

# Syntax, semantics, ambiguity

LIN 313 Language and Computers  
August 29, 2025

# Administrivia

- Canvas Updates
  - [Homework 1 Posted](#) (Due Friday 9/12 at noon)
  - discussion board for Qs
- Website updates: <https://gchronis.github.io/lin313.html>
  - Link Forum for AI (FAI) talks
  - Links to slides and handouts
- Check back for 9/5 readings
- Concept Map
  - will be posted over the weekend

# Overview

- types of linguistic ambiguity
- syntax, semantics
- constituency
- winograd schemas
- garden path sentences

# Crash Blossoms

# Crash Blossoms

McDonald's Fries the Holy Grail for Potato Farmers

# Crash Blossoms

McDonald's Fries the Holy Grail for Potato Farmers



# What makes language "hard"

**What does this sentence mean?**



[<http://www.clker.com/clipart-green-eyes-3.html>]



[<http://www.clker.com/clipart-3163.html>]

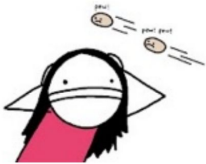


[<http://www.simonpalfrader.com/category/tournament-poker>]

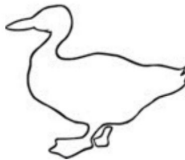
I saw her duck with a telescope.

verb

noun



[<http://casablancapa.blogspot.com/2010/05/fors.htm>]



[<http://www.supercoloring.com/pages/duck-outline/>]

# What makes language "hard"?

## Ambiguity

**What does this sentence mean?**



[<http://www.clker.com/clipart-green-eyes-3.html>]



[<http://www.clker.com/clipart-3163.html>]

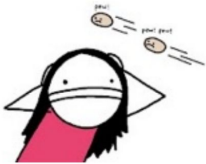


[<http://www.simonpalfrader.com/category/tournament-poker>]

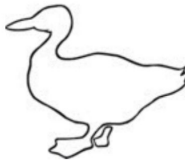
I **saw** her **duck** with a telescope.

verb

noun



[<http://casablancapa.blogspot.com/2010/05/fors.htm>]



[<http://www.supercoloring.com/pages/duck-outline/>]

A sentence is **ambiguous** if there is more than one way to interpret it.

We call these interpretations **readings**

How many readings are there of this sentence? What are they?



# What makes language "hard"?

## What does this sentence mean?



[<http://www.clker.com/clipart-green-eyes-3.html>]



[<http://www.clker.com/clipart-3163.html>]

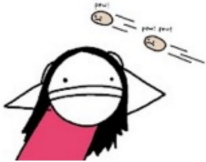


[<http://www.simonpalfrader.com/category/tournament-poker>]

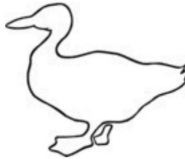
I **saw** her **duck** with a **telescope**.

verb

noun



[<http://cassblancap.blogspot.com/2010/05/fore.htm>]



[<http://www.supercoloring.com/pages/duck-outline/>]

- polysemy = 1 word, multiple meanings
- prepositional phrase attachment

# Crash Blossoms

McDonald's Fries the Holy Grail for Potato Farmers



Where does the ambiguity here come from?

# Crash Blossoms

McDonald's Fries the Holy Grail for Potato Farmers

verb

noun



# More Crash Blossoms

These **real newspaper headlines** – gathered from local, national, and international newspapers across the globe – are **ambiguous**; you can see what the journalist meant to say, but in each case there is a more amusing interpretation of the headline.

**EYE DROPS OFF SHELF**

**PROSTITUTES APPEAL TO POPE**

**KIDS MAKE NUTRITIOUS SNACKS**

**STOLEN PAINTING FOUND BY TREE**

**LUNG CANCER IN WOMEN MUSHROOMS**

**QUEEN MARY HAVING BOTTOM SCRAPED**

**DEALERS WILL HEAR CAR TALK AT NOON**

**MINERS REFUSE TO WORK AFTER DEATH**

**MILK DRINKERS ARE TURNING TO POWDER**

**DRUNK GETS NINE MONTHS IN VIOLIN CASE**

# Constituency

In syntactic analysis, a **constituent** is a word or a group of words that function as a single unit within a hierarchical structure.

