https://github.com/edupoux/MVA_2023_SL/

Speech and Natural Language Processing

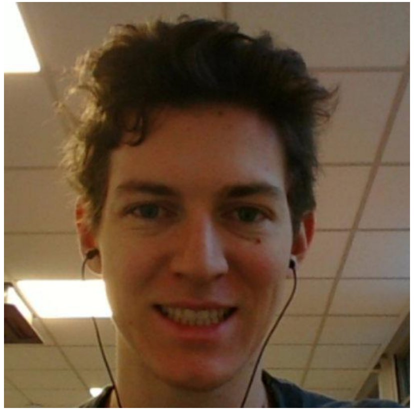
#1. Introduction

Emmanuel Dupoux & Benoît Sagot

Co-starring

Neil Zeghidour, Research Scientist,
Google Brain

Robin Algayres, PhD student,
Ecole Normale Supérieure



Holger Schwenk, Research Scientist,
Facebook AI Research

Speech processing and NLP
Overview and applications

Natural Language Processing

- *Edited text*
- *User-Generated Content*
- *Transcriptions*

Natural Language Processing (NLP)

Core tasks:
- Part-of-speech tagging
- (Syntactic) parsing
- Semantic parsing
- etc.

Analysis

Information extraction/retrieval

Sentiment analysis

Data mining

Generation

Dialogue generation (chatbots)

Image captioning

Computational creativity

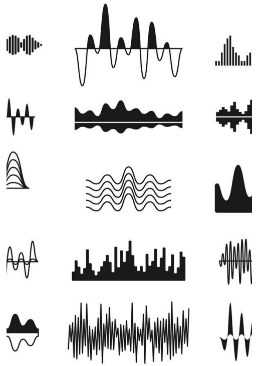
Transformation

Spelling correction

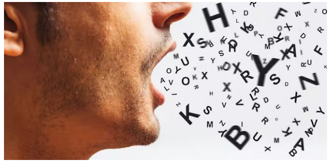
Machine translation

Automatic summarisation

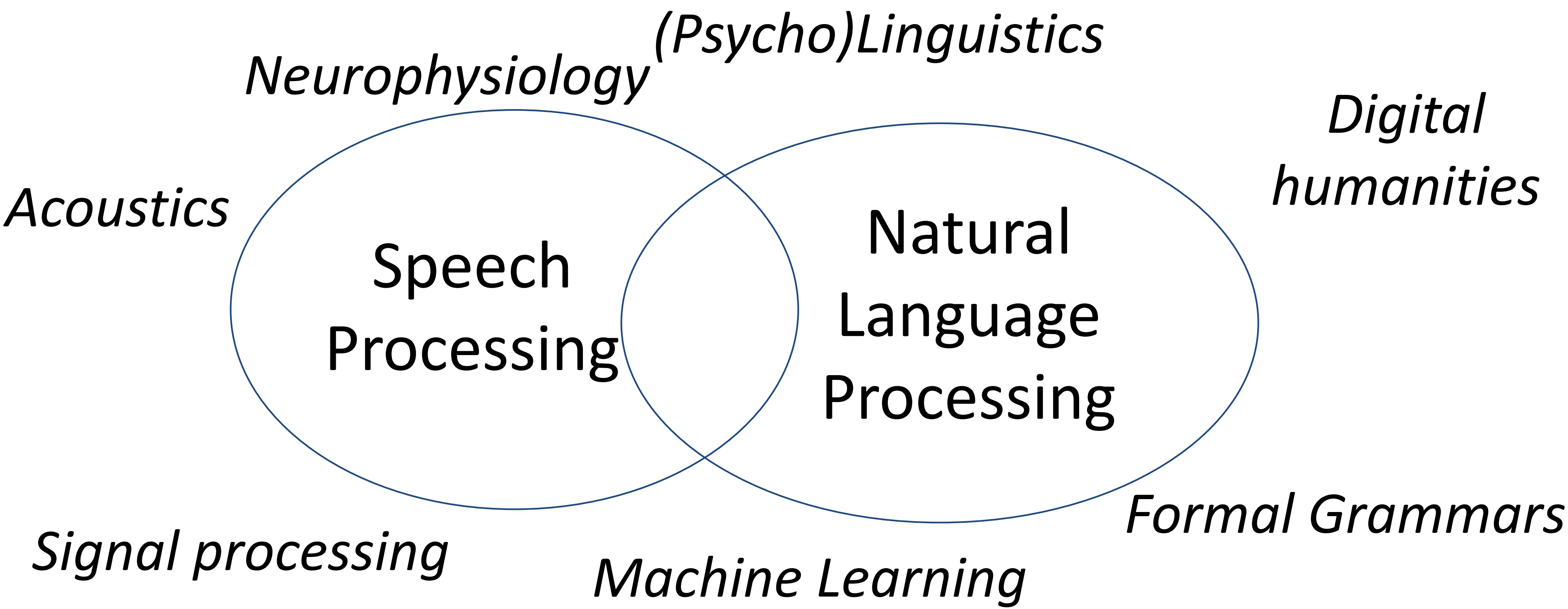
Speech Processing



- Signal Processing (analysis of audio signals)
- Phonetics (perception & production)
- Automatic Speech Recognition
- Speech Synthesis
- Speaker, Emotion and Language Identification
- Spoken Dialogue Systems



Multi-disciplinary fields



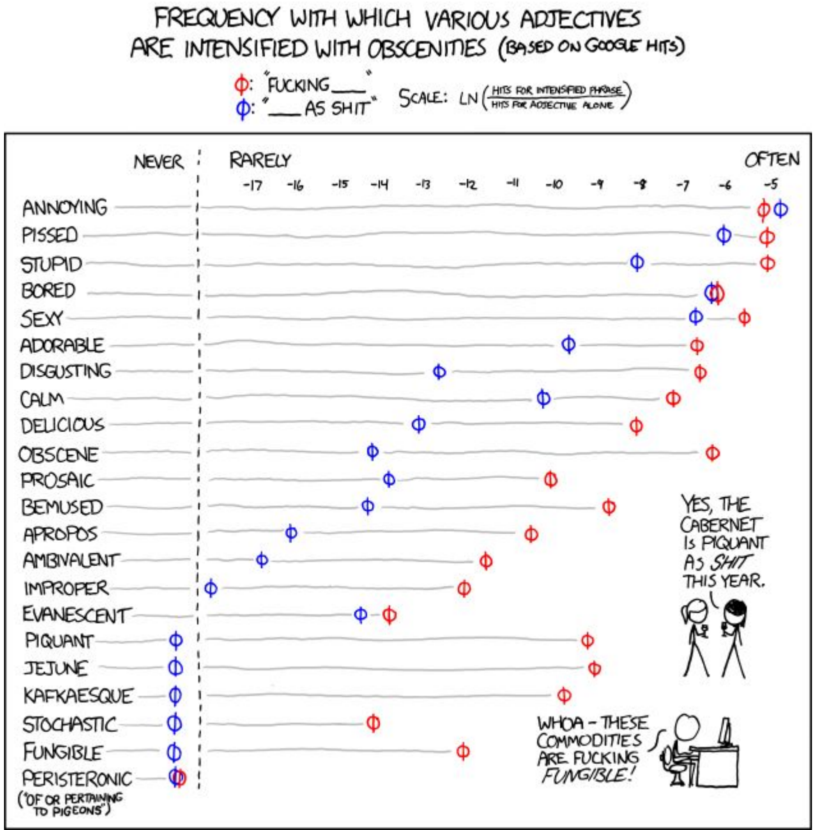
Academic applications

Academic applications: comp. linguistics

Examples

- Language modelling (synchronic, diachronic), with a number of approaches: formal, corpus-based, simulation-based, psycholinguistics, neurolinguistics

