Ling 566
Oct 11, 2022

Valence, Agreement

Announcements

- No Canvas answers from staff evenings/ weekends (but feel free to discuss amongst yourselves!)
- HW1 answer key available
- HW2 Ch 5, problem 3 is worth getting an early start on

Overview

- Review: pizza, feature structures, well-formed trees, HFP
- A problem with the Chapter 3 grammar
- Generalize COMPS and SPR
- The Valence Principle
- Agreement
- The SHAC
- Reading Questions

Pizza review

- Unification is an operation for combing constraints from different sources.
- What are those sources in the pizza example?
- Why do we need to combine information from different sources in our grammars?

Reminder: Where We Are

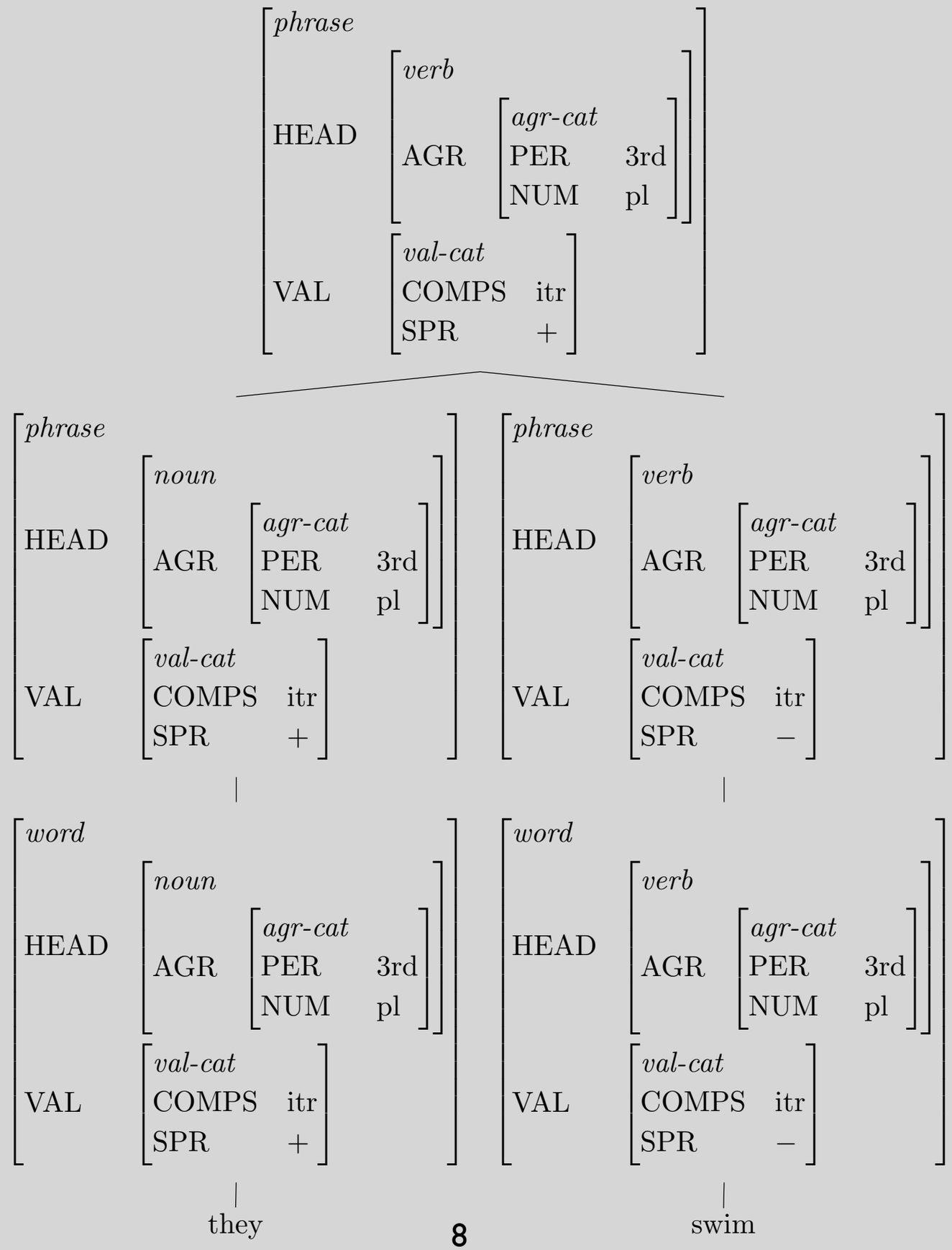
- Attempting to model English with CFG led to problems with the granularity of categories, e.g.
  - Need to distinguish various subtypes of verbs
  - Need to identify properties common to all verbs
- So we broke categories down into feature structures and began constructing a hierarchy of types of feature structures.
- This allows us to schematize rules and state cross-categorial generalizations, while still making fine distinctions.

