HG4041 Theories of Grammar

Complex Feature Values
Valence, Agreement, Case

Francis Bond
Division of Linguistics and Multilingual Studies
http://www3.ntu.edu.sg/home/fcbond/
bond@ieee.org

Lecture 3
Location: LHN-TR+36

Overview

- Last week
- A problem with the Chapter 3 grammar
- Generalize COMPS and SPR
- The Valence Principle
- Agreement
- The SHAC
- Work through problems 4.1, 4.5, 4.6

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 crosscategorial generalizations, while still making fine distinctions.

Heads

- Intuitive idea: A phrase typically contains a word that determines its most essential properties, including
  - where it occurs in larger phrases
  - what its internal structure is
- This is called the **head**
- The term **head** is used both for the head word in a phrase and for all the intermediate phrases containing that word
- NB: Not all phrases have heads

Formalizing the Notion of Head

- Expressions have a feature HEAD
- HEAD's values are of type *pos*
- For HEAD values of type *agr-cat*, HEAD's value also includes the feature AGR
- Well-formed trees are subject to the **Head Feature Principle**:

  In any headed phrase, the HEAD value of the mother and the head daughter must be identical.

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~~ compact and leave out the values irrelevant to our current point).

But it's still not quite right ...

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

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

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

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

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:

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} \quad \langle\rangle \end{bmatrix} \end{bmatrix} \rightarrow \boldsymbol{H} \begin{bmatrix} \textit{word} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} \quad \left\langle \boxed{1}, \ \ldots, \ \boxed{n} \right\rangle \end{bmatrix} \end{bmatrix} \quad \boxed{1}, \ \ldots, \ \boxed{n}$$

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

The valence comes from the lexicon

$$\left\langle \text{devour}, \begin{bmatrix} \textit{word} \\ \textsc{head} & \textit{verb} \\ \textsc{val} & \begin{bmatrix} \textsc{comps} & \langle \textsc{np} \rangle \end{bmatrix} \end{bmatrix} \right\rangle \qquad \left\langle \text{put}, \begin{bmatrix} \textit{word} \\ \textsc{head} & \textit{verb} \\ \textsc{val} & \begin{bmatrix} \textsc{comps} & \langle \textsc{np pp} \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

$$\left\langle \text{eat}, \begin{bmatrix} \textit{word} \\ \textsc{head} & \textit{verb} \\ \textsc{val} & \begin{bmatrix} \textsc{comps} & \langle (\textsc{np}) \rangle \end{bmatrix} \end{bmatrix} \right\rangle \qquad \left\langle \text{bet}, \begin{bmatrix} \textit{word} \\ \textsc{head} & \textit{verb} \\ \textsc{val} & \begin{bmatrix} \langle \textsc{np} \ (\textsc{np}) \ (\textsc{s}) \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

$$\left\langle \text{dine}, \begin{bmatrix} \textit{word} \\ \textsc{head} & \textit{verb} \\ \textsc{val} & \begin{bmatrix} \textsc{comps} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \qquad \left\langle \text{fond}, \begin{bmatrix} \textit{word} \\ \textsc{head} & \textit{adjective} \\ \textsc{val} & \begin{bmatrix} \textsc{comps} & \langle \textsc{pp}\textit{:of} \rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

Question: What if English had postpositions?

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} \quad \langle\rangle \end{bmatrix} \end{bmatrix} \rightarrow \boldsymbol{H} \begin{bmatrix} \textit{word} \\ \text{HEAD} \quad \textit{verb} \mid \textit{adj} \mid \textit{noun} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} \quad \left\langle \boxed{1}, \ \ldots, \ \boxed{n} \right\rangle \end{bmatrix} \end{bmatrix} \boxed{1}, \ \ldots, \ \boxed{n}$$

$$\begin{bmatrix} \textit{phrase} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} \quad \langle\rangle \end{bmatrix} \end{bmatrix} \rightarrow \boxed{1}, \ \ldots, \ \boxed{n} \ \boldsymbol{H} \begin{bmatrix} \textit{word} \\ \text{HEAD} \quad \textit{preposition} \\ \text{VAL} \quad \begin{bmatrix} \text{COMPS} \quad \left\langle \boxed{1}, \ \ldots, \ \boxed{n} \right\rangle \end{bmatrix} \end{bmatrix}$$

Specifiers

In English, nouns can agree with their specifiers.

