Ling 566
Oct 8, 2019

Valence, Agreement

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

$$
\begin{bmatrix}
\textit{phrase} & \\
\text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AGR} & \begin{bmatrix} \textit{agr-cat} & \\ \text{PER} & \text{3rd} \\ \text{NUM} & \text{pl} \end{bmatrix} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix} \textit{val-cat} & \\ \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix}
\end{bmatrix}
$$

$$
\begin{bmatrix}
\textit{phrase} & \\
\text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \begin{bmatrix} \textit{agr-cat} & \\ \text{PER} & \text{3rd} \\ \text{NUM} & \text{pl} \end{bmatrix} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix} \textit{val-cat} & \\ \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix}
\end{bmatrix}
\qquad
\begin{bmatrix}
\textit{phrase} & \\
\text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AGR} & \begin{bmatrix} \textit{agr-cat} & \\ \text{PER} & \text{3rd} \\ \text{NUM} & \text{pl} \end{bmatrix} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix} \textit{val-cat} & \\ \text{COMPS} & \text{itr} \\ \text{SPR} & - \end{bmatrix}
\end{bmatrix}
$$

$$
\begin{bmatrix}
\textit{word} & \\
\text{HEAD} & \begin{bmatrix} \textit{noun} & \\ \text{AGR} & \begin{bmatrix} \textit{agr-cat} & \\ \text{PER} & \text{3rd} \\ \text{NUM} & \text{pl} \end{bmatrix} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix} \textit{val-cat} & \\ \text{COMPS} & \text{itr} \\ \text{SPR} & + \end{bmatrix}
\end{bmatrix}
\qquad
\begin{bmatrix}
\textit{word} & \\
\text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AGR} & \begin{bmatrix} \textit{agr-cat} & \\ \text{PER} & \text{3rd} \\ \text{NUM} & \text{pl} \end{bmatrix} \end{bmatrix} \\
\text{VAL} & \begin{bmatrix} \textit{val-cat} & \\ \text{COMPS} & \text{itr} \\ \text{SPR} & - \end{bmatrix}
\end{bmatrix}
$$

they swim

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:

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

Head Complement Rule 2:

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

Head Complement Rule 3:

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

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:

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H}\begin{bmatrix} \textit{word} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \left\langle \boxed{1}\ , \dots\ , \boxed{n} \right\rangle \end{bmatrix} \end{bmatrix} \boxed{1}\ , \dots\ , \boxed{n}$$

- 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

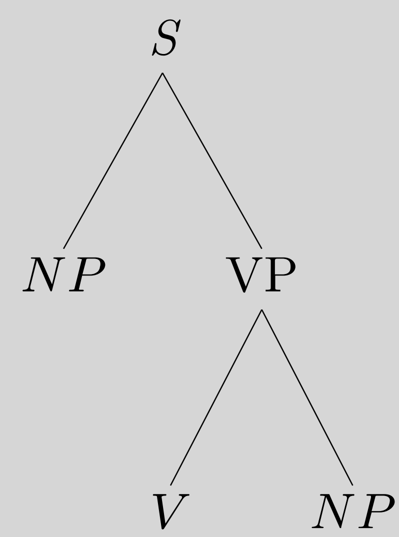
Head-Specifier Rule (Version I)

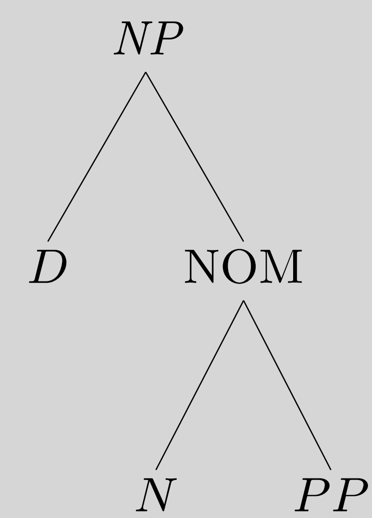
$$\begin{bmatrix} \textit{phrase} & \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \rightarrow \boxed{2} \quad \mathbf{H}\begin{bmatrix} \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle \boxed{2}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

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

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*

$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \text{NP} \rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{MOOD} & \textit{subjunctive} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\ \rangle \\ \text{SPR} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

