Features

CAS LX 422 ∼ GRS LX 722 Intermediate Syntax

Lecture 2

Morphosyntactic features

Syntax is in large part the study of principles of sentence formation. But what it cares about isn't exactly words.

(1) Three dogs are here. One dog is here.

(2) Three geese are here. One goose is here.

(3) Three deer are here. One deer is here.

Words have properties (being a verb, or being plural). More abstract than surface morphology.

Agreement

Subject and verb need to agree in number and (for *be*) person.

If the subject is plural, the verb must take a "plural" form. Crosslinguistically common.

(4) The dog wants food. The dogs want food.

(5) The dog is hungry. The dogs are hungry.

(6) I am hungry. We are hungry.

Intuitively, the plural feature is part of the meaning of the subject. It "belongs there." It is *interpretable* there. The (agreeing) plural verb morphology is more a "reflection." It is *uninterpretable* on the verb.

What are the features?

Some features seem to matter, some don't. The identity of the features is part of the theory. What is the minimal set that suffices to explain the grammar?

Some things have the [invented in early September] property. We have no reason to think syntax reacts to this property. We *do* have reason to think syntax reacts to [plural].

So, maybe things can be either singular or plural. If they're singular they have [sg] and if they're plural they have [pl].

Science

Hypothesis

[sg] and [pl] are features a word can have.

Prediction

Four classes of words: [sg], [pl], [sg,pl]. []

So are the predictions borne out?

Not exactly

There seem only to be two number-based classes of nouns in English, not four. At least there's no evidence supporting any more than just singular and plural. Thus: better to represent singular as lack of plural, or vice-versa.

Hypothesis

[pl] is a feature a word can have.

Prediction

Two classes of words: [] and [pl]

That matches what we see in English. It is possible that languages with a "dual" might require specifying both.

Category features

Syntax is concerned with *distribution*. Words seem to come in distributional classes. We call them noun, verb, adjective, preposition, etc.

- They have no noun
- They can verb
- They are adjective
- Very adverb, very adjective (where gradable)
- Right proposition (*right over the house*)
- Etc.

Category generalizations

The distributional diagnostics above can cut things up in terms of distributional classes, but how about this? Things that the prefix *un* can attach to.

(7) untie, unfold, unwrap, unpack

(8) unhappy, unfriendly, undead

(9) * uncity, uncola, unconvention

(10) * unupon, unalongside, unat

You can attach *un* to verbs and adjectives. To describe that class they must have something in common. Let's say it is the feature [X] that verbs and adjectives share.

Category generalizations

Here is an example of case marking in Russian. Relevant point here is that both adjectives and nouns get marked for case.

(11) krasivaya dyevushka vsunula chornuyu koshku v pustuyu
beautiful girl put black cat in empty
korobku
box
'The beautiful girl put the black cat in the empty box.'

To describe that class of things that get case-marked, they must have something in common. Let's say it is the feature [Y] that nouns and adjectives share.

Category features

So: verbs have [X], nouns have [Y], adjectives have [X, Y]. Well, ok, what if X=V and Y=N?

- Nouns have an [N] feature
- Verbs have a [V] feature
- Adjectives have both: [N, V]
- Prepositions have neither: []

It's not immediately obvious *why* it is P that winds up having no features and A that winds up having both, but the evidence pointed that way.

This would seem to predict that P wouldn't be expected to form a natural class with N, V, or A individually. Something to look for, that has now taken on significance.

Lexical vs functional

- N, V, A, P: These are *lexical* categories. Significant, arbitrary meaning, and generally *open-class* (new ones can be invented).
- Not all words are of this kind, we also have the more grammatical categories. The *functional* categories. Determiners, complementizers, pronouns, modals, auxiliaries.

The functional categories are a bit more like the syntactic glue that holds a sentence together. They contribute to the meaning, but in a narrow way.

Functional categories

- D: determiners (*the*, pronouns)
- C: complementizers (*that*, *for*)
- auxiliaries: *have*, *be*, *must*
- T: tense

If you squint, D is kind of "nouny"—it goes with nouns. Maybe these are all "functional versions" of the lexical categories? That is, the category D is represented by the features [N, Fn]?

Will explore this briefly here next, but otherwise we will just treat category features as atomic, along with a constraint that a syntactic object must have exactly one.

Category features?

- P: []
- N: [N]
- V: [V]
- A: [N, V]
- C: [Fn]
- D: [Fn, N]
- T: [Fn, V]
- Aux: [Fn, N, V]

Maybe. Could be. Would be an area to look for natural classes.

- Maybe [Lx] instead of [Fn]?
- Maybe [+N] vs. [–N] etc., instead of [N] vs. []. This makes weaker predictions though.
- There will be more functional categories, may need yet another dimension (feature), leading to 16. The *v* we meet later might be a functional V, yet it differs from T. Will need something for Neg (Pol), agreement? Interesting to ponder.

Agreement

English speakers agree about this:

(12) Three dogs are here. One dog is here.

(13) * Three dogs is here. One dog are here.

What's wrong with (13)? The subject and verb need to *agree* but don't.

There are two kinds of subjects (plural and nonplural) and there are two kinds of verbs. One kind of verb goes with plural subjects, one kind of verb goes with nonplural subjects. The main thing that differentiates the two kinds of verbs is the kind of subject they go with.