# Constituency

pancakes or bacon and eggs

# Constituency

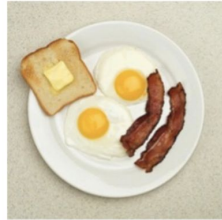
pancakes or bacon and eggs

could mean:

- a. pancakes or (bacon and eggs)



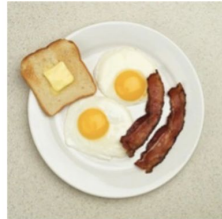
or



- b. (pancakes or bacon) and eggs



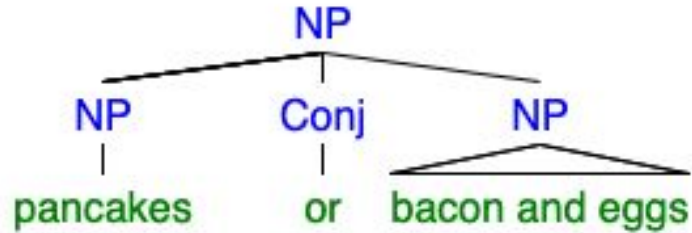
or



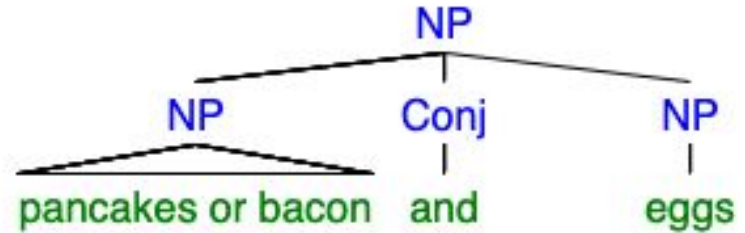


# One noun phrase, different structures

pancakes and bacon or eggs



a. pancakes or (bacon and eggs)



b. (pancakes or bacon) and eggs

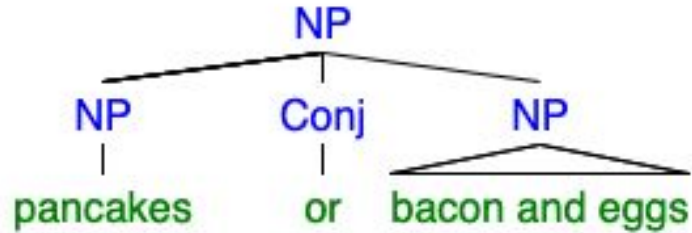


# One noun phrase, different structures

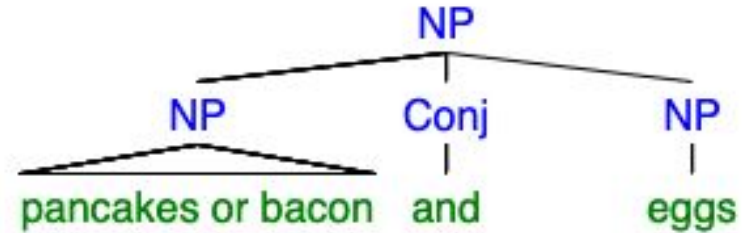
NP = "noun phrase"

Conj = "conjunction"

pancakes and bacon or eggs



a. pancakes or (bacon and eggs)



b. (pancakes or bacon) and eggs

# Transformational Grammar

The idea of the constituent is the basis for **transformational grammar** invented by Noam Chomsky in the 1950s

# Transformational Grammar

The idea of the constituent is the basis for **transformational grammar** invented by Noam Chomsky in the 1950s

## THREE MODELS FOR THE DESCRIPTION OF LANGUAGE\*

Noam Chomsky

Department of Modern Languages and Research Laboratory of Electronics  
Massachusetts Institute of Technology  
Cambridge, Massachusetts

