Informatics 1 Cognitive Science

Lecture 3: Introduction to Language

Frank Keller

17 January 2025

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

Slide credits: Frank Mollica, Chris Lucas, Mirella Lapata

Overview

Words

Rules

The Anatomy of Language

Words

What Are Words?

Pairing between a form and a meaning (arbitrary and memorized).

form:	*rõz*
meaning:	
part of speech:	N

- A language community tacitly agrees to use a particular form to convey a particular idea.
- The word *rose* does not smell sweet or have thorns, but we can use it to convey the idea of a rose.

What Are Words?

Pairing between a form and a meaning (arbitrary and memorized).

form:	*rõz*
meaning:	
part of speech:	N

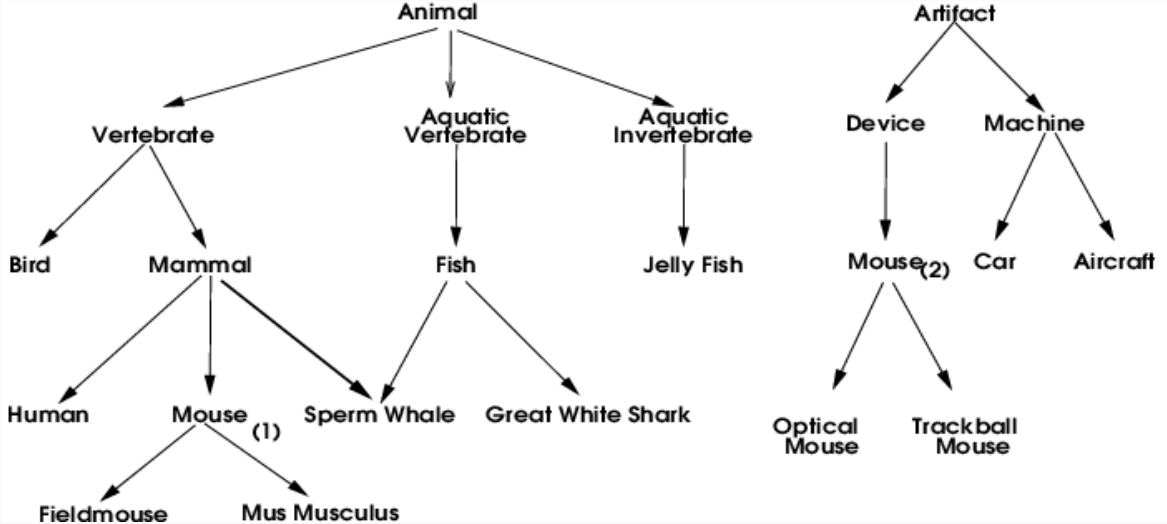
- A language community tacitly agrees to use a particular form to convey a particular idea.
- The word *rose* does not smell sweet or have thorns, but we can use it to convey the idea of a rose.
- Onomatopoeia (e.g., *oink*, *crash*) and sound symbolism (e.g., *mellifluous*, *cantankerous*) do exist, but won't get you far.

Words: The Mental Lexicon

Ferdinand de Saussure

- Arbitrary sign: conventional pairing of meaning and sound.
- The mental lexicon is the set of all the words in the language.
- Often assumed to form a hierarchy or a network.
- Speakers of the same language have mutually intelligible lexicon entries.

Example: WordNet Hierarchy



Words: The Mental Lexicon

Amazing facts about how the mental lexicon is acquired and used:

Words: The Mental Lexicon

Amazing facts about how the mental lexicon is acquired and used:

- Between birth and adulthood, children learn about 40,000 unique words and idioms (*fit as a fiddle*).

Words: The Mental Lexicon

Amazing facts about how the mental lexicon is acquired and used:

- Between birth and adulthood, children learn about 40,000 unique words and idioms (*fit as a fiddle*).
- During the second postnatal year, children's productive vocabulary accelerates dramatically: vocabulary explosion.

Words: The Mental Lexicon

Amazing facts about how the mental lexicon is acquired and used:

- Between birth and adulthood, children learn about 40,000 unique words and idioms (*fit as a fiddle*).
- During the second postnatal year, children's productive vocabulary accelerates dramatically: vocabulary explosion.
- People recognize and produce words extremely quickly: meaning of spoken word is accessed by listener in 0.2 seconds.

