CAS LX 522
Syntax I

Week 1. Introduction

Some things we know

- Is this English?
  - Pat the book lifted.
  - Pat lifted the book.
  - Lifted Pat the book.
  - Pat book the lifted.
- Why?

It’s surprisingly complicated

1) Tony threw out the couch.
2) Tony threw the couch out.

- Prepositions can go on either side of the object.

3) Tony stormed out the door.
4) * Tony stormed the door out.

It’s surprisingly complicated

5) What did Mary say John bought?
6) What did Mary say that John bought?

- Ok, that is optional.

7) Who did Mary say bought coffee?
8) *Who did Mary say that bought coffee?

How do people know this?

- All native speakers of English know this.

- Little kids weren’t told these rules (or punished for violating them)...
  - “You can’t question a subject in a complement embedded with that”
  - “You can’t use a proper name as an object if the subject is co-referential.”
Two questions

- What do people know about their language?
  - Including things we know “unconsciously”
- If we don’t know we know it, how did we come to know it?

Systematicity

- What people eventually end up with is a system with which they can produce (and rate) sentences. A grammar.
- Even if you’ve never heard these before, you know which one is “English” and which one isn’t:
  16) Eight very lazy elephants drank brandy.
  17) Eight elephants very lazy brandy drank.

Positive and negative evidence

- Adults know if a given sentence S is grammatical or ungrammatical. This is part of the knowledge kids gain through language acquisition.
  - Kids hear grammatical sentences (*positive evidence*)
  - Kids are not generally told which sentences are ungrammatical (*no negative evidence*)

Positive and negative evidence

- One of the striking things about child language is how few errors they actually make.
- For negative feedback to work, the kids have to make the errors (so that it can get the negative response).
- But they don’t make the errors.

The “Language instinct”

- The linguistic capacity is part of being human.
- Like having two arms, ten fingers, a vision system, humans have a language faculty.
- The language faculty (tightly) constrains what kinds of languages a child can learn.
- =“Universal Grammar” (UG).

But languages differ

- English, French: Subject Verb Object (SVO)
  - John ate an apple.
  - Pierre a mangé une pomme.
- Japanese, Korean: Subject Object Verb (SOV)
  - Taroe-va ringo-o tabeta.
  - Chelsuwu-ka sakweo-lul mekessta.
- Irish, Arabic (VSO), Malagasy (VOS), …
But languages differ

- English: Adverbs before verbs
  - Mary quickly eats an apple.
    - (also: Mary ate an apple quickly)
  - *Mary eats quickly an apple.

- French: Adverbs after verbs
  - Geneviève mange rapidement une pomme.
  - *Geneviève rapidement mange une pomme.

Parameters

- We can categorize languages in terms of their word order: SVO, SOV, VSO.
- This is a parameter by which languages differ.
- The dominant formal theory of first language acquisition holds that children have access to a set of parameters by which languages can differ; acquisition is the process of setting those parameters.
- What are the parameters? What are the “universal” principles of grammar?

The enterprise

- The data we will primarily be concerned with are native speaker intuitions.
- Native speakers, faced with a sentence S, know whether the sentence S is part of their language or isn’t. These intuitions are highly systematic.
- We want to uncover the system (which is unconscious knowledge) behind the intuitions of native speakers—their knowledge of language.

I-language

- We are studying the system behind one person’s pattern of intuitions.
- Speakers growing up in the same community have very similar knowledge, but language is an individual thing (“I-language”).
- I-languages of a community is can be characterized, but it is external to the speaker (“E-language”), not any one person’s knowledge, a generalization over many people’s I-languages.
  - For example, Parisian French.

Competence

- We are also concerned with what a person knows. What characterizes a person’s language competence. We are in general not concerned here with how a person ends up using this knowledge (performance).
- You still have your language competence when you are sleeping, in the absence of any performance. Being drunk doesn’t make you think “bought some John coffee” is English, though perhaps one might say it.