### Abstract

We investigate several conceptions of linguistic structure to determine whether or not they can provide simple and "revealing" grammars that generate all of the sentences of English and only these. We find that no finite-state Markov process that produces symbols with transition from state to state can serve as an English grammar. Furthermore, the particular subclass of such processes that produce  $n$ -order statistical approximations to English do not come closer, with increasing  $n$ ,

observations, to show how they are interrelated, and to predict an indefinite number of new phenomena. A mathematical theory has the additional property that predictions follow rigorously from the body of theory. Similarly, a grammar is based on a finite number of observed sentences (the linguist's corpus) and it "projects" this set to an infinite set of grammatical sentences by establishing general "laws" (grammatical rules) framed in terms of such hypothetical constructs as the particular phenomena, words, phrases, and so on, of the

# Kinds of ambiguity

- structural ambiguity
  - e.g., what does
- referential ambiguity
  - co-reference: what do pronouns **refer back** to?
    - e.g. Corrine told Sarah to give Flannery their favorite syntax textbook.
  - metaphor, metonymy, and figurative language
    - e.g. The White House is patrolling the streets of Washington D.C.
- semantic or lexical ambiguity
  - e.g. They went down to the bank.
- "Situated truth" (a.k.a indexicality or *deixis*)
  - e.g. I live in Texas
  - e.g. Tomorrow is a holiday.

# Referential Ambiguity: Winograd Schemas

Consider the following example from Terry Winograd (1972). Who does "they refer to"?

- (a) "The city councilmen refused the demonstrators a permit because **they** feared violence."
- (b) "The city councilmen refused the demonstrators a permit because **they** wanted violence."

# Terry Winograd & SHRDLU

COGNITIVE PSYCHOLOGY 3, 1–191 (1972)

## Understanding Natural Language

TERRY WINOGRAD<sup>1</sup>

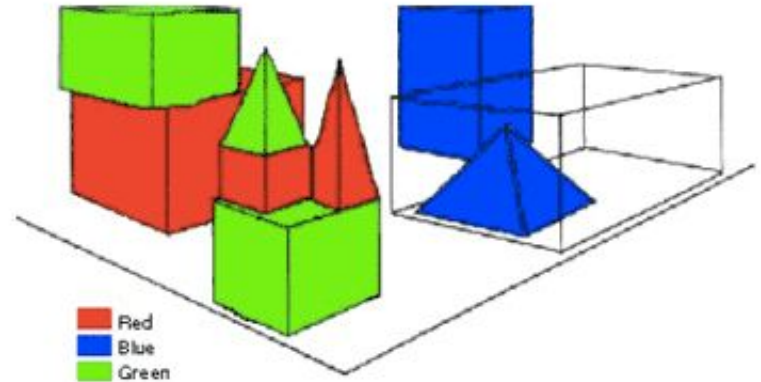
*Massachusetts Institute of Technology  
Cambridge, Massachusetts*

This paper describes a computer system for understanding English. The system answers questions, executes commands, and accepts information in an interactive English dialog.

It is based on the belief that in modeling language understanding, we

# Terry Winograd & SHRDLU

COGNITIVE PSYCHOLOGY 3, 1-191 (1972)



## Understanding Natural

TERRY WINOGRAD

*Massachusetts Institute of Technology  
Cambridge, Massachusetts*

**Person:** Pick up a big red block.

**Computer:** OK.

**Person:** Grasp the pyramid.

**Computer:** I don't understand which pyramid you mean.

This paper describes a computer system for understanding English. The system answers questions, executes commands, and accepts information in an interactive English dialog.

It is based on the belief that in modeling language understanding, we

# WINOGRANDE: An Adversarial Winograd Schema Challenge at Scale

Keisuke Sakaguchi\*, Ronan Le Bras\*, Chandra Bhagavatula\*, Yejin Choi\*<sup>†</sup>

\*Allen Institute for Artificial Intelligence <sup>†</sup>University of Washington  
{keisukes, ronanlb, chandrab, yejinc}@allenai.org

## Abstract

The Winograd Schema Challenge (WSC) (Levesque, Davis, and Morgenstern 2011), a benchmark for commonsense reasoning, is a set of 273 expert-crafted pronoun resolution problems originally designed to be unsolvable for statistical models that rely on selectional preferences or word associations. However, recent advances in neural language models have already reached around 90% accuracy on variants of WSC. This raises an important question whether these models have truly acquired robust commonsense capabilities or whether they rely on spurious biases in the datasets that lead to an overestimation of the true capabilities of machine commonsense.

