MSc Introduction to Syntax

Lecture 1: What syntax is about

1. *What is this course about?*
This course is about syntactic theory. Syntactic theory is about the rules and principles that determine how people speaking a particular language know how to distinguish the possible word orders in that language from the impossible ones. But what exactly do people know when they 'know the syntax of English'?

A *prescriptivist* grammarian will answer that such a person knows how to properly apply certain rules that are considered to reflect 'good grammar', like "don't split your infinitives" (so don't write *how to properly apply*), "don't end a sentence with a preposition" or "use *whom* rather than *who* as the object of a verb or preposition".

(1) a. #...to boldly go where no man has gone before
a'. to go boldly where no man has gone before

(2) a. #This is the kind of bank which you cannot rely on.
a'. This is the kind of bank on which you cannot rely.

(3) a. #Who did you see?
a'. Whom did you see?

Such rules are mostly learned consciously, and require explicit instruction via grammar books and/or school teachers. These rules can be arbitrary, in the sense that they do not reflect what speakers of English actually can or cannot do syntactically. Rather, usually they reflect the fact that, in cases where speakers of English can use two different forms to express the same thing, there is a strong tendency to think one of these must be 'better' than the other (of course, the same holds for other languages, for which all sorts of different prescriptive rules have been proposed as well). Consequently, grammarians, writers and/or language policy makers come up with various arguments to proclaim one of these forms the correct one. (Sometimes such rules reflect the fact that one of the forms was really the only possible one in older stages of the language, such as in the case of the distinction in (3)). Whatever the merits or otherwise of rules that, by fiat, weed out alternative forms in this way, the fact that they are the result of conscious, often rather arbitrary, decisions on what people *should* do, rather than reflecting the actual state of things (what people *are* doing), means that they are not a very interesting subject for scientific inquiry. Rules like these are not dissimilar to rules for table etiquette or traffic rules, and while noone would want to dispute the usefulness of traffic rules at least, there is not much point in trying to make an in-depth study of why the British drive on the left of the road, while people on the Continent drive on the right – this was just the result of people in charge of such things at the time arbitrarily deciding on one or the other.

A *descriptivist* grammarian will approach the above question ('what do people know when they know the syntax of English') by pointing out that people who have learned English as a first language can tell for any string of words, such as those in (4)-(10), whether it is a possible English sentence or not, without having had explicit

instruction for this or being conscious of the rule system that he or she uses to decide on this. (In (4)-(10), an asterisk in front of a sentence indicates that the sentence is not a well-formed one. This is standard notation in linguistics, which we will follow throughout this course).

(4) a. Mary read a book.
b. * Mary read which book? (except as an 'echo question')
c.Which book did Mary read?
d. Mary read a book and the newspaper.
e. *Which book did Mary read and the newspaper?

(5) a. Who did Mary say that she saw yesterday?
b. *Who did Mary regret that she saw yesterday?
c. *Who does Mary wonder if she will see today?

(6) a. John does not like those novels.
b. Those novels, John does not like.
c. John does not like those novels by Jones.
c. *Those novels, John does not like by Jones.

(7) a. Jane likes reading novels.
a'. What does Jane like reading?
b. Reading novels upsets Jane.
b. *What does reading upset Jane?

(8) a. John is eager to please.
a'. What is John eager to do?
b. John is easy to please.
b'. *What is John easy to do?

(9) John wants to shave himself/him. (*himself* = John, *him* ≠ John)
John wants Bill to shave himself/him. (*himself* ≠ *John, him* = John)

(10) etc. etc.

Note that it is unlikely that a child learns how to distinguish all the possible from all the impossible sentences just by explicit instruction, because:
(i) how many children will ever be explicitly instructed about data such as those in (4)-(9)? Yet English speakers will agree on their relative grammaticality. (It is even questionable whether all children will encounter sentences like these at all before being able to determine that they are possible or not).
(ii) all children show a similar developmental path in first-language acquisition, irrespective of the explicit instruction they might get (in fact, there is some evidence that suggests they can ignore explicit instruction).

