb + *the way* + path-naming (directional) expression. The meaning of the sentence, according to Levin and Rapoport (1988) and Jackendoff (1990), can be characterized in terms of motion (along the path) with the verb naming an activity that is either the means of the motion or some 'coextensive action or manner'. Thus (98c-d) might be paraphrased as in (99) or (100), if we interpret the relation between the two relevant events as one of means:

(99) a. He consummately made his way through a largely secret press meeting by means of ad-libbing.
b. The players made their way up the middle of the field by means of mauling (people).

The 'coextensive action or manner' interpretation associated with *X's way* sentences is more naturally illustrated by examples like the following:

(100) a. She whistled her way out of the room.
'She exited the room while whistling'
b. You drank your way through college.
'You went through college, drinking all the way'

The basic properties of this class of sentences can be analyzed by positing a derivational construction like the one in Figure 19, where ME/MA abbreviates the cover term MEANS/MANNER, intended to allow both of the interpretations just illustrated.[83]

What Figure 19 says is that an intransitive or transitive verbal lexeme (the daughter) may give rise to another lexeme (the mother) whose ARG-ST list includes the daughter's subject and whose FRAMES list includes that of the daughter. The elements added to the mother's ARG-ST list are an NP headed by the idiomatic noun *way* and a directional PP. The element added to the daughter's FRAMES list is a *going-fr* whose PATH is identified with the PP's situational index (s_2), and whose MANNER is the situational index of the daughter (s_1). In addition, the XARG of the *way* NP must be an overt pronominal which also functions as the subject (Q) of the directional PP and whose index (i) is the THEME of the *going-fr*. This construction thus gives rise to constructs like the one shown in Figure 20.[84]

[83]This shorthand collapses both uses of this construction into a single display. A more careful treatment might posit two sister constructions with a common supertype expressing their common properties.

[84]It is commonly thought that the overt pronominal XARG (the genitive pronoun modifying *way*) must be coindexed with the subject of *ad-lib*, a constraint that could of course be added to Figure 19. However, examples like the following (brought to my attention by Paul Kay) appear to be inconsistent with such a constraint:

$$\textit{verb-way-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \textit{derived-intrans-v-lxm} & \\ \text{FORM} & \langle X \rangle \\ \text{ARG-ST} & \left\langle \mathrm{Z}, \begin{bmatrix}\text{NP} & \\ \text{LID} & \langle \textit{i-way-fr}\rangle \\ \text{XARG} & Q:\textit{pron}_i\end{bmatrix}, \begin{bmatrix}\text{PP}[\textit{dir}] & \\ \text{VAL} & \langle Q\rangle \\ \text{IND} & s_2\end{bmatrix}\right\rangle \\ \text{SYN} & Y \\ \text{SEM} & \begin{bmatrix}\text{IND} & s_1 \\ \text{FRAMES} & \left\langle \begin{bmatrix}\textit{going-fr} & \\ \text{THEME} & i \\ \text{PATH} & s_2 \\ \text{ME/MA} & s_1\end{bmatrix}\right\rangle \oplus L\end{bmatrix}\end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix}\text{FORM} & \langle X\rangle \\ \text{ARG-ST} & \langle Z\rangle \oplus \langle(\text{NP})\rangle \\ \text{SYN} & Y:[\text{CAT } \textit{verb}] \\ \text{SEM} & \begin{bmatrix}\text{IND} & s_1 \\ \text{FRAMES} & L\end{bmatrix}\end{bmatrix}\right\rangle \end{bmatrix}$$

FIGURE 19 *Verb-Way Construction*
(↑*deriv-cxt*)

Given that constructs like the one in Figure 20 are well-formed, it follows from the Sign Principle that the mother in such a construct is licensed. These lexemes give rise to words that can combine with the appropriate nonsubject valents (by the Predicational Head-Complement Construction discussed in section 8 below) to build a VP like (101):

(101) $\{^{VP}$ [ad-libbed] [his way] [through a largely secret press meeting] $\}$

And this can combine with an adverb and a subject NP to form sentences like (98c) above.[85]

Readers steeped in the CxG tradition sometimes have an adverse initial reaction to this analysis, perceiving it as 'too lexical', presumably because the information being constrained is associated with the verb, rather than the phrases in which the verb appears. I urge such readers to reconsider their

(i) She participated in the Miss Femina India contest and won the title, which paved her way to the Miss Universe Crown.

[85]For an SBCG treatment of the analog of this construction in Dutch, see Poss 2010.

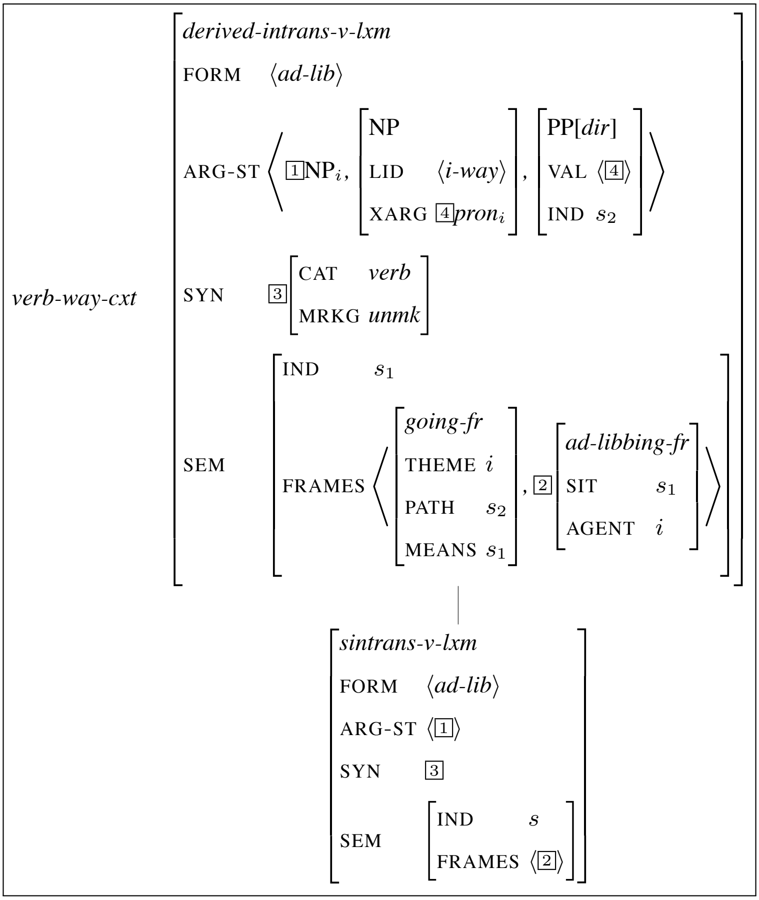


FIGURE 20 A Construct Instantiating the Verb-*Way* Construction

reaction. In general, the composition of phrases is largely determined by lexical information organized in large part by lexical classes. Crosslinguistically, there are many morpholexical regularities that also play an important role in phrasal construction. These include passivization, causativization, applicative formation, and other lexical phenomena that are typically analyzed via derviational constructions introducing appropriate affixation. The general principles underlying the construction of VPs, APs, PPs, and the like are the highly schematic head-complement constructions discussed in section 8.3 below, which allow a lexical head to combine with an appropriate set of lexically specified elements (members of the head's ARG-ST list). Hence a construction like Figure 19, which licenses a new lexical class of verbs with extended ARG-ST lists, has the effect of allowing verbs that are intrinsically intransitive to appear within a transitive-like VP, as long as they cooccur with the appropriate expressions, i.e. an NP headed by *way* and a directional PP. The boundary between the phrasal and the lexical is subtle in SBCG, as the form-meaning constraints imposed by listemes, lexical class constructions, and combinatoric constructions are inextricably intertwined.

The three constructions illustrated in (97) exhibit a 'family resemblance'. For example, they all involve a unary construct whose mother's valence extends that of the daughter. They differ from one another in terms of what the valence extensions can be and what constructional meaning is involved (though these meanings also resemble one another). It is thus natural to analyze them in terms of a constructional family: three construct types with a common supertype, just as Goldberg (1995) and P. Kay (2005) propose for the 'Argument Structure Constructions'. This approach allows each individual construction to be simplified slightly, leaving the supertype construction to capture all properties of the family at large. I will not explore the details of this analysis here. Note, however, that if the Passive Construction requires its daughter to be of type *trans-v-lxm*, then the individual extended valence constructions can specify whether their mother belongs to *derived-intrans-v-lxm* or *derived-trans-v-lxm*. Only in the latter case will a lexeme licensed by that construction be able to serve as the daughter of a construct of type *passive-v-lxm*. This provides one way of controlling the feeding relations among derivational constructions so as to account for contrasts like (102):

(102) a. The napkin was sneezed off the table (by Pat). [Caused Motion]
b.*The men were drunk silly (by themselves). ['Fake' Reflexive]
c.*Themselves were drunk silly (by the men). ['Fake' Reflexive]
d.*His way was lied into the meeting (by Chris). [*X's Way*]

Analyses of this sort make another point about lexical constructions in SBCG. It might be thought that lexical constructions lead to redundancy because they create lexical signs (as 'output') that look like members of a *bona* *fide* lexical class which they do not belong to, thus missing a lexical generalization. But generalizations of this kind are naturally expressed in SBCG, simply by using appropriate types in the formulation of the relevant lexical constructions. If construction A specifies that its MTR must be of type τ, then the lexical signs that are licensed by A must satisfy all the constraints that the grammar places on feature structures of type τ. Thus, some members of the lexical class associated with τ are basic, while others may be derived from a basic sign by a lexical construction whose formulation introduces no unwanted redundancy. For example, once the grammar includes a lexical class of *ditrans-v-lxm* (see (88a) above), an extended valence construction can be stated to augment the ARG-ST of strict-transitive verbal lexemes like *kick* and *throw* by specifying that the construction's MTR value is of type *ditrans-v-lxm*. Extended valence constructions in SBCG allow relevant lexical generalizations to be expressed perspicuously.[86]

8 Licensing Phrases

Phrasal (syntactic) constructs work in the same way as lexical constructs, except that they empower the grammar to build phrases from overt expressions (words or other phrases), as shown in the following declaration for type *phr-cxt*:

(103) *phr-cxt*: $\begin{bmatrix} \text{MTR} & \textit{phrase} \\ \text{DTRS} & \textit{list(overt-expr)} \end{bmatrix}$

(The mother of a phrasal construct must be a phrase and the daughters must be overt expressions, i.e. words or phrases.)

An important subtype of *phr-cxt* is *headed-construct* (*headed-cxt*), which introduces the further feature HEAD-DAUGHTER (HD-DTR):

(104) *headed-cxt*: [HD-DTR *overt-expr*]

(Headed constructs have a head daughter, which is an overt expression i.e. a word or phrase.)

And an important constraint associated with headed constructs in HPSG is the Head Feature Principle (HFP), which requires the mother's syntactic HEAD value (in our terms, the CAT value) to match that of its head daughter. The purely monotonic (non-default) version of SBCG presented here has no need for the HFP, however, since the various constraints expressing broader generalizations make it redundant.

[86] I am grateful to Adele Goldberg and Stefan Müller for useful discussion of these matters.

8.1 The Subject-Predicate Construction

Simple declarative clauses are licensed by the Subject-Predicate Construction, sketched in (105) (*subj-head-cxt* is an immediate subtype of *hd-cxt*):[87]

(105) **Subject-Predicate Construction** ($\uparrow$*subj-head-cxt*)**:**

$$\textit{subj-pred-cl} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } Y \,!\, [\text{VAL } \langle\ \rangle\,]\,] \\ \text{DTRS} & \left\langle X,\ Z : \begin{bmatrix} \text{SYN } Y : \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{VF} & \textit{fin} \\ \text{INV} & - \\ \text{AUX} & - \end{bmatrix} \\ \text{MRKG} & \textit{unmk} \\ \text{VAL} & \langle X \rangle \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{HD-DTR} & Z \end{bmatrix}$$

This construction says that two signs can combine as long as the second is a finite (and hence verbal) sign that selects the first via the VAL feature. The mother of a subject-predicate clause and its head (second) daughter will both be specified as [VF *fin*], [INV −], [AUX −], and [MRKG *unmk*]. This is because the SYN values of mother and head daughter are identified, except for the VAL value. The [INV −] specification prevents VPs headed by [INV +] verbs from heading subject predicate constructs (hence **I aren't included*) and the [AUX −] specification prevents VPs headed by [AUX +] verbs from doing the same (hence **I dŏ listen to you*). Finally, the mother of a subject-predicate construct, unlike its head daughter, must be [VAL ⟨ ⟩]. That is, the sign built by this construction is a finite sentence – a verbal projection that has 'consumed' all the valents of its lexical head.

The Subject-Predicate Construction says nothing about semantics because there is a general principle – the Principle of Compositionality (Sag et al. 2003) – that imposes the requirement that the FRAMES list of the daughters in a given construct be merged to form the mother's FRAMES list.[88] This principle and (105) together give rise to constructs like the one in Figure 21,

[87] In fact, the Subject-Predicate Construction should be simplified by moving certain of its constraints to the Subject-Head Construction, which defines the common properties of all subject-head clauses. I ignore such refinements here.

[88] Following Copestake et al. 2005, there is a further feature CONSTRUCTIONAL-CONTENT, whose value is a possibly empty list of frames that is also included in the mother's FRAMES list. This treatment (see the appendix) systematically allows for constructional meaning. The Declarative-Clause Construction, or perhaps a construction characterizing a different supertype of *subj-pred-cl*, may be formulated so as to contribute a propositionalizing frame (*proposition-frame*) as constructional meaning.

whose mother is just the phrasal sign illustrated in Figure 7 above.[89]

8.2 Issues of Locality

Although SBCG constructions, like rules in a context-free grammar, can only make reference to a mother and her daughters, they can nevertheless accommodate grammatical dependencies that are nonlocal in nature. In particular, SBCG builds on work in the GPSG/HPSG tradition that has used feature specifications to locally encode information about long-distance dependencies. Just as the featural representation of a category like 'NP' encodes the information that the phrase's head word is a noun, other feature specifications can encode key grammatical information about an element that is present in (or absent from) a phrase. For example, the VF value of a verbal phrase (VP or S) encodes a morphosyntactic property of the phrase's head word.[90] Similarly, the feature GAP[91] is used to encode the absence (or 'extraction') of an element (or, as syntacticians often put it, the 'presence of a gap') within a given phrase. By developing a theory of such feature specifications and the principles that govern their distribution throughout phrasal structures (constructs), we are *ipso facto* developing a theory of what nonlocal information can be lexically selected at a higher level of structure.

As has been recognized at least since Chomsky 1965,[92] lexical restrictions (broadly construed) are found only in circumscribed domains; that is, they are localized in a way that must be delimited by grammatical theory. Behind the search for the precise characterization of the relevant notion of lexical locality is the clear intuition that no language has, for example, a verb that requires a clausal complement that must contain an overt direct object that is feminine, or singular, etc. Early accounts of locality excluded subjects, but since idiosyncratic case assignment in numerous languages (perhaps most famously in Icelandic[93]) clearly involves the subjects of verbs, the most likely first approximation of the relevant locality domain can be formulated as follows:

---

[89] Figure 21 also illustrates the effect of two further constraints: the first, applying to most headed constructs, identifies the mother's LTOP with that of the head daughter; the second, applying to all constructs, requires the mother's BCKGRND list to include all the frames of the daughters' BCKGRND lists.

[90] Note that the lexical head can appear at arbitrary depth within a phrase (presuming that English grammar imposes no upper bound on the number of phrasal modifiers). Hence, even the dependencies mentioned so far are unbounded.

[91] This feature name is due to Sag et al. (2003). For more detailed, explicit analyses that employ the feature name 'SLASH', see Gazdar 1981, Gazdar et al. 1985, Pollard and Sag 1994, Bouma et al. 2001, Levine and Hukari 2006, and Chaves 2012.

[92] See also Kajita 1968 and Sag 2010b.

[93] See for example Thráinsson 1979, Andrews 1982, 1990 and Barðdal 2011.

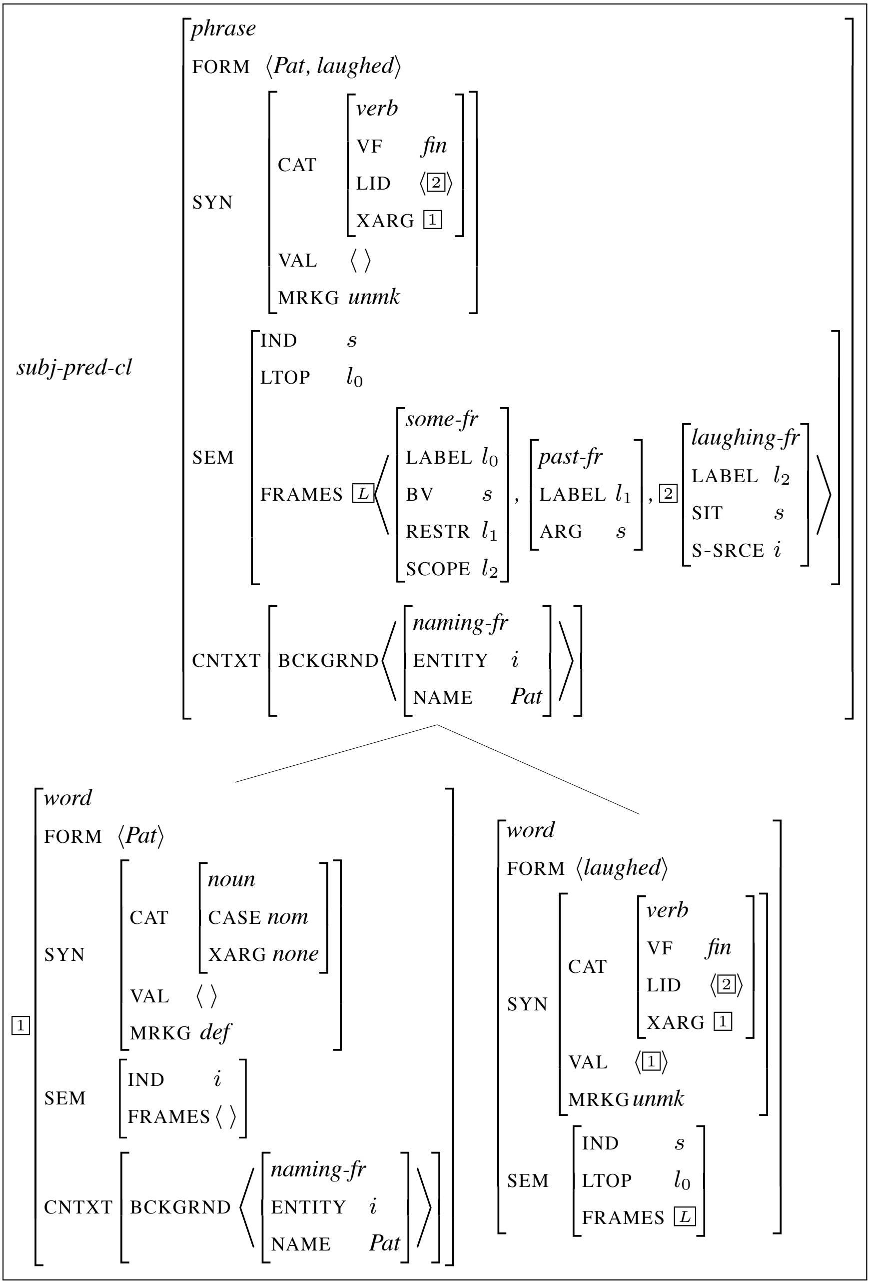


FIGURE 21 A Subject-Predicate Clause

(106) **Selectional Localism**

For purposes of category selection (subcategorization), case assignment, (non-anaphoric) agreement, and semantic role assignment, a lexical head has access only to the signs it selects via some feature (e.g. ARG-ST or SELECT), i.e. the elements that it is connected to via a grammatical relation (subject of, modifier of, etc.).