Prescriptive rules

- Another thing we need to be cautious of are prescriptive rules. Often prescriptive rules of “good grammar” turn out to be impositions on our native grammar which run counter to our native competence.
- After all, why did they need to be rules in the first place?
Prescriptive rules

- Prepositions are things you don’t end a sentence with.
- We want to successfully complete this course.
- Remember: Capitalize the first word after a colon.
- Don’t be so immodest as to say *I and John left; say John and I left instead.*
- *Impact* is not a verb.
- The book which you just bought is offensive.

Judgments

- Another complicating fact is that a sentence can be bad for any number of reasons, only some of which we are interested in at a given point.
- *Student the meditated happily.*
- The pebble meditated happily.
- A Sun rose in the East.
- John wondered who to go with.

Syntax as science

- Here, we will study syntax *scientifically.* This means, in particular, approaching syntax using the scientific method.
- Step 1: Gather observations (data)
- Step 2: Make generalizations
- Step 3: Form hypotheses
- Step 4: Test predictions made by these hypotheses, returning to step 1.

A simple introductory example

1) Bill kissed himself.
2) Bill kissed herself.
3) Sally kissed himself.
4) Sally kissed herself.

- Try these out. Which ones sound good, which ones don’t?

A simple introductory example

1) Bill kissed himself.
2) *Bill kissed herself.*
3) *Sally kissed himself.*
4) Sally kissed herself.

- Hypothesis: An anaphor must have an antecedent which agrees with it in *gender.*
Hypothesis: An anaphor must have an antecedent which agrees in gender.

Let's test the hypothesis against more data.

5) The robot saw itself in the mirror.
6) John and Bill saw himself in the mirror.
7) The boys saw himself in the mirror.
8) Mary and Jane saw herself in the mirror.
9) John and Bill saw themselves in the mirror.
10) Mary and Jane saw themselves in the mirror.
11) The boys saw themselves in the mirror.

Our hypothesis only explains (5). What is the generalization?

Hypothesis: An anaphor must have an antecedent which agrees in gender.

5) The robot saw itself in the mirror.
6) John and Bill saw himself in the mirror.
7) The boys saw himself in the mirror.
8) Mary and Jane saw herself in the mirror.
9) John and Bill saw themselves in the mirror.
10) Mary and Jane saw themselves in the mirror.
11) The boys saw themselves in the mirror.

Hypothesis: An anaphor must agree in gender and number with its antecedent

12) The executives gave themselves a raise.
13) I gave myself a cookie.
14) I gave myself a cookie.
15) You gave yourself a cookie.
16) You gave herself a cookie.
17) You gave himself a cookie.
18) You gave yourself a cookie.

Again, our hypothesis doesn’t successfully predict which of these are grammatical and which aren’t. What’s the generalization?

Person

- I is “first person singular”
- You is “second person singular” (you left) or “second person plural” (you left = y’all left)
- He, She is “third person”
- We is “first person plural”
- They is “third person plural”

Anaphors seem to agree with person. Myself for first person singular, ourselves for first person plural, yourself for second person, himself, herself, or itself for third person singular, themselves for third person plural.

Hypothesis about anaphors

- An anaphor must agree in gender, number, and person with its antecedent.

- This is the hypothesis we will end with, although there will be more to do with anaphors.

Incidentally, gender, number, and person very often go together. Very rarely will you see agreement with one and not the others as well.

Levels of adequacy

- If our hypotheses can predict the existence of the grammatical sentences in a corpus (a set of grammatical sentences), it is observationally adequate.
- If our hypotheses can predict the native-speaker intuitions about which sentences are grammatical and which are ungrammatical, it is descriptively adequate.
Levels of adequacy

- If we can take a descriptively adequate set of hypotheses one step further and account not only for the native speaker judgments but also for how children come to have these judgments, our hypotheses are **explanatorily adequate**.
- It’s this last level that we are hoping to achieve.