Note also that it is not possible to give a child a list of all possible sentences in a language, since the number of possible sentences is literally infinite (see below). Apparently, every healthy child spontaneously learns a set of rules/principles with which it can distinguish the possible sentences in the language it hears around him/her from the impossible ones, and it can apply these principles continuously without

having to make a conscious effort. This is a faculty with which all humans are endowed. It seems that there is a *critical period* for this faculty to be active – after a certain age, this spontaneous language learning ability disappears, so that learning a language after this age is not always equally successful and does require a conscious effort. The question of *how* a child acquires these rules/principles (and whether this learning mechanism is specific to language or not) is a hotly debated one, but we will not go into that in this course. It is some of the properties of the subconsciously learned syntactic rule system that we are going to investigate. Note that for most if not all varieties of present-day English, this rule system will deem the sentences in (1a), (2a) and (3a) possible sentences of English, just like the ones in (1a'), (2a') and (3a'). Henceforth, when we talk about 'grammatical' or 'ungrammatical' sentences, we mean this in the descriptivist sense, not the prescriptivist one.

(NB: Although we will mainly discuss English examples, the human language faculty is of course such that children can acquire any other language with equal ease when this is the language in their environment. The syntax of other languages can differ in all sorts of interesting ways from that of English, but some of the fundamentals, such as the fact that sentences seem to be made of constituents rather than being just strings of words, may hold more universally. Unfortunately, we do not have the time to go into syntactic language variation in any great detail).

2. *More than just strings*

The first thing to realize when we start looking at the kind of principles that distinguish possible word orders from impossible ones is that sentences are not just linear (or, rather, temporal) strings of words. They have an internal *structure*. Children acquiring a language already appear to be aware of this, since (it has been claimed) they only ever seem to use *structure-sensitive* rules in trying to account for the language data they hear around them -- even though these rules might not yet be the same rules of the target adult language. Children make mistakes (meaning they produce sentences that are not possible sentences in the adult language), but not random ones.

The following example, based on Santorini and Krochs' (henceforth S&K) (19) to (25) in chapter 1, illustrates this. It concerns the way so-called yes/no-questions are formed in English. Suppose a child hears (11b) and (12b) and realizes they are the yes/no-question counterparts to declarative (11a) and (12a), respectively.

(11) a. The girl is tall.
b. Is the girl tall?

(12) a. The tall girl can see the boy who is holding the plate.
b. Can the tall girl see the boy who is holding the plate?

If sentences are just strings of words, an entirely plausible hypothesis for the child to entertain is that these data indicate that yes/no-questions in English are formed thus:

(13) *Rule 1*: Find the first auxiliary verb in the clause and put this up front.

This rule would give results like (14b).

(14) a. The boy who is holding the plate can see the tall girl.
b. *Is the boy who holding the plate can see the girl?
c. Can the boy who is holding the plate see the girl?

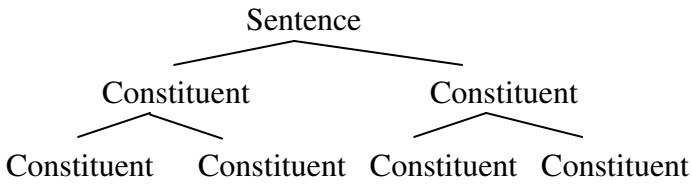
(14c) indicates that the correct rule for adult English is:

(15) *Yes/no question formation in English*
Find the first auxiliary verb *after a particular group of words belonging together, namely the subject of the sentence,* and put this up front.

The example shows that the rule of yes/no-question formation does not simply count words (what is the *first* auxiliary verb), but is sensitive to the fact that a sentence is divided into groups of words that belong closer together than others. Such a group of words that cling together is called a *constituent* of the sentence (more on how to distinguish constituents next week). S&K cite work by Crain and Nakayama that appears to show that, while children do make errors in producing yes/no-questions, they never make errors that indicate that they do not treat the sentence as consisting of constituents, including a subject constituent (i.e. they do not make errors that would be on a par with (14b)).