In SBCG, this amounts to a restriction that the only nonlocal elements that can be selected are those whose grammatical information is encoded by phrases that contain them. That is, a verb can require the noun that follows it to be accusative because accusativity is a property of the noun phrase projected from that noun. Our various features and the particular choices made about the nature of their values, taken together with general constraints on how information is percolated as phrasal signs are constructed, constitute a precise formulation of the basic idea embodied in (106).

This mode of analysis is implicit in $\overline{X}$-Theory, which since its inception has involved the percolation of category information from lexical heads to the phrases they project. Work in GPSG and HPSG extended the set of HEAD features to include features like CASE, VF, and AGR, and thus expanded the scope of $\overline{X}$-Theory to include the percolation of case, verb inflection, and agreement information through exactly the same domains. The HPSG/SBCG theory of VALENCE (SUBCAT), SELECT (MOD), and GAP (SLASH) extends the range of feature-based analysis to include the selection of valents by heads, the selection of modified elements by modifiers, and the cooccurrence of fillers and gaps.

And by adding a specific feature like XARG to systematically propagate certain information about elements embedded within dependents, we have localized certain nonlocal information, making information about a clause's subject NP locally accessible to a construction or word that imposes constraints on that clause at a higher level of structure. For example, Bender and Flickinger (1999) analyze agreement in English tag questions by allowing the subject's agreement information to percolate up to the top of the clause via XARG. When a clause is combined with the tag, the XARG values of the two daughters must be compatible. This induces the familiar tag question agreement pattern illustrated in (107):

(107)
$$\text{[They left,] didn't} \left\{ \begin{array}{l} \text{they} \\ \text{*(s)he} \\ \text{*we} \\ \text{*you} \\ \text{*I} \end{array} \right\} ?$$

The problem here is not selectional locality, but rather the related issue of constructional locality, about which we may formulate the following hypothesis:

(108) **Constructional Localism:**
Constructions license mother-daughter configurations without reference to embedding or embedded contexts.

Notice that Constructional Localism is an immediate consequence of the feature geometry assumed in SBCG, which, unlike earlier work in HPSG, draws a fundamental distinction between signs and constructs. Constructional Localism does not preclude an account of nonlocal dependencies in grammar, it simply requires that all such dependencies be locally encoded in signs in such a way that information about a distal element can be accessed locally at a higher level of structure.

On the basis of data like (109), it is clear that the agreement between the two subjects here is semantic in nature, whereas the agreement between each verb and its subject is intuitively syntactic in nature:[94]

(109) a. Sears is having a sale, aren't they?
b. At least one of us is sure to win, aren't we?
c. The crowd is getting agitated, aren't they?

Notice, however, that in an analysis along the lines shown in Figure 22, the seemingly nonlocal agreement relation between the two subject NPs is localized. That is, by positing XARG values that are identified with a clause's subject, we make it possible to treat the agreement in tag questions via a constraint requiring the relevant semantic relation between the XARG value of the main clause and the pronominal XARG value of the tag clause (the NPs that are shaded and coindexed in Figure 22).

There is independent motivation for the feature XARG. A case in point is the English 'copy raising' construction (Rogers 1974, Potsdam and Runner 2001, Asudeh 2002), illustrated in (110):[95]

(110) a. There looks like { there's going to be a storm / *it's going to rain / *Kim's going to win } .

[94] See Oehrle 1987, Culicover 1992, and P. Kay 2002a for discussion.

[95] These judgments are nuanced, as *looks* (*like*) is systematically ambiguous between a copy raising verb and a dyadic predicator where the subject is assigned a semantic role. For this reason, the examples in (110b) are possible on the latter reading, though the acceptability is slightly degraded.

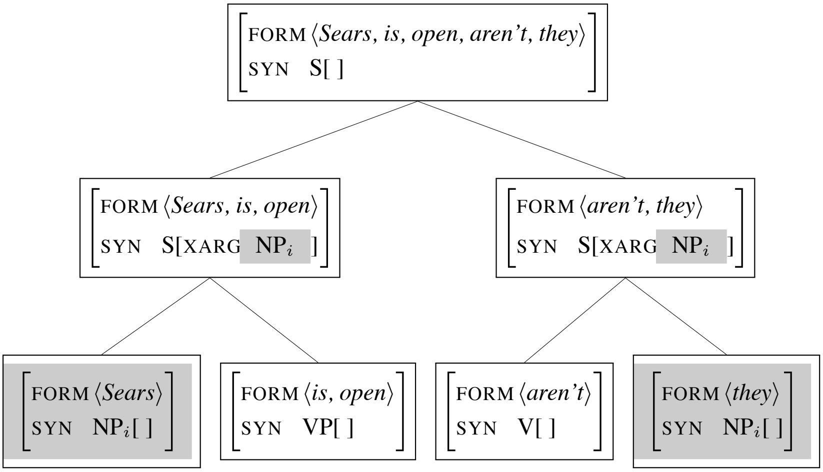


FIGURE 22 A Tag Question Analysis Tree

b. ?Kim looks like { there's going to be a storm / it's going to rain / Pat's going to win }.

Assuming, following Pollard and Sag (1994), that there are three subtypes of the type *index*: *ref* (*referential-index*), *it* and *there* – contrasts like these can be treated simply by providing the relevant *look* lexemes with the ARG-ST list in (111):[96,97]

(111)
$$\begin{bmatrix} \textsc{form} & \langle \textit{look} \rangle \\ \textsc{arg-st} & \left\langle \text{NP}_i, \begin{matrix}\textsc{prt}\\ [\textit{like}]\end{matrix}, \begin{bmatrix} \text{S} & \\ \textsc{xarg} & \langle \text{NP}_i[\textit{pron}] \rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

As noted above (see (52)), the XARG value of a verbal lexeme is also the first member of the verb's ARG-ST and VAL lists. Hence the constraints in (111) (enforced in large part by lexical class constructions) guarantee that the clausal complement of *look* will have a pronominal subject coindexed with the subject of *look* – a nonlocal dependency that is encoded locally.

[96] An alternative is to treat *like* as a marker that combines with a finite clause.

[97] Also relevant are controlled pronominal subjects in Serbo-Croatian (Zec 1987), Halkomelem Salish (Gerdts and Hukari 2001) and other languages, where control verbs also include the ARG-ST specification in (110). The analytic problems of raising across Polish prepositions (Przepiórkowski 1999, Dickinson 2004) and complementizer agreement in Eastern Dutch dialects (Höhle 1997) are similar, and submit to similar analysis.

8.3 The Head-Complement Constructions

With these lexical contrasts in place, we can now discuss the basic analysis of VPs, APs, PPs, and $\overline{\text{N}}$s. There are two basic patterns for complement realization in English and many other languages. The first, which is typical of predicative expressions of all categories and also of VPs, requires all complements except the subject to be realized within a head-complement construct, as illustrated in Figure 23. Note that the mother in such a construct has the head daughter's subject valent on its VAL list; this NP is its XARG value as well.[98]

The second pattern of complement realization is utilized by 'case-marking' prepositions of the sort we have already seen in section 7.2 above. According to this 'saturating' mode of realization, all the head's valents are realized as sisters of the lexical head, as shown in Figure 24.

These two patterns can be analyzed in terms of the following two phrasal constructions:

(112) **Predicational Head-Complement Construction** (↑*headed-cxt*)**:**

$$pred\text{-}hd\text{-}comp\text{-}cxt \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X \,!\, [\text{VAL } \langle Y \rangle]] \\ \text{DTRS} & \langle Z \rangle \oplus L : nelist \\ \text{HD-DTR} & Z : \begin{bmatrix} word \\ \text{SYN } X : \begin{bmatrix} \text{CAT} & [\text{XARG } Y] \\ \text{VAL} & \langle Y \rangle \oplus L \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(113) **Saturational Head-Complement Construction** (↑*headed-cxt*)**:**

$$sat\text{-}hd\text{-}comp\text{-}cxt \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X \,!\, [\text{VAL } \langle\, \rangle]] \\ \text{DTRS} & \langle Z \rangle \oplus L : nelist \\ \text{HD-DTR} & Z : \begin{bmatrix} word \\ \text{SYN } X : \begin{bmatrix} \text{CAT} & \begin{bmatrix} prep \\ \text{XARG} & none \end{bmatrix} \\ \text{VAL} & L \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

What (112) says is that a word specifying an external argument may combine with all its valents except the first to form a phrase. (113) allows a preposition lacking an external argument to combine with all its valents. The mother's

[98] Here and throughout, EXPRNCR abbreviates EXPERIENCER.

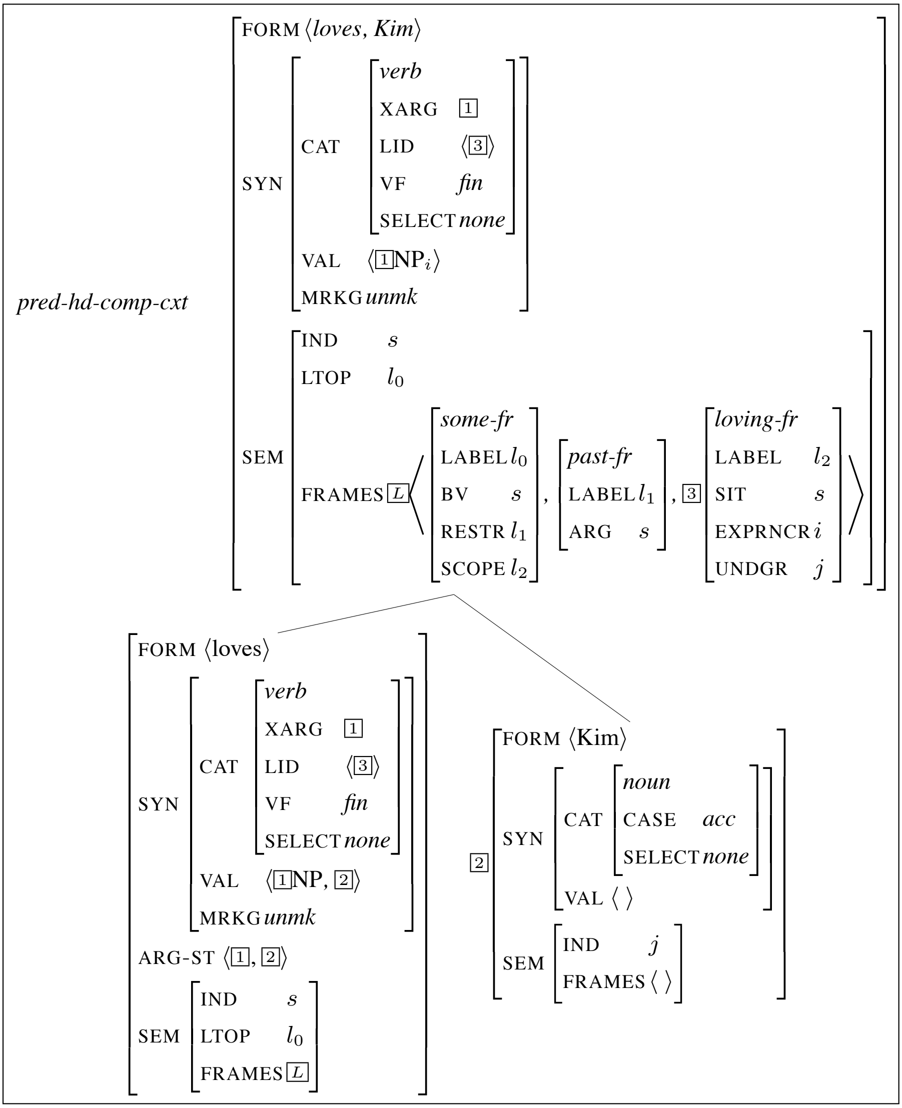


FIGURE 23 A Predicational Head-Complement Construct

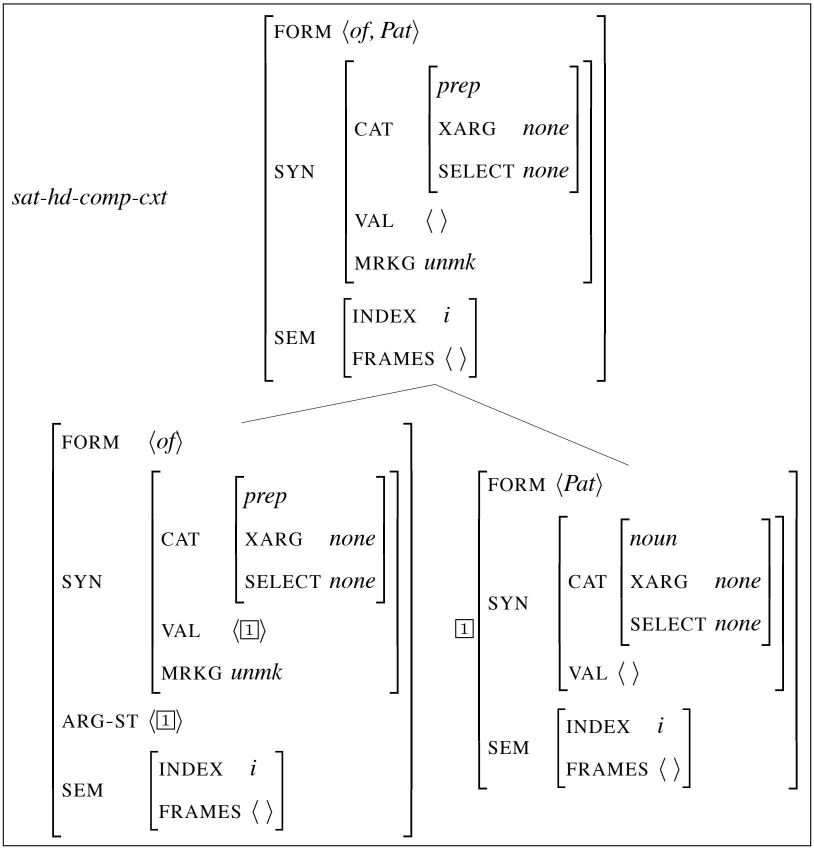


FIGURE 24 A Saturational Head-Complement Construct

SYN value in both constructions must match that of the first daughter (the head daughter), except for the VAL value. This analysis thus presupposes the existence of lexemes like those shown in (114) and (115):[99]

[99]These constructs of course also obey further constraints on linear order that I will not discuss here. For convenience, I have omitted discussion of linear ordering, assuming that the order of elements on the DTRS list determines the order of elements on the mother's FORM list. This is a simplification of a complex set of issues that have motivated ID/LP format (the separation of constructions and the principles that order their daughters) and 'Linearization Theory', the augmentation of sign-based grammar to allow interleaving of daughters as an account of word order freedom. On ID/LP grammars, see Gazdar and Pullum 1981, Gazdar et al. 1985, and Pollard and Sag 1987, among others. On Linearization Theory, see Reape 1994, Müller 1995, 1999, Donohue and Sag 1999, Kathol 2000, 2002, 2004, and Daniels and Meurers 2004.

(114)

$$\left[\begin{array}{ll} \textit{strans-v-lxm} & \\ \text{FORM} & \langle \textit{love} \rangle \\ \text{SYN} & \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \textit{verb} & \\ \text{XARG} & \boxed{1} \\ \text{SELECT} & \textit{none} \end{array}\right] \\ \text{MRKG} & \textit{unmk} \\ \text{VAL} & \langle \boxed{1}\text{NP}_i, \boxed{2}\text{NP}_j \rangle \end{array}\right] \\ \text{ARG-ST} & \langle \boxed{1}, \boxed{2} \rangle \\ \text{SEM} & \left[\begin{array}{ll} \text{IND} & s \\ \text{FRAMES} & \left\langle \left[\begin{array}{ll} \textit{loving-fr} & \\ \text{LABEL} & l_2 \\ \text{SIT} & s \\ \text{EXPRNCR} & i \\ \text{UNDGR} & j \end{array}\right] \right\rangle \end{array}\right] \end{array}\right]$$

(115)

$$\left[\begin{array}{ll} \text{FORM} & \langle \textit{of} \rangle \\ \text{SYN} & \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \textit{prep} & \\ \text{XARG} & \textit{none} \\ \text{SELECT} & \textit{none} \end{array}\right] \\ \text{VAL} & \langle \boxed{1}\text{NP}_i \rangle \\ \text{MRKG} & \textit{unmk} \end{array}\right] \\ \text{ARG-ST} & \langle \boxed{1} \rangle \\ \text{SEM} & \left[\begin{array}{ll} \text{INDEX} & i \\ \text{FRAMES} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

8.4 The Head-Functor Construction

We now turn to the Head-Functor Construction. I follow the essential insights of Van Eynde (1998, 2006, 2007), who argues that significant generalizations in the grammar of nominals are missed by analyses based on so-called 'functional categories'.[100] In their place, he develops a unified analysis of markers (including determiners) and modifiers in terms of a simple, direct combination of a 'functor' expression and the head that it selects, based on the SELECT

[100] See also Allegranza 1998b, 2007; Hudson 2000, 2004; and Newmeyer 2008a,b, 2009.

feature, discussed in section 3.3 above.

