

## 1. Some more thoughts on noun class systems

After our discussion of noun classes last class, it occurred to me that perhaps Navajo was a counterexample to what I was saying. What I was saying was:

1. (a) Noun class systems are semantically inert (mostly?)
- (b) Gender systems are a type of noun class system
- (c) In a gender system, every noun is assigned a gender, even if it has no natural/semantic gender. Further, some nouns that do have a semantic gender are assigned a morphosyntactic gender which doesn't match its semantic gender. A famous example: German *Mädchen*, 'maiden', which is neuter, in a gender system with masculine, feminine and neuter groups.

Question for German speakers: someone last time indicated that for the purposes of pronominal reference, German speakers seem to have a choice between semantic gender (feminine pronoun) and morphosyntactic gender (neuter pronoun). Is that also true for, e.g. adjectival agreement? If we were talking about 'a red maiden' (in, say, Nominative case), is there a choice between *rotes Mädchen* ('red.Neuter maiden') and *rote Mädchen* ('red.Fem maiden')? My guess is no.

- (d) In a gender system, and hence in classificatory systems generally, the classificatory features are morphosyntactic, and are semantically inert.
- (e) Therefore, class systems are not a counterexample to the assertion that the syntax has access to only a very limited amount of semantic information -- if in fact it has access to *any* semantic information at all.

2. (a) Consider, however, Navajo classificatory verbs of motion, discussed in a paper by Mary Willie called "Individual and Stage Level Predication and the Navajo Classificatory Verbs".

(b) These verbs are used to describe the position of something in space (stative), something's change of position or movement to another place (intransitive), and caused change of position or movement to another place (transitive and ditransitive). The object whose movement or position is being described, in the default case, specifies the choice of which verb to use.

(c) There is a medium-sized closed set of these verbs, and the choice of which verb to use seems roughly to correspond to certain physical properties of nouns, however, they're like other classificatory systems in that it's difficult to characterize exactly their semantic content. For example, there are 11 verbs which are used to describe the location of things at rest, listed below:

3. a. **si'á** (default category): can include roundish/squarish things (bottle, book), large small things (hogan, ring).
- b. **siká** mostly things held in by a container, often liquid: coffee in a cup, puddle of water, dishes in a dishpan (without water).
- c. **silà** rope-like objects or things that come in pairs, also songs.

- d. **siyí** large loads or bundles of something: a 25-lb bag of flour, load of firewood, or a large body of water like a lake.
- e. **shijool** wool-like or fluffy objects, as well as dried-up plants.
- f. **shijaa'** a group of multiple items of the same type, like a bunch of houses, a pile of beads or coins, or even a litter of kittens
- g. **sití** animate things in a reclining position: man or horse lying down.
- h. **sitá** stick-like or flat objects: a cane, an unattached door, a sheet of ice.
- i. **sitléé'** moist or mushy stuff: oatmeal, playdoh, mashed potatoes
- j. **sittsooz** flexible, foldable things: paper, blankets, clothing
- k. **sinil** like (f), can refer to multiple items, or like (c), things that come in pairs.

So, there is some consistency, but also some strange combinations, and, the giveaway for a morphosyntactic system, there is a default class. When these verbs are used with their normal associates, there's no particular semantic implication other than existence or location. So, e.g. (4) is just equivalent to "There's a bowl of soup".:

- 4. atoo' siká  
 soup 3sgNOM.sits-liquid-in-container (class b above)  
 "There's soup"

However, it seems like the verb classes do carry semantic information, in that choice of verb might tell you which, e.g., of the two kinds of money you picked up (these verb is not one of the stative positional ones above, but ones indicating caused-change-of-position, related to the stative classes j and k above), or which of two kinds of rocks you're talking about (classes ? and f above):

- 5. a. béeso **ndiitsoos**  
 money 3sgACC.1sgNOM.found-paper-like  
 "I found a bill"
- b. béeso **ndiinil**  
 money 3sgACC.1sgNOM.found-multiple  
 "I found some coins"
- c. tsé shizhood  
 rock 3sgNOM.sits-large-object  
 "There's a boulder"
- d. tsé shijaa'  
 rock 3sgNOM.sits-multiple  
 "There's some pebbles"