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*

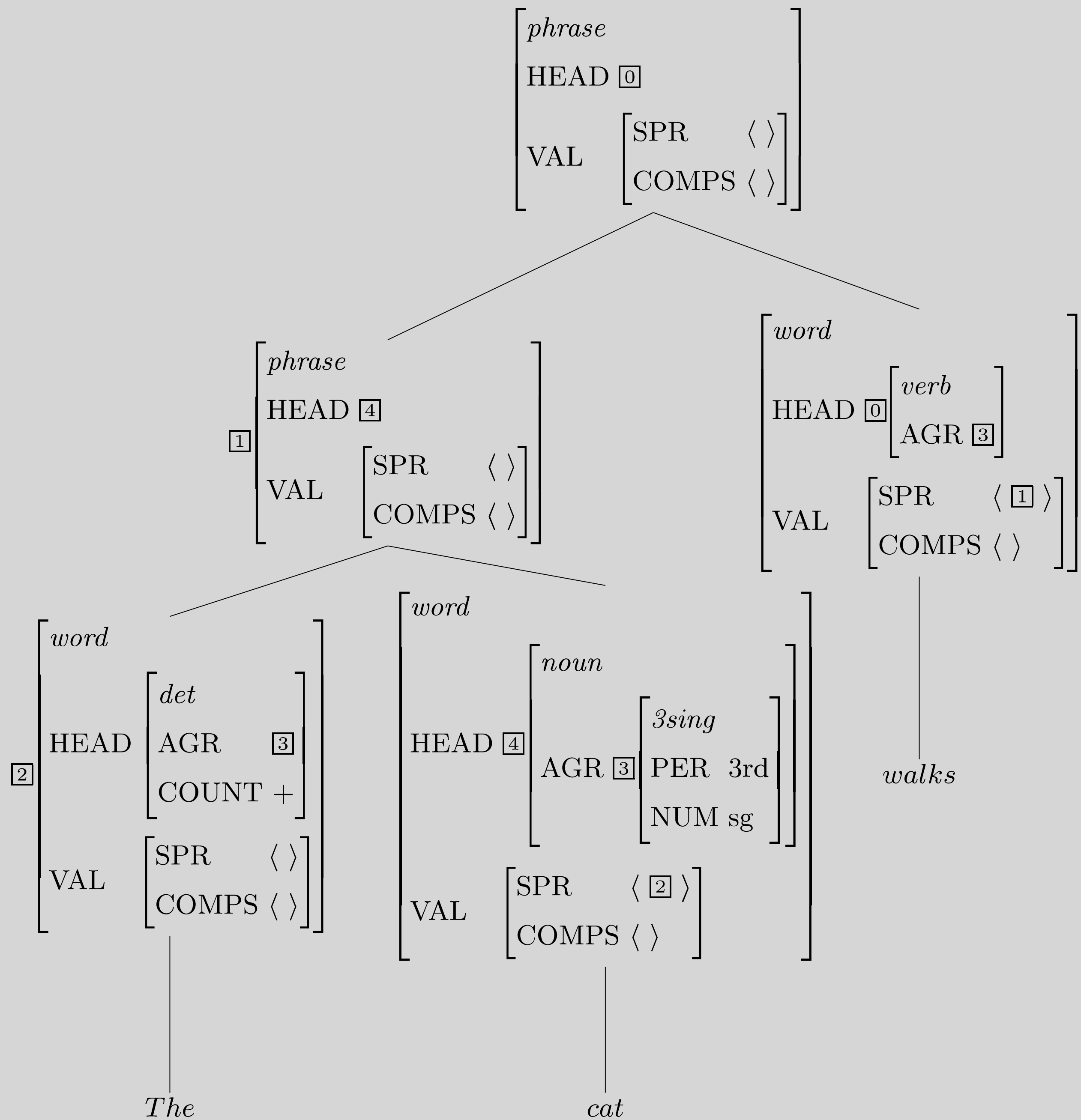
Agreement

- 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:

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


phrase
HEAD 0
VAL SPR ⟨ ⟩ COMPS ⟨ ⟩
1 phrase
HEAD 4
VAL SPR ⟨ ⟩ COMPS ⟨ ⟩
word
HEAD 0 verb AGR 3
VAL SPR ⟨ 1 ⟩ COMPS ⟨ ⟩
2 word
HEAD det AGR 3 COUNT +
VAL SPR ⟨ ⟩ COMPS ⟨ ⟩
word
HEAD 4 noun AGR 3 3sing PER 3rd NUM sg
VAL SPR ⟨ 2 ⟩ COMPS ⟨ ⟩
walks
The
cat

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

- Confusing: "The effect of the Valence Principle is that: appropriate elements mentioned in particular rules are canceled from the relevant valence specifications of the head daughter in head-complement or head-specifier phrases”
- "we mean simply that the Valence Principle is enforced unless a particular grammar rule specifies both the mother's and the head daughter's value for some valence feature." Is there any real - world example of this? Does any human language specifies the difference in the values of valence features between mother and head daughter?

