

Chapter 4

Complex Feature Values

4.1 Lecture notes

Chapter Four, Lecture One

I. Pep talk

- We're getting into the hardest part of the course. In chapters 3-8, we develop a lot of machinery, which we use to analyze more interesting phenomena in chapters 10-15. It's important to master the technical stuff we present in this part of the course, so you have the tools to do the fun stuff later on.

II. Review: The story so far

- In trying to use CFG to model English, we ran into the problem of finding the right level of granularity in classifying words into categories.
- ? What factors led to making many separate categories for verbs? (Valence)
- ? What factors led to wanting one general category for verbs?
- The solution we proposed was to enrich the system of categories by giving categories internal structure.
- ? What are the advantages of this? (Simpler rules, accommodation of lexical idiosyncrasy)

? What are the disadvantages? (Redundancy in the lexicon – we’ll get to a way to deal with that in Chapter 8.)

III. Valence features

- A big step in the process of moving information from the grammar to the lexicon is the replacement of VAL (which took atomic values) with COMPS and SPR (which take lists of feature structures as their values).
- This move collapses the rules even more: The various VP rules are all replaced with one, general Head-Complement Rule. [Slides:1-2]
- The Head-Complement Rule can also handle cases that weren’t covered by the old rules, and would have required many separate rules.

? What are some examples of valence patterns not covered by the old rules?

- This approach to grammar makes a prediction about the world’s languages, because the grammar is simpler if different phrase types (VPs, NPs, PPs) have parallel internal structure.
- Suppose English were almost exactly as it is, except that it had postpositions instead of prepositions. Instead of saying (1) we would say (2).

(1) Chris talked to Pat about Sandy.

(2) Chris talked Pat to Sandy about.

? What would we have to change in our grammar to model this version of English? (We would need a separate rule for PPs, distinct from the main Head-Complement Rule. [Slides:3])

- So the schematization of the grammar rules means that languages which don’t have parallelism across phrase types require more complex grammars.
- If we assume that languages with simpler grammars should be favored over those with more complex grammars, then this predicts that languages with parallel structure across phrase types should be more common than languages without that parallelism.

This prediction is borne out: human languages are predominately head-initial or head-final. By this we mean that lexical heads of all categories **tend** to uniformly come first in their phrase or last in their phrase.

- When a feature of grammar that makes things simpler is also very common across the world’s languages, linguists call it ‘unmarked’. It is unmarked for languages to display parallelism across phrase types.

IV. How do COMPS and SPR work?

- The COMPS and SPR features of a lexical item specify a list of feature structures indicating what other words the lexical item requires.
- The Head-Complement and Head-Specifier Rules specify how those lists correspond to local trees in which the specifier and complements are realized. [Slides:4]
- ? Why does the SPR feature take a list of feature structures as its value instead of just a feature structure? (For a uniform representation of unfulfilled valence requirements as non-empty lists and fulfilled valence requirements as empty lists.)
- ? Which rule applies lower in the tree? (HCR) Why? (Because the HCR requires a *word* as its head daughter and head specifier phrases will always be of type *phrase*. [Slides:5])
- ? How does the SPR requirement get transmitted to the mother of a Head-Complement Rule? (The Valence Principle: [Slides:6])
- The Valence Principle expresses the generalization that a word’s valence requirements stick around until they get realized by a word or phrase in the tree.
- This is why the textbook discusses the metaphor of ‘cancellation’.
- It is important to note, however, that this does not mean that the grammar only works bottom-up. The Valence Principle, like everything else in the grammar, is only a constraint on local trees.

V. Complex feature values and CFG equivalence

- Feature structure categories are a generalization of context-free grammar.
- The grammar of Chapter 3 was in fact equivalent to a CFG – one with many different categories and many different rules. (You have to compile out the schematized rules into the different possibilities.)
- The COMPS and SPR features change this, however, because they allow for recursive feature structures. [Slides:7] (In general, it is the possibility of feature-structure valued features that allows for recursive feature structures, but in our grammar COMPS and SPR are the first place where this actually happens.)
- It is possible that if we assumed some further constraints, e.g. (1) that COMPS lists are always finite, (2) that elements on SPR lists are all [SPR ⟨ ⟩] and [COMPS ⟨ ⟩] and (3) that elements on COMPS lists are all [COMPS ⟨ ⟩], that we could put an upper bound on the set of categories used by our grammar. In this case, we would be able to show that, for any grammar we write, there is always a CFG that is strongly equivalent to it. By ‘strongly equivalent’, we mean a grammar that would assign to every grammatical sentence the same tree structure, but would have an appropriate atomic category labelling the node (as a CFG must) instead of our feature structure category. Such CFGs would have a lot of categories, though...