And, most interestingly, when used with the "wrong" items, the verb can be used to imply something about the manner or appearance of the item in question (what Willie calls the 'extended' use):

6. a. asdzáá léi' sitléé'  
 woman certain 3NOM.sits-mushy  
 "A certain woman is sitting all sprawled out."
- b. t'áá 'aaníí lééchaa'í léi sità  
 truly dog certain 3sgNOM.sit-sticklike/rigid  
 There's a truly dead/skinny dog.

Willie points out that the latter example demonstrates that the semantic contribution of the classifier verb is really there: you can modify it (with 'truly', in this case). Interestingly, she notes that this modification is impossible when used with appropriate objects:

7. #t'áá 'aaníí gish sità  
 truly cane 3sgNOM.sit-sticklike/rigid  
 "There's a truly (sticklike) cane"

The moral of the story, as I see it, is that perhaps in their non-extended use, the classifiers have/contribute no semantic content, and hence are indicative of a gender-type nominal classification system in action. But in their extended use, it's the mismatch between noun and classifier that creates a flag to look for other possible semantic implications, and the natural classes, or "basic" meaning, of the classifiers re-emerge (just like inappropriate pronominal reference in German, e.g.). So, paper topic, anyone? Examining the existence of non-literal or inappropriate uses of classifiers in some nominal-classifier language? Apparently gender mistakes by non-native speakers are very funny to native speakers of gender-system languages — this is perhaps a place to start.

## 2 The Lexicon: what does it look like?

8. *A fairly conventional view*

A List of Words, and special operations for making new words

Phonological Word = Syntactic Word = Semantic Word

/kæt/	N, +animate? +count?	+animate? +count? <i>everything we know about catness</i>
-------	----------------------------	---

Operations: compounding, causative/inchoative, etc.

9. Problematic cases:

(a) Phonological Word syntactic terminal  
 Root+Inflection

<i>cat+s</i>	N+Num	
<i>kiss+ed</i>	V+T	
<i>ik+ase+ta</i>	V+V+T	(Japanese)
go.make.past		

