



Ling 566  
Oct 13, 2020  
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 combining 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

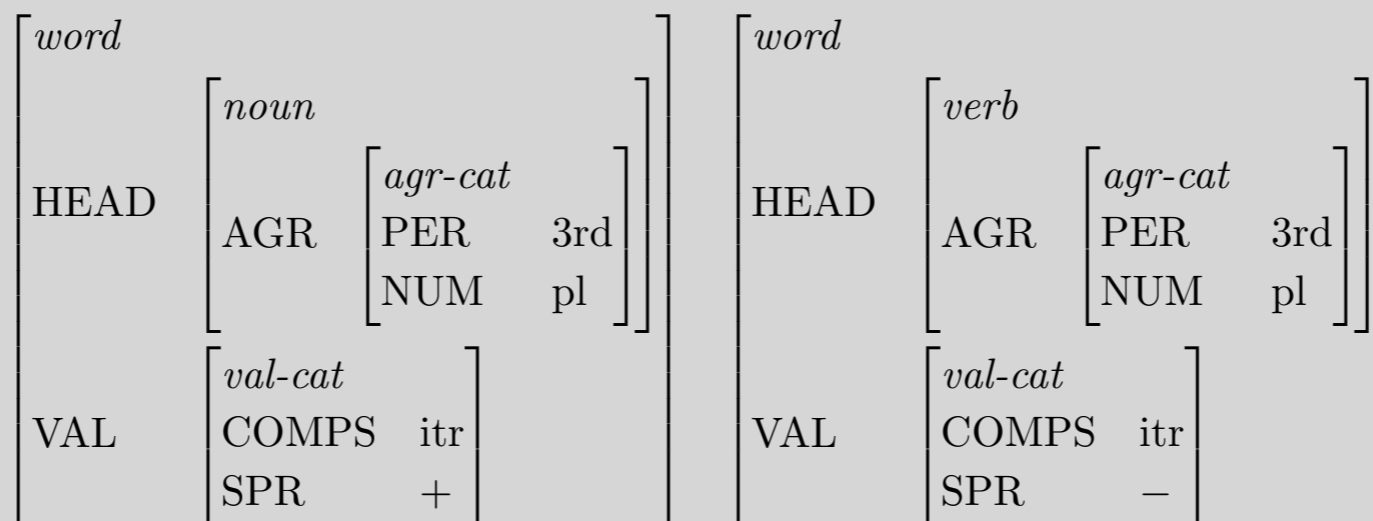
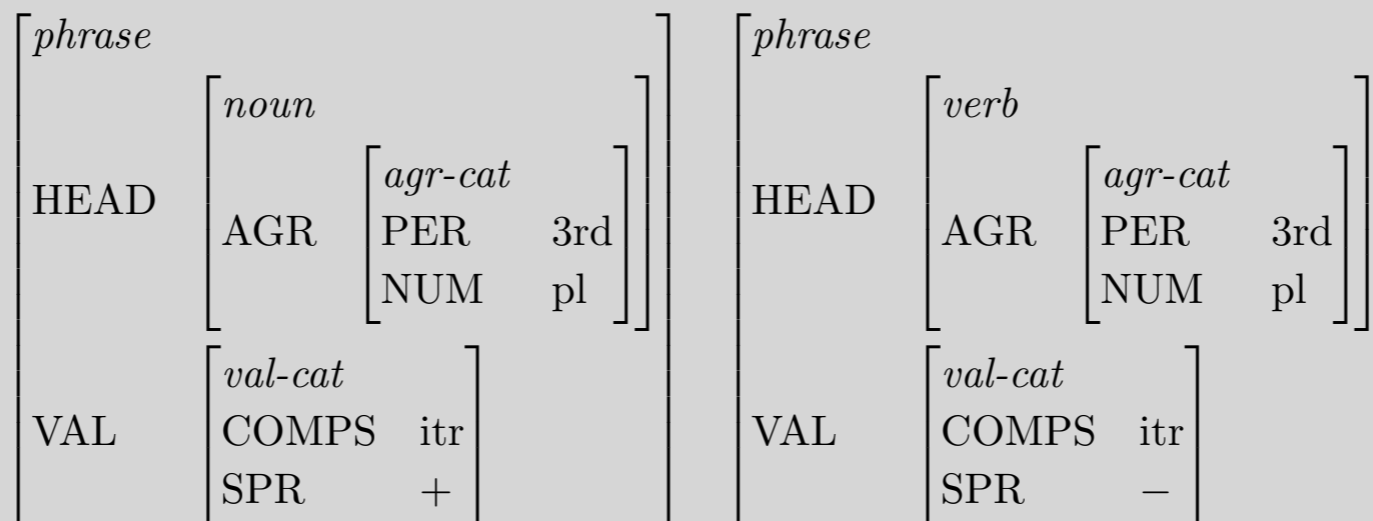
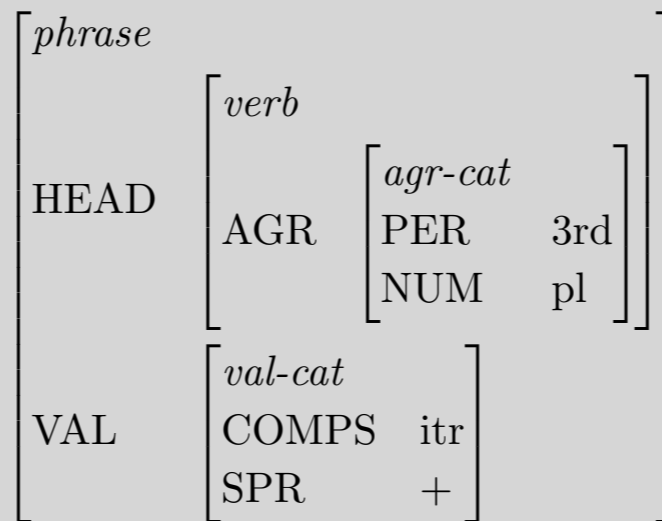
- Attempting to model English with CFG led to problems with the granularity of categories, e.g.
  - Need to distinguish various subtypes of verbs
  - Need to identify properties common to all verbs
- So we broke categories down into feature structures and began constructing a hierarchy of types of feature 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



