Stanford University
11 November 2003

Some Foundational Issues for Construction Grammar:
Mutual Definition and Cluster Concepts

Arnold M. Zwicky, Stanford University (zwicky@csli.stanford.edu)

1. The puzzle of "parts of speech": (i) categories (i.e., extensionally, sets) of lexemes; (ii) small in number; (iii) mutually exclusive; (iv) exhaustive; (v) with members sharing syntactic, morphological, semantic, and possibly discourse properties.

Huddleston & Pullum (2002:22):
Noun: *tree, pig, sugar, hatred, union, Picasso, London*
Verb: *do, fly, melt, think, damage, give, have, be, must*
Adjective: *good, nice, big, easy, ugly, helpful, reddish, fond*
Adverb: *obviously, easily, helpfully, frankly, soon, so, too*
Preposition: *of, to, by, into, between, over, since, toward(s)*
Determinative: *the, this, that, a(n), some, all, every, each*
Subordinator: *that, for, to, whether, if*
Coordinator: *and, or, but, nor*
Interjection: *ah, damn, gosh, hey, oh, ooh, ouch, whoa, wow*

2. General question: What are the natural classes of lexemes in a language? Syntax and morphology.

3. "Distribution": what lexemes can fill some slot. Sophisticated view of distribution: what lexemes can fill a particular slot *in a particular construction* (where a construction is one sort of (complex) sign, a pairing of conditions on form/content with meaning/use). Here's where meaning comes in.

4. Methodological question: How do you discover constructions? Any way you can; there are no discovery procedures. And there can be serious issues about how constructions are individuated.

Accompanying circumstance adverbials: WITH + [Subject + X]
Category question: What lexemes can serve as the head word of X?
Apparent answer: V, P, Adj, or N (at least) lexemes.
Two types of these adverbials.

Type I: X is licensed as Predicate in combination with Subject, hence has a V lexeme as head; construction requires VForm:N (the "-*ing* form") of this head:
with Kim weighing so little
with Sandy being desperately sick with the flu
with Terry having no useful ideas

Type II: "missing BE": X is licensed as Predicative complement of BE in combination with Subject, hence can have head lexemes of a variety of categories:

with Kim undergoing surgery today (Progressive: V in VForm:N)
with Sandy snubbed by her friends (Passive: V in VForm:P, the "*-en* form")
with Terry soon to ascend to the throne (Prospective BE TO)
with Kim in Tokyo all week (transitive P)
with Kim away this week (intransitive P)
with Sandy desperately sick with the flu (Adj)
with Terry our leader for the next week (N)

Note: Modals like MUST (even peripheral modals like OUGHT) are disallowed – in Type I because modals have no VForm:N, in Type II because modal VPs (indeed, finite VPs in general) are not licensed as Predicative complements, of BE or any other V.

5. At least two sorts of constructions, differing in what kinds of constructs they license:
      (1) Constituents (including linear ordering of the parts)
      (2) Valency sets

6. Crossreferences between these two types:
      (1) The expressions combined in a constituent construction must satisfy conditions on their syntactic functions (e.g., Subject, Predicate, and Predicative above), conditions that refer to valency sets.
      (2) The expressions filling slots in a valency set are either lexemes or phrases; the phrases are constituents licensed by constituency constructions.

7. Involvement of lexemes:
      (0) The central problem of syntax is how the composition of an expression is related to its distribution. All the conceptual apparatus of syntax arises from trying to describe how the internal syntax and the external syntax of expressions fit together. Internally, expressions are constructed around central words (heads, in Bloomfield's sense); the category of an expression is the category of its head word. Externally, the distribution of an expression is a matter of the syntactic functions it can serve.
      (1) Default category-function relationship: constituents are built up around head words, which are instances of lexemes belonging to particular categories; phrases with heads of category X have a *characteristic function set*, a set of syntactic functions they are licensed to occur in (e.g., NPs are associated with a characteristic function set that includes Subject, Direct Object, Prepositional Object, Vocative, at least).
      (2) But constructions can also stipulate non-characteristic functions; hence, PP Subjects, etc.
      (3) Both constituency and valency constructions can have lexical slots; for each such slot, say Slot i in Construction j, there is a class of lexemes (the *L-category* (i, j)) that can fill this slot. Correspondingly, for each phrasal slot there is a class of expressions (the *F-category* (i, j)) that can fill this slot. (An F-category is then the extension of a syntactic function.)
      (4) In addition, for each phrasal slot, there is an L-category (Head, i, j) of lexemes that can serve as the head of a phrase filling that slot.
      (5) In addition, some Condition k for Construction j can require the presence of a word, not necessarily the head, within an expression filling Slot i in this construction (requiring, say, a

WH word a particular contributing expression); to such a condition there corresponds an L-category (Cond-k, i, j).