Phonological Word *sometimes* = syntactic terminal, sometimes non-constituent

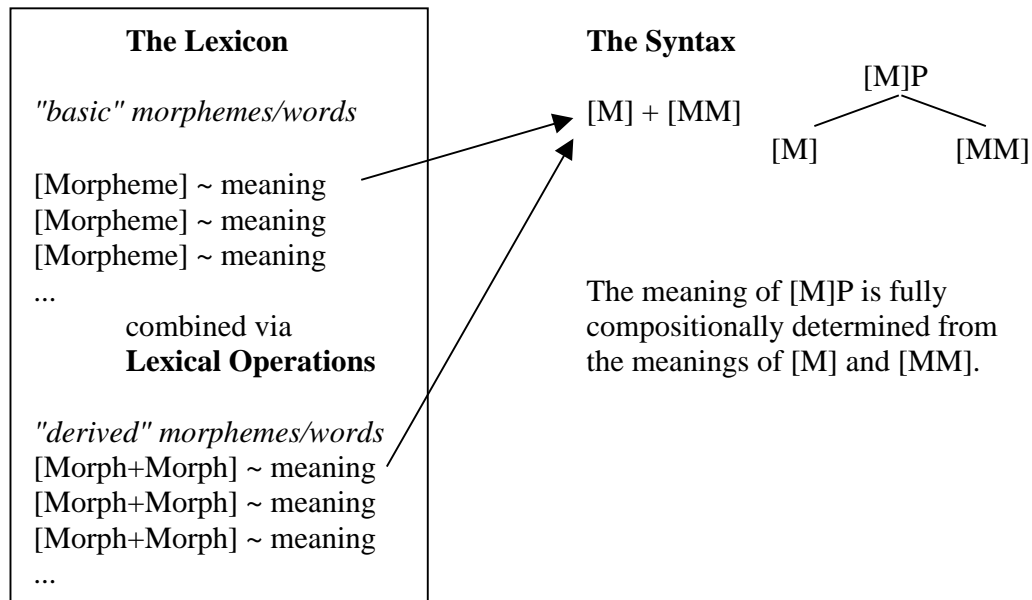
Clitics

I did/I'd  
 do not/don't  
 [the girl sitting next to Stephen (Pinker)]'s pants

10. Ok, so maybe the lexicon doesn't contain phonological words. Let's say it contains *morphemes*, which then may be combined either

- (a) through lexical operations or
- (b) through syntactic operations.

(The results of (a) can and do get passed on to (b)), as in the diagram in (c):



The schema for morphemes is going to be basically the same as outlined above.

11. *No Escape from Syntax*: Marantz 1997

(a) Question: What's the difference between *lexical* combining operations and *syntactic* combining operations?

Generative Lexicalist's answer: the results of *lexical* combining operations can have an idiosyncratic meaning. Syntactic operations result in compositional interpretation, lexical operations have compositional consequences (e.g. for inference patterns) but basically produce meanings that are in some sense semantic atoms, with idiosyncratic characteristics. (This is why, according to lexicalists, *kill* isn't the same as *cause to die*).

(b) *Marantz's point (from "Cat as a phrasal idiom")*

It's perfectly possible for the result of combining operations that everybody recognizes as syntactic to have an idiosyncratic meaning — in fact, it happens all the time! They're called idioms. In the absence of other evidence, then, we should assume that morphemes are put together by syntactic operations, but that some combinations of morphemes get idiosyncratic meanings associated with them?

(i) *idioms:*           kick the bucket  
                          take a leak  
                          take a break  
                          take heart

(ii) the effects of meaning associated with syntactic structure still "show through" *kick the bucket* doesn't exactly mean *die*, rather, it means something like "croak" or "pass on": #He's been kicking the bucket for three days (Compare: He's been dying for three days).

Ditto for morphemically complex words: "transmission":

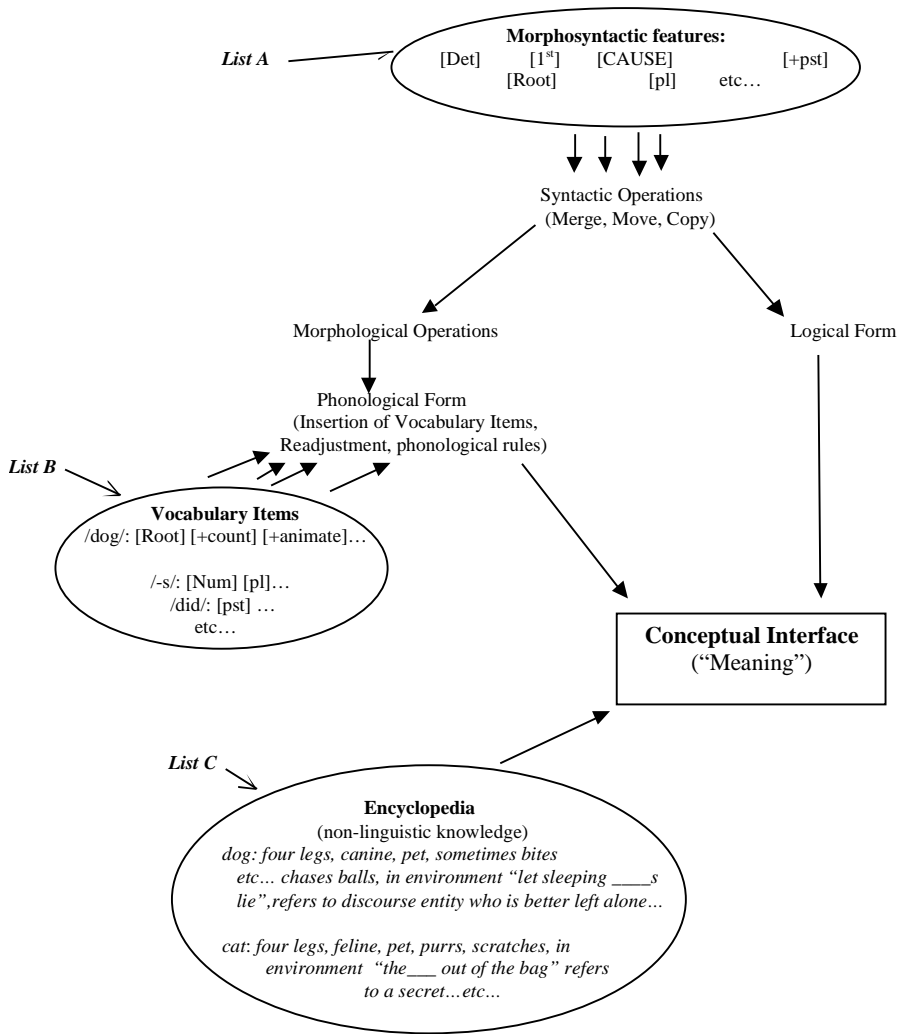
trans+mit+ion = carries information about a process, the result of its preverb+verb stem structure being followed by a nominalizing suffix. Its idiosyncratic meaning is that it refers to a *thing*, rather than an actual process — but it's crucially a thing *for accomplishing a process*: "transmission"-type nouns don't mean the same thing as monomorphemic "cat"-type nouns. That is, the fabled "idiosyncrasy" of lexemes is an illusion; it's the same idiosyncrasy we see at work in bigger-than-word chunks.

