Chapter 22

Categories, Prototypes, and Default Inheritance

Robert Malouf

Abstract. Researchers in typological and functional approaches to language have argued that the cross-linguistic properties of nominalizations follow from the univeral semantic properties of syntactic categories. Specifically, nouns denote objects and verbs denote actions in the prototypical case. In this paper I argue that this prototypical association of semantic type with syntactic category can be represented in an HPSG lexical type hierarchy as a default constraint. Unmarked words will inherit from the categorial default, while words that do not inherit from the prototype will be structurally or behaviorally marked. Thus, default constraints provide a way to import insights about grammatical prototypes from more cognitively oriented work on language into a formal theory of grammatical structure.

It has long been noted [Comrie, 1976] that certain types of nominalizations are quite common in the world's languages, while other types are rare or nonexistent. For example, in English there are gerunds that combine with a subject like a noun does, but with an object like a verb does but no gerunds that combines with a subject like a verb does but an object like a noun does:

(22.1) a. Pat's often watching television [POSS-*ing*]
b. Pat often watching television [ACC-*ing*]
c.*Pat often watching of television [??]

[Croft, 1991, 83] summarizes the pattern of cross-linguistic variation with his Deverbalization Hierarchy (DH): "If a verbal form inflects for tense-aspect-modality like a predicated verb, then it will take subject and object dependents like a predicated verb. If a verbal form takes a subject dependent like a predicated verb, then it will take an object dependent like a predicated verb." This generalization is confirmed by [Koptevskaja-Tamm, 1993]'s much more extensive and systematic cross-linguistic survey of nominalization types.

Previous analyses of gerunds like *watching* in (22.1a–b) in Head-Driven Phrase Structure Grammar (e.g., [Malouf, to appear]) have treated them as a kind of mixed lexical category that show a combination of noun-like and verb-like lexical properties, derived by the Gerund Formation Lexical Rule (GFLR) in Figure 22.1. Like a typical verb, gerunds derived by the GFLR select for a subject and complements. But, like a typical noun, they select for a specifier, and since *gerund* is a subtype of *noun*, they project a phrase with the external distribution of an NP. While this analysis can account for the mixed properties of gerunds without recourse to highly

Proceedings of FHCG'98
Gosse Bouma, Geert-Jan Kruijff & Richard Oehrle (editors).
Chapter 22, Copyright ©1998, Robert Malouf.

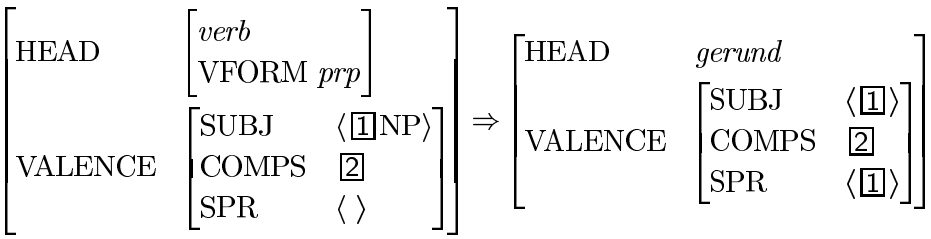


Figure 22.1: Gerund Formation Lexical Rule

abstract or unmotivated syntactic structures, nothing in this approach restricts what kinds of noun/verb hybrids can be formed. From this standpoint, the three construction types in (22.1) should be equally likely, as they all involve a deverbal noun with some combination of noun-like and verb-like lexical properties. If one were to make the (highly dubious) assumption that across languages all possible feature values are equally likely, then this analysis predicts that one should find many kinds of mixed categories that in fact do not occur.