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:

$$\left[ \begin{array}{l} \textit{phrase} \\ \text{VAL} \end{array} \left[ \begin{array}{ll} \text{COMPS} & \textit{itr} \\ \text{SPR} & - \end{array} \right] \right] \rightarrow \mathbf{H} \left[ \begin{array}{l} \textit{word} \\ \text{VAL} \end{array} \left[ \begin{array}{ll} \text{COMPS} & \textit{itr} \\ \text{SPR} & - \end{array} \right] \right]$$

Head Complement Rule 2:

$$\left[ \begin{array}{l} \textit{phrase} \\ \text{VAL} \end{array} \left[ \begin{array}{ll} \text{COMPS} & \textit{itr} \\ \text{SPR} & - \end{array} \right] \right] \rightarrow \mathbf{H} \left[ \begin{array}{l} \textit{word} \\ \text{VAL} \end{array} \left[ \begin{array}{ll} \text{COMPS} & \textit{str} \\ \text{SPR} & - \end{array} \right] \right] \text{ NP}$$

Head Complement Rule 3:

$$\left[ \begin{array}{l} \textit{phrase} \\ \text{VAL} \end{array} \left[ \begin{array}{ll} \text{COMPS} & \textit{itr} \\ \text{SPR} & - \end{array} \right] \right] \rightarrow \mathbf{H} \left[ \begin{array}{l} \textit{word} \\ \text{VAL} \end{array} \left[ \begin{array}{ll} \text{COMPS} & \textit{dtr} \\ \text{SPR} & - \end{array} \right] \right] \text{ NP NP}$$



# 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:

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

- 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)