All major categories specify values for SELECT in Van Eynde's theory: nouns, adjectives, adverbs, prepositions, and verbs. For some of these, e.g. finite verbs, the value is *none*. Attributive adjectives, by contrast, select unmarked heads of category *noun* and are themselves unmarked ([MRKG *unmk*]). Determiners are similar, but have a specific MRKG value, as illustrated in (116):

(116) a.

$$\begin{bmatrix} \text{FORM} & \langle \textit{happy} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{adj} & \\ \text{SELECT} & \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & \textit{noun} \\ \text{MRKG} & \textit{unmk} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{MRKG} & \textit{unmk} \end{bmatrix} \end{bmatrix}$$

b.

$$\begin{bmatrix} \text{FORM} & \langle \textit{the} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{det} & \\ \text{SELECT} & \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & \textit{noun} \\ \text{VAL} & \langle\, \rangle \\ \text{MRKG} & \textit{unmk} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{MRKG} & \textit{def} \end{bmatrix} \end{bmatrix}$$

Given lexical specifications like these,[101] we can formulate the Head-Functor Construction as follows:

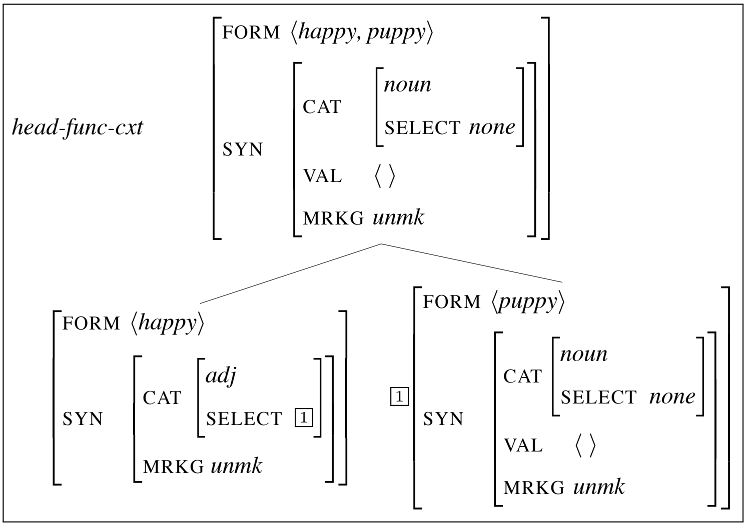
(117) **Head-Functor Construction: (↑*headed-cxt*):**

$$\textit{head-func-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X \,!\, [\text{MRKG } M\,]] \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & [\text{SELECT } Y\,] \\ \text{MRKG} & M \end{bmatrix} \end{bmatrix},\ Y\!:\![\text{SYN } X] \right\rangle \\ \text{HD-DTR} & Y \end{bmatrix}$$

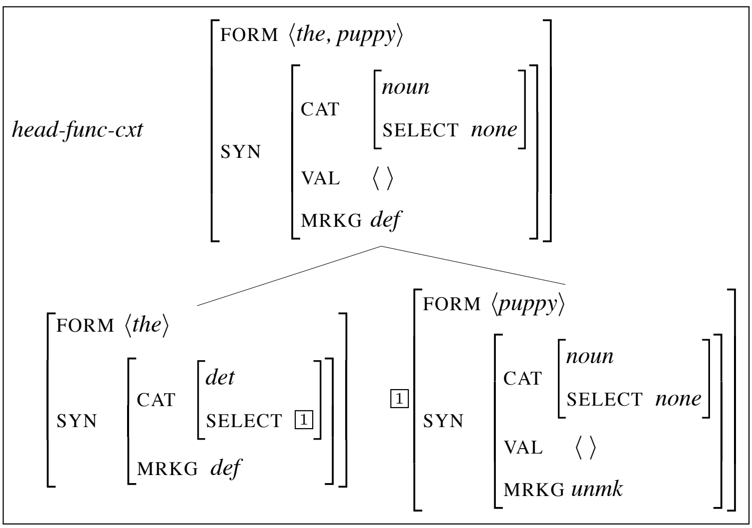
This construction allows us to construct both modified and determined phrases, as shown in (118)–(119). Expressions like **happy the puppy* are blocked because attributive adjectives (e.g. *happy*) select unmarked expressions whose category is *noun*. Note that in each of these constructs, the mother's SELECT specification is inherited from the head daughter, in accordance with the Head-Functor Construction.

[101] The details in (116) need not all be specified in individual listemes. I am here conflating the effect of lexical class constructions and listemes for ease of exposition.

(118)



(119)



9 Auxiliaries

As is well known, English makes an important distinction between auxiliary and nonauxiliary verbs. The data in (120) illustrate what have been referred to (see Quirk et al. 1985; Warner 1993a) as the NICE properties, which distinguish the two classes:

(120) **The NICE Properties**:

Negation (Finite): Lee will not eat apples / *Kim eats not apples.

Inversion: Has Lee eaten apples? / *Eats Lee apples?
Contraction of *not*: didn't, shouldn't / *eatn't,...
Ellipsis (of VP): Kim isn't kicking the ball, but Lee is __ / *but Lee likes __ .

Following Sag (to appear), we add a fifth property of (finite) auxiliaries: their ability to perform a 'rebuttal' function either by being prosodically focused or by combining with the particles *too* or *so* (in American varieties). These combinations are typically used to reaffirm the truth of a proposition that has just been denied by the addressee:

(121) A: Kim won't read it.
B: Kim wíll read it.

(122) A: Kim won't read it.
B: Kim will so/too read it.

The resulting **NICER** properties constitute the essential empirical domain that must be treated by any adequate analysis of the English Auxiliary System (EAS).

The basis of my analysis of the EAS is the contrast between the listemes licensing auxiliary and non-auxiliary verbal elements. Auxiliary verbs belong to the type *aux-v-lxm* (a subtype of *sraising-v-lxm*; see (51) above) and (with various lexical exceptions, to be noted) are unspecified for both AUX and INV. Nonauxiliary verbs, by contrast, instantiate the type *main-v-lxm* and are required to be [AUX −] and [INV −], as was shown in (53) above. This allows listemes to be simplified, as illustrated in (123):

(123) a.
$$\begin{bmatrix} \textit{aux-v-lxm} & \\ \text{FORM} & \langle \textit{can} \rangle \\ \text{SYN} & [\text{CAT} \; [\text{VF} \; \textit{fin}]] \\ \text{ARG-ST} & \left\langle \text{X}, \begin{bmatrix} \text{VP} & \\ \text{VF} & \textit{base} \end{bmatrix} \right\rangle \\ \text{SEM} & [\text{FRAMES} \; \langle [\textit{can-fr}] \rangle] \end{bmatrix}$$

b.
$$\begin{bmatrix} \textit{nonaux-sraising-v-lxm} & \\ \text{FORM} & \langle \textit{continue} \rangle \\ \text{ARG-ST} & \left\langle \text{X}, \begin{bmatrix} \text{VP} & \\ \text{VF} & \textit{prp} \end{bmatrix} \right\rangle \\ \text{SEM} & [\text{FRAMES} \; \langle [\textit{continuing-fr}] \rangle] \end{bmatrix}$$

This gives rise to lexeme contrasts like the following:

(124)

$$\begin{bmatrix} \textit{aux-v-lxm} & \\ \text{FORM} & \langle \textit{can} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{XARG} & \boxed{1} \\ \text{VF} & \textit{fin} \\ \text{AUX} & +/- \\ \text{INV} & +/- \end{bmatrix} \\ \text{VAL} & \boxed{L} \end{bmatrix} \\ \text{ARG-ST} & \boxed{L} \left\langle \boxed{1}\text{NP}, \begin{bmatrix} \text{VP} & \\ \text{VF} & \textit{base} \\ \text{VAL} & \langle \boxed{1} \rangle \\ \text{LTOP} & l \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{can-fr} & \\ \text{ARG} & l \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

(125)

$$\begin{bmatrix} \textit{nonaux-sraising-v-lxm} & \\ \text{FORM} & \langle \textit{continue} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{XARG} & \boxed{1} \\ \text{VF} & \textit{fin} \\ \text{AUX} & - \\ \text{INV} & - \end{bmatrix} \\ \text{VAL} & \boxed{L} \end{bmatrix} \\ \text{ARG-ST} & \boxed{L} \left\langle \boxed{1}\text{NP}, \begin{bmatrix} \text{VP} & \\ \text{VF} & \textit{inf} \\ \text{VAL} & \langle \boxed{1} \rangle \\ \text{LTOP} & l \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{continuing-fr} & \\ \text{ARG} & l \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Four of the NICER properties (negation, contraction, ellipsis, and rebuttal) are analyzed in terms of lexical constructions whose daughter must be specified as [AUX +], hence excluding all nonauxiliary verbs. The mother in all the constructs allowed by these constructions, however, may be [AUX −], as we saw in the analysis of finite negation in section 6.5 above. This is illustrated for VP-Ellipsis in Figure 25.

The resulting [AUX −] words then project [AUX −] VPs, in accordance with the Predicational Head-Complement Construction presented in section 8.3 above. Since the Subject-Predicate Construction in section 8.1 above requires a head daughter that is [AUX −], this provides an analysis of sentences like those in (126):

(126) a. Lee can not eat apples.
b. Lee can't eat apples.
c. Lee can't __ .
d. Lee can so/too eat apples.
e. Lee cán eat apples.

In a sentences like (127), by contrast, the listemically unspecified AUX value of the finite auxiliary has simply been resolved to '−':

(127) a. Kim can eat apples.
b. Kim is eating an apple.

Since finite forms of the auxiliary verbs, but not those of nonauxiliary verbs, can be resolved to [INV +], it also follows that only auxiliary verbs will be able to function as the head daughter in the various aux-initial constructs, some of which were discussed in section 2 above (see also Fillmore et al. this volume). The literature on auxiliaries discusses both positive and negative exceptions to 'inversion'. In our terms, specifying a listeme as [INV +] guarantees that the words licensed via that listeme appear only in aux-initial contexts (**I* **aren't** *the right choice*), while a listemic [INV −] specification prevents an element from occurring in such contexts (***Better** *they do that?*). This provides an analysis covering a range of data that has never been systematized in any transformational treatment, as far as I am aware.

Finally, Sag (to appear) provides an account of the exceptional auxiliary verb *do*, which has required considerable machinery within previous transformational analyses. The proposed treatment involves nothing more than a lexical exception: auxiliary *do* is lexically specified as [AUX +]. Because it is so specified, it can appear in any of the NICER environments, but it cannot appear in an environment requiring that its AUX value simply resolve to '−'. That is, it cannot appear in examples like (128):

(128) *Kim dĭd eat apples.

The essential ingredient of this analysis of *do* is the reinterpretation of the feature AUX. In previous analyses, the specification [AUX +] designated the

$$
\textit{vpe-cxt}\quad
\begin{bmatrix}
\textit{word} & \\
\text{FORM} & \langle \textit{will} \rangle \\
\text{SYN} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & - \\ \text{VF} & \textit{fin} \end{bmatrix} \\
\text{VAL} & \langle \boxed{1} \rangle \\
\text{MRKG} & \boxed{4}
\end{bmatrix} \\
\text{ARG-ST} & \boxed{L}\left\langle \boxed{1}\text{NP}_i[\textit{nom}],\ \boxed{2}\begin{bmatrix} \text{VP}[\textit{pro}] & \\ \text{VF} & \textit{bse} \\ \text{LTOP} & l \end{bmatrix} \right\rangle \\
\text{SEM} & \begin{bmatrix} \text{IND} & s \\ \text{FRAMES} & \langle \boxed{5} \rangle \end{bmatrix}
\end{bmatrix}
$$

$$
\Big|
$$

$$
\begin{bmatrix}
\textit{word} & \\
\text{FORM} & \langle \textit{will} \rangle \\
\text{SYN} & \begin{bmatrix}
\text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & + \\ \text{VF} & \textit{fin} \end{bmatrix} \\
\text{VAL} & \langle \boxed{1}, \boxed{2} \rangle \\
\text{MRKG} & \boxed{4}\textit{unmk}
\end{bmatrix} \\
\text{ARG-ST} & \boxed{L} \\
\text{SEM} & \begin{bmatrix} \text{IND} & s \\ \text{FRAMES} & \left\langle \boxed{5}\begin{bmatrix} \textit{will-fr} & \\ \text{SIT} & s \\ \text{ARG} & l \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

FIGURE 25 A Finite Auxiliary Verb Undergoes VP-Ellipsis property of being an auxiliary verb, while in the present proposal, this specification indicates the property of being an auxiliary construction (i.e. one of the NICER constructions). The difference in interpretation allows a straightforward account of the exceptionality of auxiliary *do*. For a fuller discussion, including a treatment of the problem of nonfinite auxiliary *do* in British English, e.g. (129), see Sag to appear.

(129) For one thing, a postponement will be seen worldwide as a declaration that Britain is in crisis; tourism would suffer even more than it is **doing** already. (*The Guardian*, 24/03/2001, cited in Miller 2002, q.v.)

10 Filler-Gap Constructions

There are numerous 'filler-gap' constructions in English, including the following, all of which are surveyed in Sag 2010a (q.v.):

(130) ***Wh*-Interrogative Clause:**
a. {[How foolish] [is he]}?
b. I wonder {[how foolish] [he is]}.

(131) ***Wh*-Exclamative Clause:**
a. {[What a fool] [he is]}!
b. It's amazing {[how odd] [they are]}.

(132) **Topicalized Clause:**
{[The bagels,] [I like]}.

(133) ***Wh*-Relative Clause:**
a. I met the person {[who] [they nominated]}.
b. I'm looking for a bank {[in which] [to place my trust]}.

(134) ***The*-Clause:**
a. The more people I met, {[the happier] [I became]}.
b. {[The more people] [I met]}, the happier I became.

All kinds of clauses exhibit a filler-gap dependency between a clause-initial filler phrase and a gap located within the sentential head daughter. However there are a number of parameters of variation distinguishing these varieties of clause from one another, including the following:

(135) **Parameters of Variation in FG Clauses:**
a. Is there a distinguished *wh* element in the filler daughter, and if so, what kind?

b. What are the possible syntactic categories of the filler daughter?
c. What are the possible syntactic categories of the head daughter?
d. Can the head daughter be inverted/finite? Must it be?
e. What is the semantics and/or syntactic category of the mother?
f. What is the semantics and/or syntactic category of the head daughter?
g. Is the clause an island? Must it be an 'independent clause'?

The analysis of filler-gap dependencies naturally breaks down into three problems: (1) the binding environment, where the filler is introduced, (2) the filler-gap dependency path, and (3) the realization of the gap. Building on a long tradition, beginning with Gazdar's (1981) pioneering work and including Pollard and Sag 1994, Bouma et al. 2001, Levine and Hukari 2006, and Chaves 2012, the presence of a gap is encoded in terms of a nonempty specification for the feature GAP (e.g. [GAP ⟨NP⟩]). By contrast, an expression containing no unbound gaps is specified as [GAP ⟨ ⟩].

Here I follow Ginzburg and Sag (2000), whose traceless analysis allows a lexical head to appear without a valent (subject, object, or other complement) and its GAP list contains an element corresponding to that valent. That is, a word's VAL list is shorter than its ARG-ST list just in case the missing element is on the word's GAP list. These GAP lists must also include elements that are on the GAP lists of the word's valents, as shown in (136):

(136) a. No Gap (*They like Lou*):

$$\left[\begin{array}{ll} \text{FORM} & \langle \textit{like} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1}\begin{bmatrix}\text{NP} \\ \text{GAP} \ \langle\ \rangle\end{bmatrix}, \boxed{2}\begin{bmatrix}\text{NP} \\ \text{GAP} \ \langle\ \rangle\end{bmatrix} \right\rangle \\ \text{SYN} & \begin{bmatrix}\text{VAL} & \langle \boxed{1}, \boxed{2} \rangle \\ \text{GAP} & \langle\ \rangle\end{bmatrix} \end{array}\right]$$

b. Object Gap (*that they like* __)

$$\left[\begin{array}{ll} \text{FORM} & \langle \textit{like} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1}\begin{bmatrix}\text{NP} \\ \text{GAP} \ \langle\ \rangle\end{bmatrix}, \boxed{2}\begin{bmatrix}\text{NP} \\ \text{GAP} \ \langle\ \rangle\end{bmatrix} \right\rangle \\ \text{SYN} & \begin{bmatrix}\text{VAL} & \langle \boxed{1} \rangle \\ \text{GAP} & \langle \boxed{2} \rangle\end{bmatrix} \end{array}\right]$$

c. Gap within Object (*that they like* [*your review of* __]):

$$\begin{bmatrix} \text{FORM} & \langle \textit{like} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1}\begin{bmatrix}\text{NP} & \\ \text{GAP} & \langle\, \rangle\end{bmatrix}, \boxed{2}\begin{bmatrix}\text{NP} & \\ \text{GAP} & \langle \boxed{3}\text{NP}\rangle\end{bmatrix}\right\rangle \\ \text{SYN} & \begin{bmatrix}\text{VAL} & \langle \boxed{1}, \boxed{2}\rangle \\ \text{GAP} & \langle\boxed{3}\text{NP}\rangle\end{bmatrix}\end{bmatrix}$$

d. Gaps within Subject and Object
(*that* [*proponents of* __] *like* [*reading about* __]):

$$\begin{bmatrix} \text{FORM} & \langle \textit{like} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1}\begin{bmatrix}\text{NP} & \\ \text{GAP} & \langle \boxed{3}\text{NP}\rangle\end{bmatrix}, \boxed{2}\begin{bmatrix}\text{NP} & \\ \text{GAP} & \langle \boxed{3}\text{NP}\rangle\end{bmatrix}\right\rangle \\ \text{SYN} & \begin{bmatrix}\text{VAL} & \langle \boxed{1}, \boxed{2}\rangle \\ \text{GAP} & \langle\boxed{3}\text{NP}\rangle\end{bmatrix}\end{bmatrix}$$

In (136a), neither of the verb's valents contains a gap; hence the GAP value of both the subject and the object is the empty list, which in turn is registered as the verb's GAP value. The entire object NP is identified with the verb's GAP value in (136b), and in (136c) it is the GAP value of the object NP that is identified with the verb's GAP value. Finally, in (136d), the GAP values of the subject and object arguments are merged and identified with the verb's GAP value. This gives rise to (so-called) parasitic gaps, where two gaps are associated with a single filler. Note that in all cases shown in (136), the verb registers the information about what unbound gaps appear in its local syntactic context. This information is passed up to higher syntactic contexts by simple constraints. For example, in non-gap-binding constructs, a head daughter's GAP list must be the same as its mother's GAP list. Thus GAP specifications are inherited precisely as indicated in the structure shown in Figure 26.