One way to augment the analysis in Figure 22.1 to take the Deverbalization Hierarchy into account would be to impose the following constraints on possbile lexical entries:

(22.2) a. $\begin{bmatrix} \textit{word} \\ \text{VFORM} \quad \textit{fin} \end{bmatrix} \supset \begin{bmatrix} \text{SUBJ} \quad \langle \textit{synsem} \rangle \end{bmatrix}$

b. $\begin{bmatrix} \textit{word} \\ \text{SUBJ} \quad \langle \textit{synsem} \rangle \end{bmatrix} \supset \textit{predicator}$

If a word is finite (i.e. shows tense/aspect/modality marking), then it selects for a subject like a finite verb does. If a word takes a subject, then it is a predicator. That is to say, it inherits all the constraints on the linking of argument positions to grammatical functions that apply to verbs [Davis, 1996]. However, this is merely a representation of the Deverbalization Hierarchy and in no way explains why these particular constraints hold of gerund universally.

In contrast, [Croft, 1991] argues that the Deverbalization Hierarchy follows from the universal nature of syntactic categories. He proposes that syntactic categories are prototypical pairings of semantic type and pragmatic function, where the three basic universal semantic types are *action*, *object*, and *property*, and the three pragmatic functions are *predication*, *reference*, and *modification*. Croft's notion of semantic type corresponds well with the HPSG notion of semantic type. Pragmatic functions are more or less what [Searle, 1969] called **propositional acts**: referring indicates what one is talking about, and predicating indicates what is being said about it.

On the basis of a survey of twelve languages (along with informal observation of many more), Croft goes on to argue for the following three universal syntactic categories:

(22.3)

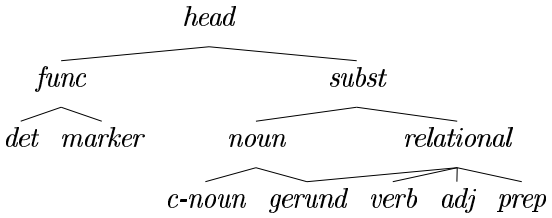
Verb	Action	$\longleftrightarrow$	Predication
Noun	Object	$\longleftrightarrow$	Reference
Adjective	Property	$\longleftrightarrow$	Modification

This is not to say that other associations cannot exist. The claim is only that these associations are prototypical. The prediction is that non-prototypical associations will be structurally or behaviorally marked, and that " ... nonprototypicality involves an 'unnatural' or at least 'imperfect' correlation of lexical semantic root with syntactic construction (referring expression, predication, attribution). In those cases, there will be a mixture of properties, some associated with the syntactic category expressed and some associated with the syntactic construction that the lexical semantic root is naturally correlated with" [Croft, 1991, 100].

In contrast to most generative treatments of syntactic categories, the approach taken by [Croft, 1991] is much like traditional grammars and modern school grammars in offering a semantic characterization of the parts of speech. In light of work on prototype categories, many of the shortcomings of the traditional approach to word classes can be seen as shortcomings of the classical Aristotelean approach to categories and classification in general. While it is not possible to state necessary and sufficient semantic conditions for nounhood, that does not mean the class of nouns has no semantic coherence.

In HPSG terms, pragmatic functions would have to be associated with types of constructions. Clausal constructions are used for predicating and certain head/specifier constructions are used for referring. However, the categories in (22.3) are *lexical* categories. The link between a word and the constructions it can be a part of is its HEAD value. A partial hierarchy of HEAD values is given in (22.4).

(22.4)



In this context, *c-noun* corresponds to Croft's *reference* and *verb* is Croft's *predication*. A straightforward way to express Croft's basic insight in the context of an HPSG lexical hierarchy is as pairings of a semantic type with a head type:

(22.5) a. $v \rightarrow$

$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{CONT} & \textit{psoa} \end{bmatrix}$$

b. $n \rightarrow$

$$\begin{bmatrix} \text{HEAD} & \textit{c-noun} \\ \text{CONT} & \textit{nom-obj} \end{bmatrix}$$

The content type *psoa* stands for 'parameterized state of affairs' in the situation theoretic sense, and corresponds to Croft's *action*, and the content type *nom-obj* corresponds to Croft's *object*.