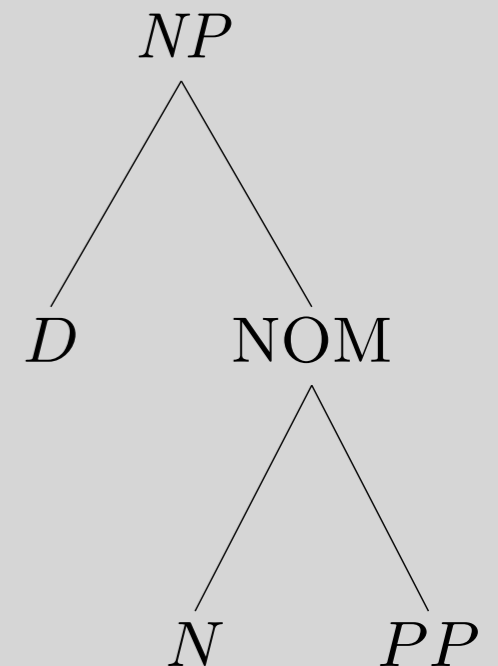
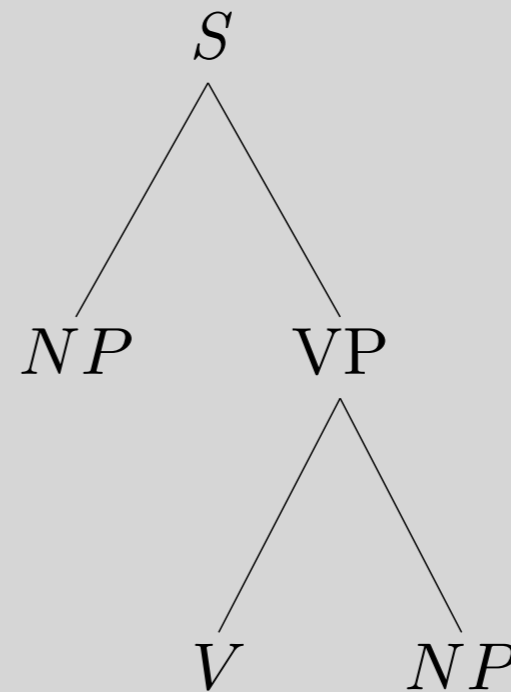
$$\left[ \begin{array}{l} \textit{phrase} \\ \text{VAL} \left[ \begin{array}{l} \text{COMPS} \langle \rangle \\ \text{SPR} \langle \rangle \end{array} \right] \end{array} \right] \rightarrow \boxed{2} \mathbf{H} \left[ \begin{array}{l} \text{VAL} \left[ \begin{array}{l} \text{COMPS} \langle \rangle \\ \text{SPR} \langle \boxed{2} \rangle \end{array} \right] \end{array} \right]$$

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