There are three other features that play a role in the analysis of filler-gap constructions: WH, REL, and STORE. Specifications for all of these features percolate up through the filler daughter to provide an account of the 'pied-piping' phenomenon. For example, in (137), the filler NP *whose friend* has a nonempty WH value that percolated up from the interrogative *wh*-word it contains:

(137) {[**Whose** *suggestion*] [*do you think Kim likes?*]}

The percolation here obeys the same general constraints as those governing the feature GAP: when the daughters all have an empty specification for the

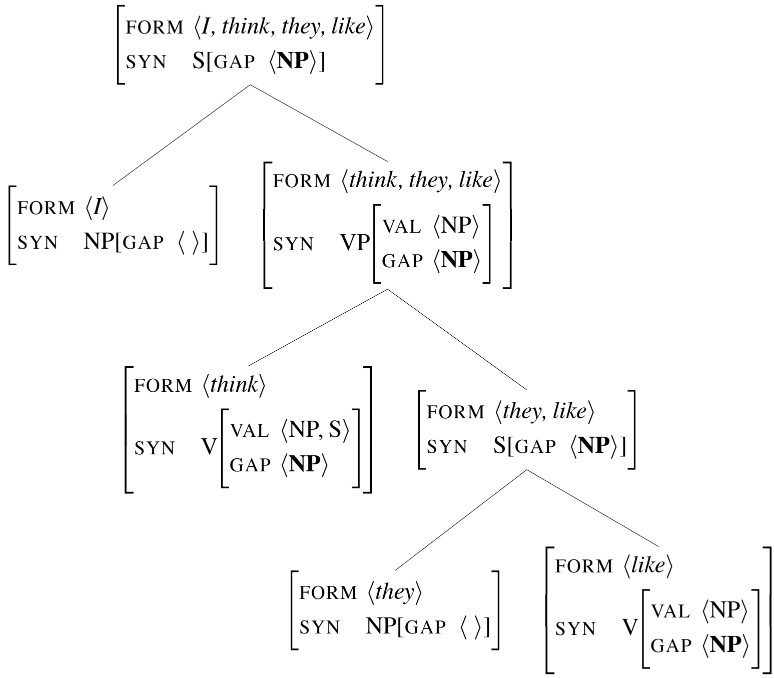
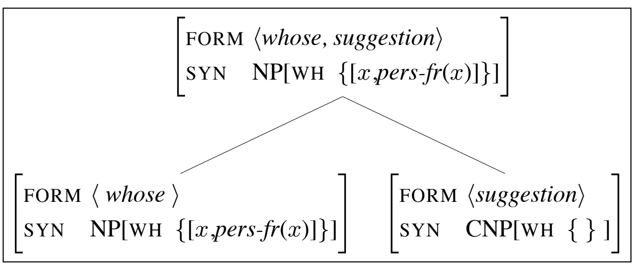


FIGURE 26 An Incomplete Derivation Showing 'Percolation' of GAP Specifications

feature in question, so does the mother; but when one of the daughters bears a nonempty specification (e.g. the GAP specification of each verb in Figure 26 or the WH specification of the word *whose* in the phrase *whose suggestion*, as shown in (138)), then the mother bears the same nonempty specification.[102]

(138)



[102]The expression '$[x, \textit{pers-fr}(x)]$' designates a parameter, described more fully in Ginzburg and Sag 2000. A parameter is here taken to be a pair consisting of an index and a restriction that must hold of its values. I use '*pers-fr*(x)' as a shorthand for '$\begin{bmatrix} \textit{person-fr} \\ \text{ENTITY} \;\; x \end{bmatrix}$'.

Just as nonempty specifications for GAP indicate the presence of an unbound gap, nonempty WH specifications mark the presence of an unbound interrogative or exclamative *wh*-word within the filler. Nonempty REL-specifications play a similar role, marking the presence of a relative *wh*-word ({[**whose** *mother*] [*I like* _]}) or *the*-word ({[**the** *more*] [*you read*]}).

Though space limitations prevent me from spelling out all the details of this analysis here (see Sag 2010a), the essentials of the treatment of *wh*-constructions can be sketched. The common properties of the various filler-gap clauses enumerated earlier are in part expressed in terms of the common construct type *filler-head-construct* (*filler-head-cxt*), whose instances are constrained by the following (nonmaximal) construction:[103]

(139) **Filler-Head Construction** (↑*headed-cxt*)**:**

$$\textit{filler-head-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X_1 \,!\, [\text{GAP } L]] \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SYN } X_2 \,!\, \begin{bmatrix}\text{WH} \\ \text{REL}\end{bmatrix} \\ \text{SEM } [\text{IND } \alpha] \\ \text{STORE } \Sigma \end{bmatrix}, H \right\rangle \\ \text{HD-DTR} & H : \begin{bmatrix} \text{SYN } X_1 : \begin{bmatrix} \text{CAT } \textit{verbal} \\ \text{GAP } \left\langle \begin{bmatrix} \text{SYN } X_2 \\ \text{SEM } [\text{IND } \alpha] \\ \text{STORE } \Sigma \end{bmatrix} \right\rangle \oplus L \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Filler-head constructs thus require exactly two daughters: a filler and a head daughter. The construction in (139) links the INDEX and STORE values of the filler and the filler's SYN value (except values for the features WH and REL) to the corresponding values of the first element of the head daughter's GAP list. This GAP element is in turn identified with a gap within the head daughter, in the manner just illustrated in Figure 26 and in (136). Any remaining elements on the head daughter's GAP list (members of the list *L*) must become part of the GAP list of the mother, which allows unbound gaps to be 'passed up' to a higher binder in the case of sentences with overlapping filler-gap dependencies, e.g. (140):

(140) [*Problems this hard*], I never know {[*who*] [to talk to _ about _]}.

The syntactic category of the head daughter (and hence that of its mother) is required to be *verbal*, which must resolve (see Figure 3 above) to one of its

[103] Here and throughout, Σ variables refer to sets of feature structures.

two subtypes, i.e. to *verb* or *complementizer*. Accordingly, the head daughter of a filler-gap construction must always be a verbal projection (S or VP) or a CP.

The theory of filler-gap constructions is articulated in terms of the construct hierarchy shown in Figure 27. Each type shown in this figure is associated with a construction that specifies the defining properties of a given class of constructs – the class of topicalized clauses (*top-cl*), the class of finite *wh*-relative clauses (*f-wh-rcl*), infinitival *wh*-relative clauses (*i-wh-rcl*), etc. Thus, the Filler-Head Construction enforces only those constraints that apply to all filler-gap clauses.[104]

One of the filler-gap constructions discussed in both Ginzburg and Sag 2000 and Sag 2010a (q.v.) is the Nonsubject *Wh*-Interrogative Construction, which I formulate here as (141):

(141) **Nonsubject *Wh*-Interrogative Construction** (↑*wh-int-cl*)

ns-wh-int-cl ⇒

$$
\left[\begin{array}{ll}
\text{MTR} & \left[\text{SEM}\left[\text{FRAMES}\ \left\langle\left[\begin{array}{ll}\text{PARAMS} & \{\pi,\ldots\}\\ \text{PROP} & l\end{array}\right]\right\rangle \oplus L\right]\right]\\
\text{DTRS} & \left\langle \left[\text{SYN}\left[\begin{array}{ll}\text{CAT} & \textit{nonverbal}\\ \text{WH} & \{\pi\}\end{array}\right]\right], \left[\begin{array}{ll}\text{SYN} & \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{INV} & X\\ \text{IC} & X\end{array}\right]\\ \text{VAL} & \langle\,\rangle\end{array}\right]\\ \text{SEM} & \left[\begin{array}{ll}\text{LTOP} & l\\ \text{FRAMES} & L\end{array}\right]\end{array}\right]\right\rangle
\end{array}\right]
$$

In constructs defined by (141), the head (second) daughter and the mother must include matching specifications for the features IC and INV.[105] Hence, it follows (for nonsubject *wh*-interrogatives) that an aux-initial head daughter is possible just in case the construct is an independent clause:

(142) a. {[Who] [will you visit __]}?

[104] The types *interrogative-clause* (*int-cl*), *relative-clause* (*rel-cl*), *exclamative-cl* (*excl-cl*), and *declarative-cl* (*decl-cl*) are motivated by the existence of general properties that are characteristic of each type of clause, e.g. properties that hold of *wh*- and non-*wh*-relatives alike.

[105] Since *ns-wh-int-cl* is a subtype of *wh-interrogative-clause* (*wh-int-cl*), which in turn is a subtype of *filler-head-cxt*, the identity of the SYN value of mother and head daughter is guaranteed by the Filler-Head Construction formulated in (139) above. Note that since *wh-int-cl* is also a subtype of of *interrogative-cl* (see Figure 27), constraint inheritance will ensure that instances of *ns-wh-int-cl* also have the general properties of interrogative clauses (see (146) below).

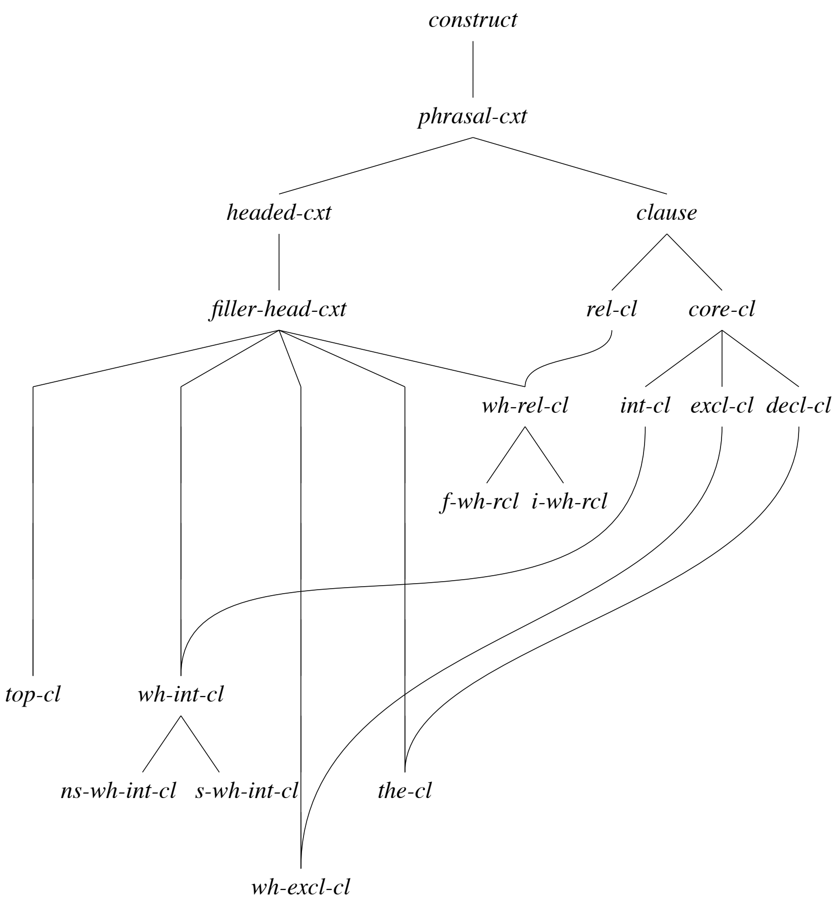


FIGURE 27 Hierarchy of Filler-Gap Clauses (after Sag 2010a)

b.*{[Who] [you will visit _]}?
c. They don't know {[who] [you will visit _]}.
d.*They don't know {[who] [will you visit _]}.

Moreover, (141) permits the range of filler constituents in *wh*-interrogatives to be quite broad – NP, PP, AP, and AdvP fillers are all possible:

(143) a. {[Whose suggestion] [do you like _]}?
b. {[To whom] [did you send the letter _]}?
c. {[How happy] [are they _]}?
d. {[How quickly] [do you think you can do that _]}?

Finally, (141) specifies an appropriate meaning for a *wh*-interrogative clause, based on the semantics developed in Ginzburg and Sag 2000. The basic components are a proposition (PROP), here determined by the LTOP of the head daughter and an associated set of parameters that must include the parameter (the π in the PARAMS value) of the interrogative *wh*-word within the filler daughter (e.g. the parameter $[x, \textit{pers-fr}(x)]$ introduced by *who* in (142a), *whose* in (143a), and *whom* in (143b)). The meaning of (144a), which Ginzburg and Sag render as (144b), is analyzed in terms of the equivalent list of resolved MRS expressions in (144c):[106]

(144) a. Who do you like?

b.
$$\begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \{[\text{x}, \textit{pers-fr}(x)]\} \\ \text{PROP} & \begin{bmatrix} \textit{proposition} & \\ \text{SIT} & s \\ \text{SOA} & \begin{bmatrix} \text{QUANTIFIERS} & \langle\,\rangle \\ \text{NUCLEUS} & \begin{bmatrix} \textit{like-rel} & \\ \text{EXPRNCR} & \textit{you} \\ \text{UNDGR} & x \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

c.
$$\left\langle \begin{bmatrix} \textit{question-fr} & \\ \text{PARAMS} & \{[\text{x}, \textit{pers-fr}(x)]\} \\ \text{PROP} & l \end{bmatrix}, \begin{bmatrix} \textit{like-fr} & \\ \text{LABEL} & l \\ \text{EXPRNCR} & \textit{you} \\ \text{UNDGR} & x \end{bmatrix} \right\rangle$$

The Nonsubject *Wh*-Interrogative Construction interacts with our earlier constraints governing headed constructs and filler-head constructs to license filler-head structures like the one shown in Figure 28.

[106]This general approach to the semantics of interrogatives is discussed in considerable detail in Ginzburg and Sag 2000.

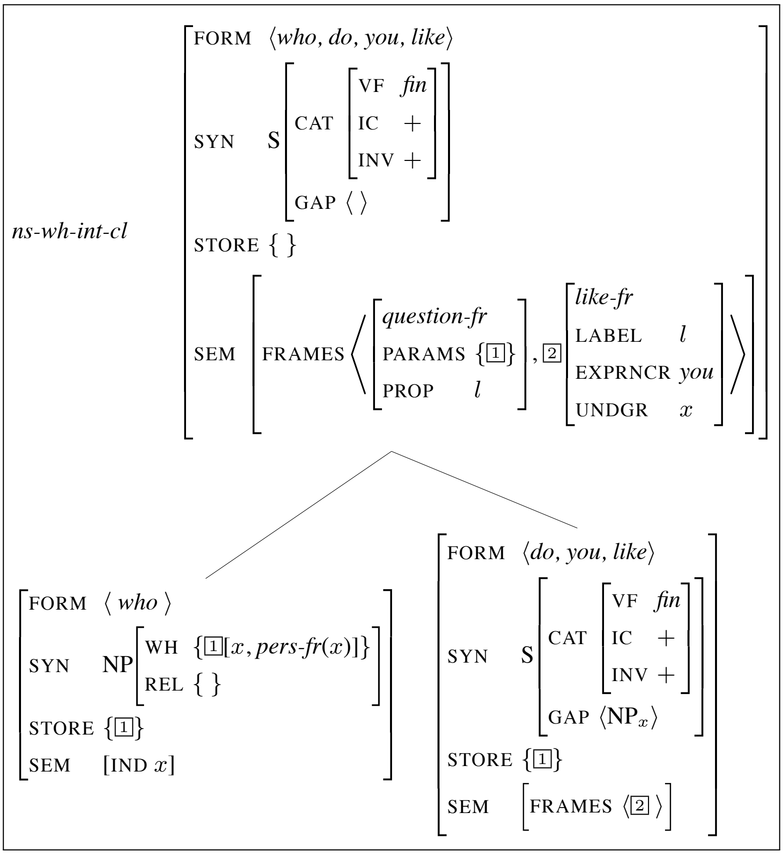


FIGURE 28 A Construct Licensed by the Nonsubject *Wh*-Interrogative Construction

This analysis relies crucially on the information specified in interrogative *wh*-words, which is illustrated in (145):

(145)
$$\begin{bmatrix} \text{FORM} & \langle \textit{who} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{noun} & \\ \text{SELECT} & \textit{none} \end{bmatrix} \\ \text{WH} & \Sigma : \{([x, \textit{pers-fr}(x)])\} \\ \text{REL} & \{\ \} \end{bmatrix} \\ \text{STORE} & \Sigma \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & x \\ \text{FRAMES} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$$

Each *wh*-interrogative word may thus introduce a parameter[107] into its WH value and in this case the parameter will also be in its STORE set. These 'stored' parameters are then 'passed up' through the analysis tree so that they can be integrated into the semantic composition at a higher level of analysis. The syntactic feature WH is of particular interest in regard to the 'pied-piped' interrogative clauses in (143). In these examples, the nonempty WH value provided by the filler-internal interrogative *wh*-word is passed up to the top of the filler expression (as shown in (138) above), where it must enter the semantic composition.

The constraint interaction between The Nonsubject *Wh*-Interrogative Construction and its superordinate constructions is quite subtle here. For example, the Interrogative Construction, which characterizes the general properties of all interrogative clauses, plays an important role:[108]

(146) **Interrogative Construction (↑*core-cl*):**

$$\textit{interrogative-cl} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{SEM} & \begin{bmatrix} \textit{question-fr} & \\ \text{PARAMS} & \Sigma_1 \\ \text{PROP} & \textit{proposition} \end{bmatrix} \\ \text{STORE} & \Sigma_2 \dot{-} \Sigma_1 \end{bmatrix} \\ \text{DTRS} & \textit{list}([\text{REL}\ \{\ \}]) \\ \text{HD-DTR} & [\text{STORE}\ \Sigma_2] \end{bmatrix}$$

[107] The specification '$([x, \textit{pers-fr}(x)])$' means that the indicated parameter is optional, i.e. that the set indicated in (145) may be either empty or singleton. It follows from Ginzburg and Sag's analysis that when the WH value is the empty set, it is required to be 'in situ', rather than in the filler daughter of the *wh*-interrogative clause.

[108] '$\dot{-}$' is a 'contained' set difference operation that removes elements from a set nonvacuously. That is, the result of '$\dot{-}$' is defined only if the elements to be removed are members of the set in question, i.e. if Σ_1 is a subset of Σ_2 in (146). Thus $\{x, y\} \dot{-} \{y\} = \{x\}$, but $\{x, y\} \dot{-} \{z\}$ is undefined.

First, the construction in (146) requires that each daughters' REL value be the empty set. This, taken together with Ginzburg and Sag's constraint requiring a lexical head to amalgamate the RELvalues of its syntactic arguments and pass the result up to its mother, prevents unbound relative *wh*-words from appearing anywhere within an interrogative clause; that is, relative *wh*-words are barred from appearing in situ. Second, according to (141), the parameter π in the filler daughter's WH value must be included in the mother's PARAMS set. And because of (146), this parameter must also be included in the head daughter's STORE value, but absent from the mother's STORE, i.e. contained in $\Sigma_0 - \Sigma_1$. That is (thinking in terms of traversing the tree from the 'bottom-up'), π and possibly some other parameters, associated with in situ interrogative *wh*-words, are retrieved from the head daughter's STORE value and the remaining parameters are passed up, becoming the mother's STORE value. As a consequence, the inheritance of stored parameters must proceed as shown in Figure 29. This analysis, worked out in detail in Ginzburg and Sag 2000, provides a comprehensive account of ambiguities like (147), first brought to light by C. L. Baker (Baker 1970):

(147) Who remembers where we bought what?

In fact, Ginzburg and Sag provide a comprehensive, construction-based treatment of the syntactic and semantic properties of interrogative and exclamative clauses in English that is executed in unusual detail.