However, the constraints in (22.5) only capture half of Croft's claim: categories are not just associations of meaning and function, they are *prototypical* associations

of meaning and function. Without a notion of prototype, this classification is either as easily falsified as the traditional semantic classification, or it is as empty of predictive power as the generative classification.

How then do we represent prototypicality in the HPSG lexical hierarchy? A standard HPSG type hierarchy is essentially a classification by **schema**, where a schema is "an abstract characterization that is fully compatible with all the members of the category it defines" [Langacker, 1987, 371]. A monotonic type hierarchy is a special case of what Langacker calls a **schematic network**, one that includes only elaboration links and in which the 'elaborates' relation is transitive. To generalize our type hierarchy to represent prototypicality then all we need to do is to introduce extension links, i.e., to allow **default inheritance** [Lascarides & Copestake, 1997]. In a default constraint, feature values marked with '/' can be overridden by conflicting constraints on more specific types.

(22.6) a. $v \rightarrow$

$$\begin{bmatrix} \text{HEAD} & \textit{relational / verb} \\ \text{CONT} & \textit{psoa} \end{bmatrix}$$

b. $n \rightarrow$

$$\begin{bmatrix} \text{HEAD} & \textit{noun / c-noun} \\ \text{CONT} & \textit{nom-obj} \end{bmatrix}$$

All things being equal, objects of type v will be constrained to have a HEAD value of type *verb*. Some subtype of v however could impose a conflicting constraint, say that the HEAD value be an object of type *gerund*. Rather than creating an inconsistency, this additional constraint **overrides** the default information in (22.6a).

Any more specific constraints that override the defaults in (22.6) must still satisfy the hard constraints in (22.6). By default the head value of a word of type v is of type *verb*, but no matter what its head value must be compatible with *relational*. This rules out a subtype of v with a head value of *det*, for example. As adverbs modify *relational* phrases, (22.6a) is consistent with the intuition that there is a connection between adverbial modification and event-denoting semantics.

Default inheritance as appealed to by, e.g., [Sag, 1997], is an abbreviatory device that helps simplify the construction of lexical type hierarchies. When used in this way, defaults add nothing essential to the analysis. They simply provide a mechanism for minimizing the number of types required. Any type hierarchy that uses defaults can be converted into an equivalent one that does not use defaults, but is perhaps undesirable for methodological reasons. Either way, the constraints that ultimately are inherited by phrasal and lexical constructions are the same, and the analysis will make the same predictions.

What I am proposing here is a rather different use of defaults. If default inheritance is used to model extension from a prototype, then two hierarchies that yield the same set of lexical constraints via a different arrangement of types would be substantively different. Different hierarchies implicitly make different claims about the prototypicality of certain constraints, and claims about prototypicality have empirical consequences. So rearranging a type hierarchy to eliminate defaults actually results in a theory that makes different predictions. This does not present a problem, since default unification operation as defined by [Lascarides & Copestake, 1997] avoids many of the shortcomings generally associated with multiple inheritance with default overriding (see [Lascarides et al., 1996]).

Following [Greenberg, 1966], [Croft, 1991] identifies several morphosyntactic criteria for identifying marked and unmarked forms. These properties follow more or less directly from a hierarchical representation of prototypes. For example, Croft points out that "an unmarked form will display at least as great a range of grammatical behavior as the marked form" (56). By definition, less marked types in the hierarchy are less specific than more marked types, and the constraints on a more marked type must be at least as specific as the constraints on a less marked type. Marked forms, which are of a more specific type and which are subject to potentially more specific constraints, will naturally occur in more specific environments.

Default inheritance provides a convenient way to model prototypicality in a lexical type hierarchy. In a narrow sense, every non-maximal type (that is, every type that has at least one subtype) is a prototype with respect to its subtypes. In a more general sense, we can say that the more subtypes a type has, the greater its prototypicality. What defaults do not explain is the connection between prototypicality and productivity. Nothing so far accounts for the fact that the prototype in effect provides a template for new members of a class.