- In Number:

  (1) *This dog barked.*
  (2) **This dogs barked.*
  (3) **These dog barked.*
  (4) *These dogs barked.*

- In Countability

  (5) *Much furniture was broken.*
  (6) * *A furniture was broken.*
  (7) * *Much chair was broken.*
  (8) *A chair was broken.*

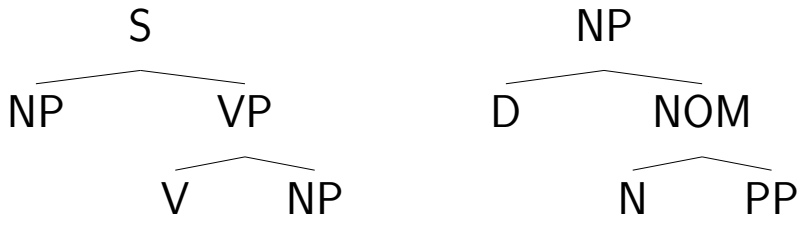
Head-Specifier Rule (Version I)

$$\begin{bmatrix} \textit{phrase} \\ \textit{VAL} \quad \begin{bmatrix} \textit{comps} & \langle\rangle \\ \textit{SPR} & \langle\rangle \end{bmatrix} \end{bmatrix} \rightarrow \boxed{2} \quad \boldsymbol{H} \begin{bmatrix} \textit{VAL} \quad \begin{bmatrix} \textit{comps} & \langle\rangle \\ \textit{SPR} & \langle \boxed{2} \rangle \end{bmatrix} \end{bmatrix}$$

- 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?

- **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 keep the information on the COMPS and SPR lists until it gets "canceled" by the Head-Complement and Head-Specifier rules.

The Parallelism between S and NP

- Motivation:
  - pairs like
    - (9) *Chris lectured about syntax*
    - *(10)* *Chris's lecture about syntax*
  - both S and NP exhibit agreement
    - (11) *The bird sings/*sing*
    - *(12)* *The birds sing/*sings*
    - *(13)* *this/*these bird*
    - *(14)* *these/*this birds*
- So we treat NP as the saturated category of type *noun* and S as the saturated category of type *verb*.

Any other reason to treat V as the head of S?

- In standard English, sentences must have verbs.
  (How about non-standard English or other languages?)

- Verbs taking S complements can influence the form of the verb in the complement:

  (15) *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

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

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.

$$S = \begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

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* $\not\Rightarrow$ *X did something PP*, then the PP is a complement.

  (16) *Pat relied on Chris* $\not\Rightarrow$ *Pat did something on Chris*
  (17) *Pat put nuts in a cup* $\not\Rightarrow$ *Pat did something in a cup*
  (18) *Pat slept until noon* $\Rightarrow$ *Pat did something until noon*
  (19) *Pat ate lunch at Bytes* $\Rightarrow$ *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.