10.1 What's X Doing Y?

The analysis of *wh*-interrogatives just sketched also provides a natural home for a treatment of the 'What's X Doing Y?' (WXDY) Construction, discussed by Kay and Fillmore (1999) and illustrated in (148):

(148) a. What are they doing being so polite to Bo and Pat?
'Why are they being so polite to Bo and Pat?'
b. What is your name doing in my book? (Kay and Fillmore 1999: 3)
'How come your name is in my book?'

The semantic fact of particular importance in Kay and Fillmore's discussion is the unexpected causal interpretation paraphrasable in terms of *why*, *how come* or *what is the reason that*, as indicated in (148).

The essential ingredients of WXDY, according to Kay and Fillmore are the following:

(149) a. an interrogative filler *what* participating in a *wh*-interrogative construction,
b. a form of the copula governing *doing*,

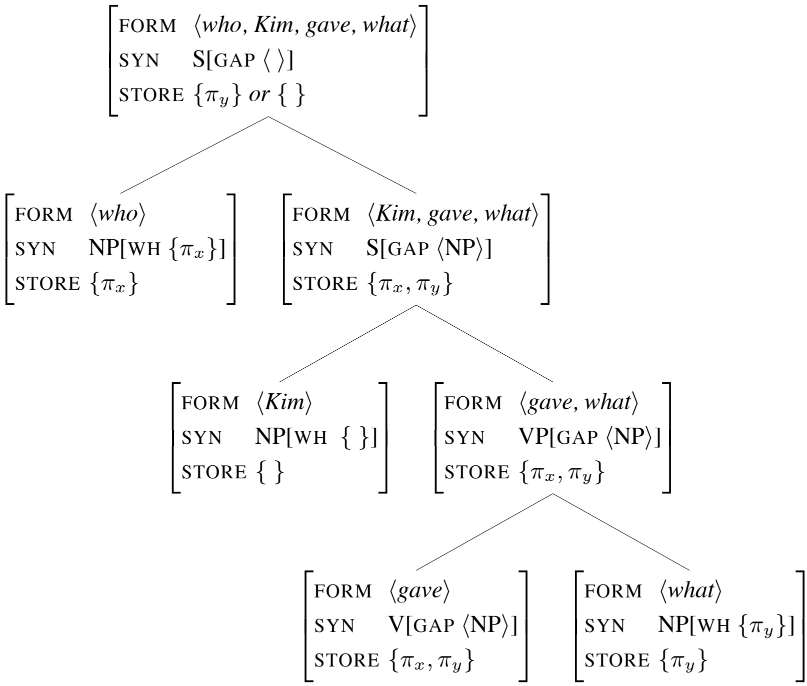


FIGURE 29 Stored Parameters in a Multiple *Wh*-Interrogative Derivation

c. a gap associated with the object of the progressive participle of the verb *do*,
d. a predicative XP following *doing*, forming a constituent with it,
e. the impossibility of negation, either of *be* or of *do*,
f. a causal interrogative semantics, and
g. a pragmatic attribution of incongruity of the proposition whose cause is being questioned.

These points are illustrated by the following examples:

(150) a. I wonder what the salesman will say this house is doing without a kitchen. (Kay and Fillmore 1999: 3)
b.*What does your name keep doing in my book?
c.*What will your name (be) do in my book?
d. What is he doing? (lacks WXDY semantics)
e.*What aren't they doing being so polite to Bo and Pat? [see glosses in (148)]
f.#What is he doing drunk, which everyone knew he would be?

Example (150a) is of particular importance, for it shows that the scope of the causal operator is not necessarily the same as the clause following the *what*. That is, though the position of *what* demarcates the top of the interrogative clause, it is the embedded structure *this house is doing without a kitchen* whose causality is to be explained by the salesman. (150a) does not mean 'I wonder why it is that the salesman will say that this house lacks a kitchen'.

WXDY finds a simple analysis within SBCG. Perhaps surprisingly, this analysis is purely lexical in nature.[109] First, in order to account for the role of *be* in WXDY, I posit a listeme like the following:

(151) $$\begin{bmatrix} \textit{copula-lxm} \\ \text{ARG-ST} \quad \left\langle X, \begin{bmatrix} \text{VP} \\ \text{LID} \quad \langle \textit{i-doing-fr} \rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

This listeme, which arguably inherits all of its remaining properties from the Copula Construction (a lexical class construction), selects a subject (X) and a VP complement whose LID is the idiomatic *i-doing-frame*. Like other copula *be*-lexemes, this is an auxiliary verb with subject-raising properties. And because its FRAMES list is empty, it makes no contribution to the semantics.

The lexicon contains only one listeme whose LID is *i-doing-fr*, and hence only one lexeme that gives rise to words that can head the VP complement of the *be* in (151). This listeme, because it exceptionally includes the specification [VF *prp*], will give rise to only one kind of word – present participles like the one sketched in Figure 30. The ARG-ST list of the verb in Figure 30 contains three elements: a subject ($\boxed{1}$) a direct object ($\boxed{2}$) and a predicational phrase ($\boxed{3}$). The direct object, however, is absent from the VAL list and present on the verb's GAP list. Moreover, this GAP element must be specified as [INDEX i] and [STORE $\{[i, \textit{thing-fr}(i)]\}$]. Since the gap's STORE and INDEX values are identified with those of the filler daughter in all Filler-Gap clauses (see (139) above), the filler daughter will also be so specified. A consequence of this is that the filler daughter must simply be *what*, since there is no other way that its INDEX can be identical to the index of the indicated parameter. Relevant lexical properties of interrogative *what* are shown in (152):

[109] An earlier version of this analysis was developed together with Susanne Riehemann.

$$\left[\begin{array}{ll}
\textit{word} & \\
\text{FORM} & \langle \textit{doing} \rangle \\
\text{SYN} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll} \textit{verb} & \\ \text{LID} & \langle \textit{i-doing-fr} \rangle \\ \text{VF} & \textit{prp} \end{array}\right] \\
\text{VAL} & \langle \boxed{1}, \boxed{3} \rangle \\
\text{GAP} & \left\langle \boxed{2} \left[\begin{array}{ll} \text{NP} & \\ \text{INDEX} & x \\ \text{STORE} & \{[x, \textit{thing-fr}(x)]\} \end{array}\right] \right\rangle
\end{array}\right] \\
\text{ARG-ST} & \left\langle \boxed{1}, \boxed{2}, \boxed{3} \left[\begin{array}{ll} \text{VP} & \\ \text{VAL} & \langle \boxed{1} \rangle \\ \text{LTOP} & l \end{array}\right] \right\rangle \\
\text{SEM} & \left[\begin{array}{ll}
\text{INDEX} & s \\
\text{FRAMES} & \left\langle \left[\begin{array}{ll} \textit{justification-fr} & \\ \text{EXPLICANS} & x \\ \text{EXPLICANDUM} & l \end{array}\right] \right\rangle
\end{array}\right]
\end{array}\right]$$

FIGURE 30 *doing*

(152)

$$\left[\begin{array}{ll}
\text{FORM} & \langle \textit{what} \rangle \\
\text{SYN} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll} \textit{noun} & \\ \text{SELECT} & \textit{none} \end{array}\right] \\
\text{STORE} & \Sigma \\
\text{WH} & \Sigma : \{[x, \textit{thing-fr}(x)]\} \\
\text{REL} & \{\ \}
\end{array}\right] \\
\text{SEM} & \left[\begin{array}{ll} \text{INDEX} & x \\ \text{FRAMES} & \langle\ \rangle \end{array}\right]
\end{array}\right]$$

Finally, by couching the semantics in terms of a *justification-frame*, we may be able to predict the pragmatic incongruity effect observed by Kay and Fillmore without further stipulation. Observe that the pragmatic effect of (153a) and (153b) seem comparable.[110]

(153) a. What is your name doing in my book?
b. What is the justification for your name being in my book?

[110] I thank Paul Kay for this observation.

In sum, the WXDY Construction is analyzed in terms of a *be* listeme that selects for a complement whose LID is *i-doing-fr*. Because words like the participle in Figure 30 are the only kind that mention *i-doing-fr*, these are the only words available to serve as lexical heads of the VP complement of *be* in WXDY. But since words like Figure 30 have a nonempty GAP list, they must appear at the bottom of a filler-gap dependency. Moreover, since the first member of their GAP list must have the STORE and INDEX properties of *what*, that filler-gap dependency must be a *wh*-interrogative clause (main or embedded) whose filler daughter is the word *what*. Figure 30 also links things together to produce a semantics asking about the justification for a certain proposition, where that proposition is constructed from the subject of *do* (which is also the subject of *be*) and its final complement. A construct illustrating WXDY, a FS of type *ns-wh-int-cl*, is shown in Figure 31.

11 Conclusion

In this chapter, I have attempted to present and explain the basic concepts of Sign-Based Construction Grammar without excessive formalization. At the same time, I have taken pains to illustrate how SBCG may be applied to a number of important grammatical problems that have been addressed in the literature – both the generative transformational literature and the now extensive bodies of work in both Construction Grammar and Head-Driven Phrase Structure Grammar. I believe that the approach sketched here will 'scale up' to provide consistent, comprehensive linguistic descriptions. In addition, I have tried to show that many ideas developed within the Construction Grammar tradition fit naturally within SBCG. Any treatise that tries to achieve such a goal is by necessity programmatic in nature, and mine is no exception. But the initial results seem promising. The other chapters in this volume, I hope, will convince the reader that there is every reason to believe that SBCG provides a natural framework in which to develop a construction-based theory of grammar that is both descriptively and theoretically satisfying.

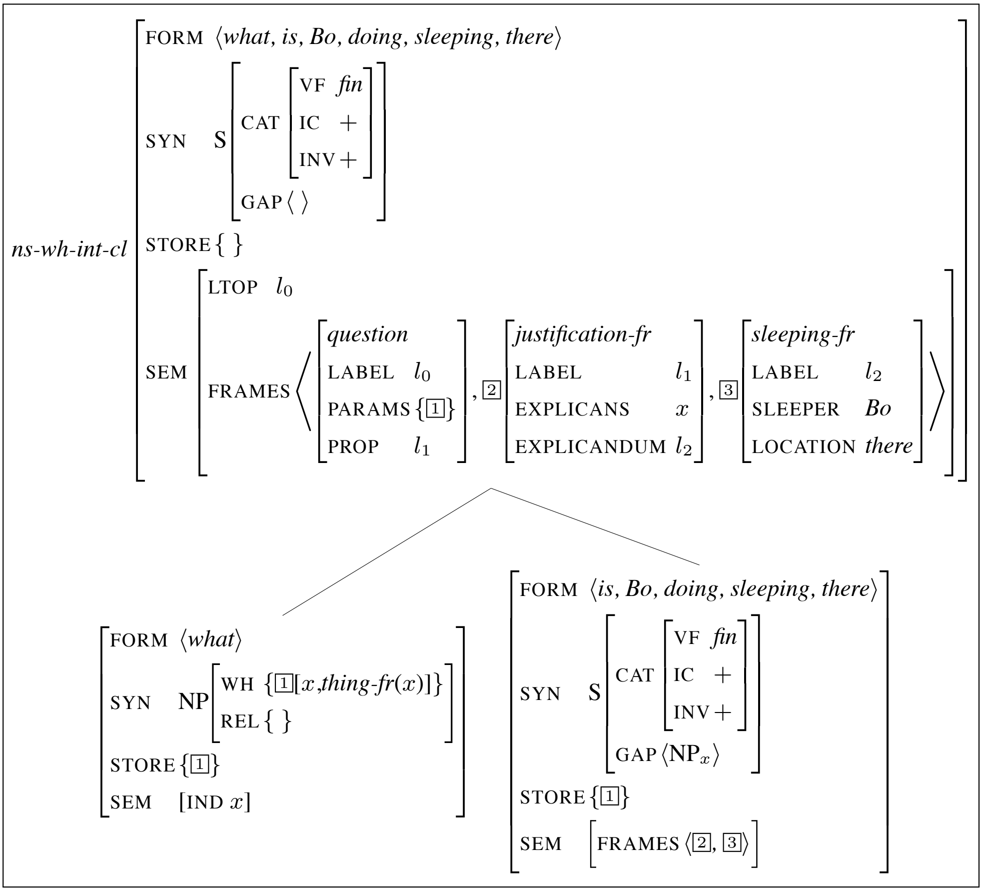
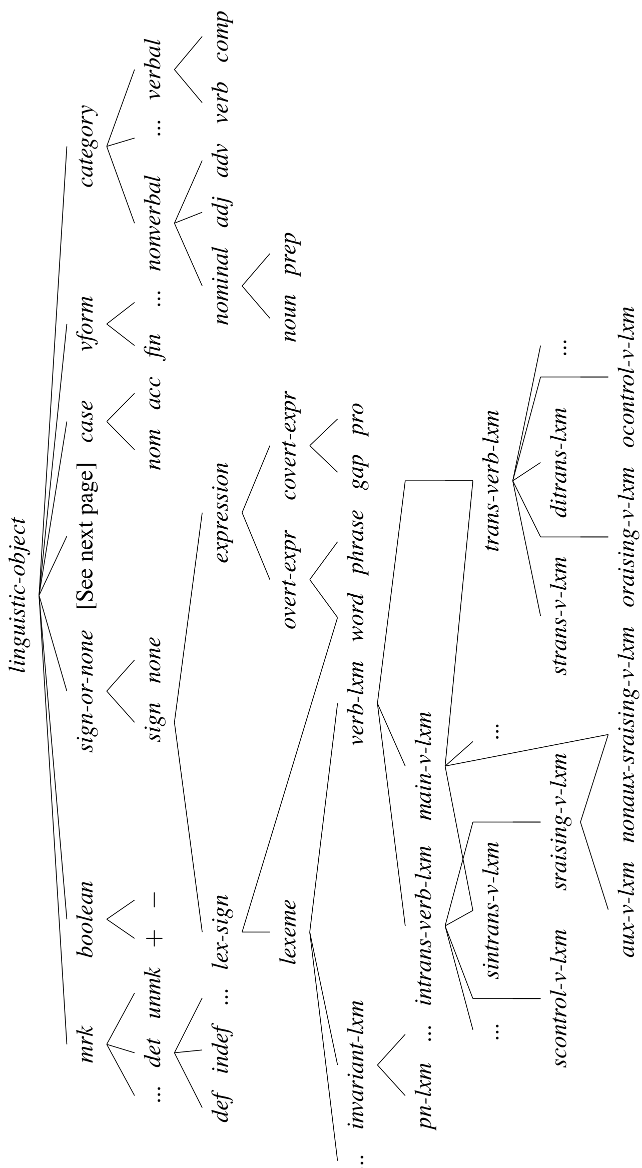
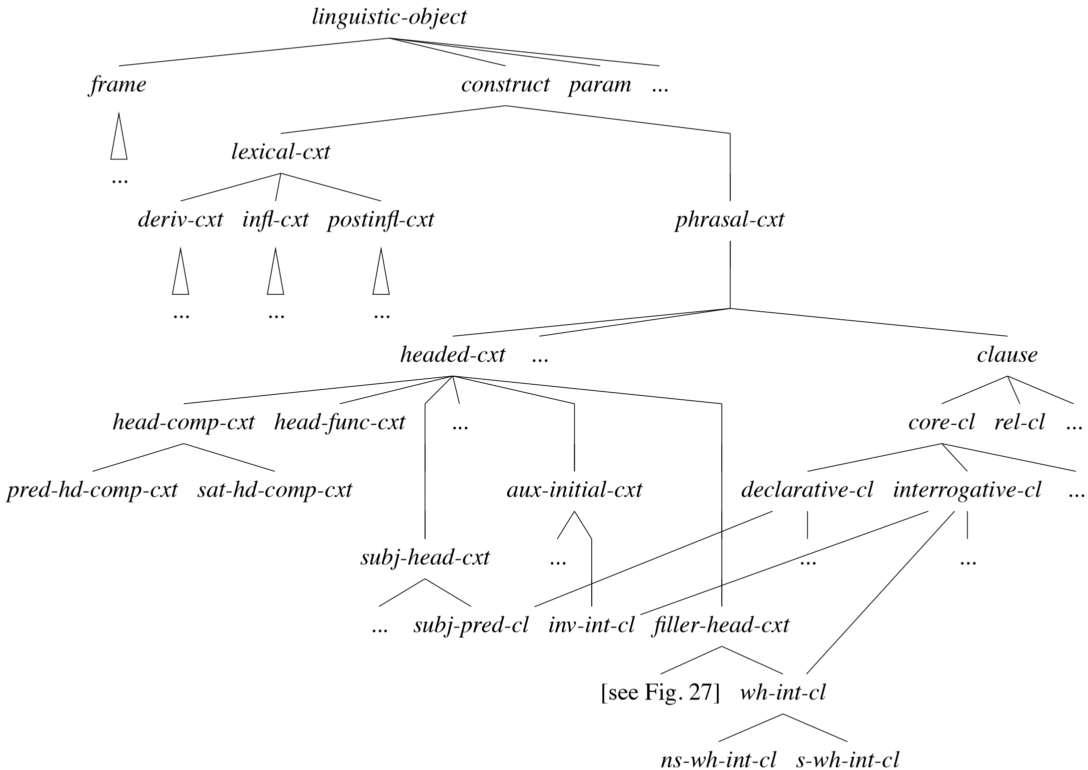


FIGURE 31 A Construct Illustrating the WXDY Phenomenon

Appendix: English Grammar

A1: Grammar Signature: A Partial Type Hierarchy




linguistic-object
frame
construct
param
...
lexical-cxt
deriv-cxt
infl-cxt
postinfl-cxt
phrasal-cxt
headed-cxt
clause
head-comp-cxt
head-func-cxt
core-cl
rel-cl
pred-hd-comp-cxt
sat-hd-comp-cxt
aux-initial-cxt
declarative-cl
interrogative-cl
subj-head-cxt
subj-pred-cl
inv-int-cl
filler-head-cxt
[see Fig. 27]
wh-int-cl
ns-wh-int-cl
s-wh-int-cl