As it is usually presented, the HPSG hierarchical lexicon is completely closed: all lexical types are explicitly listed and located within the type hierarchy. However, as [Koenig & Jurafsky, 1994] observe, such type hierarchies make it difficult to model productivity. In a closed hierarchy, every type is listed, whether or not its existence is predictable. Since no distinction is made between irregular types and regular types that follow a productive pattern, one must appeal to some external mechanism to explain why it is that new types, with a few interesting exceptions, always get assimilated to the regular pattern. Koenig & Jurafsky argue that the lexical hierarchy should only include types which are not predictable; all other types should be implicitly introduced by a process of **on-line type creation**. This same on-line type construction also applies to new forms, giving them the same behavior as existing but predictable forms.

Using a combination of default inheritance and on-line type construction, we can now give a picture of lexical organization that can account for the Deverbalization Hierarchy. Suppose every word is classified along four dimensions, according to linking, realization, inflection, and category properties.

The first two dimensions are familiar from much earlier work in virtually all syntactic theories. Within HPSG, linking has been explored in depth by [Wechsler, 1995] and [Davis, 1996]. The realization dimension has received less attention in the HPSG literature, but it is no less important. [Pollard & Sag, 1994] assume that realization is trivial, that is, that at least for verbs the first element on the argument structure is always the subject and any remaining arguments are complements. More recently, efforts have turned to exploring alternative realization possibilities: [Miller & Sag, 1997] and [Abeillé et al., to appear] propose that French allows certain arguments to be realized either as complements or as clitics, and [Bouma et al., 1998] analyze long-distance dependency constructions as a realization alternation.

So, semantic roles are associated with positions in an abstract syntactic argument structure, which are in turn realized as elements in the 'surface' syntax. Each of these associations allows a certain amount of variation but is constrained by general principles. So far, this model of the lexicon should be fairly uncontroversial. While the details are sometimes hotly contested, in its general outline this is the picture that has emerged as a standard view in HPSG and in many other

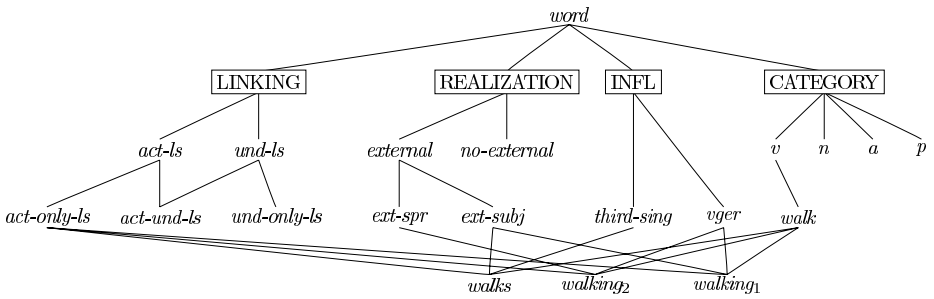


Figure 22.2: A completed type hierarchy

frameworks as well.

For concreteness, we can assume the linking constraints in [Davis, 1996], although the details will not be important. Realization types map argument structure to valence. In the realization dimension, the detailed formalization of these constraints is not vital to my argument, but to be specific I will assume constraints along the lines of those proposed by [Bouma et al., 1998]. Function words, like determiners and coordinating conjunctions, don't have any argument structure and so the realization is trivial. Substantive words make a distinction between internal and external arguments. Internal (i.e., non-initial) arguments are always mapped to the COMPS list, at least in languages with nominative/accusative case marking. The non-initial arguments of a substantive word are either realized as complements or are realized via a long-distance dependency construction. Substantives can be further divided according to how their external argument is realized. Relational words (e.g., verbs, prepositions, adjectives) realize their external argument as a subject, while nouns realize their external argument as a specifier.

Inflectional classes determine the phonological form of a word. For example, there would be an inflectional class that relates a present tense verb with a third person singular AGR value with the *-s* form. The GFLR can now be recast as an inflectional class:

(22.7) *vger* →

$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{ROOT} & \boxed{1} \\ \text{I-FORM} & f_{ing}(\boxed{1}) \end{bmatrix} \\ \text{HEAD} & \textit{gerund} \end{bmatrix}$$

When the constraint in (22.7) combines with a subtype of *v*, the head value *gerund* will override the default head value *verb*.

