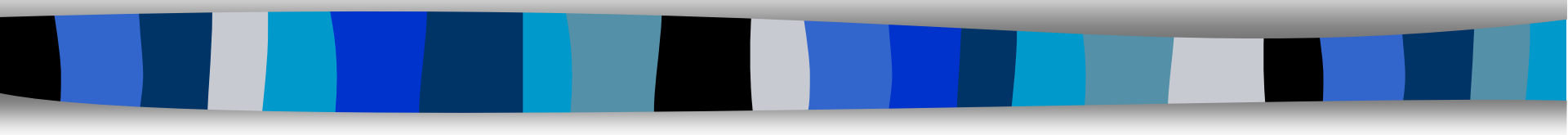
Syntax



Andrew Carnie

The web page for this textbook

http://www.blackwellpublishers.co.uk/carnie

Topic 1: Syntax: some background

- What is syntax?
- Syntax as a (cognitive) science
- Rules
- prescriptivism vs. descriptivism
- Evaluating Grammars
- Language as an instinct

Q. What is Syntax??

- The scientific study of sentence structure
- Perspective: The psychological (or cognitive) organization of sentence structure in the mind.

Q. What is a sentence??

- A hierarchically organized structure of words that maps sound to meaning and vice versa.

sounds ⇔ sentences ⇔ meaning

Scientific Method

- Study of syntax is a *science*.
- Uses the scientific method
  - Observe some data
  - Make some generalizations
  - Develop a hypothesis
  - Test against more data

Scientific method

Anaphor: A noun that refers back to a previously mentioned noun: "self" nouns.

1) John loves himself

2) Mary loves herself

3) John and Mary love themselves

Generalization: The form of the "Xself" seems to be dependent upon the gender/number of the noun they refer to.

Hypothesis: Anaphors (Xself) agree with the noun they refer to in number and gender.

4) The boy loves himself/*herself/*themselves

Rules: A kind of hypothesis

- In this class, we will encode our hypotheses about sentence structure using rules.
- A group of rules are called a Grammar.
- Grammar is a scary word. But it doesn't mean what you think it does. A grammar in the linguistic sense is a cognitive structure. It is the part of the mind that generates and understands language.

Prescriptive vs. Descriptive Rules

We are always told to never split infinitives.

Who(m) did you give the book to?

Hopefully, we'll never learn the rules of grammar!

Prescriptive vs. Descriptive

- Prescriptive rules prescribe how we should speak
- Descriptive rules describe how we actually speak.

Which is more scientific?

Prescriptive Rules

- These are made up by so called language mavens

These are made up by so called language mavens! Who are they to tell you how to speak?!?
Prescriptive rules are often based on the rules of Latin or "logic". Who says Latin is so great? Why should language be logical?

Descriptive rules are the way to go!

Descriptive Rules

- The rules we will use are said to generate the sentences of the languages we are looking at. They actually build the sentences we produce. They are sentence building rules.
- The kind of grammar we are looking at is called generative grammar (=group of rules that generate the sentences of a language)

Sources of Data

- Corpora of Spoken & Written Language
  - Collections of recorded real world speech
  - Telephone recordings (LDC)
  - Newspapers, Books, Magazines
  - Folk tales etc recorded in the field.

Sources of Data

- *Where do you wonder if he lives?
  - How do you know this is ungrammatical?
  - Have you ever heard this sentence uttered?
  - Will the fact that this sentence is ungrammatical appear in any corpus?
- Every day, you produce grammatical sentences that have never been uttered before.

Sources of Data

- Corpora are not sufficient. They don't contain negative information (such as what sentences are ungrammatical), and they can never contain all the sentences of a language.
- We need to access our mental knowledge (also called "competence") about sentences.

Sources of Data

- A special experimental technique for tapping our syntactic knowledge.
- This technique is called the acceptability judgement.
- In the psychology literature, this is sometimes also called magnitude estimation

Acceptability Judgements

- Unfortunately, sometimes acceptability judgements are called intuitions.
- The term ‘intuition’ has a negative connotation: makes us think of fortune tellers and psychics.
- However, acceptability judgements are both *experimentally valid* and *statistically sound*.

Acceptability Judgements

- We will apply acceptability judgements in this class non-statistically. For the most part this will give us the right results. Statistical proof of judgements is possible, but we won't bother.

Performance vs. Competence

- Performance refers to what we do
- Competence refers to what we know about the language
- Our scientific concern: Both
- Our focus in this course: Competence

Evaluating Grammars

- Observationally Adequate Grammar: A grammar that accounts for all the observed (corpus/performance) data.
- Descriptively Adequate: Accounts for observations and acceptability judgements (competence). And generalizations
- Explanatorily Adequate: Accounts for observations, acceptability, AND accounts for language acquisition.

we aspire to Explanatorily Adequate Grammars.

Observationally Adequate Grammar: A grammar that accounts for all the observed corpus data.

- All and only the sentences in the data
  - Allow only sentences that have been seen
  - Exclude any sentences that have not been seen
- A problem: Any corpus both over and undergenerates
- A solution: Competence-based observational adequacy

Descriptively Adequate: Accounts for all observed data and all acceptability judgements (competence).

- Account for grammaticality **intuitions**
- Capture descriptive generalizations

Explanatorily Adequate: Explain why things are the way they are

- Identify the Laws of Nature at work
- Heavenly Bodies
  - Tyco: Described motions of planets with unprecedented accuracy (Observation)
  - Kepler: Determined that all planets have elliptical orbits (Descriptive Generalization)
  - Newton: Deduced the elliptical orbits of the planets from the laws of motion and gravitation (Explanation)

Chomsky's Conception of an Explanatorily Adequate Grammar

- The Laws of Grammar: Universal Grammar
- What's being explained by the laws: the miracle of language acquisition

Learning vs. Acquisition

- *Learning* involves conscious gaining of knowledge
- *Acquisition* involves subconscious gaining of knowledge

Chemistry is learned. Languages are acquired.