- Sociolinguistics

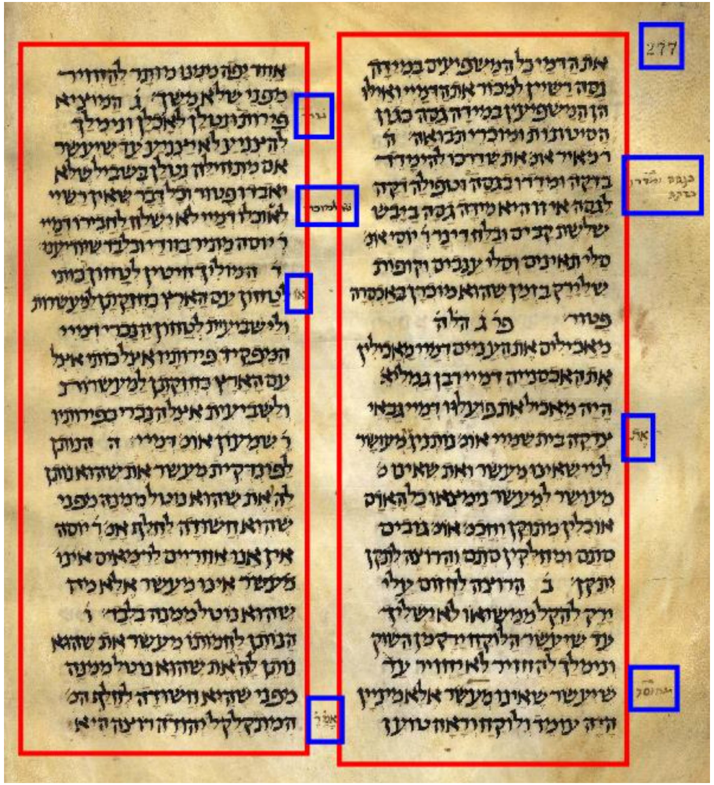


xkcd.com

Academic applications: digital humanities

Examples

- Exploitation of textual data for research in other domains (history, philology…)
- Computational epigraphy
(in collaboration with image processing specialists)



Academic applications: audio corpus analysis

- Analysis of infant-directed speech
- Analysis of pathological speech
- Analysis of animal vocalization

Industrial applications

Applications of NLP

- Information extraction, information retrieval, question answering
- Text mining (ex.: opinion surveys)
- Text generation, text simplification, automatic summarisation
- Spelling correction (writing aid, post-OCR, normalisation of noisy/non-canonical texts)
- Machine translation, computer-aided translation
- Chatbots, conversational agents, question answering systems
- Medical applications (early diagnosis, language-based medical monitoring…)

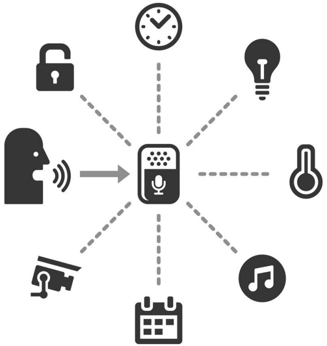
Applications of speech processing

- Speech compression, enhancement
- Cochlear implants, Speech Aids
- Forensics, Biometrics, Diagnostics
- Intelligence, Information Retrieval
- Dictation systems, Automatic Captioning
- Chatbots, call center automation
- Smart Home/Car devices (Echo, Alexa, Siri)









Advanced applications: Chatting with a computer

Dialogue systems in films

Dave: Open the pod bay doors, HAL.
HAL: I'm sorry Dave, I'm afraid I can't do that.

(2001: A Space Odyssey, 1968)

Theodore: Do you talk to someone else while we're talking?
Samantha: Yes.
Theodore: Are you talking with someone else right now? People, OS, whatever...
Samantha: Yeah.
Theodore: How many others?
Samantha: 8,316.
Theodore: Are you in love with anybody else?
Samantha: Why do you ask that?
Theodore: I do not know. Are you?
Samantha: I've been thinking about how to talk to you about this.
Theodore: How many others?
Samantha: 641.

(Her, 2013)

Dialogue systems in real life

- ELIZA program (Weizenbaum, 1966)
  - early natural language processing system capable of carrying on a limited form of conversation
  - makes use of pattern-matching to process the input and translate it into suitable outputs
  - **ELIZA doesn't need to *know* anything to mimic a psychotherapist**
- Loebner Prize competition, since 1991
  - An event has attempted to put various computer programs to the Turing test

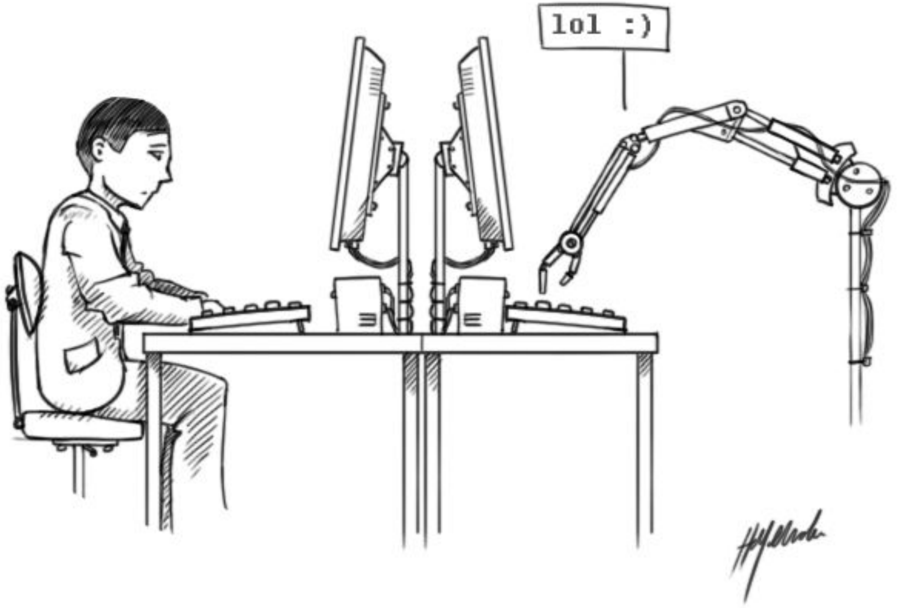
User1: You are like my father in some ways.
ELIZA1: WHAT RESEMBLANCE DO YOU SEE
User2: You are not very aggressive but I think you don't want me to notice that.
ELIZA2: WHAT MAKES YOU THINK I AM NOT AGGRESSIVE
User3: You don't argue with me.
ELIZA3: WHY DO YOU THINK I DON'T ARGUE WITH YOU
User4: You are afraid of me.
ELIZA4: DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU

Turing's test (1950)

To which extent can a machine simulate the linguistic behaviour of a human?

However:

- Not a test of intelligence (humans are not always intelligent, machine's intelligence is often not very human)
- Some behaviours are easy to reproduce
- Some behaviours can fool the human because of his/her attempts (spelling errors…)
- Results can be influenced by the human communicating with the machine



Today's chatbots and virtual assistants

A new generation of chatbots

- M (Facebook), ~~Tay~~ and Zo (Microsoft), Siri (Apple), Alexa (Amazon)…

Major limits:

- The machine learning component can be fooled/perverted
- Level of understanding and appropriateness of reaction
- Language variety is a major issue

Today’s chatbots and virtual assistants

A new generation of chatbots

- M (Facebook), ~~Tay~~ and Zo (Microsoft), Siri (Apple), Alexa (Amazon)…

Major limits:

- The machine learning component can be fooled/perverted
- Level of understanding and appropriateness of reaction
- Language variety is a major issue

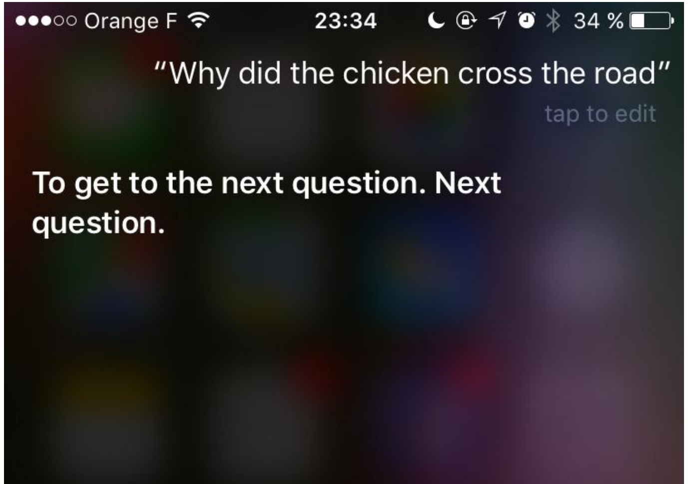
Today’s chatbots and virtual assistants

A new generation of chatbots

- M (Facebook), ~~Tay~~ et Zo (Microsoft), Siri (Apple), Alexa (Amazon)…

Major limits:

- The machine learning component can be fooled/perverted
- Level of understanding and appropriateness of reaction
- Language variety is a major issue



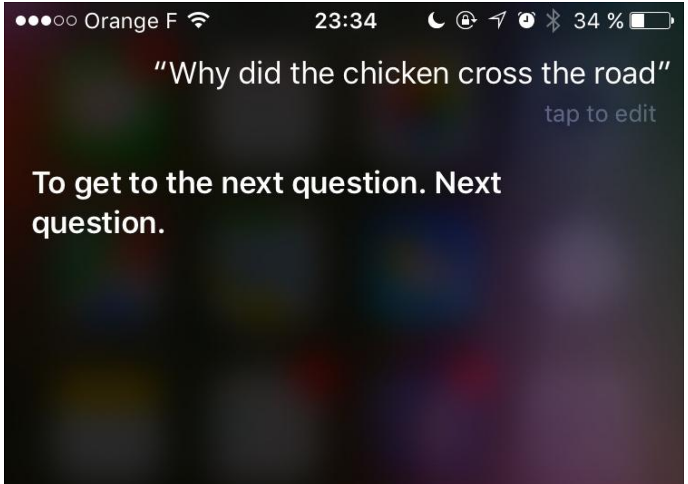
Today's chatbots and virtual assistants

A new generation of chatbots

- M (Facebook), ~~Tay~~ et Zo (Microsoft), Siri (Apple), Alexa (Amazon)…

Major limits:

- The machine learning component can be fooled/perverted
- Level of understanding and appropriateness of reaction
- Language variety is a major issue



Fast forward to ChatGPT

- Generative language model fine-tuned in a complex way (incl. Reinforcement learning) to behave in the way the developers want a conversational AI to behave
  - Produced by OpenAI, a Microsoft-supported company
  - Free… for how long? Really free? Are we not free data providers and free “white hats”?
- Multiple mechanisms to prevent ChatGPT from producing harmful, offensive, racist content
  - It still can produce such content
- Capable of behaviours with a high level of abstraction
  - E.g. it can behave, if it is asked to, as a UNIX terminal
- No guarantee of any kind on the quality/accuracy of its answers

Fast forward to ChatGPT

BE

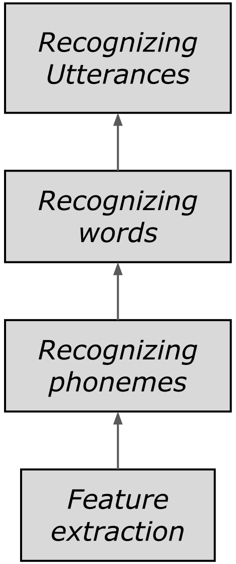
How is ChatGPT better than Siri?

ChatGPT is a large language model that has been trained on a vast amount of text data, whereas Siri is a virtual assistant developed by Apple that uses natural language processing to perform various tasks. ChatGPT has the ability to generate human-like text, whereas Siri is designed to understand and respond to specific voice commands and queries. Additionally, ChatGPT is a general-purpose model, meaning it can perform a wide range of language-related tasks, whereas Siri is designed for specific tasks such as setting reminders, making phone calls, and playing music.

Speech and language processing over time

Computational approaches to speech processing over time

Hand engineered system



1952 Audrey: single voice, single digits
1960: Shoebox: 16 words
1970: Harpy: 1000 words

Problems :
Labor intensive (procedures, hard coded heuristics)
Brittle (early errors catastrophic)

Computational approaches to speech processing over time

Hand engineered system

Probabilistic system

Recognizing Utterances

↑

Recognizing words

↑

Recognizing phonemes

↑

Feature extraction

Acoustic model

Language model

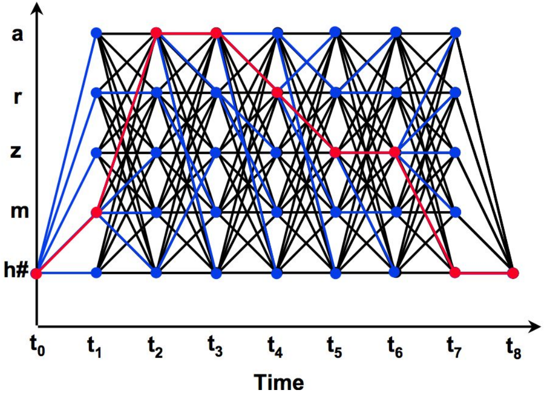
P(W|X)~P(X|W).P(W)

Decoding

$\mathbf{W} = \text{argmax}_W \, P_\theta(X|W).P(W)$

Learning

$\boldsymbol{\theta} = \text{argmax}_\theta \, P_\theta(X)$



<u>Advantage</u>: robust to error

<u>Problem</u>: lots of annotated data

<u>Algorithms</u>: State search space

Dynamic programming, EM

Computational approaches to speech processing over time

Hand engineered system

Probabilistic system

Deep learning systems

Recognizing Utterances

Recognizing words

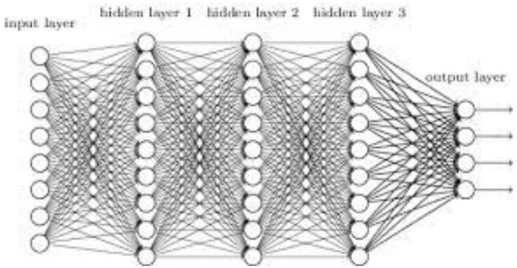
Recognizing phonemes

Feature extraction

Acoustic model

Langage model

$$P(W|X) \sim P(X|W).P(W)$$



Decoding:

Distributed intermediate representations

Probabilistic interpretation of final layer

Learning:

Stochastic Gradient Descent

<u>Advantage</u>: more robust, scalable

<u>Problem:</u> less interprtabe, needs even more (less) annotated data

Computational approaches to NLP over time

Symbolic approaches

- **Computational expertise:**

Formal grammars (algebraic grammars, mildly context-sensitive grammars, polynomial languages…), parsing algorithms, dynamic programming

- **Comp. linguistics expertise:**

Formal and descriptive linguistics, grammar engineering, development of lexical resources

Computational approaches to NLP over time

Symbolic approaches

- **Computational expertise:** Formal grammars (algebraic grammars, mildly context-sensitive grammars, polynomial languages…), parsing algorithms, dynamic programming

- **Comp. linguistics expertise:** Formal and descriptive linguistics, grammar engineering, development of lexical resources

Statistical approaches

- **Computational expertise:** (statistical) machine learning, supervised, semi-supervised and non-supervised (PCFG, CRF, MEMM, discriminative algorithms…), hybrid approaches

- **Comp. linguistics expertise:** development of annotated corpora (training dataset), development of lexical resources

Computational approaches to NLP over time

Symbolic approaches

- **Computational expertise:** Formal grammars (algebraic grammars, mildly context-sensitive grammars, polynomial languages…), parsing algorithms, dynamic programming
- **Comp. linguistics expertise:** Formal and descriptive linguistics, grammar engineering, development of lexical resources

Statistical approaches

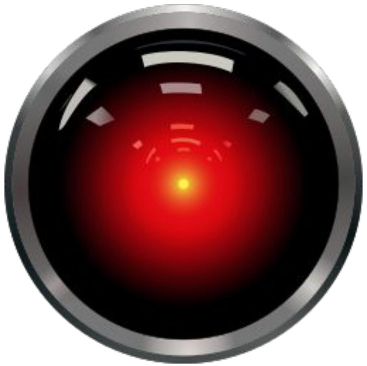
- **Computational expertise:** (statistical) machine learning, supervised, semi-supervised and non-supervised (PCFG, CRF, MEMM, discriminative algorithms…), hybrid approaches
- **Comp. linguistics expertise:** development of annotated corpora (training dataset), development of lexical resources