Chapter Four, Lecture Two

I. Determiners and Subjects as Specifiers

- ? Why should we say that determiners and subjects are parallel?
1. Nominalizations: [Slides:8]
 - (3) a. The enemy destroyed the city.
 - b. The enemy’s destruction of the city
 2. Similar cooccurrence restrictions (number agreement): [Slides:9]
 - (4) a. They swim.
 - b.*They swims.
 - c. Several cars drove up.

d.*A cars drove up.

- If determiners and subjects are parallel, then NP and S are also parallel.
- ? Why? (They both represent the SPR and COMPS saturated level.)
- This means that we are treating the VP as the head daughter of S, and not the subject NP.
- ? What is the evidence for this?
 - In general, the head is the obligatory element of a phrase. So we find imperative Ss without any subject (*Get out of here!*) and other kinds of S where the subject has been dislocated (*Who do you think [__ will win the race]?*) [We return to this kind of S in Chapter 15.]. It is the essential element whose presence in the phrase determines what kind of phrase it is.
 - Selectional restrictions:
 - The information about the head is shared by the whole phrase. So if we have elements that select for a particular form of something down inside one of their complements, then we generally want that thing to be the head of the phrase.
 - For example, verbs may select singular or plural NPs, but they don't select for whether the subject NP has a determiner or a complement. So we say that the noun is the head of the NP, and number is a HEAD feature, but SPR and COMPS aren't.
 - Now some verbs take an S complement that has a subjunctive verb in it: [Slides:10]
 - (5) a. Kim insists (that) Pat leave.
 - b.*Kim insists (that) Pat leaves.
 - c.*Kim knows (that) Pat leave.
 - d. Kim knows that Pat leaves.
 - We can model this by saying that MOOD (subjunctive v. indicative) is a HEAD feature, and that V is the head of S. Then the mood information is on the S node, and the COMPS value of the upstairs verb can select for a complement S with a particular mood. [Slides:11]

- If N were the head of the S, we wouldn't have the information about the downstairs verb's mood at the S level.

II. PPs—Complements or Modifiers?

- As we start looking at more varied valence patterns, we run into a data problem: How do we tell when a PP after a verb is a modifier and when it is a complement?
- ? What did the textbook say?
 1. Obligatory PPs or ones with idiosyncratically selected prepositions are complements.
 2. PPs that denote essential participants are complements.
- Here's another test (based on work by Jack Hawkins): If X Ved (NP) PP does not entail X did something PP, then the PP is a complement. [Slides:12–15]
- The test doesn't go in both directions: (6) does entail (7) but the *to*-phrase is obligatory. [Slides:16]
 - (6) Pat handed a note to Chris
 - (7) Pat did something to Chris
- So this test is a sufficient but not necessary condition for complementhood.

III. Two kinds of agreement

- ? What are the two kinds of agreement in our fragment of English? (subject-verb and determiner-noun)
- ? Are they handled the same way? (No, NP-V agreement is all in the SPR value of the verb; D-N agreement is handled via a principle, the NSA.)
- In fact, the basic idea is the same in both cases: The SPR value of a lexical head specifies the AGR value of the specifier.
- Nouns serve as the head in one of these relations (D-N) and as the (head of the) specifier in the other (NP-V). So they need an AGR value (unlike V), but they also need a SPR value with an AGR in it (unlike D).

-
- That is, the N will have both an AGR value and a SPR value with an AGR value in it. If they aren't the same, we'll get things like: [Slides:17]
 - (8) a.*These fish eats plankton.
 - b.*This fish eat plankton.
 - These data show that although *fish* is underspecified for NUM, it can only have one value in any given sentence.
 - The lexical entry for *fish* without the NSA looks like this: [Slides:18]
 - In the first sentence, this gets filled in a bit like this: [Slides:19]
 - The NSA says that the two AGR values in that structure have to be the same, so the structure is bad. Nothing else rules it out.

IV. Practice

- For each of the following sentences, have the students explain how it is ruled out in terms of the Head-Specifier Rule, the NSA, the HFP and lexical entries (including count v. mass nouns when appropriate). [Slides:20]
 - (9) a.*These dog barked.
 - b.*Many chair was/were broken.
 - c.*These furniture are broken.
- ? How many people say: These data are murky?
 How many people say: This data is murky?
 How many people say: These data is murky?
 [Slides:21]
- ? How can we explain the variation? (*Data* is either a mass or a count noun. For many people, it is both. That is, many people have two lexical entries for *data*.)
- ? How can we account for the lack of variation with respect to **These data is...*
- Consider some data from Hindi. [Slides:22–25]
- ? What small change do we have to make to the grammar for Hindi VPs? (Hindi is head-final. Otherwise the same rules and same principles work.) [Slides:26]

- The only difference really is in the lexical entries. Hindi has underspecification in different places from English.
- ? What are the lexical entries for the words in these sentences?