A2: Grammar Signature: Some Type Declarations

$$\textit{sign}: \begin{bmatrix} \text{PHON} & \textit{phon-obj} \\ \text{FORM} & \textit{morph-obj} \\ \text{SYN} & \textit{syn-obj} \\ \text{SEM} & \textit{sem-obj} \\ \text{CNTXT} & \textit{context-obj} \\ \text{STORE} & \textit{set(param)} \end{bmatrix}$$

$$\textit{lex-sign}: \begin{bmatrix} \text{ARG-ST} & \textit{list(expression)} \end{bmatrix}$$

$$\textit{context-obj}: \begin{bmatrix} \text{C-INDS} & \textit{contextual-index} \\ \text{BCKGRND} & \textit{list(proposition)} \end{bmatrix}$$

$$\textit{sem-obj}: \begin{bmatrix} \text{LTOP} & \textit{label} \\ \text{IND} & \textit{index} \\ \text{FRAMES} & \textit{list(frame)} \end{bmatrix}$$

$$\textit{syn-obj}: \begin{bmatrix} \text{CAT} & \textit{category} \\ \text{VAL} & \textit{list(expression)} \\ \text{MRKG} & \textit{mark} \\ \text{GAP} & \textit{list(expression)} \\ \text{WH} & \textit{set(expression)} \\ \text{REL} & \textit{set(expression)} \end{bmatrix}$$

$$\textit{contextual-index}: \begin{bmatrix} \text{SPKR} & \textit{index} \\ \text{ADDR} & \textit{index} \\ \text{UTT-LOC} & \textit{index} \\ \ldots & \end{bmatrix}$$

$$\textit{construct}: \begin{bmatrix} \text{MTR} & \textit{sign} \\ \text{DTRS} & \textit{nelist(sign)} \\ \text{CXT-CONTENT} & \textit{list(frame)} \end{bmatrix}$$

$$\textit{lex-cxt}: \begin{bmatrix} \text{DTRS} & \textit{list(lex-sign)} \end{bmatrix}$$

$$\textit{infl-cxt}: \begin{bmatrix} \text{MTR} & \textit{word} \\ \text{DTRS} & \textit{list(lexeme)} \end{bmatrix}$$

$$\textit{pinfl-cxt}: \begin{bmatrix} \text{MTR} & \textit{word} \\ \text{DTRS} & \textit{list(word)} \end{bmatrix}$$

$$\textit{phr-cxt}: \begin{bmatrix} \text{MTR} & \textit{phrase} \\ \text{DTRS} & \textit{list(overt-expr)} \end{bmatrix}$$

$$\textit{headed-cxt}: \begin{bmatrix} \text{HD-DTR} & \textit{overt-expr} \end{bmatrix}$$

$$\textit{category}: \begin{bmatrix} \text{SELECT} & \textit{sign-or-none} \\ \text{XARG} & \textit{sign-or-none} \\ \text{LID} & \textit{list(frame)} \end{bmatrix}$$

$$\textit{verbal}: \begin{bmatrix} \text{VF} & \textit{vform} \\ \text{IC} & \textit{boolean} \end{bmatrix}$$

$$\textit{verb}: \begin{bmatrix} \text{AUX} & \textit{boolean} \\ \text{INV} & \textit{boolean} \end{bmatrix}$$

$$\textit{noun}: \begin{bmatrix} \text{CASE} & \textit{case} \end{bmatrix}$$

A3: Some Lexical-Class Constructions

Proper Noun Construction (↑*invariant-lxm*):

$$
\textit{pn-lxm} \Rightarrow \left[\begin{array}{ll}
\text{FORM} & L \\
\text{SYN} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{noun}} \\ \text{SELECT} & \textit{none} \\ \text{XARG} & \textit{none} \end{array}\right] \\
\text{VAL} & \langle\,\rangle \\
\text{MRKG} & \textit{def}
\end{array}\right] \\
\text{SEM} & \left[\begin{array}{ll} \text{IND} & i \\ \text{FRAMES} & \langle\,\rangle \end{array}\right] \\
\text{CNTXT} & \left[\begin{array}{ll} \text{BCKGRND} & \left\langle \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{naming-fr}} \\ \text{ENTITY} & i \\ \text{NAME} & L \end{array}\right] \right\rangle \end{array}\right]
\end{array}\right] \quad (= \text{Fig. 9})
$$

Verb Lexeme Construction (↑*lexeme*):

$$
\textit{verb-lxm} \Rightarrow \left[\begin{array}{ll}
\text{ARG-ST} & \langle X\,, \ldots \rangle \\
\text{SYN} & \left[\begin{array}{ll}
\text{CAT} & \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{verb}} \\ \text{LID} & L \\ \text{SELECT} & \textit{none} \\ \text{XARG} & X \end{array}\right] \\
\text{MRKG} & \textit{unmk}
\end{array}\right] \\
\text{SEM} & \left[\begin{array}{ll} \text{LTOP} & l_{0=_q 1} \\ \text{FRAMES} & L : \langle ([\text{LABEL} \;\; l_1]) \rangle \end{array}\right]
\end{array}\right] \quad (= (52))
$$

Transitive Locative Construction: (↑*trans-verb-lxm*):

$$
\textit{trans-loc-v-lxm} \Rightarrow \left[\begin{array}{ll}
\text{ARG-ST} & \left\langle \text{NP}_x\,, \text{NP}_y\,, \left[\begin{array}{l} \text{PP}_s[\textit{dir}] \\ \text{VAL} \; \langle \textit{pro}_y \rangle \end{array}\right] \right\rangle \\
\text{SEM} & \left[\begin{array}{ll} \text{FRAMES} & \left\langle \left[\begin{array}{ll} \multicolumn{2}{l}{\textit{loc-motion-fr}} \\ \text{AGENT} & x \\ \text{THEME} & y \\ \text{PATH} & s \end{array}\right] \right\rangle \end{array}\right]
\end{array}\right] \quad (= (96a))
$$

Applicative Construction: (↑*trans-verb-lxm*):

$$\textit{trans-with-v-lxm} \Rightarrow \begin{bmatrix} \text{ARG-ST} & \langle \text{NP}_x, \text{NP}_z, \text{PP}_s[\textit{with}] \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{loc-with-fr} & \\ \text{AGENT} & x \\ \text{THEME} & z \\ \text{MEANS} & s \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \quad (= \ (96b))$$

A4: Some Listemes:

$$\begin{bmatrix} \textit{pn-lxm} & \\ \text{FORM} & \langle \textit{Kim} \rangle \end{bmatrix} (= \ (49)) \qquad \begin{bmatrix} \textit{cn-lxm} & \\ \text{FORM} & \langle \textit{book} \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \langle [\textit{book-fr}] \rangle \end{bmatrix} \end{bmatrix} \quad (= \ (23))$$

$$\begin{bmatrix} \textit{sintrans-v-lxm} & \\ \text{FORM} & \langle \textit{laugh} \rangle \\ \text{SEM} & [\text{FRAMES} \ \ \langle [\textit{laughing-fr}] \rangle \ ] \end{bmatrix} \quad (= \ (50))$$

$$\begin{bmatrix} \textit{strans-v-lxm} & \\ \text{FORM} & \langle \textit{love} \rangle \\ \text{SEM} & [\text{FRAMES} \ \ \langle \textit{loving-fr} \rangle \ ] \end{bmatrix}$$

$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \text{FORM} & \langle \textit{give} \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \langle [\textit{giving-fr}] \rangle \end{bmatrix} \end{bmatrix} \quad (= (90a))$$

$$\begin{bmatrix} \textit{trans-verb-lxm} & \\ \text{FORM} & \langle \textit{spray} \rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \langle [\textit{spray-fr}] \rangle \end{bmatrix} \end{bmatrix} \quad (= \ (93))$$

$$\begin{bmatrix} \textit{aux-v-lxm} & \\ \text{FORM} & \langle \textit{can} \rangle \\ \text{ARG-ST} & \left\langle \text{X}, \begin{bmatrix} \text{VP} & \\ \text{VF} & \textit{base} \end{bmatrix} \right\rangle \\ \text{SYN} & [\text{CAT} \ \ [\text{VF} \ \ \textit{fin}]] \\ \text{SEM} & [\text{FRAMES} \ \ \langle [\textit{can-fr}] \rangle] \end{bmatrix} \quad (= \ (123a))$$

$$\begin{bmatrix} \textit{nonaux-sraising-v-lxm} & \\ \text{FORM} & \langle \textit{continue} \rangle \\ \text{ARG-ST} & \left\langle \text{X}, \begin{bmatrix} \text{VP} & \\ \text{VF} & \textit{prp} \end{bmatrix} \right\rangle \\ \text{SEM} & [\text{FRAMES } \langle [\textit{continuing-fr}] \rangle] \end{bmatrix} \quad (=\ (123\text{b}))$$

$$\begin{bmatrix} \textit{pseudo-trans-v-lxm} & \\ \text{FORM} & \langle \textit{kick} \rangle \\ \text{ARG-ST} & \left\langle X_i, \begin{bmatrix} \text{NP} & \\ \text{LID} & \langle \textit{i-bucket-fr} \rangle \\ \text{MRKG} & \textit{i-the} \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{kicking}^{dying}\textit{-fr} & \\ \text{PROTAGONIST} & i \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \quad (=\ (75))$$

$$\begin{bmatrix} \textit{cn-lxm} & \\ \text{FORM} & \langle \textit{bucket} \rangle \\ \text{SYN} & [\text{CAT } [\text{LID } \langle \textit{i-bucket-fr} \rangle]] \\ \text{SEM} & \begin{bmatrix} \text{IND} & \textit{none} \\ \text{FRAMES} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \quad (=\ (76))$$

$$\begin{bmatrix} \textit{strans-v-lxm} & \\ \text{FORM} & \langle \textit{lose} \rangle \\ \text{ARG-ST} & \left\langle X_i\ , \begin{bmatrix} \text{NP}_j & \\ \text{XARG} & [\textit{pron}]_i \\ \text{LID} & \langle \textit{i-cool}^{composure}\textit{-fr} \rangle \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \textit{losing-fr} & \\ \text{AGENT} & i \\ \text{ENTITY} & j \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \quad (=\ (80))$$

$$\begin{bmatrix} \text{FORM} & \langle \textit{who} \rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{noun} & \\ \text{SELECT} & \textit{none} \end{bmatrix} \\ \text{WH} & \Sigma : \{([x, \textit{pers-fr}(x)])\} \\ \text{REL} & \{\ \} \end{bmatrix} \\ \text{STORE} & \Sigma \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & x \\ \text{FRAMES} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \quad (=\ (145))$$

$$\begin{bmatrix} \textit{copula-lxm} & \\ \text{ARG-ST} & \left\langle X, \begin{bmatrix} \text{VP} & \\ \text{LID} & \langle \textit{i-doing-fr} \rangle \end{bmatrix} \right\rangle \end{bmatrix} \quad (=\ (151))$$

$$\begin{bmatrix} \textit{excp-do-lxm} & \\ \text{FORM} & \langle \textit{do} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1}, \boxed{2}, \boxed{3} \begin{bmatrix} \text{VP} & \\ \text{VAL} & \langle \boxed{1} \rangle \\ \text{LTOP} & l \end{bmatrix} \right\rangle \\ \text{SYN} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{verb} & \\ \text{LID} & \langle \textit{i-doing-fr} \rangle \\ \text{VF} & \textit{prp} \end{bmatrix} \\ \text{VAL} & \langle \boxed{1}, \boxed{3} \rangle \\ \text{GAP} & \left\langle \boxed{2} \begin{bmatrix} \text{NP} & \\ \text{INDEX} & i \\ \text{STORE} & \{[i, \textit{thing-fr}(i)]\} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & s \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{justification-fr} & \\ \text{EXPLICANS} & x \\ \text{EXPLICANDUM} & l \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

A5: Some Combinatoric Constructions:

Principle of Compositionality:

$$\textit{construct} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SEM [FRAMES } L_0 \oplus \ldots \oplus L_n]] \\ \text{DTRS} & \langle [\text{SEM [FRAMES } L_1]], \ldots, [\text{SEM [FRAMES } L_n]] \rangle \\ \text{CXT-CONTENT} & L_0 \end{bmatrix}$$

Inflectional Construction (↑*lex-cxt*):

$$\textit{infl-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{ARG-ST} & L \\ \text{CNTXT} & X \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{ARG-ST} & L \\ \text{CNTXT} & X \end{bmatrix} \right\rangle \end{bmatrix}$$

Preterite Construction (↑*infl-cxt*):

$$\textit{preterite-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{FORM} & \langle \mathbf{F}_{pret}(X) \rangle \\ \text{SYN} & Y : [\text{CAT [VF } \textit{fin}]] \\ \text{SEM} & \begin{bmatrix} \text{IND} & s \\ \text{LTOP} & l_{2 =_q 0} \\ \text{FRAMES} & \left\langle \begin{bmatrix} \textit{some-fr} & \\ \text{LBL} & l_0 \\ \text{BV} & s \\ \text{RESTR} & l_1 \end{bmatrix}, \begin{bmatrix} \textit{past-fr} & \\ \text{LBL} & l_1 \\ \text{ARG} & s \end{bmatrix} \right\rangle \oplus L \end{bmatrix} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{FORM} & \langle X \rangle \\ \text{ARG-ST} & \langle \text{NP}[\textit{nom}], \ldots \rangle \\ \text{SYN} & Y \\ \text{SEM} & \begin{bmatrix} \text{IND} & s \\ \text{LTOP} & l_2 \\ \text{FRAMES} & L \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \quad (= (57))$$

Zero Inflection Construction (↑*infl-cxt*):

$$\textit{zero-infl-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & X \; ! \; \textit{word} \\ \text{DTRS} & \langle X : \textit{invariant-lxm} \rangle \end{bmatrix} \quad (= (60))$$

***Un*-Verb Construction (↑*deriv-cxt*):**

$$
\textit{un-verb-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{FORM} & \langle \mathbf{F}_{un}(X) \rangle \\ \text{ARG-ST} & L_1 \\ \text{SYN} & Y \\ \text{SEM} & [\text{FRAMES } L_2 \oplus \ \ldots] \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \textit{strans-v-lxm} & \\ \text{FORM} & \langle X \rangle \\ \text{ARG-ST} & L_1 \\ \text{SYN} & Y \\ \text{SEM} & \begin{bmatrix} \text{FRAMES} & L_2 \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \quad (= (62))
$$

Negative Auxiliary Construction (↑*post-infl-cxt*):

$$
\textit{neg-aux-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{FORM} & \langle W \rangle \\ \text{SYN} & X \, ! \begin{bmatrix} \text{CAT} & Y \, ! \, [\text{AUX } -] \\ \text{VAL} & \langle \text{Adv}[\textit{neg}] \rangle \oplus L \end{bmatrix} \\ \text{SEM} & Z \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{FORM} & \langle W \rangle \\ \text{SYN} & X : \begin{bmatrix} \text{CAT} & Y : \begin{bmatrix} \text{AUX} & + \\ \text{VF} & \textit{fin} \end{bmatrix} \\ \text{VAL} & L \end{bmatrix} \\ \text{SEM} & Z \end{bmatrix} \right\rangle \end{bmatrix} \quad (= \text{Fig. } 16)
$$

Verb-*Way* Construction (↑*deriv-cxt*):

$$
\textit{verb-way-cxt} \Rightarrow \begin{bmatrix}
\text{MTR} & \begin{bmatrix}
\multicolumn{2}{l}{\textit{derived-intrans-v-lxm}} \\
\text{FORM} & \langle X \rangle \\
\text{ARG-ST} & \left\langle Z, \begin{bmatrix} \text{NP} \\ \text{LID } \langle \textit{i-way-fr} \rangle \\ \text{XARG } \textit{pron}_i \end{bmatrix}, \begin{bmatrix} \text{PP}[\textit{dir}] \\ \text{VAL } \langle \textit{pron} \rangle \\ \text{IND } \ s_2 \end{bmatrix} \right\rangle \\
\text{SYN} & Y \\
\text{SEM} & \begin{bmatrix} \text{IND} & s_1 \\ \text{FRAMES} & \left\langle \begin{bmatrix} \multicolumn{2}{l}{\textit{going-fr}} \\ \text{THEME} & i \\ \text{PATH} & s_2 \\ \text{ME/MA} & s_1 \end{bmatrix} \right\rangle \oplus L \end{bmatrix}
\end{bmatrix} \\
\text{DTRS} & \left\langle \begin{bmatrix}
\text{FORM} & \langle X \rangle \\
\text{ARG-ST} & \langle Z \rangle \oplus \langle (\text{NP}) \rangle \\
\text{SYN} & Y : [\text{CAT } \textit{verb}] \\
\text{SEM} & \begin{bmatrix} \text{IND} & s_1 \\ \text{FRAMES} & L \end{bmatrix}
\end{bmatrix} \right\rangle
\end{bmatrix} \quad (= \text{ Fig. } 19)
$$

Subject-Predicate Construction (↑*subj-head-cxt*):

$$
\textit{subj-pred-cl} \Rightarrow \begin{bmatrix}
\text{MTR} & [\text{SYN } Y \,!\, [\text{VAL } \langle\,\rangle\,]\,] \\
\text{DTRS} & \left\langle X,\ Z : \begin{bmatrix} \text{SYN } Y : \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{VF} & \textit{fin} \\ \text{INV} & - \\ \text{AUX} & - \end{bmatrix} \\ \text{MRKG} & \textit{unmk} \\ \text{VAL} & \langle X \rangle \end{bmatrix} \end{bmatrix} \right\rangle \\
\text{HD-DTR} & Z
\end{bmatrix} \quad (= (105))
$$

Predicational Head-Complement Construction (↑*headed-cxt*):

$$
\textit{pred-hd-comp-cxt} \Rightarrow \begin{bmatrix}
\text{MTR} & [\text{SYN } X \,!\, [\text{VAL } \langle Y \rangle]] \\
\text{DTRS} & \langle Z \rangle \oplus L : \textit{nelist} \\
\text{HD-DTR} & Z : \begin{bmatrix} \textit{word} \\ \text{SYN } X : \begin{bmatrix} \text{CAT} & [\text{XARG } Y] \\ \text{VAL} & \langle Y \rangle \oplus L \end{bmatrix} \end{bmatrix}
\end{bmatrix} \quad (= (112))
$$

Saturational Head-Complement Construction (↑*headed-cxt*):

$$\textit{sat-hd-comp-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X \,!\, [\text{VAL } \langle\,\rangle]] \\ \text{DTRS} & \langle Z\rangle \oplus L{:}\textit{nelist} \\ \text{HD-DTR} & Z : \begin{bmatrix} \textit{word} & \\ \text{SYN} & X : \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{prep} & \\ \text{XARG} & \textit{none} \end{bmatrix} \\ \text{VAL} & L \end{bmatrix} \end{bmatrix} \end{bmatrix} \quad (= (113))$$

Head-Functor Construction: (↑*headed-cxt*):

$$\textit{head-func-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X \,!\, [\text{MRKG } M\,]] \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{CAT} & [\text{SELECT } Y\,] \\ \text{MRKG} & M \end{bmatrix} \end{bmatrix}, \; Y : [\text{SYN } X] \right\rangle \\ \text{HD-DTR} & Y \end{bmatrix} \quad (= (117))$$

Aux-Initial Construction: (↑*headed-cxt*):

$$\textit{aux-initial-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{S} & \\ \text{VAL} & \langle\,\rangle \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{V} & \\ \text{AUX} & + \\ \text{INV} & + \\ \text{VAL} & L \end{bmatrix} \right\rangle \oplus L \end{bmatrix}$$

Filler-Head Construction (↑*headed-cxt*):

$$\textit{filler-head-cxt} \Rightarrow \begin{bmatrix} \text{MTR} & [\text{SYN } X_1 \,!\, [\text{GAP } L\,]] \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{SYN } X_2 \,! \begin{bmatrix} \text{WH} \\ \text{REL} \end{bmatrix} \\ \text{SEM } [\text{IND } \alpha] \\ \text{STORE } \Sigma \end{bmatrix}, H \right\rangle \\ \text{HD-DTR} & H : \begin{bmatrix} \text{SYN} & X_1 : \begin{bmatrix} \text{CAT} & \textit{verbal} \\ \text{GAP} & \left\langle \begin{bmatrix} \text{SYN} & X_2 \\ \text{SEM} & [\text{IND } \alpha] \\ \text{STORE} & \Sigma \end{bmatrix} \right\rangle \oplus L \end{bmatrix} \end{bmatrix} \end{bmatrix} \quad (= (139))$$

Interrogative Construction (↑*core-cl*):

$$\textit{interrogative-cl} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{SEM} & \begin{bmatrix} \textit{question} & \\ \text{PARAMS} & \Sigma_1 \\ \text{PROP} & \textit{proposition} \end{bmatrix} \\ \text{STORE} & \Sigma_2 \dot{-} \Sigma_1 \end{bmatrix} \\ \text{DTRS} & \textit{list}([\text{REL}\ \{\ \}]) \\ \text{HD-DTR} & [\text{STORE}\ \Sigma_2] \end{bmatrix} \quad (=\ (146))$$

Nonsubject *Wh*-Interrogative Construction (↑*wh-int-cl*)

$$\textit{ns-wh-int-cl} \Rightarrow \begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{SEM} & \begin{bmatrix} \text{FRAMES} & \left\langle \begin{bmatrix} \text{PARAMS} & \{\pi, \ldots\} \\ \text{PROP} & l \end{bmatrix} \right\rangle \oplus L \end{bmatrix} \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{CAT} & \textit{nonverbal} \\ \text{WH} & \{\pi\} \end{bmatrix}, \begin{bmatrix} \text{INV} & X \\ \text{IC} & X \\ \text{VAL} & \langle\ \rangle \\ \text{LTOP} & l \\ \text{FRAMES} & L \end{bmatrix} \right\rangle \end{bmatrix} \quad (=\ (141))$$

References

Abeillé, Anne, Olivier Bonami, Danièle Godard, and Jesse Tseng. 2006. The Syntax of French *à* and *de*: An HPSG Analysis. In P. Saint-Dizier, ed., *Syntax and Semantics of Prepositions*, pages 147–162. Dordrecht: Springer.