Neural approaches

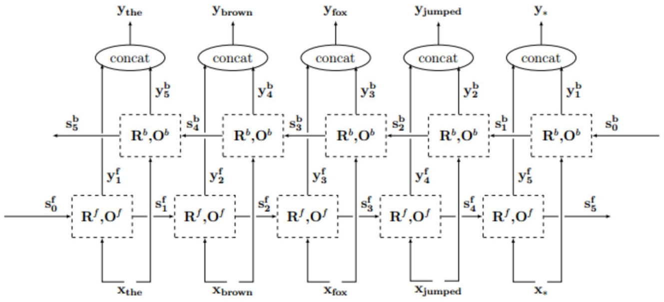
- **Comp. expertise:** neural networks, deep learning, end-to-end training, neural language modelling
- **Comp. ling. exp.:** same as for statistical approaches

NLP and AI

- NLP is one of the areas within “old” AI
  - AI = computationally simulate human behaviours requiring intelligence
  - Among them: understanding, producing and transforming speech / language



- One of the areas taking advantage of the “new” AI
  - In particular: deep learning
  - *confusion between objectives and means*



NLP and AI

- Neural approaches have resulted in major improvements, esp. in:
  - Machine translation
  - Semantic analysis
  - Conversational agents (chatbots)
- However,
  - Technically complex approaches (training times…)
  - It is difficult to “correct” a neural model
  - Often require huge amounts of training data
  - Models are dependent to the characteristics of training data

A very quick introduction to linguistics

Sentence-level analysis

Phonological level

International Phonetic Alphabet

[aɪ pʰiː eɪ]

Sentence-level analysis

Phonological level

International Phonetic Alphabet

[aɪ p^{h}iː eɪ]

Graphemic level

enough, cough, draught, although, brought, through, thorough, hiccough

Analysis in context

Sentence-level analysis

Morphological level

brav+itude, bio+terror-isme/-iste, skype+(e)r

mang-er-i-ons = MANGER+cond+1pl

Phonological level

International Phonetic Alphabet

[aɪ pʰiː eɪ]

Graphemic level

enough, cough, draught, although, brought, through, thorough, hiccough

Analysis in context

Sentence-level analysis

Syntactic level

John saw a dog yesterday which was a Yorkshire Terrier

Morphological level

brav+itude, bio+terror-isme/-iste, skype+(e)r

mang-er-i-ons = MANGER+cond+1pl

Phonological level

International Phonetic Alphabet

[aɪ pʰiː eɪ]

Graphemic level

enough, cough, draught, although, brought, through, thorough, hiccough

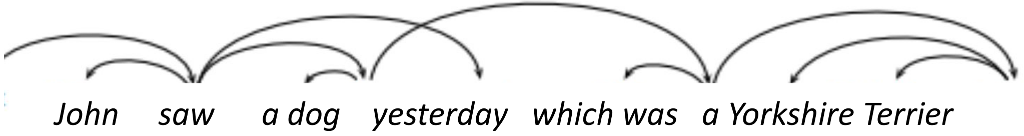
Analysis in context

Sentence-level analysis

Semantic level

The landlord$_{\text{SPEAKER}}$ has not yet **REPLIED**$^{\textbf{Communication\_response}}$ in writing$_{\text{MEDIUM}}$ to the tenant$_{\text{ADRESSEE}}$ objecting the proposed alterations$_{\text{MESSAGE}}$. *DNI*$_{\text{TRIGGER}}$

Syntactic level



Morphological level

brav+itude, bio+terror-isme/-iste, skype+(e)r

mang-er-i-ons = MANGER+cond+1pl

Phonological level

International Phonetic Alphabet

[aɪ pʰiː eɪ]

Graphemic level

enough, cough, draught, although, brought, through, thorough, hiccough

Analysis in context

Sentence-level analysis

Linguistic context

— *You know what? John gave Peter a Christmas present yesterday*
— *Wow, was **he** surprised? What was **it** like?*
— *Surprisingly good. **He** spent quite a bit on it.*

Semantic level

The landlord$_{\text{SPEAKER}}$ has not yet **REPLIED**$^{\text{Communication\_response}}$ in writing$_{\text{MEDIUM}}$ to the tenant$_{\text{ADRESSEE}}$ objecting the proposed alterations$_{\text{MESSAGE}}$. *DNI*$_{\text{TRIGGER}}$

Syntactic level

John saw a dog yesterday which was a Yorkshire Terrier

Morphological level

brav+itude, bio+terror-isme/-iste, skype+(e)r

mang-er-i-ons = MANGER+cond+1pl

Phonological level

International Phonetic Alphabet

[aɪ pʰiː eɪ]

Graphemic level

enough, cough, draught, although, brought, through, thorough, hiccough

Analysis in context

Sentence-level analysis

Extra-linguistic context

*Found **him** in the street inside a bag. I think **he** is happy with his new life*

http://9gag.com/gag/azVnEwp/found-him-in-the-street-inside-a-bag-i-think-he-is-happy-with-his-new-life

Linguistic context

— *You know what? John gave Peter a Christmas present yesterday*
— *Wow, was **he** surprised? What was **it** like?*
— *Surprisingly good. **He** spent quite a bit on it.*

Semantic level

The landlord$_{\text{SPEAKER}}$ has not yet **REPLIED**$^{\text{Communication\_response}}$ in writing$_{\text{MEDIUM}}$ to the tenant$_{\text{ADRESSEE}}$ objecting the proposed alterations$_{\text{MESSAGE}}$. *DNI*$_{\text{TRIGGER}}$

Syntactic level

John saw a dog yesterday which was a Yorkshire Terrier

Morphological level

brav+itude, bio+terror-isme/-iste, skype+(e)r

mang-er-i-ons = MANGER+cond+1pl

Phonological level

International Phonetic Alphabet

[aɪ pʰiː eɪ]