Agreement, asymmetry

We can call the verbs that go with plural subjects "plural." Why not? We need to call them something. We could even indicate them with the feature [pl]. The verb is not really "plural" in any meaningful way, but it is *reflecting* the pluralness that originated with the subject.

Said that way, "agreement" is really identity. The subject and verb must have the same number feature. That is, the verb needs to take on the same number feature as the subject had.

Technical note: This is easy if number is represented as [±pl]. But there is nuance if it is represented as [pl] vs []. In the latter case, we can't just compare subject and verb to see if they match (since a nonplural verb would match with any kind of subject). Privative features likely require assuming the verb is nonplural by default, and changed to plural only if the subject has [pl].

Person agreement

English pronouns make several distinction over and above the singular/plural distinction. One distinction in in *person*, which is sensitive to who is talking and to whom. English distinguishes three persons.

	singular	plural
first person	I	we
second person	you	you
third person	he/she/it	they

We could model person with [1], [2], and [3]—except that predicts eight distinctions and we only have evidence for three. Better to use two features, then we predict four—which is closer at least.

Person agreement

By eliminating [3], we can predict a system that has four persons. The ones below, plus the [1,2] person. Not morphologically distinguished in English, but still semantically kind of sensible if we consider [1] to represent "speaker" and [2] to represent "addressee." Third person is neither, [1,pl] is "we" (excluding you) and [1,2,pl] is "we" (including you). Some languages do distinguish inclusive *we* from exclusive *we*.

	singular	plural
first person [1]	I	we
second person [2]	you	you
third person	he/she/it	they

Don't know about [1,2] (nonplural). Maybe it's exactly the two of us, acting with singleness of mind.

ϕ features

Collectively, person, number, and gender features are referred to as ϕ-features.

These are the features that are generally involved in subject-verb agreement, across languages. We group them together because they seem to (usually) have their effects together (not separately).

Case features

English pronouns change form depending on where they are.
Information about syntactic position is encoded by case features.

(14) They left. I saw them. They saw me.

- In English, case is only visible on pronouns
- In other languages, case is visible on all nouns (and sometimes on noun-associates, like adjectives or determiners).
- Lexical items are bundles of features, e.g. [Acc, 1, sg, Prn]
- The syntactic system arranges these into sentences, with some pronunciation and some meaning.
- The pronunciation of [Acc, 1, sg, Prn] is “me.”

Features and pronunciation

singular			plural		
nom	acc	gen	nom	acc	gen
I	me	my	we	us	our
you	you	your	you	you	your
he	him	his	they	them	their
she	her	her	they	them	their
it	it	its	they	them	their

Not every distinction.

- Only 3rd distinguishes gender
- 2nd does not distinguish number or between Nom and Acc
- 3rd singular feminine does not distinguish between Acc and Gen

Verbal features

Some features are specific to verbs. Like [past], which is a tense feature that separates *wrote* and *kicked* from *write* and *kick*. English does not seem to need to posit a [future] feature.

(Again, this is primarily about distribution. It does not appear that we need to use a [future] feature to explain the facts we see in English. Though for some languages we would. In English, future is expressed via modal *will* or with the verb *go*.)

Realizing verbal features

There is a kind of structure to the verbal paradigm, not every distinction results in a different form. The verb *write* only appears in three tensed forms: *wrote*, *writes*, *write*. The verb *be* gets slightly more complex inflection. We will assume in general that this is not syntax's problem; syntax provides the features, morphology is responsible for pronouncing them. Possibly including combining syntactic elements, such that the structure of *smarter* and *more intelligent* wind up being the same syntactically.

Rules that seem to work for pronuncation are:

	if [1] or [2] treat as [pl]	if [2] treat as [pl]
[past, pl]	↓	were
[past]	wrote	was
[pl]	write	are
[1]	↓	is
else	writes	am

Participles

Another form that verbs can take are their participle forms, like *writing* and *written*. These forms in English are not expressing tense, but rather aspect.

- The *-ing* form (“present participle”) comes after *be*, indicating a continuing event.
- The *-en* form (“past participle”) comes after *have*, indicating a completed event.

Do not be fooled by the tense-like traditional names, they are independent of tense. *I have written, I had written, I am writing, I was writing*. I will generally refer to them as “progressive” [prog] and “perfective” [perf]. They do not generally agree with the subject (and are kind of structurally too far away to do so.)

Bare verbs/infinitive

The last main form verbs can take is the bare/infinitive form. Basically the citation form. These often come after *to*, and we will indicate this with [inf]. E.g.: *to write*.

Might be somewhat of a coincidence that it is the bare verb. Other languages do mark the infinitive with its own inflection, much like the inflection you see on participles.

Infinitives and participles (across languages) often do not show any subject agreement (some show object agreement).

Feature summary

Features we have so far:

- Categories:
  - Atomic: [N], [V], [Adv], [Adj], [P], [D], [T], [Aux], [C], [Neg]?
  - Or maybe: [N], [V], [Fn], [Lx]?
- Nominal features:
  - case: [nom], [acc], [gen]
  - ϕ-features:
    - person:[1], [2]
    - number: [pl]
    - gender: [fem]
- Verbal features:
  - tense: [inf], [past]
  - aspect: [perf], [prog]