The last dimension is category. The relevant constraints on category types were given in (22.6). Individual lexemes, defined as subtypes of a lexical category type, need only specify non-predictable information. At a minimum, this will include a phonological form and a semantic relation.

Given these constraints, we can now apply on-line type construction to yield a virtual type hierarchy. The result is given in Figure 22.2. Note that this is a

classification of *sign*s, and so is orthogonal to the classification of HEAD values in (22.4). Each word inherits from at least one of the maximal types in each of the four dimensions: any word which satisfies this requirement is predicted to be a valid word. Furthermore, the words in Figure 22.2 are the *only* words that will be generated. All other combinations of types will be ruled out by conflicting constraints. The only combinations of types that do not have conflicting constraints on them are *walks, walking*$_1$*,* and *walking*$_2$*.* These three types correspond to the third person singular form, the ACC-*ing* gerund, and the POSS-*ing* gerund, respectively.

The effect of the type *vger* is the same as that of the GFLR. Since *gerund* is a subtype of *noun* in the hierarchy in (22.4), a phrase projected by a gerund will be able to occur anywhere an NP is selected for. Thus, verbal gerund phrases will have the external distribution of NPs. Adverbs modify objects of category *relational,* which include verbs, adjectives, and verbal gerunds, among other things. Since adjectives only modify *c(ommon)-nouns,* verbal gerund phrases will contain adverbial rather than adjectival modifiers. Since *verb* is a distinct subclass of *relational* disjoint from *gerund,* verbal gerund phrases will not have the distribution of true VPs.

Because a gerund has a *psoa* content just like a verb does, its semantic roles will get linked to argument positions just like a verb's do. *Gerund* is a subtype of *subst*, so its non-initial arguments will get mapped to the COMPS list, and a gerund will take the same complements as the verb from which it is derived. But, *gerund* is a subtype of both *noun* and *relational*, so a gerund's external argument could get mapped either to the subject or to the specifier. Thus we get the effect of the old gerund lexical rule, without having to stipulate the exceptional argument mapping.

This kind of analysis is much more restrictive than the lexical rule analysis. One could write a lexical rule that produces a gerund which takes a subject like a verb does, but cannot take complements, even though cross-linguistically gerunds of that type do not seem to occur. Under the approach sketched here, the only way for a word to take a subject is for it to be a subtype of *ext-subj*. Since *ext-subj* is a subtype of *external*, anything that takes a subject will potentially take complements too. Similarly, VFORM values are only appropriate for objects of type *verb*. That means that any gerund that inflects like a finite verb (and thus is [VFORM fin]) will also by necessity inherit from *ext-subj*.

The implicational structure of the Deverbalization Hierarchy is directly reflected by the subsumption ordering in the lexical hierarchy in Figure 22.2. This means that given this lexical layout, there is no way to derive a gerund of the type seen in (22.1c). In a weak sense, this explains the behavior of the English verbal gerund. The properties of the English verbal gerund follow from the structure of the English lexical hierarchy, which in turn is motivated independently by other lexical constructions in English. There is no need to impose additional unmotivated constraints like those in (22.2).

The Deverbalization Hierarchy however is a universal claim about gerunds in all languages. It would be absurd to try to explain a universal pattern in terms of parochial facts about the structure of the English lexicon. If the lexical patterns in Figure 22.2 are to serve as an explanation for a universal property of language, they themselves must be universal. Obviously, the specific details of Figure 22.2 cannot be universal. At best only the general outlines of the organization of the lexicon could be universal, and even then there is considerable variation. [Manning, 1996] shows that nominative/accusative and ergative/absolutive

languages have very different realization constraints, and [Manning & Sag, 1995] analyze a number of diathesis alternations as alternative realizations of a basic argument structure.