A Tree is Well-Formed if …

- It and each subtree are licensed by a grammar rule or lexical entry
- All general principles (like the HFP) are satisfied.
- NB: Trees are part of our model of the language, so all their features have values (even though we will often be lazy and leave out the values irrelevant to our current point).

The Head Feature Principle

- Intuitive idea: Key properties of phrases are shared with their heads
- The HFP: In any headed phrase, the HEAD value of the mother and the head daughter must be identical.
- Sometimes described in terms of properties "percolating up" or "filtering down", but this is just metaphorical talk



But it's still not quite right…

- There's still too much redundancy in the rules.
- The rules and features encode the same information in different ways.

Head-Complement Rule 1:

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & - \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H} \begin{bmatrix} \textit{word} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & - \end{bmatrix} \end{bmatrix}$$

Head Complement Rule 2:

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & - \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H} \begin{bmatrix} \textit{word} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \text{str} \\ \text{SPR} & - \end{bmatrix} \end{bmatrix} \text{ NP}$$

Head Complement Rule 3:

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \text{itr} \\ \text{SPR} & - \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H} \begin{bmatrix} \textit{word} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \text{dtr} \\ \text{SPR} & - \end{bmatrix} \end{bmatrix} \text{ NP NP}$$

Solution:
More Elaborate Valence Feature Values

- The rules just say that heads combine with whatever their lexical entries say they can (or must) combine with.
- The information about what a word can or must combine with is encoded in list-valued valence features.
  - The elements of the lists are themselves feature structures
  - The elements are "cancelled" off the lists once heads combine with their complements and specifiers.

Complements

Head-Complement Rule:

$$\begin{bmatrix} phrase \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H} \begin{bmatrix} word \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \langle \boxed{1}, \ldots, \boxed{n} \rangle \end{bmatrix} \end{bmatrix} \boxed{1}, \ldots, \boxed{n}$$

- This allows for arbitrary numbers of complements, but only applies when there is at least one.
  - Heads in English probably never have more than 3 or 4 complements
  - This doesn't apply where Head-Complement Rule 1 would. (Why?)
- This covers lots of cases not covered by the old Head-Complement Rules 1-3. (Examples?)

Specifiers

Head-Specifier Rule (Version I)

$$\begin{bmatrix} phrase \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \rightarrow \boxed{2}\ \mathbf{H}\begin{bmatrix} \text{VAL} \quad \begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \boxed{2}\ \rangle \end{bmatrix} \end{bmatrix}$$

- Combines the rules expanding S and NP.
- In principle also generalizes to other categories.
- Question: Why is SPR list-valued?