Basic principles

- Parameters of variation
- How to set the parameter from input

Refresher on syntax

- Words can be grouped into **categories** by *part of speech* like noun, verb, adjective, preposition, ...
- Parts of speech are determined *distributionally* (traditional “semantic” definitions don’t work)
  - The yinkish dripner blokred quastofically.
  - *Yinkish* is an adjective, *dripner* is a noun, *to blork* is a verb, *quastofically* is an adverb.

Constituents

- The words that make up a sentence like...
  - The students did their syntax assignment.
  - ...are grouped together into component parts, *constituents*, which function together as a unit.
  - Among them, [the students], the do-ers, and [their syntax assignment], the done.

Phrases

- A phrase is a constituent that has a central core word (called the *head* of the phrase); other words in the phrase generally relate to (or modify the meaning of) the head.
- The category of the head determines the category of the phrase.
  - *The happy students* is a noun phrase, headed by the noun *students*. *Happy* modifies *students*, the specifies which students. *Ran swiftly* is a verb phrase, *swiftly* modifies *ran*.

Sentences

- Complete sentences need to have a subject and a verb.
  - John left.
  - *John.
  - *Left.
  - The happy students left speedily.
- So sentences are made of noun phrases and verb phrases.
**Sentences**
- John left (speedily).
- The happy students left (speedily).

We could say a sentence is either:
- a noun and a verb
- a noun and a verb phrase
- a noun phrase and a verb
- a noun phrase and a verb phrase

Or we could say:
- a sentence is a noun phrase and a verb phrase
- a phrase always has a head, and sometimes that’s all.

**Trees**
- We can start by drawing the structure of a sentence like this, which means: “John left is a Sentence composed of a Noun Phrase (composed of John) and a Verb Phrase (composed of left). Note the heads.

```
    S
   / \    
  /   \   
 NP  VP   
   /     
  /      
 N      V
   /    
  /     
 John  left
```

**Of the past and the future**
- Serious scientific study of sentence structure of this kind generally began in the 50’s, driven to a great extent by the work of Noam Chomsky.

- It’s now half a century later, and we have learned a lot about how syntax works.

**Of the past and the future**
- Progress was incremental, and often required revising our assumptions about how sentences are really put together.

- Data was examined, generalizations were arrived at, hypotheses were formed, predictions were tested—and often led to revisions of the generalizations and the hypotheses, and so forth.

**Of the past and the future**
- Two goals of the class:
  - Think like a syntactician.
  - Be able to read books, articles, etc. about syntax.

- It’s not really enough to just know what people concluded, we need to understand why they concluded what they did.

**Some milestones**
- Until about the mid-70’s, phrase structure rules (like what we’ll see today).
- Mid-70’s, X-Bar Theory (a generalization about what are possible PSRs).
- In the 80’s, a fairly significant shift to Government and Binding Theory (viewing grammar a little less like a computer program). Very productive.
- In the 90’s, another shift to the Minimalist Program (an attempt at simplification, as well as a change in philosophy).
What’s different

- Progress was incremental, and very often the differences between the stages of generative linguistic theory are not as great as they seem. Just looking at the forest from a different perspective (sometimes with a microscope).
- So let’s start with the data and some generalizations, and pretend that we’re back in the 60’s for a while.

Finding constituents

- How do we find constituents in a sentence? For many of them, we can guess, but a guess isn’t evidence. If sentences and phrases have structure, we should be able to test for this structure.

Replacement test

- A constituent is a group of words which function as a unit. If you can replace part of the sentence with another constituent (the smallest constituent being a single word), this tells us that the replaced section of the sentence is a constituent.
- This isn’t foolproof, but it usually works if you try to keep the meaning as close as possible.

Replacement test

- The students left.
- They left.
- *The students* is a constituent.
- The students ate the sandwiches.
- They ate the sandwiches.
- *The students ate* them.
- The students dined.
- [The students] [ate [the sandwiches]].

Sentence fragment test

- Generally, only constituents can be used in the fragmentary response to a question.
- Who ate the sandwiches?
  - The students.
  - *The.*
- What did the students do?
  - Ate the sandwiches.
  - *Ate the.*
- What did the students eat?
  - The sandwiches.
- [The students] [ate [the sandwiches]].