The best we can say is that there seem to be very general principles that govern the structure of the lexicon of all languages. Good candidates for such principles are that the mapping of semantic roles to syntactic positions is mediated by a level of syntacticized argument structure, and that inflectional morphology is part of the interface between the lexicon and the syntax. These entail that the lexical hierarchy of every language will be divided into at least the four dimensions discussed above. Within each dimension, the best we can say is that a language can avail itself of a limited number of choices out of a universal inventory of organizational schemes (say, nominative vs. ergative case marking).

One way of looking at this kind of model of language universals is to assume that parts of the lexical type hierarchy are mandatory, perhaps due to more general properties of human cognition or perhaps simply because they are innate. Other parts of the lexical hierarchy are then chosen from a set of universally available types. These in a sense come 'for free', since their existence is entailed by something about the way language works, and they provide the cross-linguistic motivation for the language specific types that inherit from them. Similarities between grammars of different languages arise from the similar diachronic and functional pressures that shape each language. Finally, the lower reaches of the lexical hierarchy are populated by strictly parochial types that have no cross-linguistic generality and reflect the uniqueness of each human language. This is the model that is developed in considerable detail by [Ackerman & Webelhuth, 1998], and is much like what [Croft, 1995, 504] calls **typological functionalism**. Typological functionalism makes an attractive way of integrating formal and functional approaches to language without substantially weakening either. It would allow a formal grammarian to draw on, for example, the competition models proposed by [Bates & MacWhinney, 1989] or [DuBois, 1985] without abandoning the idea that grammar is an arbitrary, self-contained, symbolic system.

Now the questions is, what is the status of the particular constraints in (22.6)? Are they hard-wired as part of Universal Grammar, or are they derived from something outside of grammar? There is some evidence from other work that suggests that the archetypes in (22.6) do indeed follow from something outside of grammar. In particular, the **semantic bootstrapping hypothesis** [Grimshaw, 1981, Pinker, 82, Pinker, 89] claims that children learning language use semantic types as a basis for syntactic classification. That is:

> ... certain cognitive categories have what I will call Canonical Structural Realization (CSR): CSR(object) = N, CSR(action) = V. [Language acquisition] employs a CSR principle: a word belongs to its CSR, unless there is evidence to the contrary [Grimshaw, 1981, 174].

If we suppose, following [Green, 97], that language acquisition is a process of incrementally building a type hierarchy, then Grimshaw's CSR principle will guarantee that types more or less like those in (22.6) will be introduced into the child's grammar early on, and will be high up in the adult hierarchy.

Thus, the constraints in (22.6) may be archetypes due to a heuristic employed by language learners and not due to anything fundamental about grammatical representation itself. Note that there is no contradiction in saying the constraints

in (22.6) are archetypes and they can be derived from something outside the grammatical system. Archetypes as I use them here provide the link between language universals and grammatical particulars, but embody no specific claim as to the origin of those universals.

Bibliography

[Abeillé et al., to appear] Anne Abeillé, Danièle Godard, Philip Miller, and Ivan A. Sag. French bounded dependencies. In Luca Dini and Sergio Balari, editors, *Romance in HPSG*. CSLI Publications, Stanford, to appear. URL `ftp://ftp-csli.stanford.edu/linguistics/sag/fbd.3.ps.gz`.

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

[Bates & MacWhinney, 1989] Elizabeth Bates and Brian MacWhinney. Functionalism and the Competition Model. In Brian MacWhinney and Elizabeth Bates, editors, *The Crosslinguistic Study of Sentence Processing*, pages 3–73. Cambridge University Press, 1989.

[Bouma et al., 1998] Gosse Bouma, Robert Malouf, and Ivan A. Sag. Satisfying constraints on extraction and adjunction. University of Groningen and Stanford University, URL `ftp://csli-ftp.stanford.edu/linguistics/sag/ks-sem/bms-98-draft.ps`, 1997.