SPR as a list

Love the symmetry with COMPS

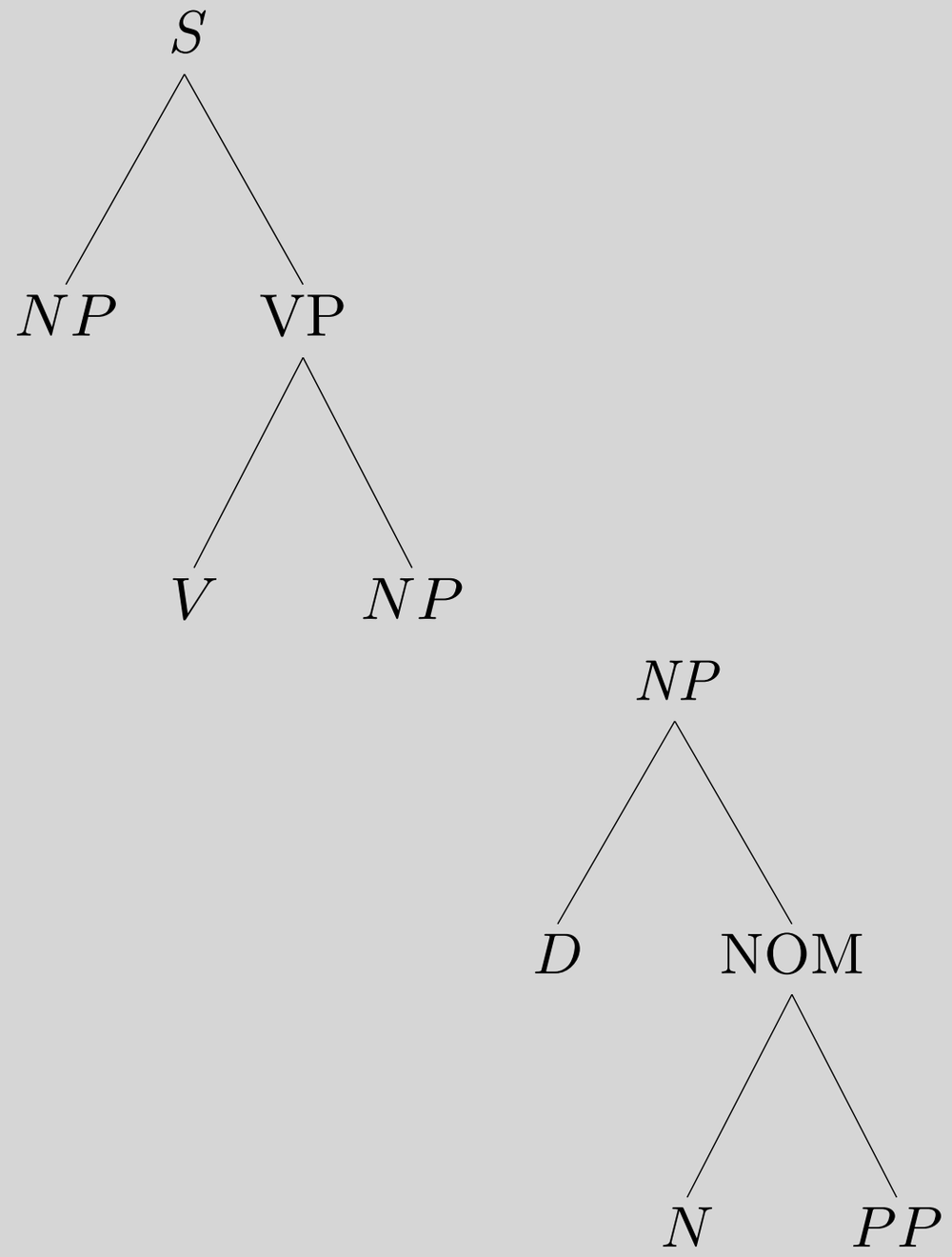
Lists that max out at 1 item are weird

I prefered +/-

None of the above

Question:

Why are these right-branching? That is, what formal property of our grammar forces the COMPS to be lower in the tree than the SPR?



Another Question…

What determines the VAL value of phrasal nodes?

ANSWER: The Valence Principle

Unless the rule says otherwise, the mother's values for the VAL features (SPR and COMPS) are identical to those of the head daughter.