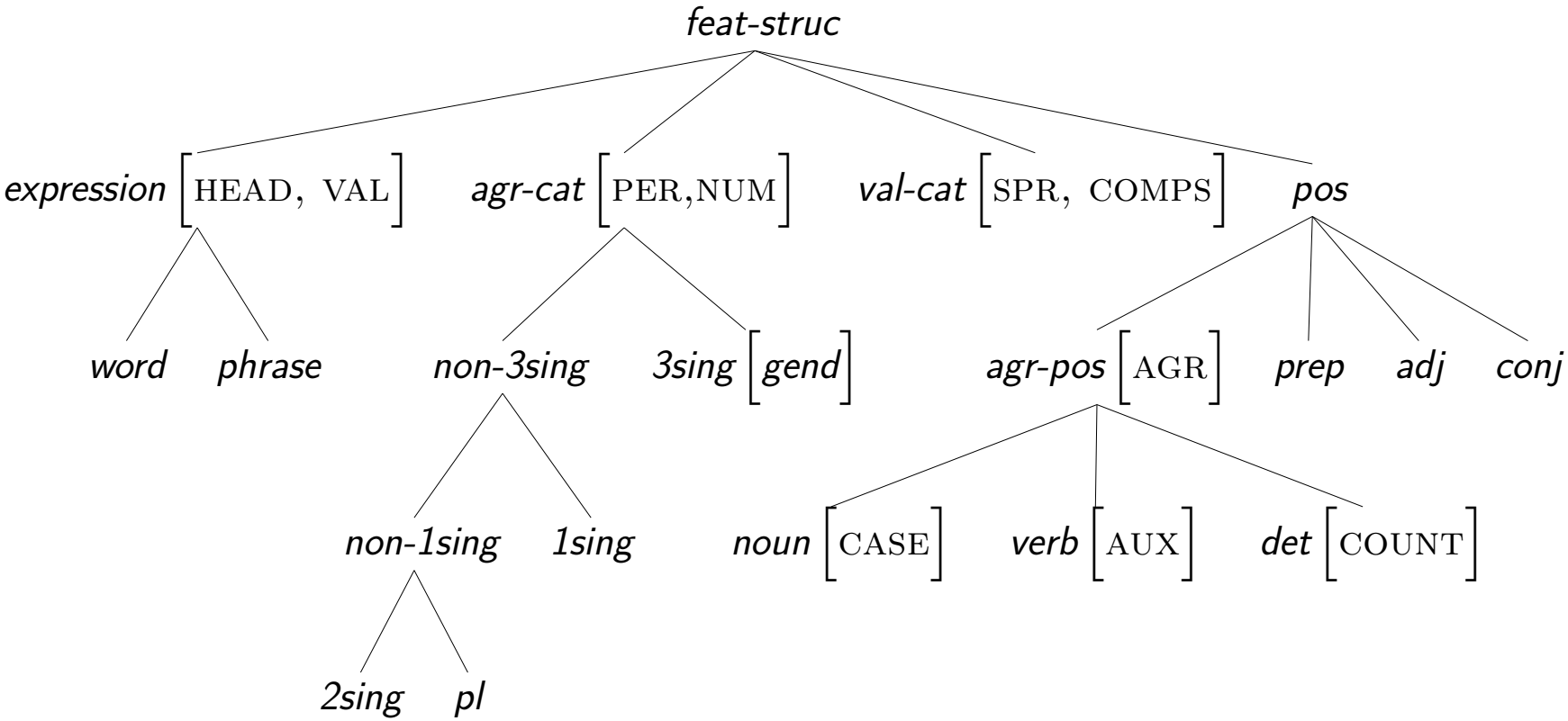
The Specifier-Head Agreement Constraint (SHAC)

Verbs and nouns must be specified as:

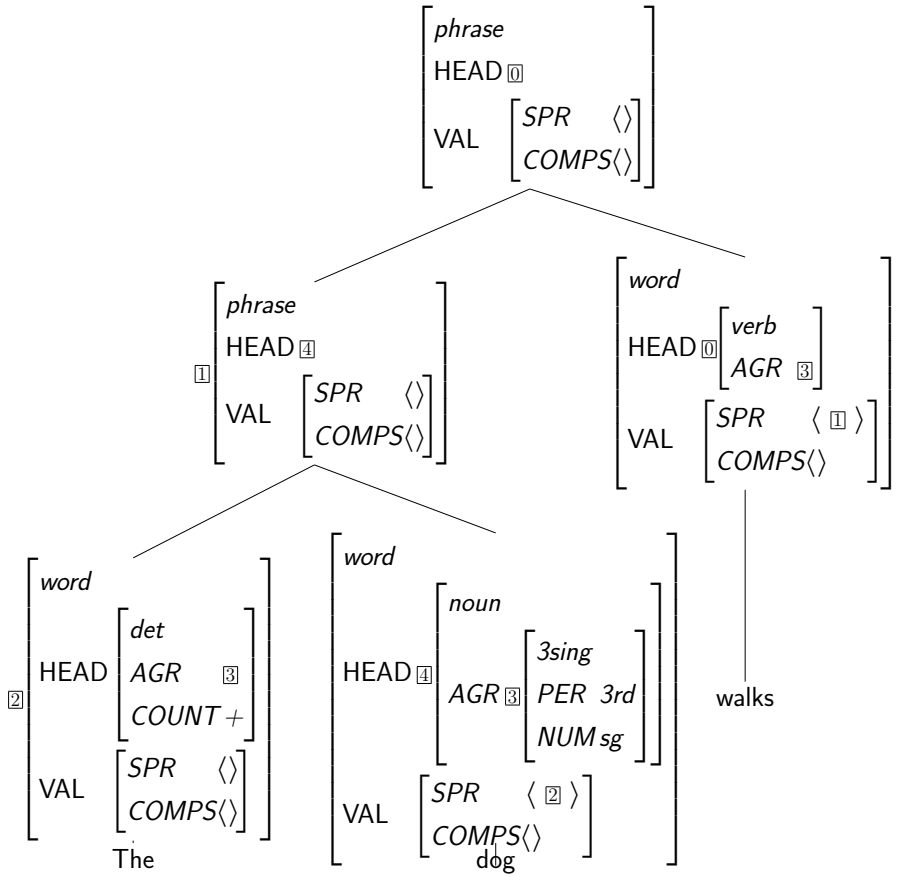
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{AGR } \boxed{1} \end{bmatrix} \\ \text{VAL} & \begin{bmatrix} \text{SPR } \left\langle \begin{bmatrix} \text{AGR } \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

