Introduction to Syntax

Phil Crone

April 9, 2014
Section locations

My section will meet in Econ 139 on Wednesdays from 11 to noon.

Dasha’s section will meet in 200-303 on Thursdays from 3 to 4.
What is syntax?

Syntax is the part of grammar that represents speakers’ knowledge of sentences and their structures.

Just as speakers have knowledge about acceptable combinations of morphemes to form words . . .

(1) Detoxify
(2) *Dehappy

. . . they also have knowledge about acceptable combinations of words to form sentences . . .

(3) David Cameron detoxified the Tory brand, but made the product radioactive.
(4) *Detoxified David Cameron Tory the brand, but the radioactive made product.
What is syntax?

Just as words have internal morphological structure . . .

usable

un-

use

-able

. . . phrases and sentences have internal syntactic structure . . .

S

NP

David Cameron

VP

V

detoxified

NP

the Tory brand
All languages have rules regarding acceptable combinations of words to form sentences, but these rules differ from language to language.

<table>
<thead>
<tr>
<th>Basic Word Order</th>
<th>Prevalence</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOV</td>
<td>45%</td>
<td>Turkish, Japanese, Latin</td>
</tr>
<tr>
<td>SVO</td>
<td>42%</td>
<td>English, Spanish, Mandarin Chinese, Russian</td>
</tr>
<tr>
<td>VSO</td>
<td>9%</td>
<td>Arabic, Irish, Zapotec</td>
</tr>
<tr>
<td>OSV</td>
<td>0%</td>
<td>Warao</td>
</tr>
<tr>
<td>OVS</td>
<td>1%</td>
<td>Apalaí, Hixkaryana</td>
</tr>
<tr>
<td>VOS</td>
<td>3%</td>
<td>Malagasy, Baure</td>
</tr>
</tbody>
</table>
Grammaticality vs. meaningfulness

(5)  # Colorless green ideas sleep furiously.

The sentence in (5) is syntactically well-formed, but doesn’t make any sense (indicated by #).

(6)  * The boy quickly in the house the ball found.

The sentence in (6) is not syntactically well-formed, but we have no trouble interpreting it.
Constituents & constituency tests

Sentences are composed of natural sub-parts called constituents. We can identify constituents through a variety of constituency tests.

(7) At a news conference in Donetsk, Ms. Tymoshenko said she was committed to strengthening the autonomy of Ukraine’s regions.

Test 1: What can stand on its own?
Q: What did Ms. Tymoshenko say she was committed to strengthening?
A: The autonomy of Ukraine’s regions.

Test 2: What can be replaced by a pronoun?

(7’) At a news conference in Donetsk, she said . . .
Constituents & constituency tests

(7) At a news conference in Donetsk, Ms. Tymoshenko said she was committed to strengthening the autonomy of Ukraine’s regions.

Test 3: What can move as a unit?

(7”) Ms. Tymoshenko said she was committed to strengthening the autonomy of Ukraine’s regions, at a news conference in Donetsk.

Test 4: What can be coordinated?

(7”) Ms. Tymoshenko said she was committed to strengthening the autonomy of Ukraine’s regions and cities.
Syntactic categories

Each constituent of a sentence belongs to a particular **syntactic category**.

(8) Phil addressed the class.

What are the constituents of (8)?

(8’) [ [ Phil ] [ addressed [ the class ] ] ]

Both *Phil* and *the class* are part of the same syntactic category. They are **Noun Phrases (NP’s)**. Each can be replaced by another NP and the sentence is still grammatical.

(8’’) The class addressed Phil.

(8’’’) Dasha addressed the class.

(8’’’’) The woman sitting in the tenth row typing on her computer addressed the man to her right wearing a V-neck sweater.

(8’’’’’) …
Syntactic categories

(9) Phil addressed the class.

The constituent *addressed the class* is a **Verb Phrase (VP)**. It can be replaced by other VP's.

(9′) Phil loves linguistics.

(9″) Phil watched *House of Cards* for three hours last night.

(9‴) ...

Constituents of the same syntactic category have the same distribution.
Syntactic categories

Every phrase contains a **head**. The category of the head determines the category of the phrase.

**Noun Phrase (NP):** men, the man, the man with a telescope

**Verb Phrase (VP):** sees, always sees, rarely sees the man

**Adjective Phrase (AP):** happy, very happy, very happy about winning

**Prepositional Phrase (PP):** over, nearly over, nearly over the hill

**Adverbial Phrase (AdvP):** brightly, more brightly, more brightly than the sun

**Noun (N):** puppy, boy, man, soup

**Verb (V):** find, run, sleep, throw

**Adjective (A):** red, big, happy, candid

**Preposition (P):** up, down, across, into

**Adverb (Adv):** again, always, brightly, often
In addition to **lexical categories**, there are also **functional categories**.

**Determiner (D):** the, a, every, that  
**Auxiliary (Aux):** have, may, can, will  
**Conjunction (Conj):** and
Phrase structure trees

We can represent information about constituency, syntactic categories, and word order in a **phrase structure tree**.

(10) Dasha loves *Breaking Bad*.

```
S
  NP  VP
    N  V  NP
      Dasha  loves  N
          "Breaking Bad"
```
Phrase structure trees

Phrase structure trees can also represent syntactic ambiguities that are not revealed if we just consider linear word order.

(11) I shot an elephant in my pajamas.

The PP *in my pajamas* could be modifying either *shot an elephant* or *an elephant.*
Phrase structure rules

Phrase structure trees can represent the structure of sentences, but by themselves they don’t tell us which structures will be grammatical and which structures won’t.

(10’) Dasha loves *Breaking Bad*.