Movement tests—”clefting”

- If you can move a sequence of words X together to another part of a sentence that means roughly the same thing, X is a constituent.
- Clefting: creating a sentence of the shape It was — who/that — out of your sentence.
- Start with: It was — who sentence.
  - It was — who/that John bought coffee.
- Pick a candidate for being a constituent.
- Put it in place of the — and cross it out in the sentence.
  - Pick who or that.
  - It was John who/that John bought coffee.
  - It was coffee who/that John bought coffee.
Movement tests—clefting

- The students ate the sandwiches.
- It was [the students] who ate the sandwiches.
- It was [the sandwiches] that the students ate.
- "It was [the] that students ate the sandwiches.
- [The students] [ate [the sandwiches]].

Movement tests—preposing

- Preposing involves creating a sentence by putting a constituent at the beginning of the sentence of this form:
  - — is/are who/what/where/… —
- The students ate the sandwiches.
- [The students] are who ate the sandwiches.
- [The sandwiches] are what the students ate.
- [Eat the sandwiches] is what the students did.
- "[The] is what students ate the sandwiches.

Coordination test

- Generally you can replace a constituent of a certain type X with another constituent of type [X and X].
  - [The students] ate the sandwiches.
  - [[John] and [the students]] ate the sandwiches.
  - The students ate [the sandwiches]
  - The students ate [[the sandwiches] and [the eggrolls]].
- This shows two things: 1) that the students and John are each constituents (as is John and the students), and 2) that the students is the same kind of constituent as John.

When constituency tests fail

- [The students] [ate [the sandwiches]].
- But consider:
  - John prepared and the students ate the sandwiches.
  - [[John prepared] and [the students ate]] the sandwiches?
  - The coordination test failed to reveal the structure.

When constituency tests fail

- **Moral:** Don’t rely on just one constituency test.
- Use several tests, assuming that occasionally any given test might yield an anomalous result.

Trees and constituency

- [The students] [ate [the sandwiches]]
Trees and constituency

- [The students] [ate [the sandwiches]]

Phrases and constituents

- The constituents we have identified are the noun phrases and the verb phrase, which have internal pieces as well.
  - *The* is a determiner (D).
  - *Students* and *sandwiches* are nouns, heading the Noun Phrase (NP).
  - *Ate* is a verb, heading the Verb Phrase (VP)

Inside the NP

- The NPs we found have a D and an N
  - The students
  - The sandwiches
- An NP needs to have a *head noun*.
- An NP can have a determiner:
  - NP: (D) N
    - Paren mean D is optional
- But there are lots of other kinds of NPs...

Inside the NP

- **The enthusiastic students** ate the sandwiches.
- **The enthusiastic syntax students** ate the sandwiches.
- **The enthusiastic syntax students in LX522** ate the sandwiches.
Inside the NP

- The enthusiastic students…

- *Enthusiastic* is an adjective (modifying *students*).
- So we need to revise our hypothesis about the components of NP.
- NP: (D) (Adj) N

Inside the NP

- The enthusiastic syntax students…

- We know *enthusiastic* is an adjective modifying *students*. What is *syntax*?
- Syntax seems to *also* be an adjective modifying *students*. You can have two adjectives in an NP.
- We need to revise our hypothesis again.

Inside the NP

- The enthusiastic syntax students…

- NP: (D) (Adj) (Adj) N

- The big red fluffy dog barked.
- The excited big red fluffy dog barked.
- NP: (D) (Adj) (Adj) (Adj) N

- This seems to miss a generalization. Instead…
- NP: (D) (Adj+) N

  + means “repeat as many times as necessary”

Inside the NP

- The enthusiastic students in LX522…

- Back to the problem, revising our hypothesis:

  - NP: (D) (Adj+) N (PP)

Inside the NP

- The enthusiastic students in LX522…

- Intermission: What’s in LX522?
- It is a prepositional phrase (PP), with a head preposition (P) in. To a pretty close approximation:
- PP: P NP

  - In the tree… in the big green tree…

Modifiers

- Golden Rule of Modifiers:
  Modifiers are always attached within the phrase they modify.

  \[
  \text{[NP The enthusiastic syntax students in LX522]}
  \]
Modifiers

- The bird in the tree on the hill…
- This is an NP…
  - The bird in the tree on the hill left.
  - I left.
  - The bird in the tree on the hill and I left.
- Trick question: Should we revise our NP rule to NP: (D) (Adj+) N (PP+)?