More on the Valence Principle

- Intuitively, the VAL features list the contextual requirements that haven't yet been found.
- This way of thinking about it (like talk of "cancellation") is bottom-up and procedural.
- But formally, the Valence Principle (like the rest of our grammar) is just a well-formedness constraint on trees, without inherent directionality.

So far, we have:

- Replaced atomic-valued VAL features with list-valued ones.
- Generalized Head-Complement and Head-Specifier rules, to say that heads combine with whatever their lexical entries say they should combine with.
- Introduced the Valence Principle to carry up what's not "canceled".

The Parallelism between S and NP

- Motivation:
  - pairs like *Chris lectured about syntax* and *Chris's lecture about syntax*.
  - both S and NP exhibit agreement

    *The bird sings/*sing* vs. *The birds sing/ *sings*

    *this/*these bird* vs. *these/*this birds*
- So we treat NP as the saturated category of type *noun* and S as the saturated category of type *verb*.

Question: Is there any other reason to treat V as the head of S?

- In mainstream American English, sentences must have verbs. (How about other varieties of English or other languages?)
- Verbs taking S complements can influence the form of the verb in the complement:

  *I insist/*recall (that) you be here on time*.
- Making V the head of S helps us state such restrictions formally

A possible formalization of the restriction on *insist*

$$\left[\begin{array}{ll}\text{HEAD} & \textit{verb} \\ \text{VAL} & \left[\begin{array}{ll}\text{SPR} & \left\langle \text{NP} \right\rangle \\ \text{COMPS} & \left\langle \left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{ll}\textit{verb} \\ \text{MOOD} & \textit{subjunctive}\end{array}\right] \\ \text{VAL} & \left[\begin{array}{ll}\text{COMPS} & \left\langle\ \right\rangle \\ \text{SPR} & \left\langle\ \right\rangle\end{array}\right]\end{array}\right] \right\rangle \backslash \end{array}\right]\end{array}\right]$$

Note that this requires that the verb be the head of the complement. We don't have access to the features of the other constituents of the complement.

An Overlooked Topic: Complements *vs*. Modifiers

- Intuitive idea: Complements introduce essential participants in the situation denoted; modifiers refine the description.
- Generally accepted distinction, but disputes over individual cases.
- Linguists rely on heuristics to decide how to analyze questionable cases (usually PPs).

Heuristics for Complements vs. Modifiers

- Obligatory PPs are usually complements.
- Temporal & locative PPs are usually modifiers.
- An entailment test: If **X Ved (NP) PP** does not entail **X did something PP**, then the PP is a complement.

Examples

– *Pat relied on Chris* **does not entail** *Pat did something on Chris*
– *Pat put nuts in a cup* **does not entail** *Pat did something in a cup*
– *Pat slept until noon* **does entail** *Pat did something until noon*
– *Pat ate lunch at Bytes* **does entail** *Pat did something at Bytes*