Graphemic level

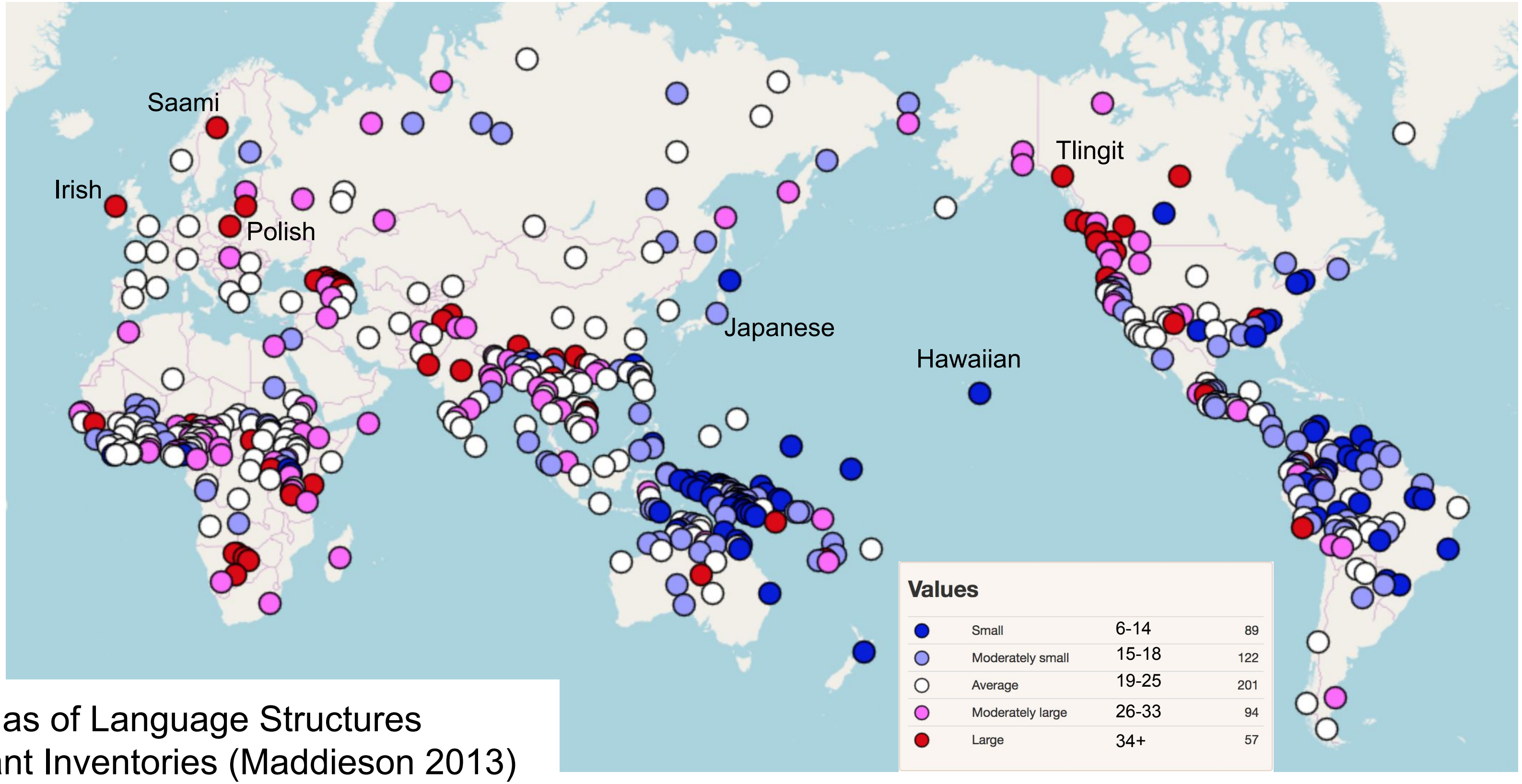
enough, cough, draught, although, brought, through, thorough, hiccough

The 4 major challenges of language processing

- Language **diversity**
- Language **variation**
- Language **ambiguity**
- Language **sparsity**

Language diversity

Phonological diversity



World Atlas of Language Structures
Consonant Inventories (Maddieson 2013)

Phonological diversity

Central Rotokas	Bilabial	Alveolar	Velar
Voiceless	p	t	k
Voiced	b ~ β	d ~ ɾ	g ~ ɣ

Ubykh		Labial		Alveolar			Postalveolar					Velar				Uvular					Glottal
							laminal closed		laminal		apical										
		plain	phar.	plain	lab.	lat.	plain	lab.	plain	lab.		pal.	plain	lab.	phar.	pal.	plain	lab.	phar.	phar. & lab.	
Plosive	voiceless	p	pˤ	t	tʷ							kʲ	k	kʷ		qʲ	q	qʷ	qˤ	qˤʷ	
	voiced	b	bˤ	d	dʷ							gʲ	g	gʷ							
	ejective	p’	pˤ’	t’	tʷ’							kʲ’	k’	kʷ’		qʲ’	q’	qʷ’	qˤ’	qˤʷ’	
Affricate	voiceless			t͡s			t͡ʃ		t͡ɕ	t͡ɕʷ	ʈ͡ʂ										
	voiced			d͡z			d͡ʒ		d͡ʑ	d͡ʑʷ	ɖ͡ʐ										
	ejective			t͡s’			t͡ʃ’		t͡ɕ’	t͡ɕʷ’	ʈ͡ʂ’										
Fricative	voiceless	f		s		ɬ	ʃ	ʃʷ	ɕ	ɕʷ	ʂ		x			χʲ	χ	χʷ	χˤ	χˤʷ	h
	voiced	v	vˤ	z			ʒ	ʒʷ	ʑ	ʑʷ	ʐ		ɣ			ʁʲ	ʁ	ʁʷ	ʁˤ	ʁˤʷ	
	ejective					ɬ’															
Nasal		m	mˤ	n																	
Approximant						l						j		w	wˤ						
Trill				r																	

Wikipedia

Phonological diversity

Syllables are formed of phoneme sequences

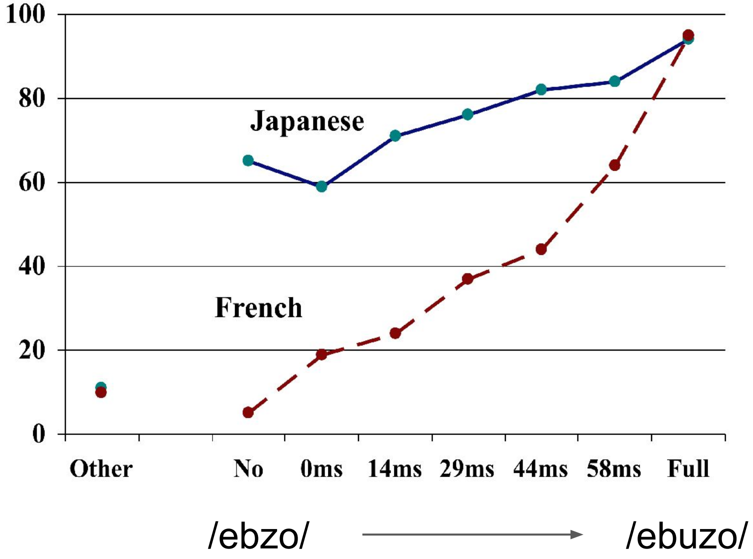
In most languages, some syllables are valid, some are not

Japanese: only V, CV, VN, CVN allowed

> phonological adaptation of borrowings:

sphinx > /sufiNkusu/

Christmas > /kurisumasu/



Dupoux et al. 1998

Phonological diversity

Different vowel/consonant frequencies and cluster usage:

Georgian /gvbrdɣvnis/ ‘he's plucking us’

Nuxalk (“Bella Coola”) *clhp’xwlhtlhplhhskwts’* /xɬp’χʷɬtʰɬpʰɬːskʷʰt͡s’/
‘he had possessed a bunchberry plant’

Hawaiian *He aha kēia?* ‘What is it?’

Morphological diversity

- Analytic and isolating languages
  - Each word carries exactly one meaning
  - Ex.: Chinese /u̯ɔ^{214} mən^{4} tʰa̠n^{35} kɑŋ^{55}t͡ɕʰin^{35} lə^{5}/ (1st_pers plur PLAY PIANO past) ‘we played the piano’