Words: The Mental Lexicon

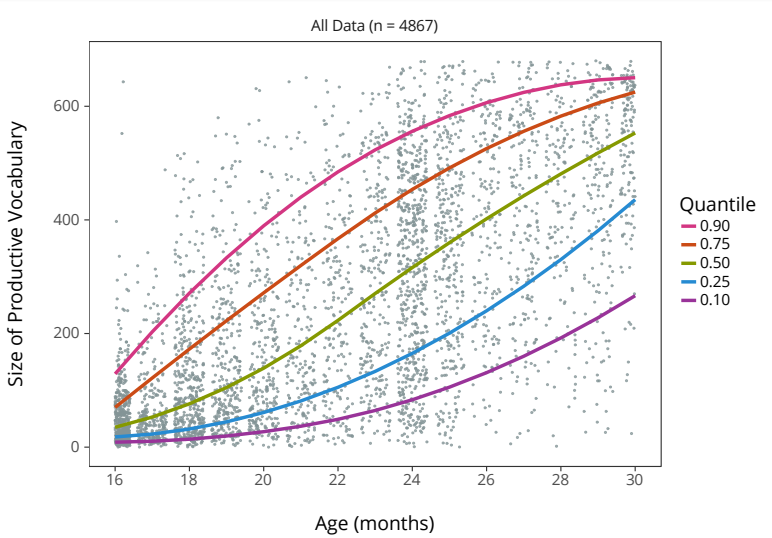
Amazing facts about how the mental lexicon is acquired and used:

- Between birth and adulthood, children learn about 40,000 unique words and idioms (*fit as a fiddle*).
- During the second postnatal year, children's productive vocabulary accelerates dramatically: vocabulary explosion.
- People recognize and produce words extremely quickly: meaning of spoken word is accessed by listener in 0.2 seconds.
- The brain takes 0.25 seconds to name an object, and further 0.25 seconds to program mouth and tongue to pronounce it.

Word Learning

Growth curve from the MacArthur-Bates Communicative Development Inventory (MCDI):

http://wordbank.stanford.edu/



Rules

Rules

- We do not just blurt out isolated words.
- Rather we combine them into phrases and sentences.
- The meaning of the combination can be inferred from the meanings of the words and the way they are arranged.

Rules

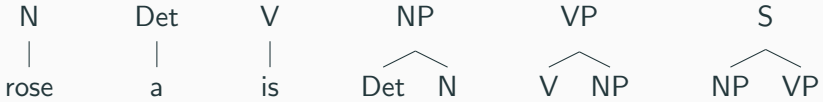
- We do not just blurt out isolated words.
- Rather we combine them into phrases and sentences.
- The meaning of the combination can be inferred from the meanings of the words and the way they are arranged.

Do these mean the same?

1. The boy saw the girl with the telescope
2. The girl saw the telescope with the boy
3. The boy with the telescope saw the girl

Rules

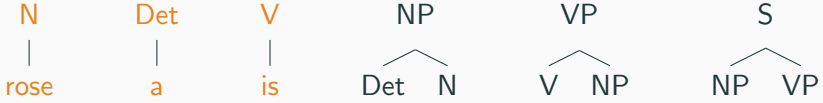
There must be a code, a set of rules that specifies how words may be arranged into meaningful combinations: the grammar.



a rose is a rose

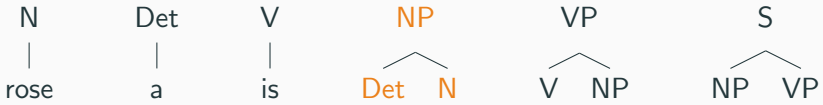
Rules

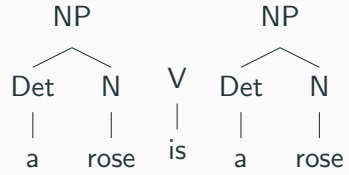
There must be a code, a set of rules that specifies how words may be arranged into meaningful combinations: the grammar.