[Comrie, 1976] Bernard Comrie. The syntax of action nominals: a cross-language study. *Lingua*, 40:177–201, 1976.

[Croft, 1991] William Croft. *Syntactic Categories and Grammatical Relations*. University of Chicago Press, 1991.

[Croft, 1995] William Croft. Autonomy and functionalist linguistics. *Language*, 71:490–532, 1995.

[Davis, 1996] Anthony R. Davis. *Lexical Semantics and Linking in the Hierarchical Lexicon*. PhD thesis, Stanford University, 1996. URL `http://www-linguistics.stanford.edu/~tdavis/thesis-ps.html`.

[DuBois, 1985] John W. DuBois. Competing motivations. In John Haiman, editor, *Iconicity in Syntax*, pages 343–365. John Benjamins, Amsterdam, 1985.

[Greenberg, 1966] Joseph H. Greenberg. *Language Universals*. Number 59 in Janua Linguarum Series Minor. Mouton, The Hague, 1966.

[Green, 97] Georgia Green. Modelling grammar growth: Universal grammar without innate principles or parameters. Paper presented at GALA97 Conference on Language Acquisition, Edinburgh, 1997. URL `http://lees.cogsci.uiuc.edu/~green/edinout.ps`.

[Grimshaw, 1981] Jane Grimshaw. Form, function, and the language acquisition device. In C.L. Baker and J. McCarthy, editors, *The Logical Problem of Language Acquisition*, pages 165–182. MIT Press, Cambridge, 1981.

[Koenig & Jurafsky, 1994] Jean-Pierre Koenig and Dan Jurafsky. Type underspecification and on-line type construction. In *Proceedings of the West Coast Conference on Formal Linguistics*, volume 13, 1994. URL `http://stripe.colorado.edu/~jurafsky/papers/wccfl94.ps.Z`.

[Koptevskaja-Tamm, 1993] Maria Koptevskaja-Tamm. *Nominalizations*. Routledge, London, 1993.

[Langacker, 1987] Ronald W. Langacker. *Foundations of Cognitive Grammar*, volume I. Stanford University Press, Stanford, 1987.

[Lascarides et al., 1996] Alex Lascarides, Ted Briscoe, Nicholas Asher, and Ann Copestake. Order independent and persistent typed default unification. *Linguistics and Philosophy*, 19:1–89, 1996.

[Lascarides & Copestake, 1997] Alex Lascarides and Ann Copestake. Default representation in constraint-based frameworks. University of Edinburgh and Stanford University, URL `http://www-csli.stanford.edu/~aac/papers/yadu.ps.gz`, 1997.

[Malouf, to appear] Robert Malouf. Mixed categories in HPSG. In Robert D. Borsley, editor, *Syntactic Categories*. Academic Press, New York, to appear. URL `http://hpsg.stanford.edu/rob/papers/gerund.pdf`.

[Manning, 1996] Christopher D. Manning. *Ergativity: Argument Structure and Grammatical Relations*. CSLI Publications, Stanford, 1996.

[Manning & Sag, 1995] Christopher D. Manning and Ivan A. Sag. Dissociations between argument structure and grammatical relations, 1995. Paper presented at 1995 HPSG Workshop, Tübingen.

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

[Pinker, 82] Steven Pinker. A theory of the acquisition of lexical interpretive grammars. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*, pages 655–726. MIT Press, Cambridge, 1982.

[Pinker, 89] Steven Pinker. *Learnability and cognition*. MIT Press, Cambridge, Ma, 1989.

[Pollard & Sag, 1994] Carl Pollard and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. University of Chicago Press and CSLI Publications, Chicago and Stanford, 1994.

[Sag, 1997] Ivan A. Sag. English relative clause constructions. *Journal of Linguistics*, 33:431–484, 1997.

[Searle, 1969] John Searle. *Speech Acts*. Cambridge University Press, Cambridge, 1969.

[Wechsler, 1995] Stephen Wechsler. *The Semantic Basis of Argument Structure*. CSLI Publications, Stanford, 1995.