(c) So, if multimorphemic words and idioms are the **output** of the syntax, semantic interpretation (including 'lexical' interpretation), must be post-syntactic. That is, in order to interpret *kick the bucket*, you need to look at the stem *kick* and see that it's in the environment [<sub>DP</sub> the bucket]. — in fact, that it's a sister to the DP *the bucket*. Then, in the *encyclopedic ("lexical")* entry for *kick*, you can see a statement to the effect that "In the environment "kick the bucket", *kick* means "die".

This is the Distributed Morphology way: the source of all generativity, in the lexicon or elsewhere, is the single module, *the syntax*. So, keep this in mind as we read generative-lexicon work (nearly all the work we'll be reading). Any proposal that a lexical generativist makes *ought to be implementable* not by lexical correspondence rules or operations in the lexicon, but rather, in the syntax. And most importantly, the "semantic primitives" proposed by lexical semanticists ought to also appear as syntactic primitives (and hence perhaps appear as overt morphology in some language?) or else correspond to some syntactic configuration that carries the meaning in question.

Note that this approach explains the existence of seemingly robust morphemes that don't have a "meaning" on their own: *cran+berry*, *de+ceive/re+ceive/con+ceive* (note that *ceive* must be a morpheme because it alternates with *cept* in nominalizations, uniformly), *kit* and *caboodle* (anyone know what a *caboodle* is?) Such morphemes have meanings — but *only in the environment of other morphemes*. Their Encyclopedia entry has no structure-independent meaning listed.

## 12. Distributed Morphology



The main thing to notice is the three, non-generative lists that take the place of the lexicon: the syntactic primitives, the Vocabulary Items, and the Encyclopedia.

(a) The syntactic primitives are largely morphosyntactic features, plus a few "basic" meaning-affecting units.

(b) The Vocabulary Items (except for Roots) are inserted deterministically, competing for insertion into syntactic terminals. This explains the phenomenon of

"default" morphemes in any given paradigm: they are radically underspecified  
Vocabulary Items.

(c) The Roots are inserted in Root spaces in the syntactic terminals according to the speaker's intent, but it has to be the case that their semantic content (Encyclopedic content) is compatible with the semantic import of the syntactic frame in which they find themselves; hence, "You falled me" (for "You made me fall") is comprehensible (as noted by Pustejovsky) but bizarre; similarly, "He drank three waters" is comprehensible but bizarre. ("Colorless green ideas sleep furiously").