**Poll!**

# 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

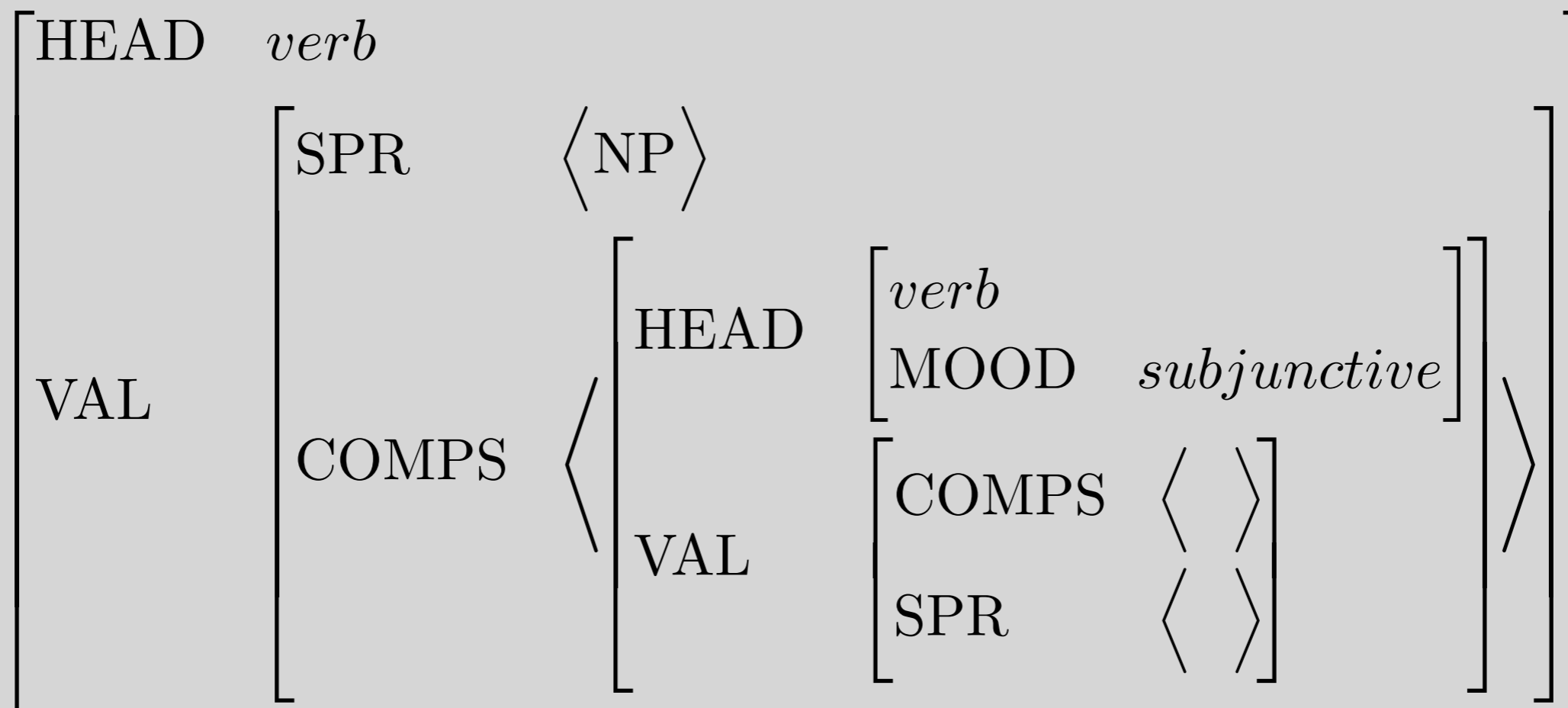
- Motivation:
  - pairs like *Chris lectured about syntax* and *Chris's lecture about syntax*.
  - both S and NP exhibit agreement
    - The bird sings/\*sing* vs. *The birds sing/\*sings*
    - this/\*these bird* vs. *these/\*this birds*
- So we treat NP as the saturated category of type *noun* and 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*