To investigate this question, we introduce **WINOGRANDE**, a large-scale dataset of 441 machine-generated Winograd

rely on statistical patterns without true capabilities of commonsense reasoning. However, recent advances in neural language models have already reported around 90% accuracy on a variant of WSC dataset.<sup>1</sup> This raises an important question:

*Have neural language models successfully acquired commonsense or are we overestimating the true capabilities of machine commonsense?*

This question about the potential overestimation leads to another crucial question regarding potential unwanted biases that the large-scale neural language models might be exploiting, essentially solving the problems *right*, but for *wrong* reasons. While WSC questions are expert-crafted, recent studies



# Winogrande

---

## Twin sentences

---

✗

The monkey loved to play with the balls but ignored the blocks because he found them exciting.

The monkey loved to play with the balls but ignored the blocks because he found them dull.

✗

William could only climb begginner walls while Jason climbed advanced ones because he was very weak.

William could only climb begginner walls while Jason climbed advanced ones because he was very strong.

✓

Robert woke up at 9:00am while Samuel woke up at 6:00am, so he had less time to get ready for school.

Robert woke up at 9:00am while Samuel woke up at 6:00am, so he had more time to get ready for school.

✓

The child was screaming after the baby bottle and toy fell. Since the child was hungry, it stopped his crying.

The child was screaming after the baby bottle and toy fell. Since the child was full, it stopped his crying.

---

# Winogrande

	Twin sentences	Options (answer)
✗	The monkey loved to play with the balls but ignored the blocks because he found them exciting. The monkey loved to play with the balls but ignored the blocks because he found them dull.	balls / blocks balls / <b>blocks</b>
✗	William could only climb begginer walls while Jason climbed advanced ones because he was very weak. William could only climb begginer walls while Jason climbed advanced ones because he was very strong.	<b>William</b> / Jason William / <b>Jason</b>
✓	Robert woke up at 9:00am while Samuel woke up at 6:00am, so he had less time to get ready for school. Robert woke up at 9:00am while Samuel woke up at 6:00am, so he had more time to get ready for school.	<b>Robert</b> / Samuel Robert / <b>Samuel</b>
✓	The child was screaming after the baby bottle and toy fell. Since the child was hungry, it stopped his crying. The child was screaming after the baby bottle and toy fell. Since the child was full, it stopped his crying.	<b>baby bottle</b> / toy baby bottle / <b>toy</b>

# Winogrande

Twin sentences		Options (answer)
✗	The monkey loved to play with the balls but ignored the blocks because he found <b>them</b> <u>exciting</u> . The monkey loved to play with the balls but ignored the blocks because he found <b>them</b> <u>dull</u> .	<b>balls</b> / blocks balls / <b>blocks</b>
✗	William could only climb beginner walls while Jason climbed advanced ones because <b>he</b> was very <u>weak</u> . William could only climb beginner walls while Jason climbed advanced ones because <b>he</b> was very <u>strong</u> .	<b>William</b> / Jason William / <b>Jason</b>
✓	Robert woke up at 9:00am while Samuel woke up at 6:00am, so <b>he</b> had <u>less</u> time to get ready for school. Robert woke up at 9:00am while Samuel woke up at 6:00am, so <b>he</b> had <u>more</u> time to get ready for school.	<b>Robert</b> / Samuel Robert / <b>Samuel</b>
✓	The child was screaming after the baby bottle and toy fell. Since the child was <u>hungry</u> , <b>it</b> stopped his crying. The child was screaming after the baby bottle and toy fell. Since the child was <u>full</u> , <b>it</b> stopped his crying.	<b>baby bottle</b> / toy baby bottle / <b>toy</b>

What makes language hard (for language models) now?

# What makes language hard (for language models) now?

- social cues
- common sense
- truth / world knowledge
- *action*
- ...?