Reading Questions

- "Because no complements or specifiers are introduced by this rule, we do not want any cancellation from either of the head daughter's valence features." Sec 4.5 under Valence Principle. What does this sentence mean exactly? Does this sentence allude to the fact that Head Specifier Rule must preserve the COMPS values and the Head Complement Rule must preserve the SPR values?
- How HFP propagates the values from the mother to the head daughter top-down seems to be clear. Then, would the Valence principle be a similar approach but in a bottom-up manner? I would like to see more examples of how it works in contrast with HFP.

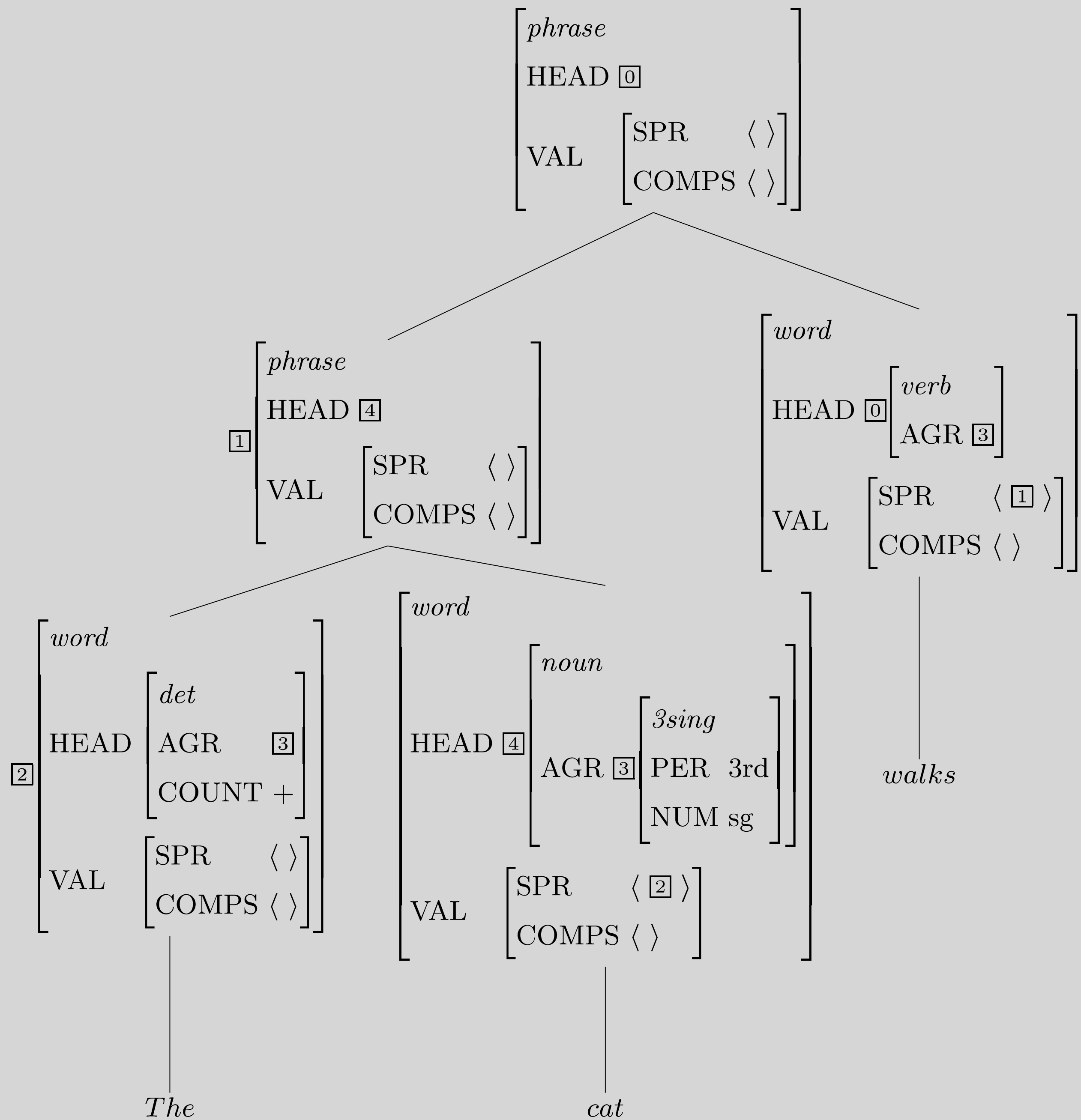
Reading Questions

- The Valence Principle states: "Unless the rule says otherwise, the mother's values for the VAL features are identical to those of the head daughter"

- However, in a well formed tree, the book states that the head daughter's VAL values should be checked off by the sister's of the node, therefore, by the time the daughter reaches the mother, the mother's VAL values should be empty lists. I am a little confused about how these constraints are compatible. Does this rule just apply to intermediate projections?