Answer to the trick question:
Not based on this evidence!

The actual structure…

Trees

- We can draw the same information in a tree.

Trees

- We can draw the same information in a tree.
Modifiers

- The book of poems with the blue cover...
- Now should we revise our NP rule to NP: (D) (Adj+) N (PP+)?
- Of poems modifies book.
- With the blue cover modifies book.
- Answer: Yes, this is evidence for the new NP rule.

Trees

- And the tree...

Inside the VP

- A VP always has a head verb (V).
  - Pat left.
  - VP: V
- A VP can sometimes have an adverb.
  - Pat left quickly.
  - VP: V (Adv)

Inside the VP

- Pat quickly left.
- Pat often left early.
- Pat cleverly rarely shouts loudly twice.
- A VP can have any number of adverbs before or after the verb.
- VP: (Adv+) V (Adv+)

Inside the VP

- The students ate the sandwiches.
  - The students ate the sandwiches hungrily.
- VP: (Adv+) V (NP) (Adv+)
  - Chris ate pizza at the café.
  - Chris ate pizza at the café hungrily.
  - Pat bought peanuts at the café for a dollar.
  - Pat bought peanuts at the café for a dollar on Tuesday triumphantly.
- VP: (Adv+) V (NP) (PP+) (Adv+)

What we’ve got...

- NP: (D) (Adj+) N (PP+)
- PP: P (NP)
- VP: (Adv+) V (NP) (PP+) (Adv+)
- The very happy students left.
- [NP D Adv Adj N ] V
- What should we do now?
AdjP and AdvP

- The very happy students left.
- The Golden Rule of Modifiers says that modifiers must attach inside the phrase they modify.
  - *Very* is modifying *happy*.
  - *Very* must be inside an Adjective Phrase (AdjP)
- AdjP: (Adv) Adj
- The students left very quickly.
  - *Very* is modifying *quickly* (an adverb).
  - *Very* must inside an Adverb Phrase (AdvP)
- AdvP: (Adv) Adv

What we’ve got now...

- NP: (D) (AdjP+) N (PP+)
- PP: P (NP)
- VP: (AdvP+) V (NP) (PP+) (AdvP+)
- AdjP: (Adv) Adj
- AdvP: (Adv) Adv
- Let’s digest this a bit.
  - Every phrase has one required element.
  - This one required element is the head.
  - Every phrase has *only one* head.

What we’ve got now...

- NP: (D) (AdjP+) N (PP+)
- PP: P (NP)
- VP: (AdvP+) V (NP) (PP+) (AdvP+)
- AdjP: (Adv) Adj
- AdvP: (Adv) Adv
- Given this, AdvP looks somewhat suspicious.
  - Which Adv is the head in *very quickly*?
- Suppose: Modifiers are *always* phrases.
  - This requires a revision to get us closer...

What we’ve got now...

- NP: (D) (AdjP+) N (PP+)
- PP: P (NP)
- VP: (AdvP+) V (NP) (PP+) (AdvP+)
- AdjP: (Adv) Adj
- AdvP: (Adv) Adv
- Hypothesis: Phrases consist of one head and modifier phrases.
  - There’s still one non-conformist in our midst.
  - Should D be a DP?
  - Put this on hold: Leave this is the sole exception.
  - Yes, D should be a DP, but there is a complication that we need to come back to.

For next time:

- Read:
  - Chapter 1-2
  - (Chapter 2 contains some material we didn’t cover this time but will address next time)
- Homework:
  - Chapter 1: problem 1.
  - Chapter 2: problems 1, 2, and 6.