8. A constructional approach to syntax allows for a new view of L-categories and F-categories, in which they are not given universally or associated with meaning directly, but instead stand in a relationship of language-particular mutual definition with each other and with constructions themselves, and receive their meanings by virtue of the semantics and pragmatics associated with particular constructions. (Cf. Croft (2001).)

9. What natural classes (L-categories) do you get? A gigantic lattice of classes, with superclasses above the "parts of speech" (e.g., the Head L-category for the phrase that is fronted in WH questions, which comprises, roughly, N, A, and P, but not V), with an amazing assortment of subclasses of these, down to tiny, even singleton, classes (the L-category for the Progressive valency construction, which has only BE as a member), and with much overlapping (e.g., the class of lexemes modifiable by MORE, which comprises both A and N – mass nouns and count nouns in the plural – lexemes).

10. Why are the traditional parts of speech so few? Pedagogical purposes of traditional grammar. But also because they represent something like a *basic* level of classification. At the basic level, most instances fit into one of a small number of categories, which are informationally/predictively rich: if X does this, then it probably does that and that…

11. Consistent with contexts or "tests" for categories that mostly line up, but don't quite line up perfectly, as with the contexts for the English Aux category. (Note "peripheral" Aux lexemes like OUGHT.) The properties cluster, but not perfectly. (Parallel in morphology: lexemes with one or two irregularities in their inflection.)

12. Consistent with a significant number of outliers, e.g. infinitival TO treated as an odd sort of V, which can head a (non-finite) clause (*for pigs to fly*) and serve as a remainder in VPE (*I don't want to*), but fails to qualify for the Head L-category in the construction that allows a Clause to serve as a Sentence (a fact that Huddleston & Pullum treat as crucial evidence against treating TO as a V. (Parallel in morphology: lexemes that are defective in inflectional forms.)

13. Consistent with there being a few items that have some properties of one basic category and some of another (or even three), as perhaps in the lexemes examined by Maling (1983) and Huddleston & Pullum (ms. 2003). (Parallel in inflectional morphology: heteroclites.) In such cases, labeling is a red herring.

14. Consistent with there being a modest number of items with quite idiosyncratic properties (H&P's Subordinators and Coordinators, perhaps). (Parallel in morphology: lexemes in small, minor paradigm classes.)

15. Paradigm classes in languages with rich inflectional systems. If we define a paradigm class in terms of the applicable morphological rules, then Latin has an enormous number of noun declensions. However, there is a very high degree of association between the applicability of different morphological rules, with the result that there is a small number of "rule clusters" (the

major declension classes), with most nouns pretty much falling in one of them.

16. Note: no actual fuzziness, clines, gradients, continua, squishes, etc. Just sets of rule-eligibilities.

17. The semantic connection: At all levels of graininess, categories are associated, at least weakly, with semantics (and information structure). ("Weakly" in the sense that the association is neither necessary nor sufficient – but pretty good.) In general, the semantic connection follows from the semantics of a construction. (But some items with the right semantics don't work, and some without it do anyway.)

References

Baker, Mark C. 2003. Lexical categories: Verbs, nouns, and adjectives. Cambridge: CUP.

Borsley, Robert D. (ed.). 2000. The nature and function of syntactic categories. Syntax and Semantics 32. San Diego CA: Academic Press.

Croft, William. 2001. Radical construction grammar: Syntactic theory in typological perspective. Oxford: OUP.

Hopper, Paul J. 1996. Discourse and the construction of categories. BLS 22.159-70.

- & Sandra A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. Lg 60.4.703-52.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. The Cambridge grammar of the English language. Cambridge: CUP.

Huddleston, Rodney & Geoffrey K. Pullum. Ms. 2003. Anomalous adjectives and prepositions in English. Paper presented at LSA Annual Meeting, Atlanta GA.

Jackendoff, Ray S. 1977. X-bar syntax. Cambridge MA: MIT Press.

Maling, Joan. 1983. Transitive adjectives: A case of categorial reanalysis. Heny & Richards 1983:1.253-89.

McCawley, James D. 1982a. The nonexistence of syntactic categories. Chapter 4 of McCawley 1982b:176-203. [Revised and enlarged version of paper at 1977 Michigan State Univ. Metatheory Conference.]

- 1982b. Thirty million theories of grammar. Chicago: Univ. of Chicago Press.

Newmeyer, Frederick J. 2000. The discrete nature of syntactic categories: Against a prototype-based account. Borsley 2000:221-50.

Pullum, Geoffrey K. 1982. Syncategorematicity and English infinitival *TO*. Glossa 16.181-215.

Zwicky, Arnold M. 1985. Clitics and particles. Lg 61.2.283-305.

- 1994. Dealing out meaning: Fundamentals of syntactic constructions. BLS 20.611-25.