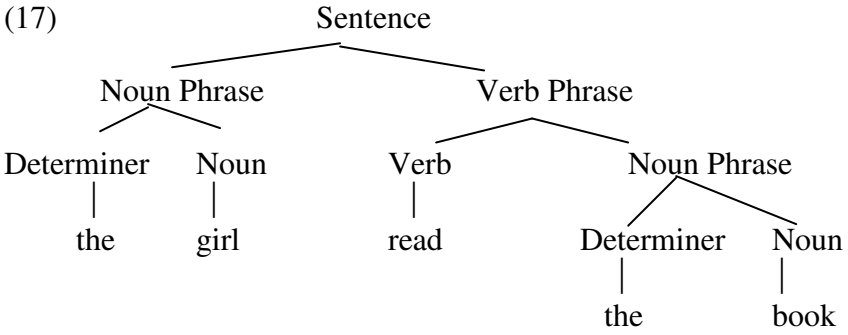
Constituents can themselves contain smaller constituents. The subject in (14a) (*the boy who is holding the plate*), for example, contains a so-called relative clause (namely *who is holding the plate*), which in turn contains constituents, such as its so-called direct object *the plate*. Such containment relationships (that is, the constituency of a sentence) can be expressed by a *tree structure:*

(16)



(16) expresses that there is a sentence that consists of two constituents, which both happen to consist of two smaller constituents. For example, a simple sentence like *The girl read a book* will have the tree structure in (17) (more on this in lectures 3 and 4):

(17)



3. *Recursion*
A fundamental property of natural (non-artificial) languages is that they allow their syntactic structures to have a property called *recursion*. This is related to what we

noted just now: constituents can contain smaller constituents. Recursion is the phenomenon that a constituent can contain a smaller constituent *of the same type as the bigger constituent itself*. It is this property that is responsible for the fact that, as mentioned earlier, the number of possible sentences in a language is infinite. For example, a sentence can contain a smaller sentence as one of its constituents. In (18a) we see that the verb *regret* can take a nominal phrase like *this fact* as its object. But in (18b) we see that it can also take a complete sentence (introduced by the function word *that*) as its object. The result is that we have a sentence that is *embedded* in a larger sentence.

(18) a. John regrets [this fact].
b. John regrets [that Mary believes these rumours].

Within the smaller sentence the same thing is possible, and so on, *ad infinitum*:

(19) John regrets [that Mary believes [that Harry has said [that the paper reported [that…

There is no principled limit to recursion. Of course, for practical reasons any sentence will come to an end in actual conversations – but *grammatically* speaking there is nothing wrong with endlessly recursive structures.

Recursion is possible for other sorts of constituents than full sentences as well. For example, a word group built around a noun, a *Noun Phrase* or *NP,* can contain within it a word group built around a preposition, a *Preposition Phrase* or *PP*:

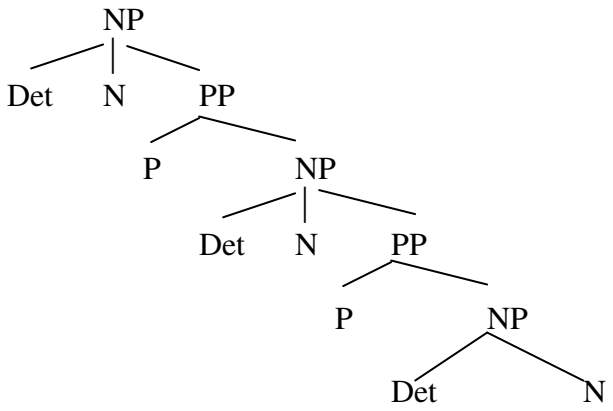
(20) [The dog [in the car]]

But as this example also shows, the PP *in the car* in turn contains another NP, namely *the car*. Nothing stops this NP from containing a PP again, just like the bigger NP in which it occurs:

(21) [The dog [in the car [outside the garage]]]

The schematic tree representation of (21) looks as follows:

(22)



We see that a Noun Phrase can contain another Noun Phrase, and a Preposition Phrase another Preposition Phrase. Again, there is no principled limit to this:

(23) The dog in the car outside the garage in the street of the town on the river …

Exercises: SK chapter 1 exercises 1, 2, 3 + explain the boxed text in exercise 4.