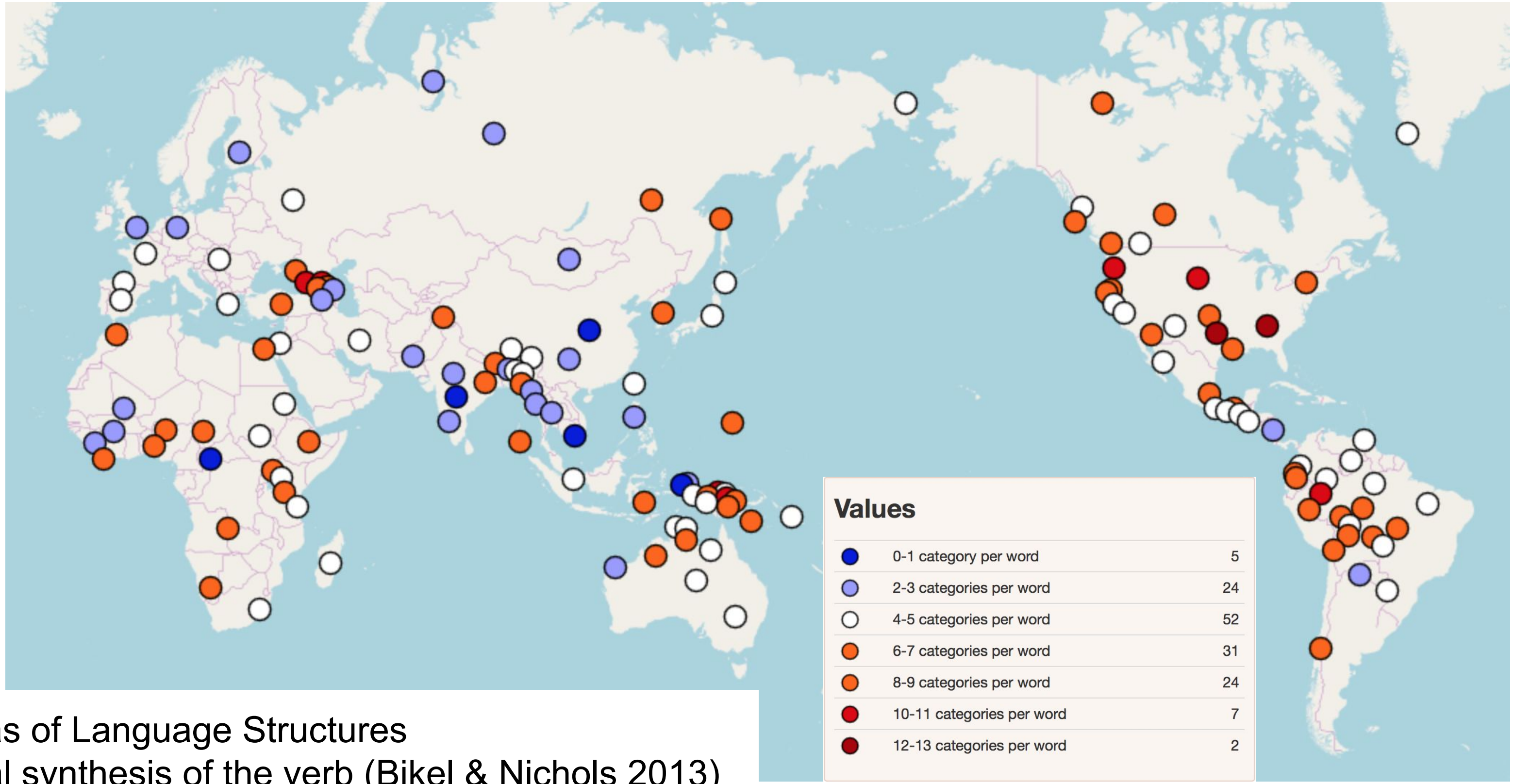
- Synthetic languages
  - Agglutinative
    - Each word can have several morphs, each carrying one meaning
    - Ex.: Turkish *el-ler-imiz-in* (HAND-pl-poss1pl-genitive) ‘of our hands’
  - Fusional
    - Each word can have several morphs, each carrying one or more meanings, of which (generally) only one lexical morph (ex.: inflectional morphology, i.e. conjugation, declension…)
    - Ex.: Latin *rexistis* /rek-s-is-tis/ (RULE-perf-perf-perf.2sg) ‘you$_{PLUR}$ ruled’
  - Polysynthetic
    - Each word can have several lexical or grammatical morphs
    - Ex.: Island Halkomelem (Salish) *hwpulqwitth’a’ustum*
      (locative-GHOST/DEATH(?)-blanket/cloth-face-transitive-passive)
      ‘to be adversely affected by a spirit entering the body through the face’

Morphological diversity

Most languages show elements of different morphological types

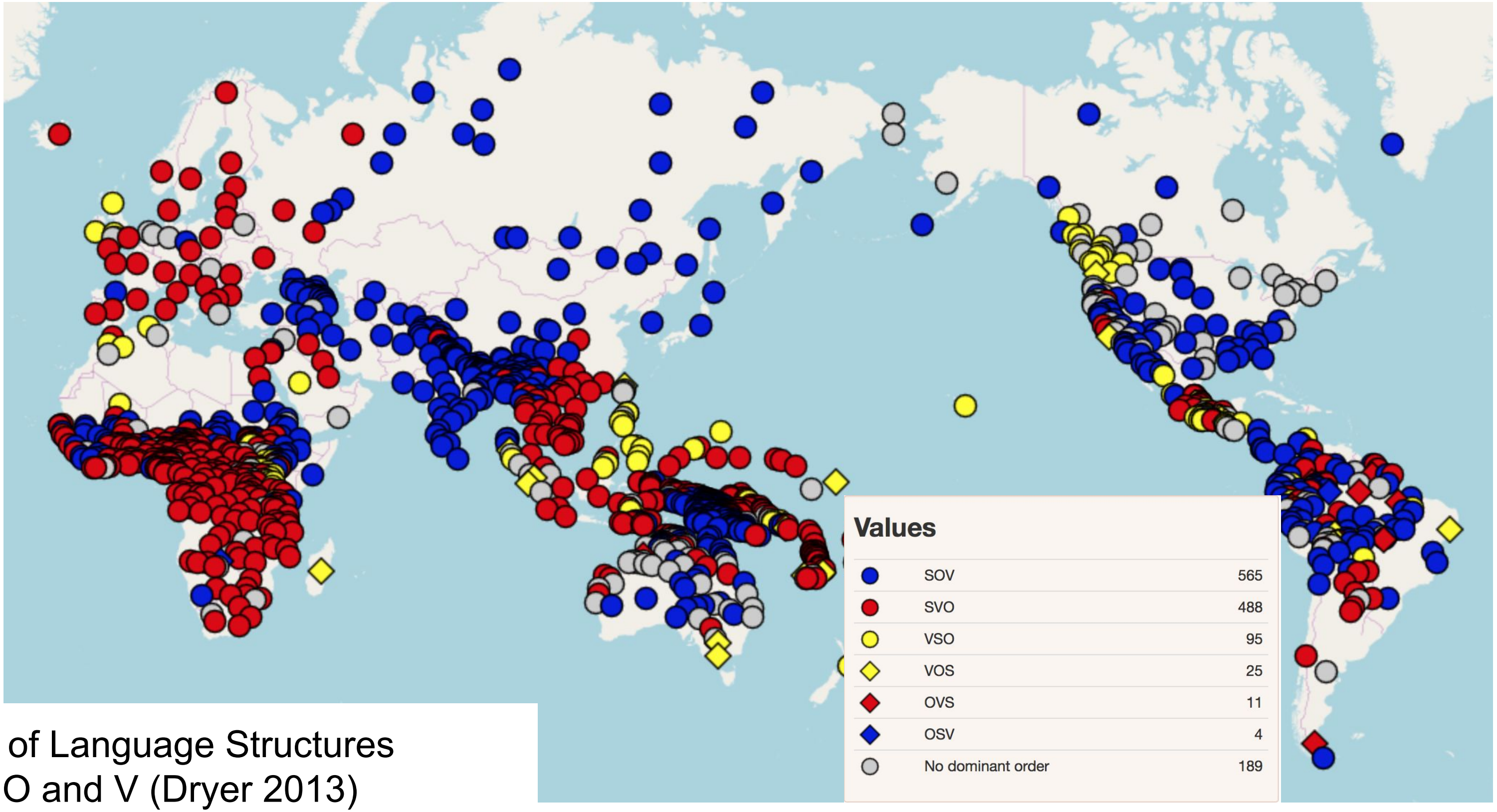
- Ex.: English!
  - *the boy will play with the dog*
  - *John’s cat eats mice*
  - *antidisestablishmentarianism* (derivational morphology)
- Other example: creating words or word-like sequences from sentences
  - French: *je-m’en-foutisme*
  - English: *You know, I can't take all this let's-be-faithful-and-never-look-at-another-person routine, because it just doesn't work* (The Boys in the Band, 1970)

Morphological diversity



World Atlas of Language Structures
Inflectional synthesis of the verb (Bikel & Nichols 2013)

Syntactic diversity



World Atlas of Language Structures
Order of S, O and V (Dryer 2013)

Syntactic diversity

Levels of configurationality

- Free word order (often with very rich morphological marking)
  - Ex.: Warlpiri
- Relatively free word order
  - Often with rich morphological marking
  - And discontinuous constituents
  - Ex.: Polish 'John went to the cinema'