How do we acquire languages?

- Obviously this question is too big to answer here, but …
- Are we instructed by our parents?
- Do we mimic our parents?

NOPE!

1) Language is infinite: We produce sentences we've never heard before

2) We know things about our language that we've never been exposed to.

Language as an instinct

Despite what they may think, parents don't teach their children to speak!

They correct content not form:
(from Marcus et al. 1992)
Adult: Where is that big piece of paper I gave you yesterday?
Child: Remember? I writed on it.
Adult: Oh that's right, don't you have any paper down here, buddy?

Language as an instinct

(from Pinker 1994, 281 – attributed to Martin Braine)
Child: Want other one spoon, Daddy
Adult: You mean, you want the other spoon.
Child: Yes, I want other one spoon, please Daddy.
Adult: Can you say "the other spoon"?
Child: Other … one … spoon
Adult: Say "other"
Child: other
Adult: "spoon"
Child: Spoon.
Adult: "other … spoon"
Child: other … spoon. Now give me other one spoon.

Language as unconscious knowledge

You know things about your language that you've never been taught:

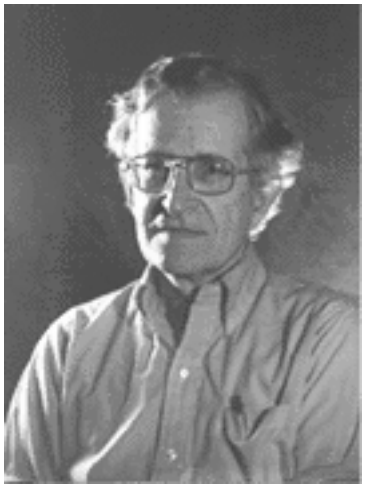
Who(m) did you think Shawn hit ?
Who(m) did you think that Shawn hit?
Who did you think hit Bill
*Who did you think that hit Bill

Language as unconscious knowledge: Things you don't know you know

- Who married his mother?
  - which person x married x's mother? (who married his own mother? Oedipus reading)
  - which person x married y's mother? (who married HIS, say Bill's, mother? who is Bill's father or stepfather? Stepfather reading)
- Who did his mother marry?
  - * which person x did x's mother marry? (no Oedipus reading)
  - which person x did y's mother marry? (stepfather reading okay)

A shocking proposal!

Noam Chomsky



The ability of humans to use language is innate (an instinct). We are prewired to use language!

Huh? languages differ?!?

How can language be an instinct if languages differ?

Proposal: Languages differ primarily in terms of what words are used, and in a set number of “parameters”

These things are learned but the rest (the basic architecture of the grammar) is innate.

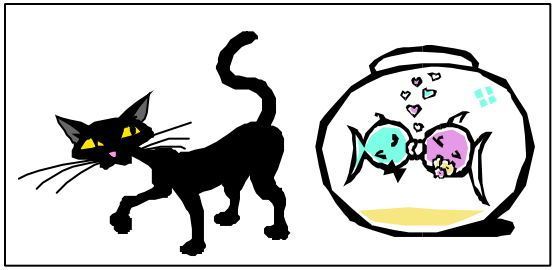
Refining Innateness

A particular language is not innate (it is acquired), but the basic tools that any given language uses are built in.

We'll be looking at these tools. Both within languages, and crosslinguistically to see what is universal (innate) and what varies among languages.

Task of a child acquiring English

- Match up a sentence that they hear with a situation in the context around them.

The cat spied the kissing fishes $=$  1

1

- To make the proof let's turn this into an algebraic operation. We'll number sentences, and we'll number situations, and look for the rule that matches them up.

What are basic building blocks?

- Example: Inferring a curve from an infinite set of points
- A grammar defines an infinite set of sentences
- The logical problem: From a finite set of data, a child must infer an infinite set of sentences
- Solution: we need a set of laws for making grammars: Universal Grammar

The content of this class

- In this class, we will be looking at the innate principles that govern sentence structure (Called Universal Grammar)
- And we will be looking at the different ways in which languages implement these innate principles.

Universal Grammar (UG)

- The building blocks that all languages use to construct the sentences of their languages.
- All languages use the same basic hardwired tools. It is the *particular implementation* of these tools that varies between languages.

Universal Grammar (UG)

- Other evidence for UG
  - Human Specificity of Language
  - Distinct area of the brain
  - Crosslinguistic similarities in language acquisition (despite cultural differences)
  - Lack of overt instruction
  - Language Universals

Summary

- Syntax: A Science, uses Scientific method, studies sentence structure
- Prescriptive/Descriptive Rules
- Generative Rules as Hypotheses

Summary

- Performance/Competence
- Evaluating Grammars:
  - Observationally Adequate
  - Descriptively Adequate
  - Explanatorily Adequate
- Learning vs. Acquisition
- Innateness of Language
- Universal Grammar: innate, hardwired building blocks of syntax.

Summary about Syntax

- Syntax is the scientific study of sentence structure
- Syntax is a branch of psychology [linguistics is a branch of psychology]
- We study competence=knowledge
- Competence is implicit knowledge

Evidence

- Corpora
- Speech
- Grammaticality judgments

Discussion Topics

- What things that we know are learned? What things are acquired?
- Language is an instinct. How is this an argument against prescriptive rules?
- There are some good reasons to keep prescriptive rules. What are they?