

Assumptions (Week 4)

1 The Lexicon

Lexical entries are complex feature structures containing (at least) syntactic, semantic and phonological information about each morpheme in the language. The syntactic features are organized as in (1).

$$(1) \quad \text{LI} = \left[\begin{array}{l} \text{CAT} \\ \text{INFL} \\ \text{SEL} \end{array} \left[\begin{array}{l} \text{TYPE} \quad \{N, V, \text{WH}, \text{etc.}\} \\ \text{INFL} \quad \left[\begin{array}{l} x\text{FORM} \quad \textit{val} \end{array} \right] \\ \text{INFL} \quad \left[\begin{array}{l} y\text{FORM} \quad \textit{val} \\ z\text{FORM} \quad _ \end{array} \right] \\ \text{SEL} \quad \left[\begin{array}{l} \text{COMP} \quad \text{CAT}_1 \\ \text{SPEC} \quad \text{CAT}_2 \end{array} \right] \end{array} \right]$$

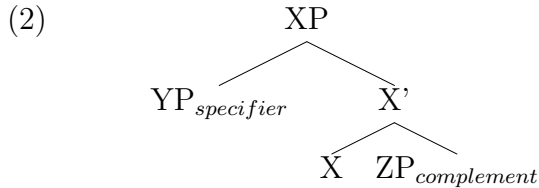
The CAT feature complex encodes information about the basic syntactic properties of the object: its basic ‘type’ (i.e., part of speech), and the kind of agreement it imposes on other objects. The latter is encoded as an INFL subfeature; the representation in (1) indicates that LI licenses valuation of another object’s *xFORM* feature as *val*. (Here *xFORM* is just a placeholder for some other feature — *vFORM*, *CASE*, whatever.) In writing these out, we omit TYPE and just write in the part of speech information, which may be a set (e.g., {N, WH}).

The INFL feature complex controls the (syntactically determined) surface shape of the lexical item. This information may be explicitly specified in a lexical entry, as with the *yFORM* value in (1), which has the value *val*, or it may be ‘unvalued’, indicated with an underline, as with the *zFORM* feature in (1). Unvalued features may become valued through the AGREE relation, defined below. We have left open the question of whether derivations that include unvalued INFL features are ungrammatical (because there are insufficient instructions about the shape that the lexical item should take), or whether they can be filled by some sort of default rule.

The SEL feature complex specifies the selectional requirements of a lexical item. It is further subdivided into a COMPLEMENT feature and a SPECIFIER feature, each of which have categories (or possibly ordered lists of categories, though I have not shown this in (1)) as their values. The difference between specifiers and complements concerns order and position within the phrasal projection of a head. Specifically:

1. Heads first combine with complements, then with specifiers. That is, all COMP features must be discharged first.
2. Languages may specify the relative ordering of head, complement, and specifier.

In English, complements are ordered to the left of the head, and specifiers are ordered to the right of the head (and any complements), giving us structures like (2).



Note that nothing rules out the possibility that different heads impose different ordering requirements on the things they select — e.g., that a head X wants its complements on the right, and a head Y wants them on the left. If we discover that this is the case, we will need to figure out some way of encoding this in the lexical entries of the relevant heads.

2 The generative component

Our generative component contains three (or maybe two) rules: AGREE, MERGE and MOVE. MOVE is defined in terms of MERGE, so is arguably not a separate principle, but to keep things as clear as possible I will define it separately here. The details of the rules as they currently stand are as stated in (3)-(5). (Note that the definition of MERGE is worded a little bit differently than the one we settled on in class, but this is only because I spelled out the definition of ‘select’; the overall requirements of the rule are identical.)

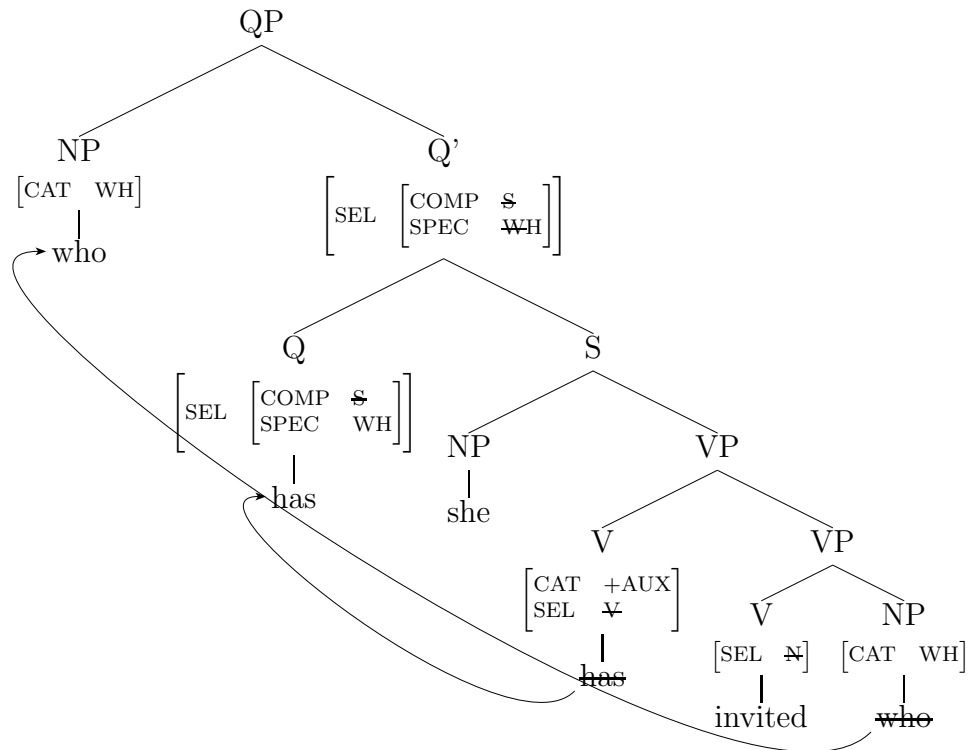
- (3) AGREE
 If syntactic object X has an unvalued INFL feature F_1 and syntactic object Y has a matching valued CAT feature F_2 , let the value of $F_1 =$ the value of F_2 .
- (4) MERGE
 If X is a syntactic object with undischarged SEL feature α and Y is a syntactic object with CAT feature α :
- i. Form Z such that Z immediately dominates X and Y.
 - ii. Discharge (delete) α on X.
 - iii. Let the features of Z = the features of X.
- (5) MOVE
 If X is a syntactic object with undischarged SEL feature α and Y is a syntactic object with CAT feature α and X contains Y:
- i. Form Z such that Z immediately dominates X and a copy of Y.
 - ii. Discharge (delete) α on X.
 - iii. Let the features of Z = the features of X.
 - iv. Delete the original occurrence of Y.