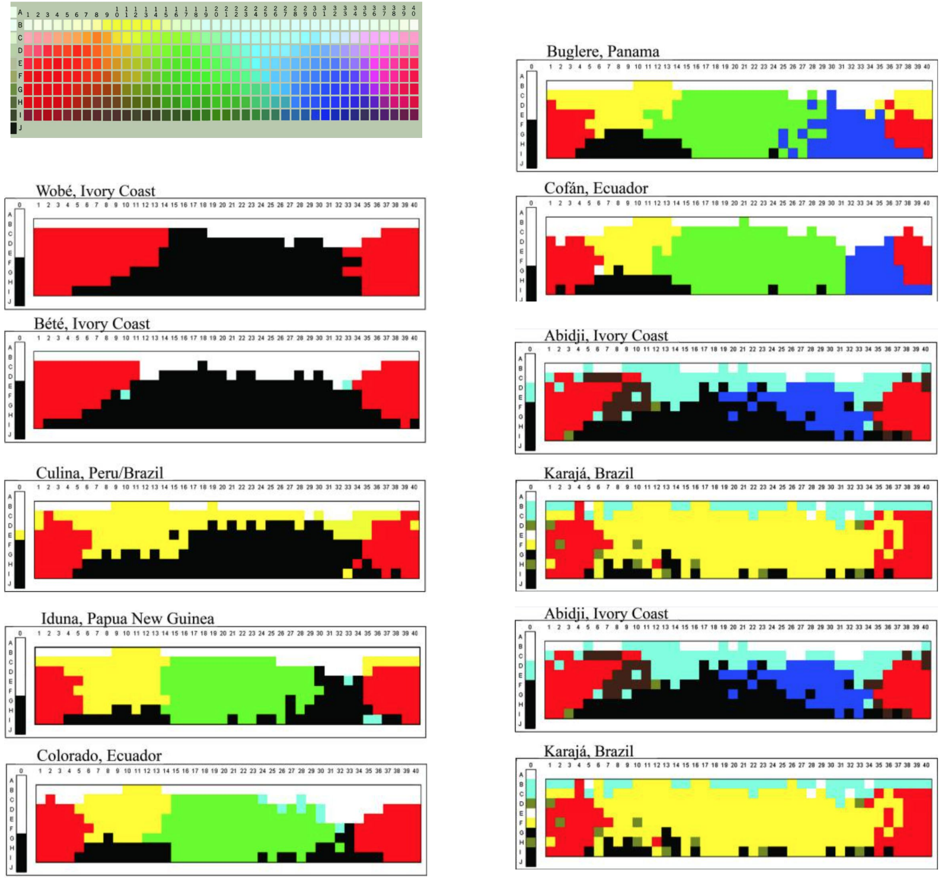
Jaś poszedł do kina.
Poszedł Jaś do kina.
Jaś do kina poszedł.
Poszedł do kina Jaś.
Do kina Jaś poszedł.
Do kina poszedł Jaś.

- Constrained word order ("configurational")
  - Ex.: English, Chinese
  - Often with limited or no morphological marking
  - Discontinuous constituents are rare

Semantic diversity

Words (fuzzily) partition the semantic space

Partitions can differ from one language to another

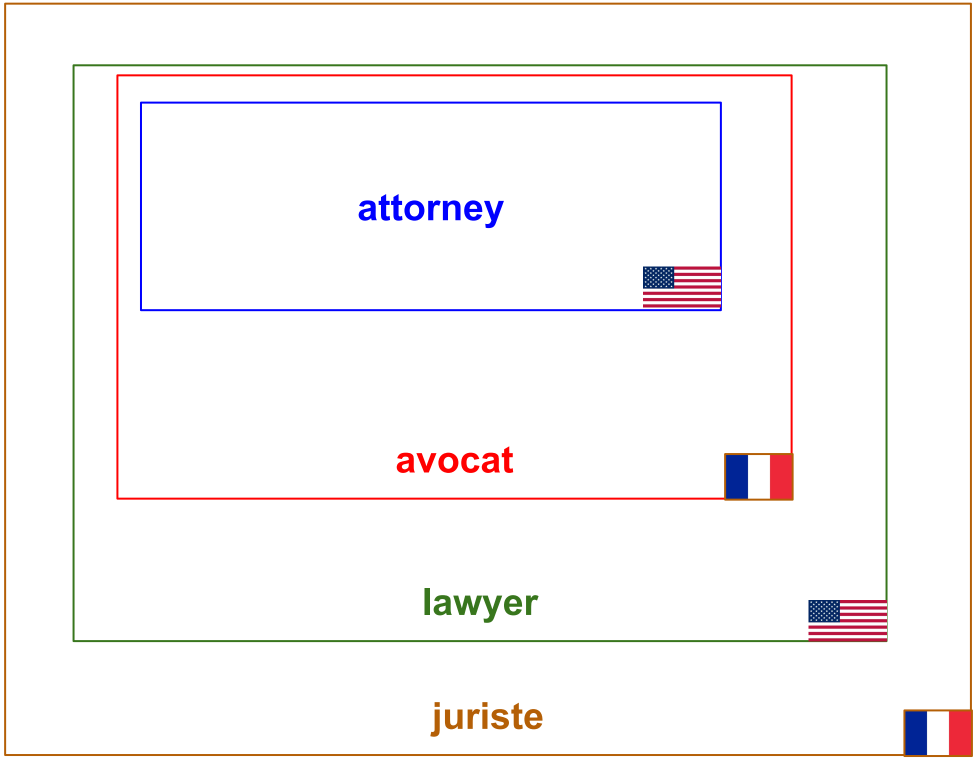


World Color Survey
(Kay et al. 2009)

Semantic diversity

Words (fuzzily) partition the semantic space

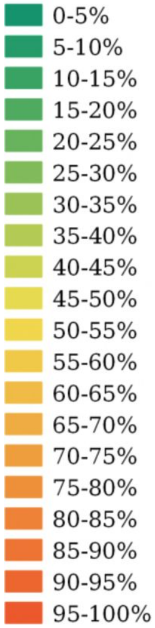
Partitions can differ from one language to another

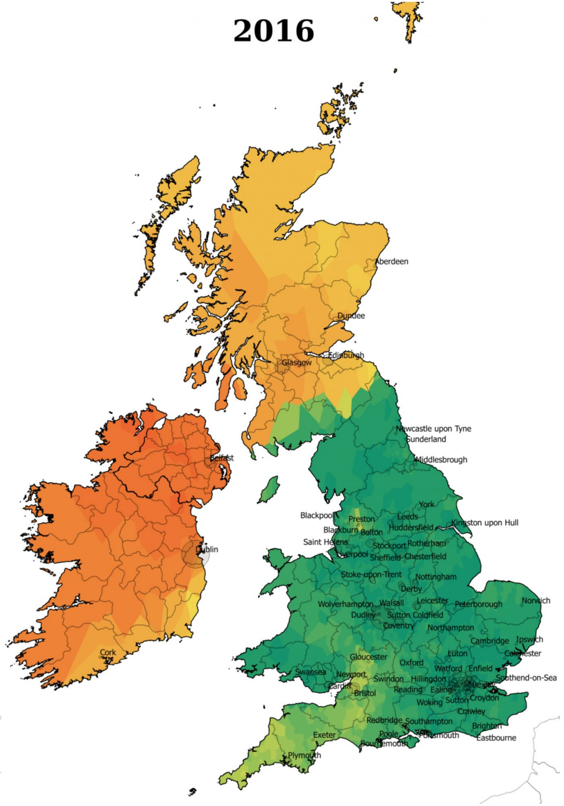


Language variation

Phonetic and phonological variation

Do you pronounce the
“r” in “arm” ?