Rules

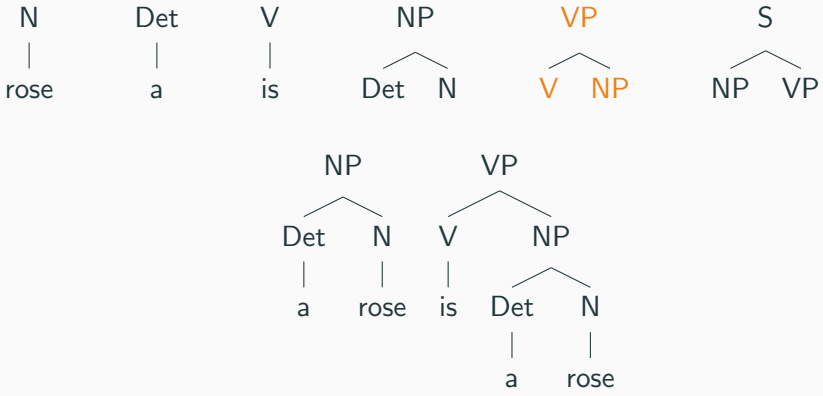
There must be a code, a set of rules that specifies how words may be arranged into meaningful combinations: the grammar.





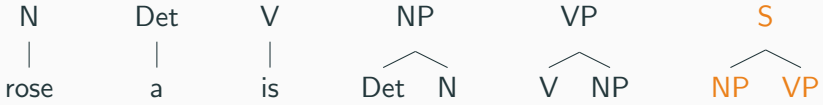
Rules

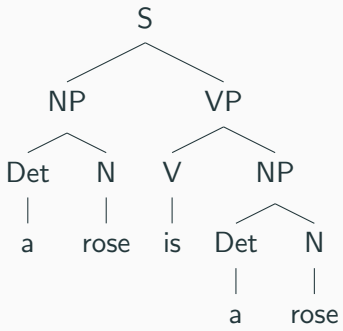
There must be a code, a set of rules that specifies how words may be arranged into meaningful combinations: the grammar.



Rules

There must be a code, a set of rules that specifies how words may be arranged into meaningful combinations: the grammar.





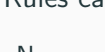
The Expressive Power of Rules

Time for a short quiz on Wooclap!

https://app.wooclap.com/GDCILA

Notation

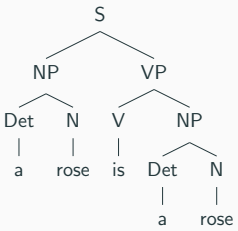
Rules can be written as trees of depth one or with arrow notation:



is the same as $N \rightarrow rose$

is the same as $S \rightarrow NP\ VP$

Sentence structures can be written as trees or with labelled brackets:



is the same as [S [NP [Det a] [N rose]] [VP [V is] [NP [Det a] [N rose]]]]

The Expressive Power of Rules

- Rules are productive, defined over kinds of words rather than actual words (we assemble new sentences on the fly).
- Symbols contained in the rules are abstract (we can talk about anything we like!)
- The rules are also combinatorial: a small inventory of elements can be assembled by rules into immense set of distinct objects.

The Expressive Power of Rules

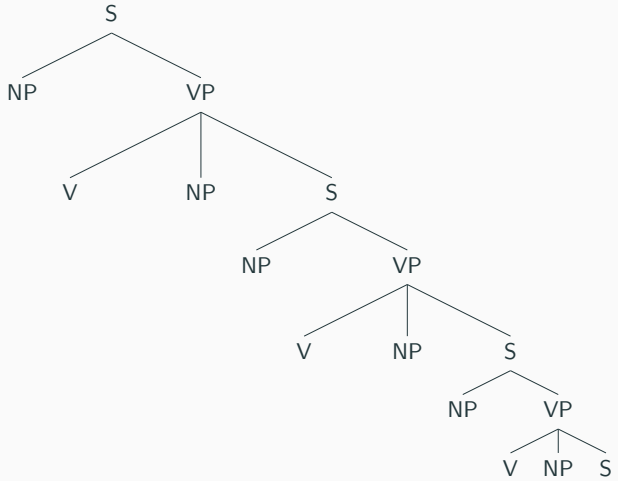
- Rules are productive, defined over kinds of words rather than actual words (we assemble new sentences on the fly).
- Symbols contained in the rules are abstract (we can talk about anything we like!)
- The rules are also combinatorial: a small inventory of elements can be assembled by rules into immense set of distinct objects.