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*

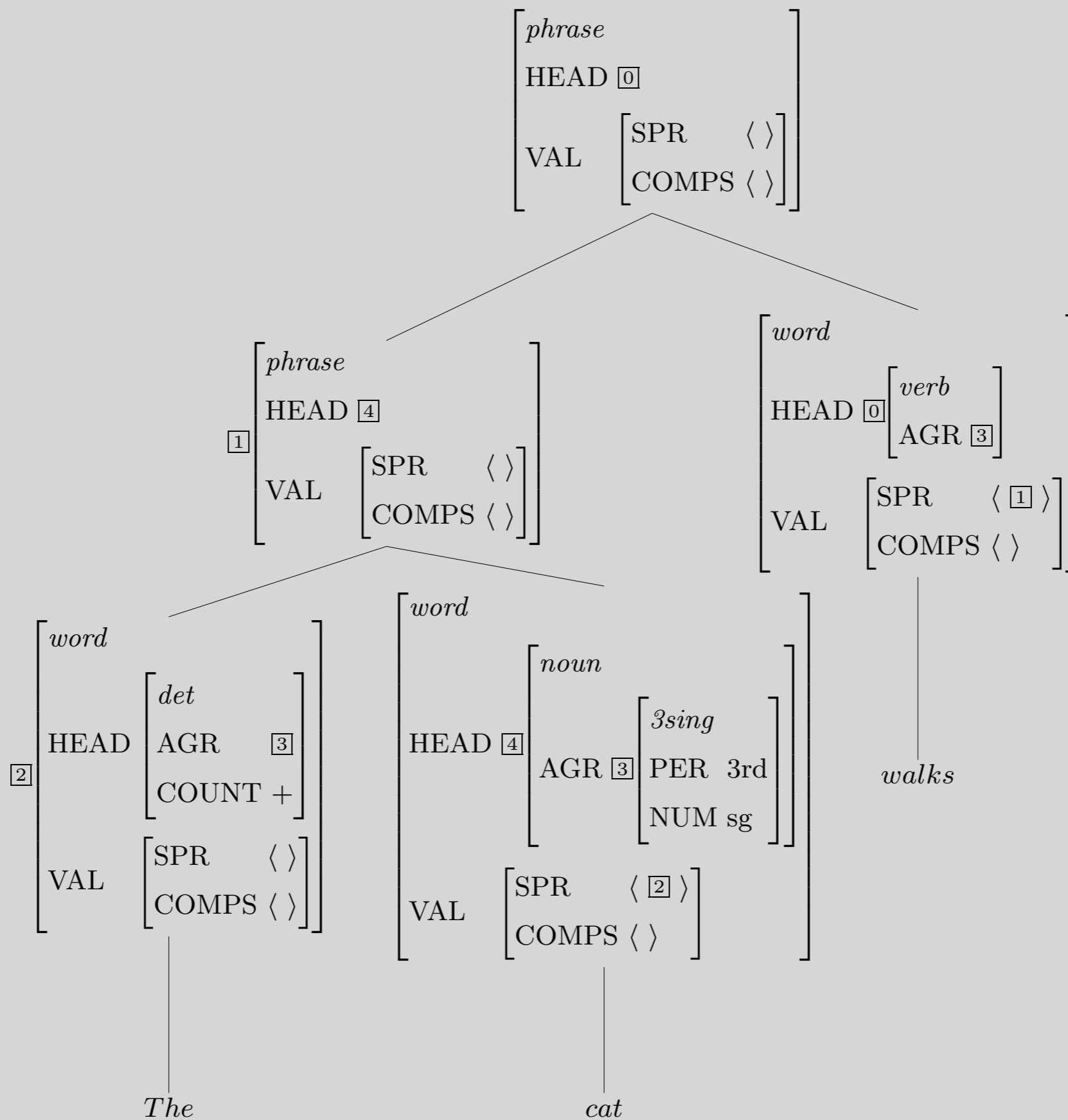
# Agreement

- 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:

$$\left[ \begin{array}{l} \text{HEAD} \left[ \text{AGR} \quad \boxed{1} \right] \\ \text{VAL} \left[ \text{SPR} \left\langle \left[ \text{AGR} \quad \boxed{1} \right] \right\rangle \right] \end{array} \right]$$



# 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

- How are lexicons generally stored? What is the lookup mechanism when searching for lexical entries?
- We've collapsed our grammar rules into 4 more general rules, but what are the consequences (good/bad) of making our lexicon more complex in the process? How can we find a good balance to make sure neither the grammar or the lexicon are over-/under-saturated?

# Reading Questions

- I'm a little confused as to why NOM has SPR<X>. What do we do in the case where (using CHP 3 grammar) SPR = -? Then there'd be an obligatory SPR X but we just said we didn't need one.
- Why do we need to pass up our AGR specification? What does this have to do with *I walk and Dan runs*?
- P102 footnote 8 says “tempted to accomplish this by making SPR a head feature” - can we give an example of why doing so would introduce unnecessary complication?

# Reading Questions

- On page 96, it is suggested that optional types of complements ([COMPS NP | S]) could be indicated with the | or operator, but then we introduce the list notation. Are optional COMPS still possible for a given lexical entry using the list notation? Or do we have separate lexical entries for each?
- I also still feel unclear about why we want the head daughter of our Head-Complement rule to be a word.

# Reading Questions

- I can appreciate that we may want to say that all values of the types of VAL should be lists, because it requires fewer codified exceptions to general rules. But if the choice is purely aesthetic, I wonder what other things we conventionally accept in HPSG that are like the decision to represent SPR's boolean structure as a list...
- In Section 4.3, it is mentioned that the book is shifting away from phrasality and a binary of words/phrases more towards "degree of saturation". I found this concept confusing. What exactly what a word or phrase be saturated in? How does this show up in our examples and why is this approach more beneficial than a word/phrase binary approach?
- Do we want to make feature COUNT and CASE as lists as well?

# Reading Questions

- One of the effects of the Valence Principle is that "the appropriate elements mentioned in particular rules are canceled from the relevant valence specifications of the head daughter in head-complement or head-specifier phrases". I am not completely sure what this means. Is this related to COMPS no longer being on the left-hand side of the rule?
- The Valence Principle: How do we know that this principle is applied when we see a grammar?

# Reading Questions

- I do not quite understand the modifications made to Head-Modifier Rule Version I to become Version II and how it solves the previously mentioned problems, such as modifiers being able to combine with a VP to build an S.