Marantz then goes on at length to discuss why the causative/inchoative alternation vs. transitive and intransitive nominalizations, as discussed by Chomsky, demonstrate that the causative form of causative/inchoative verbs is syntactically derived from the inchoative. We'll consider this type of argument again in detail later; it may be useful to come back to section 2 of this paper after we read about causation. First, however, let's look at some generative-lexicon discussions.

### 3 Pustejovsky: The Generative Lexicon

Like Levin, last week, Pustejovsky is setting out a research programme for lexical semantics, discussing all the goals he considers reasonable for a lexical semantic theory.

The first thing he says is:

13. "...without an appreciation of the syntactic structure of a language, the study of lexical semantics is bound to fail."

That is, like Levin (and Jackendoff), the study of the syntax associated with lexemes is considered central to the study of their meaning. (Note that this necessity receives an explanation in the Distributed Morphology approach).

14. (i) Semantic well-formedness, hence a theory of *possible word meaning*  
-possible result: separation of pragmatic or discourse factors  
from treatments of lexical meaning (note contrast with  
Jackendoff).
- (ii) Semantic representations need to provide a *useful semantic interpretation*  
*of a sentence*. So there will need to be **lots** of information in them.
- (iii) Look at all kinds of words.
15. Overall picture of semantic interpretation:  
Several different levels working in tandem:  
-lexical semantics  
-compositional semantics  
-discourse structure  
-temporal structure

"...The individual sources of semantic knowledge compute local inferences with a high degree of certainty. When integrated together, these inferences must be globally

coherent, ...[this results from] processes of co-operation among separate semantic modules. The basic result of such a view is that semantic interpretation proceeds in a principled fashion, always aware of what the source of a particular inference is, and what the certainty of its value is."

This position seems highly unlikely to me. Last time we saw that it's often difficult to decide if an entailment is the result of a lexical semantic factor, a compositional semantic factor, or "world knowledge". Ditto for many other aspects of meaning. Consider 16:

16. a. John pushed the cart to San Francisco.  
— inference: the cart made it to San Francisco.  
is that the result of the past tense? The verb "push"? what about:
- b. John was pushing the cart to San Francisco.  
— inference: the cart may or may not have made it to San Francisco  
past tense is still there; probably lack of inference is the result of the aspectual viewpoint. But we still don't know the source of the legitimate inference in (a)...

Point is, it's not easy, always, to pin down the systemic "source" of a piece of semantic information.

17. *Types of information which are relevant to lexical semantics:*

- a. Categorical information: N, V, Adj, Adv, Tns, Agr...  
b. Semantic type: predicative (picks out a set, e.g. *cat*, (*to*) *smoke*), relational (picks out pairs of individuals e.g. *love*, *aware*, *brother*), quantificational (relates sets of sets, e.g. *every*, *neither*), propositional (operates on propositions, e.g. *if..then*, *or*).  
c. Selectional characteristics: animate, count/mass... etc.

18. Pustejovsky on verb alternations: "...we must realize that explaining the behavior of a verb's semantic class can come only from acknowledging that the syntactic patterns in an alternation are not independent of the information carried by the arguments characterized in the very patterns themselves. In other words, the diversity of complement types that a verb or other category may take is in large part also determined by the semantics of the complements themselves."

— two messages here:

- (i) the patterns carry some meaning (c.f. *kick the bucket* above)  
(ii) the arguments carry some meaning. (c.f. *have* or *keep* last week).

19. Another class of lexical alternations: complement drop
- (a) Mary tried to start her car in the morning.  
(b) Mary tried again in the morning.  
(c) Mary attempted to start her car in the morning.  
(d) \*Mary attempted again in the morning.

"As we can see, no one semantic parameter will be sufficient to explain all complement drop cases."

hah! he didn't even try.

20. Vendler's aspectual classes:

State: *Mary knows French.*

\*Mary is knowing French.

Activity: *Mary walked.*

Mary walked for an hour/\*in an hour.