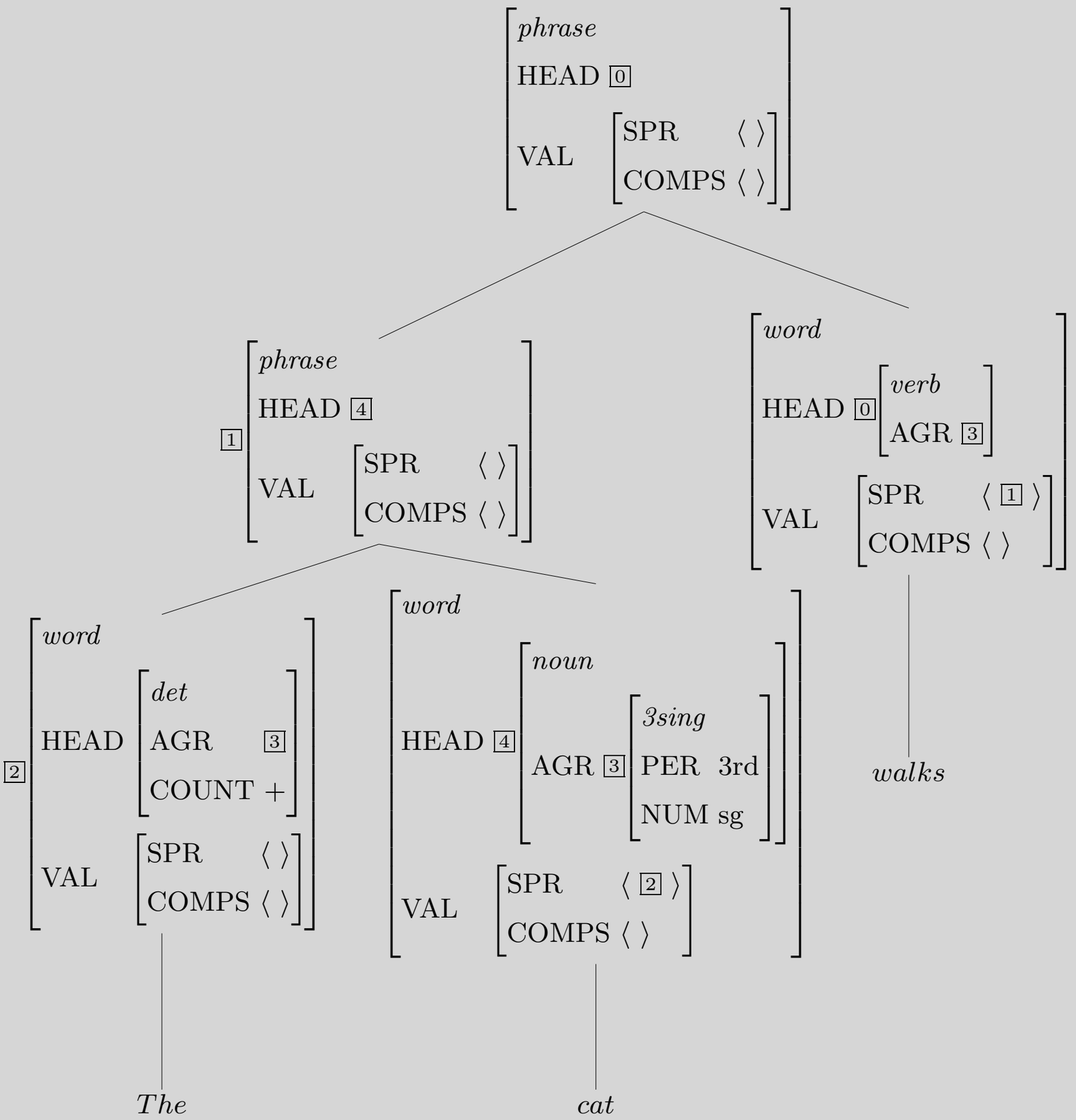
Agreement

- Two kinds so far (namely?)
- Both initially handled via stipulation in the Head-Specifier Rule
- But if we want to use this rule for categories that don't have the AGR feature (such as PPs and APs, in English), we can't build it into the rule.

The Specifier-Head Agreement Constraint (SHAC)

Verbs and nouns must be specified as:

$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \left\langle \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$



$$\left[\begin{array}{l} \textit{phrase} \\ \text{HEAD } \boxed{0} \\ \text{VAL } \left[\begin{array}{ll} \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

$$\left[\begin{array}{l} \textit{phrase} \\ \text{HEAD } \boxed{4} \\ \text{VAL } \left[\begin{array}{ll} \text{SPR} & \langle\ \boxed{1}\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

$$\left[\begin{array}{l} \textit{word} \\ \text{HEAD } \boxed{0}\left[\begin{array}{l} \textit{verb} \\ \text{AGR } \boxed{3} \end{array}\right] \\ \text{VAL } \left[\begin{array}{ll} \text{SPR} & \langle\ \boxed{1}\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