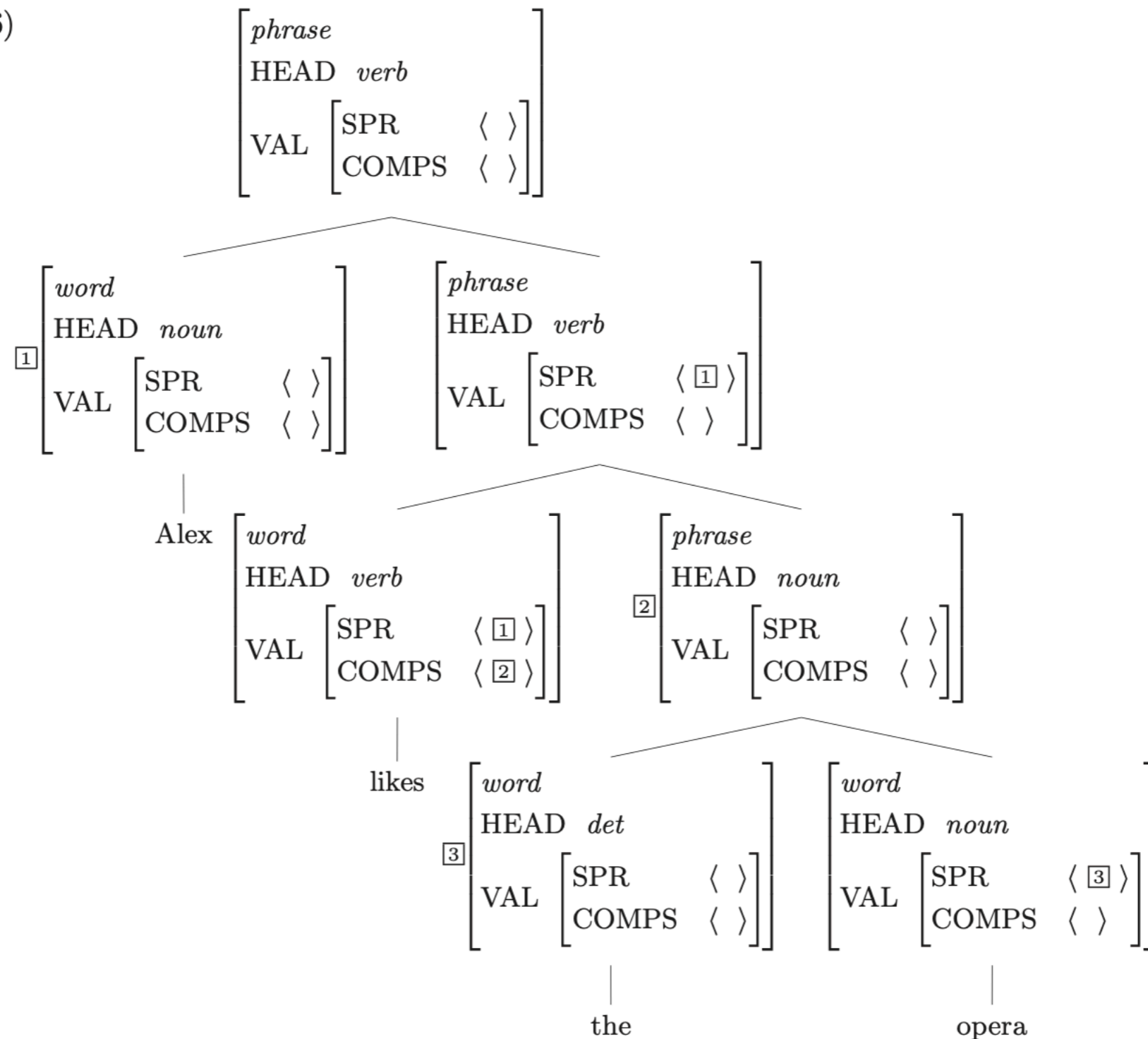
# Reading Questions

- In example 26, what did it mean by non-branching nodes? Why did Chapter 3 grammar initially cause the tree to have non-branching nodes?
- What exactly is making the tree here having no non-branching nodes? I guess the rules in this chapter is more generalized hence notation wise, some of the intermediary nodes basically end up being “merged”/generalized as one?

# Reading Questions

104 / SYNTACTIC THEORY

(26)



# Reading Questions

- Why do NOM, VP, NP, and S no longer need to mention the type phrase?