Example:

- Det $\in$ {a, any, one, the}, N $\in$ 10,000 nouns, V $\in$ 4,000 verbs
- NP $\rightarrow$ Det N allows $4 \times 10{,}000 = 40{,}000$ NPs
- VP $\rightarrow$ V NP allows $40{,}000 \times 4{,}000 = 160{,}000{,}000$ VPs
- S $\rightarrow$ NP VP allows $160{,}000{,}000 \times 40{,}000 = 6.4$ trillion Ss

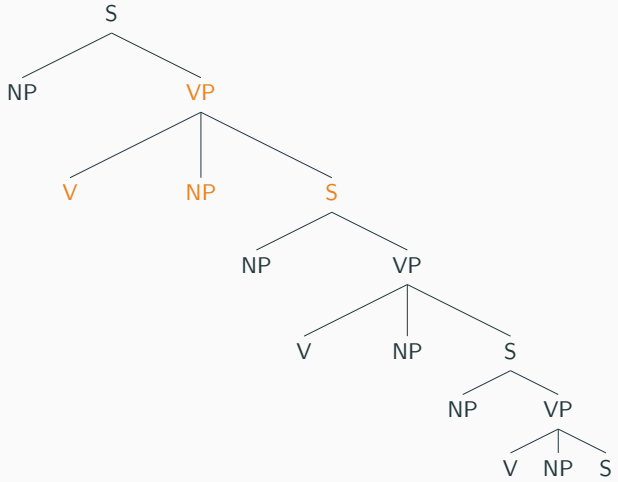
The Expressive Power of Rules

Natural languages exhibit recursion: the rules create an entity that can contain an example of itself. For example the rule VP $\rightarrow$ V NP S:



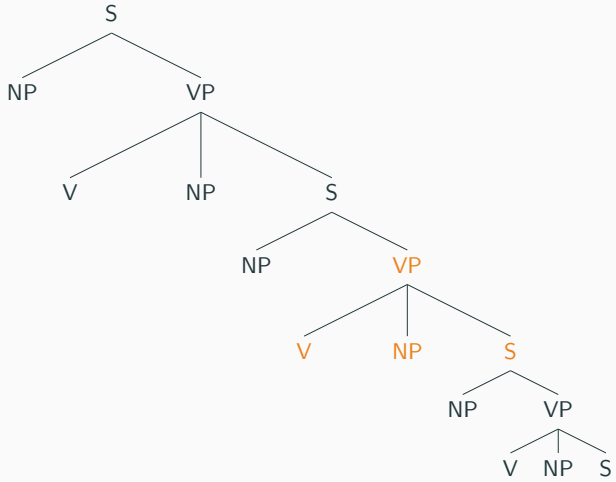
The Expressive Power of Rules

Natural languages exhibit recursion: the rules create an entity that can contain an example of itself. For example the rule VP $\rightarrow$ V NP S:



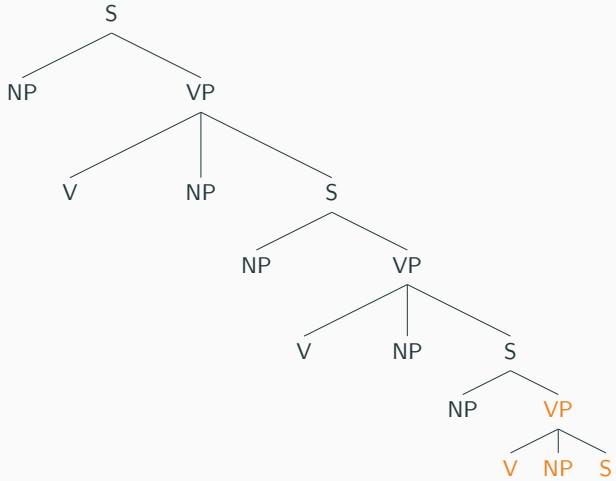
The Expressive Power of Rules

Natural languages exhibit recursion: the rules create an entity that can contain an example of itself. For example the rule VP $\rightarrow$ V NP S:



The Expressive Power of Rules

Natural languages exhibit recursion: the rules create an entity that can contain an example of itself. For example the rule VP $\rightarrow$ V NP S:



The Expressive Power of Rules

Natural languages exhibit recursion: the rules create an entity that can contain an example of itself.

