

Language: Introduction Informatics 1 CG: Lecture 2

Mirella Lapata

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

January 14, 2016

1 / 24

- Inf1CG runs weeks 1–10, with 30 slots: 29 lectures, no lectures in innovative learning week.
- <http://www.inf.ed.ac.uk/teaching/courses/inf1-cg/>
- contains contact details, time/place of lectures, tutorials, labs, software, schedule of assessments, and reading list; slides and assignments will appear there;
- course mailing list: inf1-cg-students@inf.ed.ac.uk will be used for course announcements; if you are registered for the course, you will be automatically subscribed;
- assignments: 1 and 2 will use MATLAB, 3 will be an essay;
- we will use the Piazza Q&A platform to discuss course materials and assignments;
- **You need a DICE account!** If you don't have one, apply for one through the ITO as soon as possible.

2 / 24

The Faculty of Language: What's Special about it?

What kind of biological system is language and how does it relate to other systems in our own species and other?

- ① Which aspects are learned from **environmental input** and which ones arise from the **innate design** of the brain?
- ② Which parts of a person's language ability are **specific to language** and which ones belong to more **general abilities**?
- ③ Which aspects of the language capacity are **uniquely human** and which ones **shared** with other groups of animals?

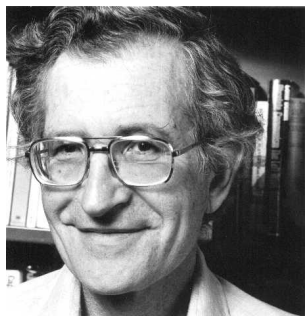
Reading:

Steven Pinker's, Words and Rules, Chapter 1

3 / 24

4 / 24

The Faculty of Language: What's Special about it?




A human language is a system of **remarkable complexity**. To come to know a human language would be an extraordinary intellectual achievement for a creature not **specifically designed** to accomplish this task. A normal child acquires this knowledge on relatively slight exposure and without specific training. He can then quite effortlessly make use of an **intricate structure of specific rules** and guiding principles to convey his thoughts and feelings to others, arousing in them novel ideas and subtle perceptions and judgments (Chomsky, 1975, p. 4).

5 / 24

What Are Words?

Pairing between a sound and a meaning (**arbitrary** and **memorized**).

sound:	rōz
meaning:	
part of speech:	N

- A language community tacitly agrees to use a particular sound 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*), sound symbolism (e.g., *mellifluous*, *cantankerous*) won't get you far.

6 / 24

Words: The Mental Lexicon



Ferdinand de Saussure

- The mental lexicon consists of a finite set of arbitrary words.
- **Arbitrary Sign**: conventional pairing of meaning and sound
- Speakers of the same language store identical entries in their mental lexicons.

- By year 2, children are learning 10 or 15 new words every day.
- A typical high school leaver knows 60,000 words.
- 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.

7 / 24

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.

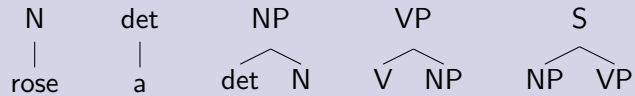
Do these mean the same?

1. Young women looking for husbands
2. Husbands looking for young women
3. Looking women husbands young for

8 / 24

Rules

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

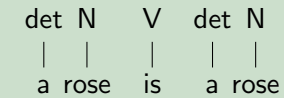
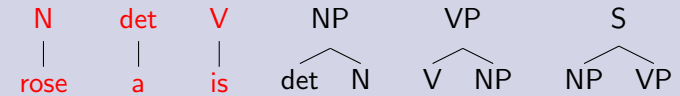


a rose is a rose

9 / 24

Rules

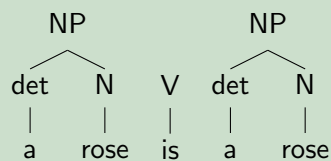
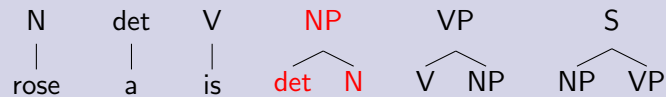
There must be a code, set of rules that specifies how words may be arranged into meaningful combinations, aka **a grammar**.



10 / 24

Rules

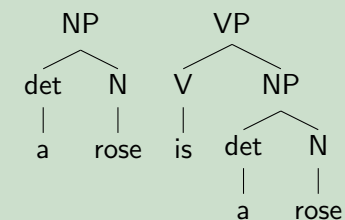
There must be a code, set of rules that specifies how words may be arranged into meaningful combinations, aka **a grammar**.



11 / 24

Rules

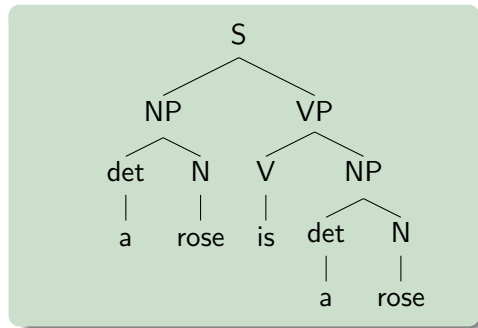
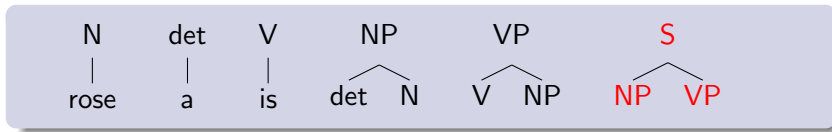
There must be a code, set of rules that specifies how words may be arranged into meaningful combinations, aka **a grammar**.