## 4.10.3 Abbreviations

(60)

$$\begin{array}{l} S = \left[ \begin{array}{l} \text{HEAD } \textit{verb} \\ \text{VAL } \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle \rangle \end{array} \right] \end{array} \right] \quad NP = \left[ \begin{array}{l} \text{HEAD } \textit{noun} \\ \text{VAL } \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle \rangle \end{array} \right] \end{array} \right] \\ \\ VP = \left[ \begin{array}{l} \text{HEAD } \textit{verb} \\ \text{VAL } \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle X \rangle \end{array} \right] \end{array} \right] \quad NOM = \left[ \begin{array}{l} \text{HEAD } \textit{noun} \\ \text{VAL } \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle X \rangle \end{array} \right] \end{array} \right] \\ \\ V = \left[ \begin{array}{l} \textit{word} \\ \text{HEAD } \textit{verb} \end{array} \right] \quad N = \left[ \begin{array}{l} \textit{word} \\ \text{HEAD } \textit{noun} \end{array} \right] \\ \\ D = \left[ \begin{array}{l} \textit{word} \\ \text{HEAD } \textit{det} \\ \text{VAL } \left[ \begin{array}{l} \text{COMPS } \langle \rangle \\ \text{SPR } \langle \rangle \end{array} \right] \end{array} \right] \end{array}$$

# Reading Questions

- Why did we bother mentioning GEND if English doesn't have grammatical gender.
- I don't understand the value of introducing *1sing*, *2sing*, and *3sing* from 4.6.1. Are they just abbreviations for the full AGR feature structure?

# Reading Questions

- What exactly is the issue with agreement described in (31)? Why would the grammar up till now exclude these sentences?
  - *With [them on our team], we'll win*
- I'm unclear about the difference between Head Specifier Rule and Specifier-Head Agreement Constraint. What specifically does each of them govern?

# Reading Questions

- Chapter 4 introduced new Value (=valence?) features. Is this mean that whether or not COMPS and SPR are needed, we still need to display both of them in every AVM, it seems very redundant and does not align with other features' behavior, such as AGR?
- When a lexical entry's SPR or COMPS values are empty lists, that means that the word is prohibited from taking a specifier/complement, right? I may be getting confused, but throughout the notation, sometimes it seems that an empty COMPS list means that a complement requirement is irrelevant/unknown, whereas other times it means that a complement is actually prohibited.

# Reading Questions

- The breakup of the AGR constraint into 3SING and NON-3SING, which in turn is further divided into 1SING and NON-1SING, while very effective in terms of paring down the size of trees, feels very specific to our English use case. To what extent are the generalizations made in this chapter applicable across languages and even across dialects of English?
- Questions I had while doing the exercises: for languages other than English, do we usually by default account for all features that appear in English? In particular, do we assume that the PER feature is universal across languages, or some languages just might not have the notion of PER?

# Reading Questions

- Throughout this chapter, there are distinctions drawn between phenomena that should be handled by grammar rules, the lexicon, or semantics. In particular, the creation of the COUNT feature was justified in 4.6.3 because the semantic analysis of mass nouns does not cover all members of the class. What are the guidelines for drawing these distinctions? In a sense, I'm wondering at what point we decide that 'exceptions' are best handled within the grammar vs. elsewhere.
- Where should we draw the line between "intuitiveness" of the grammar - something like making it more readable and in-line with our prior categorization and notions of the language - and keeping our notation as consistent and compact as possible when developing our grammar?



# Reading Questions

- Specifier-Head Agreement Constraint (SHAC) enforces the AGR value of a verb to be identical and passed to that of the NP specifier, which the verb selects. Subsequently, SHAC enforces the AGR value of the Head daughter of the NP (i.e. noun) to be identical and passed to that of the determiner, which is selected by the noun. It seems to me that the hierarchy among words and phrases within SHAC is: verb <- NP <- noun <- det. I am wondering whether it is also the order (i.e. verb<- NP <- noun <- det.) in the process of sentence generation of our model, which are built upon the four grammar rules and the three principles (i.e. HFP, Valence principle, and SHAC)? Take (52) for example, is the sentence "The dog walks" formed as: [ ] [ ] [walks] -> [ ] [dog] [walks] -> [The] [dog] [walks] ?

# Reading Questions

- In this chapter, the generalized version of the Head-Specificier Rule as well as the Head First Principle reminds me of dependency grammar (for which arrows commonly point to verbs - which can be seen as heads). What is the link between this and dependency grammar?