The two significant differences between MERGE and MOVE is that the latter states explicitly that the selected object Y is part of the structure that is already built (this is the ‘contained within X’ clause), that what gets merged to X is a copy of Y, and that the original gets deleted. It may very well be that none of these things need to be stipulated at the end of the day — that MOVE really can be viewed as a special case of MERGE where the selected object comes from structure that is

already built ('internal merge') rather than from the lexicon ('external merge'), and that the deletion requirement follows from general principles governing the mapping from syntax to phonology. For the moment, though, we can treat them as two separate (though related) principles, just to keep things as clear and explicit as possible.

Finally, MOVE as currently stated is triggered by selection. While this looks like a good way of handling phrasal movements like *wh*-movement, it is not clear that it will work for head movements, like auxiliary fronting in yes-no (and *wh*-) questions. For now you can simply assume that there is some principle that licenses movements from one head position to another, so what is going on in an example like (6a) are two instances of movement: phrasal movement of *who* to the specifier position of Q, and head movement of *has* to Q, as shown in (6b).

- (6) a. Who has she invited.
 b.



Note that this tree leaves out INFL information, and CAT information is conveyed by node labels plus features if necessary. For example, *invited* is CAT V, as indicated by its node label, but *who* has two category features: N and WH. The former is represented by the node label; the latter is listed in the feature structure.

3 Sentences and subjects

Sentences are projections of the category T(ense). English contains at least two members of the T category (and probably more): PAST and PRESENT. T selects for a complement of category V and a specifier of category N; the latter corresponds to the subject of the sentence. T also is the locus of subject agreement: it has unvalued ϕ -features (person, gender, number) which are valued under agreement with

its specifier (the subject). T heads in general select for a complement of category V and a specifier of category N, as shown in (7).

$$(7) \quad \left[\begin{array}{l} \text{CAT} \quad \text{T} \\ \text{INFL} \quad \left[\begin{array}{l} \phi \quad - \end{array} \right] \\ \text{SEL} \quad \left[\begin{array}{l} \text{COMP} \quad \text{V} \\ \text{SPEC} \quad \text{N} \end{array} \right] \end{array} \right]$$

We have not settled on a particular view about how tense information ends up on the highest verb. One coherent option is that this is done under agreement: that e.g. PAST has the feature structure in (8), which allows it to trigger agreement on a lower verb (which we must assume has an unvalued TENSE feature in its INFL array, or that TENSE is really a special type of VFORM feature).

$$(8) \quad \left[\begin{array}{l} \text{CAT} \quad \left[\begin{array}{l} \text{T} \\ \text{TENSE} \quad \textit{past} \end{array} \right] \\ \text{INFL} \quad \left[\begin{array}{l} \phi \quad - \end{array} \right] \\ \text{SEL} \quad \left[\begin{array}{l} \text{COMP} \quad \text{V} \\ \text{SPEC} \quad \text{N} \end{array} \right] \end{array} \right]$$

There are other options however. For example, it could be the case that tense morphology is directly represented in T (rather than as an inflectional property of verbs), but that it gets pronounced on an adjacent verb. We do not have to decide on these options right now.

4 Semantic arguments

We haven't actually made any assumptions regarding semantic arguments yet, so this is something that you may want to make some proposals about yourself. We did, however, hypothesize that the reason that sentences like those in the following examples are bad is because the subject isn't semantically 'linked' to the verb/VP — it is already 'semantically full', because the other constituents in the sentence are being used to express its semantic arguments: the participants in the kind of event/situation that a verb describes.

- (9) a. *The situation seems that Sam kept that promise.
 b. *The hamster appears that they have something in the box.
- (10) a. *The future is clear that your daughters book will be a success.
 b. *The result was obvious that more banks would fail.
- (11) a. *Max disturbs me that he left the room.
 b. *The event mattered to Sally that Sue was late.

We need to make some explicit assumptions about what is going on here; for now, you should assume that the basic hypothesis about why they are bad is on the right track.