$$\left[\begin{array}{l} \textit{word} \\ \text{HEAD } \left[\begin{array}{ll} \textit{det} & \\ \text{AGR} & \boxed{3} \\ \text{COUNT} & + \end{array}\right] \\ \text{VAL } \left[\begin{array}{ll} \text{SPR} & \langle\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

$$\left[\begin{array}{l} \textit{word} \\ \text{HEAD } \boxed{4}\left[\begin{array}{ll} \textit{noun} & \\ \text{AGR } \boxed{3} & \left[\begin{array}{ll} \textit{3sing} & \\ \text{PER} & \text{3rd} \\ \text{NUM} & \text{sg} \end{array}\right] \end{array}\right] \\ \text{VAL } \left[\begin{array}{ll} \text{SPR} & \langle\ \boxed{2}\ \rangle \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

walks

The Count/Mass Distinction

- Partially semantically motivated
  - mass terms tend to refer to undifferentiated substances (*air, butter, courtesy, information*)
  - count nouns tend to refer to individuatable entities (*bird, cookie, insult, fact*)
- But there are exceptions:
  - *succotash* (mass) denotes a mix of corn & lima beans, so it's not undifferentiated.
  - *furniture, footwear, cutlery*, etc. refer to individuatable artifacts with mass terms
  - *cabbage* can be either count or mass, but many speakers get *lettuce* only as mass.
  - borderline case: *data*

Our Formalization of the Count/Mass Distinction

- Determiners are:
  - [COUNT –] (*much* and, in some dialects, *less*),
  - [COUNT +] (*a, six, many*, etc.), or
  - lexically underspecified (*the, all, some, no*, etc.)
- Nouns select appropriate determiners
  - "count nouns" say SPR <[COUNT +]>
  - "mass nouns" say SPR <[COUNT –]>
- Nouns themselves aren't marked for the feature COUNT
- So the SHAC plays no role in count/mass marking.

Overview

- Review: pizza, feature structures, well-formed trees, HFP
- A problem with the Chapter 3 grammar
- Generalize COMPS and SPR
- The Valence Principle
- Agreement
- The SHAC
- Reading Questions

Reading Questions

- In section 4.3, moving away from the types 'word' vs 'phrase' is described, but most of the examples after that in the chapter still use them. Should we still include them in our feature structures for now?

- The book states that "the binary distinction between words and phrases is largely replaced by a more nuanced notion of 'degree of saturation'." The use of "largely" in this explanation has me wondering if there's something more to the word vs. phrase distinction that isn't captured by the degrees of saturation?

Reading Questions

- Section 4.3 specifically mentions under example (20) that we are trying to shift away from phrases and words, but section 4.4 notes under example (27) that a key constraint on how head-complement phrases and head-specifier phrases embed is that for the Head-Complement Rule the head daughter must be of the type word and for the Head-Specifier Rule the mother must be of the type phrase. Will we eventually end up with a set of rules that can be properly embedded without relying on the types phrase and word?

Reading Questions

- Is the removal of the non-branching nodes compulsory?
- Are all sentences head-specifier phrases that may contain sub-units/sub-trees of other phrasal types?

Reading Questions

- Is there a benefit to writing lexical items as being able to take a specifier rather than identifying words that can be specifiers?

Reading Questions

- What sorts of agreement values can appear in other (non-English) languages?
- I see that we have started to generalize and have a more or less binary category for AGR for verbs, namely 3sing and non-3sing. Cross-linguistically, this distinction doesn't hold very (for example in Spanish). Do we create a framework for subject-verb agreement on a language-by-language basis? How is this framework used for pro-drop languages, such as Spanish where the subject is optional but implied?

Reading Questions

- How do we deal with free word order languages?

Reading Questions

- It seems like an increasing amount of information is encoded in the lexical entries of words in this chapter. Does this necessitate a bottom-up approach for parsing sentences? (computationally or otherwise)