Accomplishment: *Mary built a house*

Mary built a house \*for a year/in a year.

Achievement: *Mary won the race*

Mary won the race at 3:15.

(Semelfactive (Smith 1993): *Mary kicked the door.*

\*This door kicks easily.)

21. Nominal alternations

(a) Mass/Count: water/cat

(b) To use mass nouns in a count sense, they need to be *packaged*.

Convention can package them: I'll have two coffees please. But usually, something like a container/quantity is specified: a glass of water, a stick of gum, etc. This type of packaging is often the diachronic source of nominal classifier systems.

(c) Group/individual nouns: class/student

(d) relational (vs. predicative): brother of, companion to, enemy, part of...

(e) taxonomic relations (WordNet): (Thing, Living, Animal, Vertebrate, Mammal, Feline, Cat) (Thing, Non-Living, Artificial, Transportation, Overland Vehicle, Car).

Pustejovsky on taxonomy: "The taxonomic classification of objects in the world through language can be a serious enterprise and not merely metaphysical play."

22. Adjectival information:

(a) Individual/stage level (*hungry/intelligent*)

(b) Intersective/non-intersective (*black/former*)

(c) Gradable/non-gradable (*small/former*)

(d) Relational (*aware of, envious of*)

(e) Raising/non-raising (*easy/possible*)

23. Dixon's adjective classes: Dimension, Physical Property, Color, Human Propensity, Age, Value, Speed, Difficulty, Similarity, Qualification.

(Where does an adjective like *metaphysical* or *taxonomic* come in?)

Pustejovsky on Dixon: "What is needed, I believe, is a semantic classification that captures the intuitions listed by Dixon but based on the relational and logical behavior of the predicates and not on their folk-epistemology."

24. Lexical Relations

(a) Synonymy: can be contextually defined, e.g. *plank* and *board* in carpentry but not in, e.g. piracy or home entertainment. Defined over "pairs of lexical items rather than concepts". ???

(b) Antonymy: *fast/slow, heavy/light, dark/light...* Notes *rise~ascend, fall~descend*, and says, basically, although these are synonyms, and each the antonym of its respective opposite, *rise/descend, fall/ascend* are not antonymic. This seems pretty nonsensical to me: "descend" is as much an antonym of "rise" as "ascend" is a synonym, which is to say, not much.

(c) Meronymy: part/whole relations.

(d) Entailment/presupposition: discussed earlier. Notes an additional test distinguishing one from the other:

(i) #Mary killed John, but he's not dead.

(ii) Toot is a cat, but she's not fuzzy/an animal.

Relevant presuppositions:

(iii) Ownership in selling/buying? consider a McDonald's worker, or a travel agent.

(iv) Factivity and complement type: *forget that* vs. *forget to*.

#### 4. A Concrete Proposal: Jackendoff 1992

Concepts are represented by a conceptual structure + a three-dimensional representation. The conceptual structure is constructed according to the following schema (p. 24):

(25) [Entity]  $\left( \begin{array}{l} \text{Event/Thing/Place} \\ \text{Token/Type} \\ \text{F}(\langle \text{Entity}_1, \langle \text{Entity}_2, \langle \text{Entity}_3 \rangle \rangle \rangle) \end{array} \right)$

Line 1: Feature complex denoting semantic category: 'major ontological categories': Thing, Event, State, Action, Place, Path, Property, Amount

Line 2: Feature complex denoting classification

Line 3: Argument structure

One can state rules of inference from one entry to the next:

(26) At the termination of [<sub>Event</sub> GO ([X], [<sub>Path</sub> TO ([Y])])], it is the case that [<sub>State</sub> BE ([X], [<sub>Place</sub> AT ([Y])])]

(27) The 3-D model distinguishes things that otherwise have identical Conceptual Structure: *duck* vs. *goose, chair* vs. *stool, run* vs. *jog, toss* vs. *lob*. Note that this makes a strong prediction: since it is Conceptual Structure which interacts with the syntax, no difference represented only in the 3-D structure can show a syntactic reflex.