➢ Why is this lexical?

Type Hierarchy (new Agreement!)



An example


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 ⟨⟩
The
word
HEAD 4 noun
AGR 3 3sing
PER 3rd
NUM sg
VAL SPR ⟨ 2 ⟩
COMPS ⟨⟩
dog
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 individual 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, ...*** 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, ...***)
  - lexically underspecified (***the, all, some, no, ...***)
- 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

- A problem with the Chapter 3 grammar
- Generalize COMPS and SPR
- The Valence Principle
- Agreement
- The SHAC
- (Work through problems 4.1, 4.5, 4.6)

4.1 Valence Variations

Write lexical entries (including HEAD, SPR, and COMPS values). You may use NP, VP, etc. as abbreviations for the feature structures on COMPS lists.

As you do this problem, keep the following points in mind: (1) In chapter 4 COMPS became a list-valued feature, and (2) heads select for their specifier and complements (if they have any); the elements on the SPR and COMPS lists do not simultaneously select for the head.

[*Hint: For the purposes of this problem, assume that adjectives and prepositions all have empty SPR lists.*]

A. Write lexical entries for the words *here* and *there* as they are used in (i).

(i) *Kim put the book here/there.*

[*Hint: Compare (i) to (7) on p97.*]

C. Assume that motion verbs like *jump*, *move*, etc. take an optional PP complement:

$$\left[\textit{COMPS} \quad \langle\ \textit{(PP)}\ \rangle \right]$$

Write the lexical entries for the prepositions *out*, *from* and *of*:

(i) Kim jumped out of the bushes.
(ii) Bo jumped out from the bushes.
(iii) Lee moved from under the bushes.
(iv) Leslie jumped out from under the bushes.
(v) Dana jumped from the bushes.
(vi) Chris ran out the door.
(vii) *Kim jumped out of from the bushes.
(viii) Kim jumped out.
(ix) *Kim jumped from.

D. Based on the following data, write the lexical entries for the words *grew* (in the 'become' sense, not the 'cultivate' sense), *seemed*, *happy*, and *close*.

(i) They seemed happy (to me).
(ii) Lee seemed an excellent choice (to me).
(iii) *They seemed (to me).
(iv) They grew happy.
(v) *They grew a monster (to me).
(vi) *They grew happy to me.
(vii) They grew close to me.
(viii) They seemed close to me to Sandy.

[*Note: APs have an internal structure analogous to that of VPs. Though no adjectives select NP complements (in English), there are some adjectives that select PP complements (e.g. to me), and some that do not.*]

E. Using the lexical entries you wrote for part (D), draw a tree (showing the values of HEAD, SPR, and COMPS at each node, using tags as appropriate) for *They seemed close to me to Sandy*.

4.5 Facts of English case

- For each of the following positions, determine which case the pronouns in that position must have:
  - Subject of a sentence
  - Direct object of a verb
  - Second object of a verb like ***give***
  - Object of a preposition
- Give examples

4.6 A lexicalist analysis

- Section 4.8 hinted that case marking can be limited in the same way that we handle agreement, i.e., without any changes to the grammar rules. Show how this can be done. Your answer should include lexical entries for they, us, likes, and with.

- Hint: Assume that there is a feature CASE with the values 'acc' and 'nom', and assume that English pronouns have CASE features specified in their lexical entries.

Acknowledgments and References

- Course design and slides borrow heavily from Emily Bender's course: *Linguistics 566: Introduction to Syntax for Computational Linguistics*
  `http://courses.washington.edu/ling566`