12 / 24

Rules

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



13 / 24

The Expressive Power of Rules

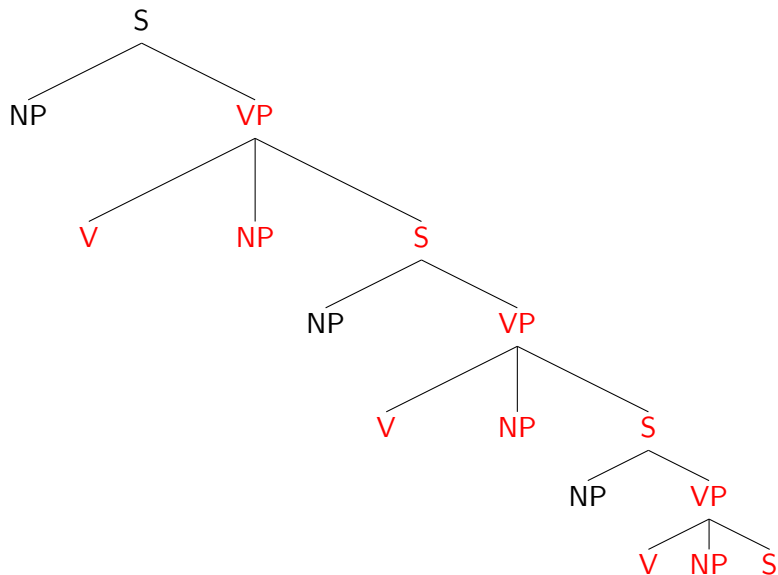
- Rules are **productive**, defined over **kinds** of words rather than **actual** words (we assemble new sentences on the fly).
- Symbols contained in the rules are **abstract** (we can talk about anything we like!)
- The rules are also **combinatorial**: a small inventory of elements can be assembled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

- Det $\in \{a, any, one, the\}$, N $\in 10,000$, V $\in 4,000$ verbs
- NP $\rightarrow det\ N$ allows $4 \times 10,000 = 40,000$ NPs
- VP $\rightarrow V\ NP$ allows $40,000 \times 4,000 = 160,000$ VPs
- S $\rightarrow NP\ V\ NP$ allows $160,000,000 \times 40,000 = 6G\ 400M$ Ss

14 / 24

The Expressive Power of Rules

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



15 / 24

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?

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

16 / 24

Rules: Their Expressive Power

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, von Humboldt. **There is no longest sentence!**

17 / 24

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**.
- And we're assuming for the time being that the word/rule duality corresponds to two distinct cognitive mechanisms:

words	↔	memory
rules	↔	computation

18 / 24

Critical Period

There is a **critical period** in development during which a language can be acquired like a native speaker.

- child vs. adult language learning
- native vs. non-native speakers
- age of immigration and language ability
- arrive before age 6 → generally pass as native speakers
- arrive after puberty → generally do not pass as native speakers
- Evidence: “feral children”, “wolf children”, “attic children”
- Wild boy of Aveyron, *L'enfant sauvage* by François Truffaut

19 / 24

Critical Period: The case of Genie

- Was tied to her potty chair for 13 years by her dad
- Only words she knew: *stop it* and *no more*
- **Could she manage to learn language?**

<https://sites.google.com/site/comperecriticalperiod/critical-period/the-story-of-genie>

- Genie learned many words but not syntax!
- Amount of language learned after critical period seems limited.
- Difficult to disentangle linguistic from other forms of deprivation.

20 / 24

Do Primates Use Language?

- Animals in the wild **communicate** with one another
- Monkeys use cries to warn each other of impending danger. (different cries for different dangers).
<http://www.bbc.co.uk/programmes/p016dgw1>
- The cries are not arbitrary but there is no use of **grammar** to arrange them into sentences or to create **new meanings**.
- Communication \neq language.

If animals don't use language naturally can we teach it to them?

21 / 24

Do Primates Use Language?



Nim Chimpsky signs "black"



Nim Chimpsky signs "I see"

- Primates comprehend meanings of limited number of symbols.
- They understand very **little in the way of syntax**, production.
- They **repeat the sentences** they were taught to produce.
- They do not possess human **generative capability**
- They do not **teach their language skills** to other members.

Nim Chimpsky (November 19, 1973–March 10, 2000) was the subject of an extended study of animal language acquisition led by Herbert S. Terrace.

22 / 24

Project Nim



<https://www.youtube.com/watch?v=yxQap9AAP0s>

23 / 24

Summary

The characteristics of language:

- It allows for **communication** between individuals.
- It consists of words, **arbitrary** pairings of a sound and meaning.
- The ordering of words is governed by a set of **rules**
- Words combine to **generate** a very large number of meanings.
- It is **dynamic** constantly changing as new words are added.
- There seems to be a **critical period** for learning language.
- Language \neq communication.

Next lecture: we will look more into Words & Rules.

24 / 24