What we need are rules to tell us which phrase structure trees are good and which ones are bad. Such rules are called **phrase structure rules**.
Phrase structure rules

Like phrase structure trees, phrase structure rules give information about word order, constituency, and syntactic categories.

(12) \( S \rightarrow \text{NP VP} \)

The rule in (12) says that a sentence consists of an NP and a VP and that the NP precedes the VP.

We can add a few more rules, and then use them to build phrase structure trees.

(13) \( \text{NP} \rightarrow \text{N} \)

(14) \( \text{VP} \rightarrow \text{V NP} \)

(15) \( \text{N} = \text{Dasha}, \text{ Breaking Bad} \)

(16) \( \text{V} = \text{loves} \)
A set of phrase structure rules generates a set of sentences and their associated phrase structure trees. The rules in (12)-(16) generate the following two trees (as well as two more):
Phrase structure rules

A **generative grammar** is a grammar that gives rules which produce all and only the grammatical sentences and their phrase structures for a given language.

But recall that there is an infinite number of possible grammatical sentences for any language.

(17) Dasha loves a show that takes place in New Mexico.

(18) Dasha loves a show that stars Bryan Cranston and takes place in New Mexico.

(19) Dasha loves a show that stars Bryan Cranston and takes place in New Mexico, which was admitted to the Union in 1912.

(20) ... 

How can we generate an infinite set of sentences (and their phrase structures) without an infinite list of rules?
Phrase structure rules

Recursion. See *Recursion*

A **recursive phrase structure rule** is a rule for a syntactic category that repeats that category on the right side. In other words, a syntactic category $XP$ can contain another instance of $XP$. Here’s an example:

(21) $VP \rightarrow VP\text{ Conj }VP$

The rule in (21) says that a $VP$ can consist of a $VP$, followed by a conjunction, followed by another $VP$. Let’s add this to our earlier set of rules, and change some of those earlier rules:

(15’) $N = \text{Dasha, Breaking Bad, The Real Housewives of Atlanta, Homeland}$

(16’) $V = \text{loves, hates}$

(22) $\text{Conj = and}$
Phrase structure rules

(12) \[ S \to NP \ VP \]
(13) \[ NP \to N \]
(14) \[ VP \to V \ NP \]
(21) \[ VP \to VP \ Conj \ VP \]
(15') \[ N = \text{Dasha, Breaking Bad, The Real Housewives of Atlanta, Homeland} \]
(16') \[ V = \text{loves, hates} \]
(22) \[ Conj = \text{and} \]

Our new set of rules can now generate sentences like the following:

(23) \[ \text{Dasha loves Breaking Bad and loves The Real Housewives of Atlanta and hates Homeland.} \]
Transformations

(24) Dasha will watch *The Real Housewives of Atlanta*.
(25) Will Dasha watch *The Real Housewives of Atlanta*?

The sentences in (24) and (25) are grammatical sentences of English. We could generate them by adding the following phrase structure rules:

(26) $S \rightarrow NP \ Aux \ VP$
(27) $S \rightarrow Aux \ NP \ VP$

But simply adding these rules would miss a generalization about (24) and (25). Both sentences are about the same type of situation, although (24) asserts that that situation will occur and (25) asks if it will occur. This difference in meaning varies systematically with the difference in structure.
Transformations

This systematicity can be explained by treating both (24) and (25) as coming from an underlying structure. This underlying structure is called a deep structure or d-structure. Deep structures are generated by phrase structure rules.

The form of the sentence that actually gets spoken is called a surface structure or s-structure. Surface structures are derived from deep structures via transformational rules.
Transformations

(24) Dasha will watch *The Real Housewives of Atlanta*.
(25) Will Dasha watch *The Real Housewives of Atlanta*?

We can say that both (24) and (25) have the same deep structure generated by our phrase structure rules. This structure is shown below.
Transformations

(24) Dasha will watch *The Real Housewives of Atlanta*.
(25) Will Dasha watch *The Real Housewives of Atlanta*?

Then we can say that the surface structure in (24) is derived from
the deep structure by applying no transformational rules.

The surface structure in (25) is derived from a transformational
rule that says a question can be formed from structures of the form
NP Aux VP by moving the Aux to precede the NP.
Recall the diversity of word orders mentioned at the beginning of the lecture. It turns out that the standard order of a verb and an object in a language is highly correlated with the order of other elements.

These orders can be broken down by phrasal category.

<table>
<thead>
<tr>
<th>Phrasal Category</th>
<th>Head Initial Language</th>
<th>Head Final Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP</td>
<td>V NP</td>
<td>NP V</td>
</tr>
<tr>
<td>VP</td>
<td>Aux VP</td>
<td>VP Aux</td>
</tr>
<tr>
<td>PP</td>
<td>P NP</td>
<td>NP P</td>
</tr>
<tr>
<td>NP</td>
<td>NP relative clause</td>
<td>relative clause NP</td>
</tr>
<tr>
<td>NP</td>
<td>N possessive NP</td>
<td>possessive NP N</td>
</tr>
<tr>
<td>NP</td>
<td>N AP</td>
<td>AP N</td>
</tr>
</tbody>
</table>
Some linguists have tried to explain these correlations via a so-called **headedness parameter**, which would specify for every language whether it is head-initial or head-final. Depending on the value of the headedness parameter for a particular language, the phrase structure rules for that language would adhere to one of the following two forms:

\[(26) \quad XP \rightarrow X YP \text{ (for head-initial languages)}\]

\[(27) \quad XP \rightarrow YP X \text{ (for head-final languages)}\]

But note that very few languages display all the canonical properties of either head-initial or head-final languages. For example, English is usually head-initial, but possessive NP’s can either proceed or follow a head noun (e.g. *my friend* and *a friend of mine*) and adjectives usually precede nouns.