Ackerman, Farrell and Gert Webelhuth. 1998. *A Theory of Predicates*. Stanford: CSLI Publications.

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1:1–27.

Allegranza, Valerio. 1998b. Determination and Quantification. In *Linguistic Specifications for Typed Feature Structure Formalisms*, vol. 10 of *Studies in Machine Translation and Natural Language Processing*, pages 281–314. Luxembourg: Official Publications of the European Communities.

Allegranza, Valerio. 2007. *The Signs of Determination. Constraint-Based Modelling Across Languages*, vol. 16 of *SABEST: Saarbrücker Beiträge zur Sprach- und Translationswissenschaft*. Frankfurt am Main: Peter Lang.

Anderson, Stephen. 1992. *Amorphous Morphology*. Cambridge: Cambridge University Press.

Andrews, Avery. 1982. The Representation of Case in Modern Icelandic. In J. Bresnan, ed., *The Mental Representation of Grammatical Relations*, pages 427–503. Cambridge, MA: The MIT Press.

Andrews, Avery. 1990. Case Structures and Control in Modern Icelandic. In J. Maling and A. Zaenen, eds., *Modern Icelandic Syntax*, pages 187–234. New York: Academic Press.

Asudeh, Ash. 2002. Richard III. In M. Andronis, E. Debenport, A. Pycha, and K. Yoshimura, eds., *Proceedings of the 38th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.

Asudeh, Ash and Ewan Klein. 2002. Shape Conditions and Phonological Context. In F. V. Eynde, L. Hellan, and D. Beermann, eds., *The Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, pages 20–30. Stanford: CSLI Publications.

Baker, C. L. 1970. *Indirect Questions in English*. Ph.D. thesis, University of Illinois at Urbana–Champaign.

Barðdal, Jóhanna. 2011. Lexical vs. Structural Case: A False Dichotomy. *Morphology* 21(3-4):619–659.

Barwise, Jon and John Perry. 1983. *Situations and Attitudes*. Bradford Books. Cambridge, MA: MIT Press. Reprinted in 1999 in the David Hume Series. Stanford: CSLI Publications.

Barwise, Jon and John Perry. 1985. Shifting Situations and Shaken Attitudes. *Linguistics and Philosophy* 8:399–452.

Bauer, Laurie. 1983. *English Word-Formation*. Cambridge: Cambridge University Press.

Bauer, Laurie and Antoinette Reynaud. 2001. A Corpus-based Study of Compounding in English. *Journal of English Linguistics* 29:101–123.

Bender, Emily and Daniel P. Flickinger. 1999. Peripheral Constructions and Core Phenomena: Agreement in Tag Questions. In G. Webelhuth, J.-P. Koenig, and

A. Kathol, eds., *Lexical and Constructional Aspects of Linguistic Explanation*, pages 199–214. Stanford: CSLI Publications.

Beyssade, Claire and Jean-Marie Marandin. 2007. French Intonation and Attitude Attribution. In P. Denis, E. McCready, A. Palmer, and B. Reese, eds., *The Proceedings of the 2004 Texas Linguistic Society Conference: Issues at the Semantics-Pragmatics Interface*, pages 1–12. Somerville, MA: Cascadilla Proceedings Project.

Bird, Steven and Ewan Klein. 1994. Phonological Analysis in Typed Feature Systems. *Computational Linguistics* 20(3):455–491.

Blackburn, P. and J. Bos. 2005. *Representation and Inference for Natural Language: A First Course in Computational Semantics*. CSLI Studies in Computational Linguistics (ed. Ann Copestake). Stanford: CSLI Publications.

Blevins, James P. 2008. The Post-Transformational Enterprise. *Journal of Linguistics* 44:723–742.

Boas, Hans C. 2003. *A Constructional Approach to Resultatives*. Stanford: CSLI Publications.

Boas, Hans C. 2005. Frame Semantics as Interlingual Representations for Multilingual Lexical Databases. *International Journal of Lexicography* 18(4):445–478.

Boas, Hans C. 2008a. Determining the Structure of Lexical Entries and Grammatical Constructions in Construction Grammar. *Annual Review of Cognitive Linguistics* 6:113–144.

Boas, Hans C. 2008b. Resolving Form-Meaning Discrepancies in Construction Grammar. In J. Leino, ed., *Constructional Reorganization*, pages 11–36. Amsterdam and Philadelphia: John Benjamins.

Boas, Hans C. 2010. The Syntax-Lexicon Continuum in Construction Grammar. *Belgian Journal of Linguistics* 24:54–82.

Bonami, Olivier and Gilles Boyé. 2002. Suppletion and Dependency in Inflectional Morphology. In F. V. Eynde, L. Hellan, and D. Beermann, eds., *The Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, pages 51–70. Stanford: CSLI Publications.

Bonami, Olivier and Gilles Boyé. 2006. Deriving Inflectional Irregularity. In S. Müller, ed., *The Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, pages 361–380. Stanford: CSLI Publications.

Bonami, Olivier and Elisabeth Delais-Roussarie. 2006. Metrical Phonology in HPSG. In S. Müller, ed., *The Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, pages 39–59. Stanford: CSLI Publications.

Bonami, Olivier and Pollet Samvelian. Submitted. The Diversity of Inflectional Periphrasis in Persian. Manuscript available at http://www.llf.cnrs.fr/Gens/Bonami/index-fr.php.

Booij, Geert. 2010. *Constructional Morphology*. Oxford: Oxford University Press.

Borsley, Robert D. and Frederick J. Newmeyer. 2009. On Subject-Auxiliary Inversion and the Notion 'Purely Formal Generalization'. *Cognitive Linguistics* 20(1):135–143.

Bouma, Gosse, Rob Malouf, and Ivan A. Sag. 2001. Satisfying Constraints on Extraction and Adjunction. *Natural Language and Linguistic Theory* 1(19):1–65.

Bouma, Gosse and Gertjan van Noord. 1994. Constraint-based Categorial Grammar. In *32th Annual Meeting of the Association for Computational Linguistics*. New Mexico.

Bresnan, Joan, ed. 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Bresnan, Joan W. 2001. *Lexical-Functional Syntax*. Oxford: Basil Blackwell's.

Carpenter, Bob. 1992. *The Logic of Typed Feature Structures with Applications to Unification-based Grammars, Logic Programming and Constraint Resolution*, vol. 32 of *Cambridge Tracts in Theoretical Computer Science*. New York: Cambridge University Press.

Chaves, Rui Pedro. 2012. On the Grammar of Extraction and Coordination. *Natural Language and Linguistic Theory* 30(2).

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1970. Remarks on Nominalization. In R. Jacobs and P. S. Rosenbaum, eds., *Readings in English Transformational Grammar*. Waltham, MA: Ginn and Co.

Chomsky, Noam. 1974. Questions of Form and Interpretation. Paper delivered at LSA Golden Anniversary Symposium, Amherst, Massachusetts, July 1974. *Montreal Working Papers in Linguistics 3*, December 1974. Also reprinted as *The Amherst Lectures*. Université de Paris VII, Département de Recherches Linguistiques, February 1975.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Clark, Eve V. and Herbert H. Clark. 1979. When Nouns Surface as Verbs. *Language* 55:776–811.

Cooper, Robin and Jonathan Ginzburg. 1996. A Compositional Situation Semantics for Attitude Reports. In J. Seligman and D. Westerståhl, eds., *Logic, Language, and Computation*, pages 151–165. Stanford: CSLI Publications.

Copestake, Ann. 1992. The Representation of Lexical Semantic Information. Cognitive Science Research Papers 280, University of Sussex.

Copestake, Ann. 2001. *Implementing Typed Feature Structure Grammars*. CSLI Lecture Notes. Stanford: CSLI Publications.

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

Copestake, Ann and Alex Lascarides. 1997. Intergrating Symbolic and Statistical Representations: The Lexicon Pragmatics Interface. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*. Madrid: ACL.

Croft, William. 2003. Lexical Rules vs. Constructions: A False Dichotomy. In H. Cuyckens, T. Berg, R. Dirven, and K.-U. Panther, eds., *Motivation in Language: Studies in honour of Günter Radden*, pages 49–68. Amsterdam: John Benjamins.

Culicover, Peter. 1992. English Tag Questions in Universal Grammar. *Lingua* 88:193–226.

Dalrymple, Mary, Ronald Kaplan, John Maxwell, and Annie Zaenen, eds. 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford: CSLI Publications.

Daniels, Mike and Detmar Meurers. 2004. GIDLP: A Grammar Format for Linearization-Based HPSG. In S. Müller, ed., *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pages 93–111. Stanford: CSLI Publications.

Davis, Anthony R. and Jean-Pierre Koenig. 2000. Linking as Constraints on Word Classes in a Hierarchical Lexicon. *Language* 76(1):56–91.

Devlin, Keith. 1991. *Logic and Information*. Cambridge: Cambridge University Press.

Di Sciullo, Anna Maria and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.

Dickinson, Markus. 2004. Polish Numeral Phrases and Predicative Modification. To appear in *The Ohio State University Working Papers in Linguistics (OSUWPL)*.

Donohue, Cathryn and Ivan A. Sag. 1999. Domains in Warlpiri. In *Sixth International Conference on HPSG–Abstracts. 04–06 August 1999*, pages 101–106. Edinburgh.

Downing, Pamela. 1977. On the Creation and Use of English Compound Nouns. *Language* 53:810–842.

Engdahl, Elisabeth and Enric Vallduví. 1994. Information Packaging and Grammar Architecture: A Constraint-Based Approach. In E. Engdahl, ed., *Integrating Information Structure into Constraint-Based and Categorial Approaches*, pages 39–79. Amsterdam: ILLC. DYANA-2 Report R.1.3.B.

Engdahl, Elisabeth and Enric Vallduví. 1996. Information Packaging in HPSG. In C. Grover and E. Vallduví, eds., *Edinburgh Working Papers in Cognitive Science, Vol. 12: Studies in HPSG*, chap. 1, pages 1–32. Scotland: Centre for Cognitive Science, University of Edinburgh.

Evans, Nicholas and Stephen Levinson. 2009. The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science. *Behavioral and Brain Sciences* 32:429–492.

Fenstad, J.E., Per-Kristian Halvorsen, T. Langholm, and J. Van Benthem. 1987. *Situations, Language and Logic*, vol. 34 of *Studies in Linguistics and Philosophy*. Dordrecht: Reidel.

Fillmore, Charles J. 1982. Frame Semantics. In *In Linguistics in the Morning Calm*, pages 111–137. Seoul: Hanshin Publishing Co.

Fillmore, Charles J. 1985. Frames and the Semantics of Understanding. *Quaderni di Semantica* 6:222–254.

Fillmore, Charles J. 1986. Pragmatically Controlled Zero Anaphora. In *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society*, pages 95–107. Berkeley Linguistics Society.

Fillmore, Charles J. 1999. Inversion and Constructional Inheritance. In G. Webelhuth, J.-P. Koenig, and A. Kathol, eds., *Lexical and Constructional Aspects of Linguistic Explanation*, *Studies in Constraint-Based Lexicalism*, chap. 21, pages 113–128. Stanford: CSLI Publications.

Fillmore, Charles J. and Collin Baker. 2010. A Frames Approach to Semantic Analysis. In B. Heine and H. Narrog, eds., *The Oxford Handbook of Linguistic Analysis*, pages 313–340. Oxford: Oxford University Press.

Fillmore, Charles J., Christopher R. Johnson, and Miriam R.L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography* 16(3):235–250.

Fillmore, Charles J. and Paul Kay. 1996. Construction Grammar Coursebook. Unpublished manuscript, UC Berkeley.

Fillmore, Charles J., Paul Kay, and Mary C. O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: The Case of *Let Alone*. *Language* 64:501–538.

Fillmore, Charles J., Russell R. Lee-Goldman, and Russell Rhomieux. This volume. The FrameNet Constructicon.

Flickinger, Daniel P. 1987. *Lexical Rules in the Hierarchical Lexicon*. Ph.D. thesis, Stanford University.

Flickinger, Daniel P. 2000. On Building a More Efficient Grammar by Exploiting Types. *Natural Language Engineering* 6(1):15–28.

Flickinger, Daniel P., Carl J. Pollard, and Thomas Wasow. 1985. A Computational Semantics for Natural Language. In *Proceedings of the Twenty-Third Annual Meeting of the ACL*, pages 262–267. ACL, Chicago, IL.

Frazier, Lyn and Keith Rayner. 1990. Taking on Semantic Commitments: Processing multiple meanings vs multiple senses. *Journal of Memory and Language* 29:181–200.

Gawron, Mark and Stanley Peters. 1990. *Anaphora and Quantification in Situation Semantics*. *CSLI Lecture Notes*. Stanford: CSLI Publications.

Gazdar, Gerald. 1981. Unbounded Dependencies and Coordinate Structure. *Linguistic Inquiry* 12:155–84.

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

Gazdar, Gerald and Geoffrey K. Pullum. 1981. Subcategorization, Constituent Order and the Notion of "Head". In Moortgat, H. van der Hulst, and Hoekstra, eds., *The Scope of Lexical Rules*, pages 107–123. Dordrecht: Foris.

Gerdts, Donna and Thomas Hukari. 2001. A-Subjects and Control in Halkomelem. In D. Flickinger and A. Kathol, eds., *The Proceedings of the 7th International Conference on Head-Driven Phrase Structure Grammar*, pages 100–123. Stanford: CSLI Publications.

Geurts, Bart. 1997. Good News about the Description Theory of Names. *Journal of Semantics* 14:319–348.

Ginzburg, Jonathan. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press.

Ginzburg, Jonathan and Robin Cooper. 2004. Clarification, Ellipsis, and the Nature of Contextual Updates. *Linguistics and Philosophy* 27(3):297–366.

Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative Investigations: The form, meaning, and use of English interrogatives*. Stanford: CSLI Publications.

Godard, Danièle and Jean-Marie Marandin. 2006. Reinforcing Negation: The case of Italian. In S. Müller, ed., *The Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, pages 174–194. Stanford: CSLI Publications.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele. 2003. Words by Default: The Persian complex predicate construction. In E. J. Francis and L. A. Michaelis, eds., *Mismatch: Form-function incongruity and the architecture of grammar*, pages 117–146. Stanford: CSLI Publications.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, Adele E. 2009. Constructions Work. *Cognitive Linguistics* 20(1):201–224.

Green, Georgia M. 1974. *Semantics and Syntactic Regularity*. Bloomington: Indiana University Press.

Green, Georgia M. 1996. The Structure of CONTEXT: The representation of pragmatic restrictions in HPSG. In J. Yoon, ed., *Studies in the Linguistic Sciences. Proceedings of the 5th annual meeting of the Formal Linguistics Society of the Midwest*.

Green, Georgia M. 2000. The Nature of Pragmatic Information. In R. Cann, C. Grover, and P. Miller, eds., *Grammatical Interfaces in HPSG*, no. 8 in *Studies in Constraint-Based Lexicalism*, pages 113–138. Stanford: CSLI Publications.

Haji-Abdolhosseini, Mohammad. 2003. A Constraint-Based Approach to Information Structure and Prosody Correspondence. In S. Müller, ed., *Proceedings of the HPSG-2003 Conference, Michigan State University, East Lansing*, pages 143–162. Stanford: CSLI Publications.

Hankamer, Jorge. 1973. Why There are Two *Than*s in English. In C. Corum, T. Smith-Stark, and A. Weiser, eds., *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, pages 179–191. Chicago: CLS.

Harris, Zellig S. 1946. From Morpheme to Utterance. *Language* 22(3):161–183.

Harrison, Annabel J. 2001. *Mice Traps: Plural Inflection in English*. Masters thesis, School of Cognitive Science, University of Edinburgh.

Haskell, Todd R., Maryellen C. MacDonald, and Mark S. Seidenberg. 2003. Language Learning and Innateness: Some Implications of Compounds Research. *Cognitive Psychology* 47:119–163.

Hobbs, Jerry. 1983. An Improper Treatment of Quantification in Ordinary English. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, pages 57–63.

Hobbs, Jerry. 1996. Monotone Decreasing Quantifiers in a Scope-Free Logical Form. In S. Peters and K. van Deemter, eds., *Semantic Ambiguity and Underspecification*, pages 55–76. Stanford: CSLI Publications.

Hockett, Charles F. 1987. *Refurbishing our Foundations*. Philadelphia: John Benjamins.

Höhle, Tilman N. 1997. Vorangestellte Verben und Komplementierer sind eine natürliche Klasse. In K. R. Durscheid, C. and M. Schwarz, eds., *Sprache im Fokus: Festschrift für Heinz Vater zum 65. Geburtstag*, pages 107–120. Tübingen: Max Niemeyer Verlag.