I think I'll tell you that I just read a news story that recounts that Stephen Brill reported that the press uncritically believed Kenneth Starr's announcement that Linda Tripp testified to him that Monica Lewinsky told Tripp that Bill Clinton told Vernon Jordan to advise Lewinsky not to testify to Starr that she had had a sexual relationship with Clinton.

How many sentences are there?

The Expressive Power of Rules

Natural languages exhibit recursion: the rules create an entity that can contain an example of itself.

I think I'll tell you that I just read a news story that recounts that Stephen Brill reported that the press uncritically believed Kenneth Starr's announcement that Linda Tripp testified to him that Monica Lewinsky told Tripp that Bill Clinton told Vernon Jordan to advise Lewinsky not to testify to Starr that she had had a sexual relationship with Clinton.

How many sentences are there? Twelve!

Potential infiniteness of the language faculty has been recognized by Galileo, Descartes, Humboldt. There is no longest sentence!

The Expressive Power of Rules

Time for a short quiz on Wooclap!

https://app.wooclap.com/GDCILA

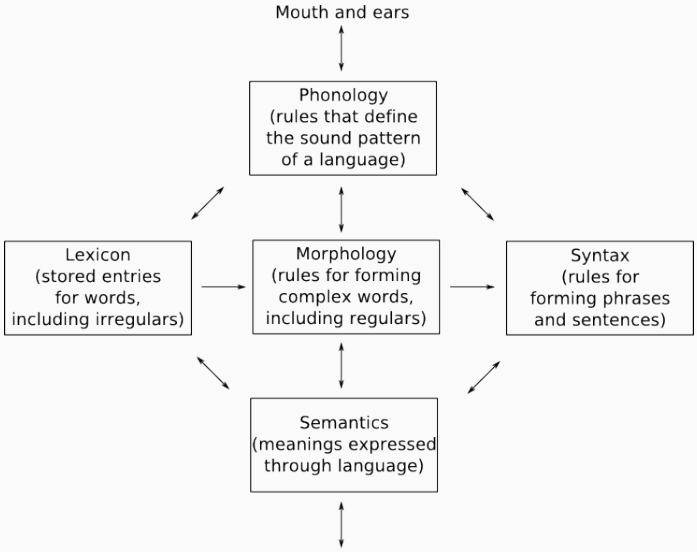
The Anatomy of Language

The Conventional Wisdom: Words plus Rules

- Human language appears to have unlimited expressive power.
- We can be led to think thoughts that have never been thought before, and that never would have occurred to us on our own.
- We need more than just individual words to manage this!
- Words combine to produce meaningful utterances.
- We describe combinations in terms of rules.
- So it makes sense to assume that the word/rule duality corresponds to two distinct cognitive mechanisms:

words	$\Longleftrightarrow$	memory
rules	$\Longleftrightarrow$	computation

The Anatomy of Language



Two Kinds of Words

Linguistics:	word-as-morphological object, as opposed to phrases and sentences
Psychology:	word-as-lexical entry, stretch of sound which has been memorized and cannot be produced by a rule Pinker calls this a listeme

- Some memorized chunks are smaller than a word in first sense.
- Others are larger than a word in the first sense.
- The second sense of word covers things which appear only as parts of words in the first sense.

Pinker's Listemes

Suffixes

-able (*cap-able*)
-ed (*wak-ed*)
-al (*refus-al*)
-ship (*fellow-ship*)

Prefixes

un- (*un-finished*)
ante- (*ante-cedent*)
co- (*co-pilot*)
re- (*re-claim*)

Pinker's Listemes

Suffixes

-able (*cap-able*)
-ed (*wak-ed*)
-al (*refus-al*)
-ship (*fellow-ship*)

Prefixes

un- (*un-finished*)
ante- (*ante-cedent*)
co- (*co-pilot*)
re- (*re-claim*)

Idioms

piece of cake
when pigs fly
like two peas in a pod
beat around the bush
see eye to eye
once in a blue moon
the last straw
the best of both worlds
costs an arm and a leg
add insult to injury

Summary

The characteristics of language:

- It consists of words, arbitrary pairings of a sound and meaning.
- Between birth and adulthood, children learn about 40,000 words.
- The ordering of words is governed by a set of rules.
- Rules are combinatorial: they combine words to generate an unlimited number of sentences.
- Cognitively, words correspond to memory, and rules correspond to computation.
- Listemes are stored in memory: words, prefixes and suffixes, idioms.