Reading Questions

- Why in example (26) "Alex likes the opera", does the node for "the" not take a comps value? Wouldn't it make sense for there to be an NP in the comps value since it'd be ungrammatical to just say "the"?
- On page 106, why is the NP ("the opera") the complement value given to the V ("likes"), whereas the VP ("likes the opera") is not the complement of the N ("Alex"), and vice versa?


phrase
HEAD 0
VAL SPR ⟨ ⟩ COMPS ⟨ ⟩
1 phrase
HEAD 4
VAL SPR ⟨ ⟩ COMPS ⟨ ⟩
word
HEAD 0 verb AGR 3
VAL SPR ⟨ 1 ⟩ COMPS ⟨ ⟩
2 word
HEAD det AGR 3 COUNT +
VAL SPR ⟨ ⟩ COMPS ⟨ ⟩
word
HEAD 4 noun AGR 3 3sing PER 3rd NUM sg
VAL SPR ⟨ 2 ⟩ COMPS ⟨ ⟩
The
cat
walks

Reading Questions

- I am still not clear what is the purpose of COMPS<> in the head(mother) node? As per the definition, it looks like it is always empty(along with SPR<> for saturated expressions). Is there a reason for it to be there even if is it going to be an empty list anyways?
- Is there ever a case where COMPS will have a non-empty value for a phrase node? In this chapter the grammar doesn't allow it, but is this an oversimplification?

Reading Questions

- In the section on Complements vs. Modifiers it was discussed how constituents like PPs can function as both and how, although it would be difficult to create a formal rule, the distinction "should be reflected in a formal theory of grammar". Like the text said, it seems that it would be difficult to include in the lexical entries, but will we later have features to distinguish whether a phrase is a complement or a modifier?

Reading Questions

- If a specific NOM could take a determiner but doesn't (e.g. *dogs* in *Dogs like me*. And you could have said *Some dogs like me*.)-- does it still have to have a non-empty value for SPR, i.e. SPR <D>?

4.10.4 The Grammar Rules

(61) Head-Specifier Rule

$$\begin{bmatrix} \textit{phrase} & \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \rightarrow \boxed{1} \quad \mathbf{H}\begin{bmatrix} \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix}$$

(62) Head-Complement Rule

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

(63) Head-Modifier Rule

$$\begin{bmatrix} \textit{phrase} \end{bmatrix} \rightarrow \mathbf{H}\begin{bmatrix} \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle\ \rangle \end{bmatrix} \end{bmatrix} \ \text{PP}$$

(64)

Coordination Rule

$$\begin{bmatrix} \text{VAL } \boxed{1} \end{bmatrix} \rightarrow \begin{bmatrix} \text{VAL } \boxed{1} \end{bmatrix}^{+} \begin{bmatrix} \textit{word} & \\ \text{HEAD} & \textit{conj} \end{bmatrix} \begin{bmatrix} \text{VAL } \boxed{1} \end{bmatrix}$$

Reading Questions

- I wonder if COUNT is only a feature of determiners of nouns. What about NOMs? Also, when we draw trees, should words like *dogs* and *water* which don't need a specifier be treated as NP-Ns or NP-NOM-Ns?

- The section on COUNT (section 4.6.3) analyzes whether count is a semantic concept or a syntactic concept. It seems like we only decide to develop a rule because we determine that it's a matter of syntax. Even if this were a purely semantic concept, wouldn't we still need to consider it in our grammar? I don't fully understand this distinction.

Reading Questions

- What is the difference between letter tags and number tags? Page 105 states that the letter tags represent lists of feature structures rather than individual feature structures, but why does SPR take a list while COMP doesn't? The examples make it seem like they take the same parameters.

Reading Questions

- We move to doing these tasks computationally, how do we attain this necessary knowledge of the lexicon? Do we have to store information about every word that's likely to occur and it's necessary complements?

Reading Questions

- Is overgeneration allowed in syntactic theory? When trying to describe the features of human language, the rules we designed either narrow or overgenerate the real language. In the feature CASE, suppose we conclude that the accusative case should come after a verb, but how does this apply to ‘me too’? How would syntactic theory deal with some special cases in language?

Reading Questions

- What’s the difference between small clauses and absolutives?
    - We want [them on our team]
    - With [them on our team], we’ll be sure to win