Hudson, Richard. 2000. Grammar without Functional Categories. In R. Borsley, ed., *The Nature and Function of Syntactic Categories*, pages 7–35. New York: Academic Press.

Hudson, Richard. 2004. Are Determiners Heads? *Functions of Language* 11:7–42.

Israel, Michael. 1996. Scalar Implicature as Lexical Semantics. *Linguistics and Philosophy* 19:619–666.

Iwata, Seizi. 2005. Locative Alternation and Two Levels of Verb Meaning. *Cognitive Linguistics* 16:355–407.

Iwata, Seizi. 2008. *Locative Alternation – A Lexical-Constructional Approach*. Amsterdam/Philadelphia: John Benjamins.

Jackendoff, Ray S. 1977. *$\overline{X}$ Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jackendoff, Ray S. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.

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

Jackendoff, Ray S. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Jackendoff, Ray S. 2002. *Foundations of Langage: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Johnson, David. 1977. Remarks and Replies: On Keenan's Definition of "Subject of". *Linguistic Inquiry* 8(4):673–692.

Kajita, Masaru. 1968. *A Generative-Transformational Study of Semi-Auxiliaries in Present-Day American English*. Tokyo: Sanseido.

Kamp, Hans and Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. No. 42 in *Studies in Linguistics and Philosophy*. Dordrecht: Kluwer.

Kaplan, David. 1989. Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In J. Almog, J. Perry, and H. Wettstein, eds., *Themes from Kaplan*, pages 481–614. Oxford: Oxford University Press.

Kaplan, Ronald M. and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan, ed., *The Mental Representation of Grammatical Relations*, pages 173–281. Cambridge, MA: MIT Press. Reprinted in Mary Dalrymple, Ronald Kaplan, John Maxwell, and Annie Zaenen, eds., *Formal Issues in Lexical-Functional Grammar, pages 29–130*. Stanford: CSLI Publications.

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

Kay, Martin. 1979. Functional Grammar. In *Proceedings of the 5th Annual Meeting of the Berkeley Linguistic Society*. Berkeley: Berkeley Linguistic Society.

Kay, Paul. 1990. EVEN. *Linguistics and Philosophy* 13:59–111.

Kay, Paul. 1997. *Words and the Grammar of Context*. Stanford: CSLI Publications.

Kay, Paul. 2002a. English Subjectless Tagged Sentences. *Language* 78.3:453–481.

Kay, Paul. 2002b. An Informal Sketch of the Formal Architecture of Construction Grammar. *Grammars* 5:1–19.

Kay, Paul. 2005. Argument Structure Constructions and the Argument-Adjunct Distinction. In M. Fried and H. C. Boas, eds., *Construction Grammar: Back to the Roots*, pages 71–98. Amsterdam/Philadelphia: John Benjamins.

Kay, Paul. To appear. The Limits of Construction Grammar. In G. Trousdale and T. Hoffmann, eds., *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.

Kay, Paul and Charles Fillmore. 1999. Grammatical Constructions and Linguistic Generalizations: The *What's X Doing Y?* Construction. *Language* 75.1:1–33.

Kay, Paul and Ivan A. Sag. 2012. A Lexical Theory of Phrasal Idioms. Unpublished manuscript, Stanford University.

Kay, Paul and Karl Zimmer. 1978. On the Semantics of Compounds and Genitives in English. In S. Tsohatzidis, ed., *Meanings and Prototypes*, pages 239–246. Routledge.

Keenan, Edward. 2002. Some Properties of Natural Language Quantifiers: Generalized Quantifier Theory. *Linguistics and Philosophy* 25:627–654.

Keenan, Edward and Bernard Comrie. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8(1):63–99.

Keenan, Edward and Dag Westerståhl. 1997. Generalized Quantifiers in Linguistics and Logic. In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*, pages 837–893. Cambridge, MA and Amsterdam: MIT Press and North Holland.

Kim, Jong-Bok. 2000. *The Grammar of Negation: A Constraint-Based Approach*. Stanford: CSLI Publications.

Kim, Jong-Bok and Ivan A. Sag. 2002. French and English 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*. Ph.D. thesis, University of Manchester.

King, Paul. 1994. An Expanded Logical Formalism for Head-Driven Phrase Structure Grammar. Arbeitspapiere des SFB 340, University of Tübingen.

Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In H. van der Hulst and N. Smith, eds., *The Structure of Phonological Representations (Part 1)*. Dordrecht: Foris.

Klein, Ewan. 2000. Prosodic Constituency in HPSG. In R. Cann, C. Grover, and P. Miller, eds., *Grammatical Interfaces in HPSG*, pages 171–203. Stanford: CSLI Publications.

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

Koenig, Jean-Pierre and Anthony R. Davis. 2001. Sublexical Modality and the Structure of Lexical Semantic Representations. *Linguistics and Philosophy* 24(1):1–54.

Kornai, Andràs and Geoffrey K. Pullum. 1990. The X-bar Theory of Phrase Structure. *Language* 66:24–50.

Lascarides, Alex and Ann Copestake. 1999. Default Representation in Constraint-Based Frameworks. *Computational Linguistics* 25:55–105.

Lee-Goldman, Russell Rafael. 2011. *Context in Constructions*. Ph.D. thesis, University of Calfornia, Berkeley.

Lees, Robert. 1957. Review of *Syntactic Structures*. *Language* 33(3):375–408.

Levi, Judith H. 1978. *The Syntax and Semantics of English Nominals*. New York: Academic Press.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Levin, Beth. 2008. Dative Verbs: A crosslinguistic perspective. *Lingvisticae Investi-*

gationes 31(2):285–312.

Levin, Beth and Tova Rapoport. 1988. Lexical Subordination. In *Papers from the Twenty-Fourth Regional Meeting of the Chicago Linguistic Society*, pages 275–289. Chicago: Chicago Linguistic Society.

Levine, Robert D. and Thomas Hukari. 2006. *The Unity of Unbounded Dependency Constructions, no. 166*. CSLI Lecture Notes. Stanford: CSLI Publications.

Malouf, Robert. 2000. A Head-Driven Account of Long-Distance Case Assignment. In R. Cann, C. Grover, and P. Miller, eds., *Grammatical Interfaces in HPSG*, pages 201–214. Stanford: CSLI Publications.

Manning, Christopher D. and Ivan A. Sag. 1998. Argument Structure, Valence, and Binding. *Nordic Journal of Linguistics* 21:107–144.

Marandin, Jean-Marie. 2008. The Exclamative Clause Type in French. In S. Müller, ed., *The Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, pages 436–456. Stanford: CSLI Publications.

Matthews, Peter H. 1991. *Morphology*. Cambridge: Cambridge University Press, 2nd ed.

Matushansky, Ora. 2008. On the Linguistic Complexity of Proper Names. *Linguistics and Philosophy* 31(5):573–627.

McCawley, James D. 1971. Where do Noun Phrases Come From? In D. D. Steinberg and L. A. Jakobovits, eds., *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, pages 217–231. New York: Cambridge University Press.

Meurers, Walt Detmar. 1999. *Lexical Generalizations in the Syntax of German Non-Finite Constructions*. Ph.D. thesis, Seminar für Sprachwissenschaft, Universität Tübingen, Tübingen, Germany. Published 2000 as Volume 145 in *Arbeitspapiere des SFB* 340, ISSN 0947-6954/00.

Meurers, Walt Detmar. 2001. On Expressing Lexical Generalizations in HPSG. *Nordic Journal of Linguistics* 24(2).

Michaelis, Laura A. and Knud Lambrecht. 1996. Toward a Construction-Based Model of Language Function: The Case of Nominal Extraposition. *Language* 72:215–247.

Miller, Philip. 2002. Les emplois non finis de *do* auxiliaire. In C. Delmas, ed., *Construire et reconstruire en linguistique anglaise: Syntaxe et sémantique*, pages 185–198. Paris: CIEREC: Travaux 107, publications de l'Université de Saint-Etienne.

Miller, Philip and Paola Monachesi. 2003. Les pronoms clitiques dans les langues romanes. In D. Godard, ed., *Les langues romanes: Problèmes de la phrase simple*, pages 67–123. Paris: Editions du CNRS.

Miller, Philip H. and Ivan A. Sag. 1997. French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15(3):573–639.

Monachesi, Paola. 2005. *The Verbal Complex in Romance*. Oxford and New York: Oxford University Press.

Müller, Stefan. 1995. Scrambling in German – Extraction into the *Mittelfeld*. In *Proceedings of the Tenth Pacific Asia Conference on Language, Information and Computation*. City University of Hong Kong.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. No. 394 in *Linguistische Arbeiten*. Tübingen: Max

Niemeyer Verlag.

Müller, Stefan. 2002. Blockaden und Deblockaden: Perfekt, Passiv und Modale Infinitive. In D. Reitter, ed., *Proceedings of TaCoS 2002*. Potsdam.

Müller, Stefan. 2004. An Analysis of Depictive Secondary Predicates in German without Discontinuous Constituents. In S. Müller, ed., *Proceedings of the HPSG-2004 Conference, Center for Computational Linguistics, Katholieke Universiteit Leuven*, pages 202–222. Stanford: CSLI Publications.

Müller, Stefan. 2006. Phrasal or Lexical Constructions? *Language* 82(4):850–883.

Newmeyer, Frederick J. 2008a. On Split CPs and the "Perfectness" of Language. In B. Shaer, P. Cook, W. Frey, and C. Maienborn, eds., *Dislocation: Syntactic, Semantic, and Discourse Perspectives*, pages 114–140. London: Routledge.

Newmeyer, Frederick J. 2008b. Review of Cedric Boeckx: *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. *Language* 84(2):387–395.

Newmeyer, Frederick J. 2009. Current Challenges to the Lexicalist Hypothesis: An overview and a critique. In W. D. Lewis, S. Karimi, and H. Harley, eds., *Time and again: Papers in Honor of D. Terence Langendoen*, pages 91–117. Amsterdam: John Benjamins.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70(3):491–538.

Oehrle, Richard T. 1975. *The Grammatical Status of the English Dative Alternation*. Ph.D. thesis, MIT.

Oehrle, Richard T. 1987. Multidimensional Categorial Grammars and Linguistic Analysis. In E. Klein and J. van Benthem, eds., *Categories, Polymorphism and Unification*, pages 231–260. Centre for Cognitive Science, University of Edinburgh.

Orgun, Cemil Orhan. 1996. *Sign-Based Morphology and Phonology*. Ph.D. thesis, University of Califonia, Berkeley.

Orgun, Cemil Orhan. 2000. Sign-Based Morphology: A Declarative Theory of Phonology-Morphology Interleaving. In B. Hermans and M. van Oostendorp, eds., *The Derivational Residue in Phonological Optimality Theory*, page 247 ff. Philadelphia/Amsterdam: John Benjamins.

Peters, Stanley and Dag Westerståhl. 2006. *Quantifiers in Language and Logic*. Oxford: Oxford University Press.

Pinker, Steven. 1989. *Learnability and Cognition*. Cambridge, MA: MIT Press.

Pinker, Steven. 1999. *Words and Rules: The ingredients of language*. London: Weidenfeld and Nicolson.

Pollard, Carl J. and Ivan A. Sag. 1987. *Information-based Syntax and Semantics, Vol. 1*. No. 13 in CSLI Lecture Notes. Stanford: CSLI Publications.

Pollard, Carl J. and Ivan A. Sag. 1992. Anaphors in English and the Scope of Binding theory. *Linguistic Inquiry* 23:261–303.

Pollard, Carl J. and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Poss, Mika. 2010. *Under Construction: Cognitive and Computational Aspects of Extended Lexical Units*. Ph.D. thesis, University of Leiden.

Potsdam, Eric and Jeffrey T. Runner. 2001. Richard Returns: Copy raising and its implications. In *Proceedings of the Thirty-Seventh Annual Meeting of the Chicago*

Linguistic Society, pages 453–468.

Przepiórkowski, Adam. 1999. *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph.D. thesis, Universität Tübingen.

Pullum, Geoffrey K. and Barbara C. Scholz. 2005. Contrasting Applications of Logic in Natural Language Syntactic Description. In L. V.-V. Petr Hájek and D. Westerståhl, eds., *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, pages 481–503. London: King's College Publications.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Quirk, Randolph, Sidney Greenbaum, Geffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rappaport Hovav, Malka and Beth Levin. 2008. The English Dative Alternation: The case for verb sensitivity. *Journal of Linguistics* 44(1):129–167.

Reape, Mike. 1994. Domain Union and Word Order Variation in German. In J. Nerbonne, K. Netter, and C. J. Pollard, eds., *German in Head-Driven Phrase Structure Grammar*, no. 46 in *CSLI Lecture Notes*, pages 151–197. Stanford: CSLI Publications.

Recanati, François. 1997. *Direct Reference: From Language to Thought*. Oxford: Blackwell.

Richter, Frank. 2004. *A Mathematical Formalism for Linguistic Theories with an Application in Head-Driven Phrase Structure Grammar*. Ph.D. thesis, (2000), Eberhard-Karls-Universität Tübingen.

Rogers, Andy. 1974. A Transderivational Constraint on Richard? In *Papers from the Tenth Regional Meeting of the Chicago Linguistic Society*, pages 551–558. Chicago Linguistic Society.

Runner, Jeffrey T. and Elsi Kaiser. 2005. Binding in Picture Noun Phrases – Implications for Binding Theory. In S. Müller, ed., *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, pages 594–613. Stanford: CSLI Publications.

Sag, Ivan A. 1997. English Relative Clause Constructions. *Journal of Linguistics* 33(2):431–484.

Sag, Ivan A. 2007. Remarks on Locality. In S. Müller, ed., *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, pages 394–414. Stanford: CSLI Publications.

Sag, Ivan A. 2010a. English Filler-Gap Constructions. *Language* 86:486–545.

Sag, Ivan A. 2010b. Feature Geometry and Predictions of Locality. In G. Corbett and A. Kibort, eds., *Features: Perspectives on a Key Notion in Linguistics*, pages 236–271. Oxford: Clarendon Press.

Sag, Ivan A. to appear. Rules and Exceptions in the English Auxiliary System. *Journal of Linguistics*.

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake, and Daniel P. Flickinger. 2002. Multiword Expressions: A pain in the neck for NLP. In *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLING 2002)*, pages 1–15. Mexico City, Mexico.

Sag, Ivan A., Hans C. Boas, and Paul Kay. This volume. Introducing Sign-Based Construction Grammar.

Sag, Ivan A., Ronald M. Kaplan, Lauri Karttunen, Martin Kay, Carl Pollard, Stuart Shieber, and Annie Zaenen. 1986. Unification and Grammatical Theory. In *Proceedings of the Fifth West Coast Conference on Formal Linguistics*, pages 228–254. Stanford: CSLI Publications.

Sag, Ivan A. and Carl J. Pollard. 1991. An Integrated Theory of Complement Control. *Language* 67:63–113.

Sag, Ivan A. and Thomas Wasow. 2011. Performance–Compatible Competence Grammar. In R. Borsley and K. Börjars, eds., *Non-Transformational Syntax - Formal and Explicit Models of Grammar*. Oxford: Wiley-Blackwell.

Sag, Ivan A., Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications, 2nd ed.

Sanford, Anthony J. and Patrick Sturt. 2002. Depth of Processing in Language Comprehension: Not Noticing the Evidence. *Trends in Cognitive Sciences* 6:382–386.

de Saussure, Ferdinand. 1916. *Cours de linguistique générale*. Paris-Lausanne: Payot. Lecture notes edited by C. Bally and A. Sechehaye. Translated into English and published in 1959 as *Course in General Linguistics*, New York: McGraw Hill.

Schwenter, Scott. 1999. Two Types of Scalar Particles: Evidence from Spanish. In J. Gutierrez-Rexach and F. Martnez-Gil, eds., *Advances in Hispanic Linguistics*, pages 546–561. Somerville, MA: Cascadilla Pres.

Schwenter, Scott A. and Shravan Vasishth. 2000. Absolute and Relative Scalar Particles in Spanish and Hindi. In *Proceedings of the 26th Annual Meeting of the Berkeley Linguistics Society*, pages 225–233. Berkeley.

Seligman, Jerry and Larry Moss. 1997. Situation Theory. In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*, pages 239–309. Cambridge, MA: MIT Press and Amsterdam: North Holland.

Steedman, Mark. 2012. *Taking Scope: The Natural Semantics of Quantifiers*. Cambridge, MA: MIT Press.

Strachey, Christopher. 1967. Fundamental Concepts in Programming Languages. Unpublished lecture notes for International Summer School in Computer Programming, Copenhagen.

Stump, Gregory T. 2001. *Inflectional Morphology*. Cambridge: Cambridge University Press.

Thráinsson, Höskuldur. 1979. *On Complementation in Icelandic*. Outstanding Dissertations in Linguistics. New York: Garland Publishing.

Tseng, Jesse. 2008. The Representation of Syllable Structure in HPSG. In S. Müller, ed., *The Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, pages 234–252. Stanford: CSLI Publications.

Tseng, Jesse L. 2003. EDGE Features and French Liaison. In J.-B. Kim and S. Wechsler, eds., *The Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*, pages 313–333. Stanford: CSLI Publications.

Tunstall, Susanne. 1998. *The Interpretation of Quantifiers: Semantics and Processing*. Ph.D. thesis, University of Massachusetts, Amherst.

Uszkoreit, Hans, Daniel P. Flickinger, Walter Kasper, and Ivan A. Sag. 2000. Deep

Linguistic Analysis with HPSG. In W. Wahlster, ed., *Verbmobil: Foundations of Speech-to-Speech Translation*, Artificial Intelligence, pages 217–238. Berlin, Heidelberg, New York: Springer-Verlag.

Van Eynde, Frank. 1998. The Immediate Dominance Schemata of HPSG: A deconstruction and a reconstruction. In P. Coppen, H. van Halteren, and L. Teunissen, eds., *Computational Linguistics in the Netherlands 1997. Selected Papers from the Eighth CLIN Meeting*, pages 119–133. Amsterdam/Atlanta: Rodopi.

Van Eynde, Frank. 2006. NP-Internal Agreement and the Structure of the Noun Phrase. *Journal of Linguistics* 42(1):139–186.

Van Eynde, Frank. 2007. The Big Mess Construction. In S. Müller, ed., *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford University*, pages 415–433. Stanford: CSLI Publications.

Warner, Anthony R. 1993a. *English Auxiliaries: Structure and History*. Cambridge: Cambridge University Press.

Warner, Anthony R. 1993b. The Grammar of English Auxiliaries: An account in HPSG. Research Paper YLLS/RP 1993-4, Department of Language and Linguistic Science, University of York.

Zec, Draga. 1987. On Obligatory Control in Clausal Complements. In M. Iida, S. Wechsler, and D. Zec, eds., *Working Papers in Grammatical Theory and Discourse Structure*, pages 139–168. Stanford: CSLI Publications.