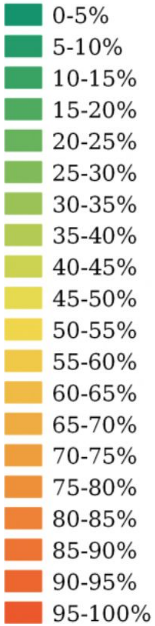


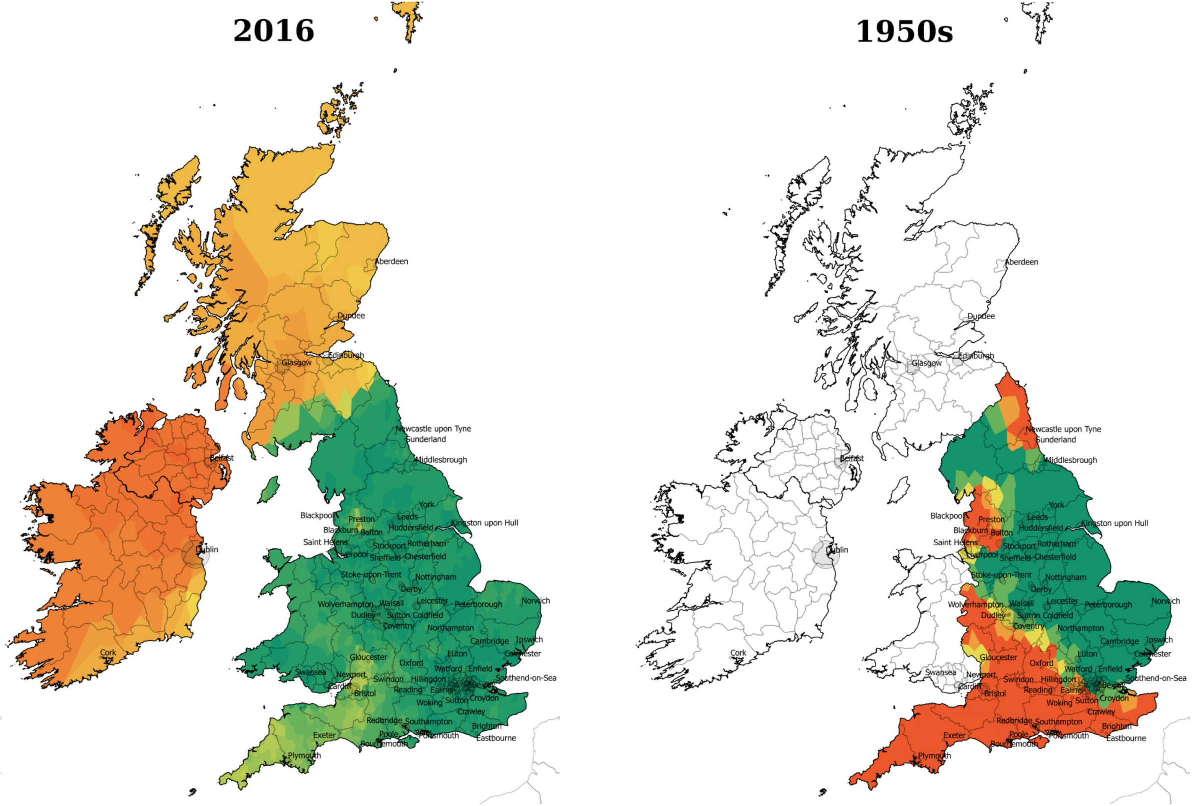


University of Cambridge (2016)

Phonetic and phonological variation

Do you pronounce the
“r” in “arm” ?





University of Cambridge (2016)

Spelling “variation”

anagement maagement maanagement
maangement magagement magement
mamagement mamangement manaagement manaement
managaement manageement manageemnt managegment
managemaent managemant managememt managemen managemenet
managementt managemet managemetn managemnent managemnet
managemnt managemrnt managemt managenent managenment managent
managerment managhement managmeent managmement managment managnment
manament manamgement mananement manangment manasgement
manegement manegment mangaement mangagement mangagment
mangament mangement manggement mangment
mangmt menagement mgmt mgnt
mnagement mngmnt mngmt

Sociolinguistic variation

Interpreting tweets produced by Chicago gang members

Tweet	Label	Youth Interpretation
If We see a opp Fuck it We Gne smoke em 😈	Aggression (Threat)	he mean like if he see opp he go kill him opp mean like the people he dont like
Dnt get caught on Dat 800 block lame ass Lil niggas Betta take Dat Shyt on stony spot	Aggression (Insult)	he saying them lil nigga better not get caught on the 800 block or they go kill them so he tell them if they wanna live they better stay on stony
Young niggas still getting shot babies still dying 🙏	Loss	he mean like teen keep die and babys and kid keep die

Blevins et al. 2016

Sociolinguistic variation

T'as vu il l'a bien cherché wsh #AperoChezRicard
> +10000, shah!
> tabuz, lavé rien fé
> ki ca ? le mec ou son chien ?
> Wtf is wrong with him ? #PETA4EVER
> ki ca ? le chien ?
> looool

Sociolinguistic variation

T'as vu il l'a bien cherché wsh #AperoChezRicard
> +10000, shah!
> tabuz, lavé rien fé
> ki ca ? le mec ou son chien ?
> Wtf is wrong with him ? #PETA4EVER
> ki ca ? le chien ?
> looool

BING translation:
You saw coming it #AperoChezRicard wsh
> +10000, shah!
> tabuz, washed anything fe
> Ki ca? the guy or his dog?
> WTF is wrong with him? #PETA4EVER
> Ki ca? the dog?
> looool

Diachronic variation

Li reis Marsilie esteit en Sarraguce.
Alez en est en un verger suz l'umbre;
Sur un perrun de marbre bloi se culchet,
Envirun lui plus de vint milie humes.
Il en apelet e ses dux e ses cuntes:
« Oëz, seignurs, quel pecchet nus encumbret :
Li emper[er]es Carles de France dulce
En cest païs nos est venuz cunfundre.
Jo nen ai ost qui bataille li dunne,
Ne n'ai tel gent ki la sue derumpet.
Cunseilez mei cume mi savie hume,
Si m(e) guarisez e de mort et de hunte. »
N'i ad paien ki un sul mot respundet,
Fors Blancandrins de Castel de Valfunde.

Chanson de Roland (XIth century)

Hwæt! Wé Gárdena in géardagum
þéodcyninga þrym gefrúnon·
hú ðá æþelingas ellen fremedon.
Oft Scyld Scéfing sceaþena þréatum
monegum maégþum meodosetla ofteáh·
egsode Eorle syððan aérest wearð
féasceaft funden hé þæs frófre gebád·
wéox under wolcnum· weorðmyndum þáh
oð þæt him aéghwylc þára ymbsittendra
ofer hronráde hýran scolde,
gomban gyldan· þæt wæs gód cyning.

Beowulf (Xth century)

Language ambiguity

Lexical ambiguity: homonymy

Homophony: same pronunciation, different words (and often spelling)

- Ex.: English *weather, wether, whether* / French: *vers, verre, ver, vert, vair*
- More extreme case = oronyms. Cf. English *ice cream* vs. *I scream*
- Even more extreme case = holoryms

  Étonnamment monotone et lasse

  Est ton âme en mon automne, hélas !

  (Louise de Vilmorin)

Homography: same spelling, different words (and sometimes pronunciation)

- Ex.: French *les poules du couvent couvent*

  English *if you have not read this book yet, read it!*

Segmentation ambiguity

Segmentation in elementary linguistic units

- *Bob | a | mangé | une |* ***pomme de terre***
- *Bob | , | sculpteur | , | a | fabriqué | une |* ***pomme | de | terre cuite***

=> distinction between **tokens** and **forms**

Token = typographic unit (*pomme de terre* is always 3 tokens)

Form (wordform) = linguistic unit (*pomme de terre* can be 1 or 3 forms)

Amalgams = several forms in one token (French *aux*, English *don't*)

Can be ambiguous! French *des* (1 token) can be *de + les* (2 forms) or *des* (1 form)

There are complex cases. Cf. French *à l' instar du* = *à_l'_instar_de + le*

Morphological ambiguity

Lemma = equivalence class of forms belonging to a same morphological paradigm

A lemma is often represented by one of its forms, the “citation form”

- Example: for a verb, the infinitive (French) or its 1st pers. prs. ind. (Latin, Greek)

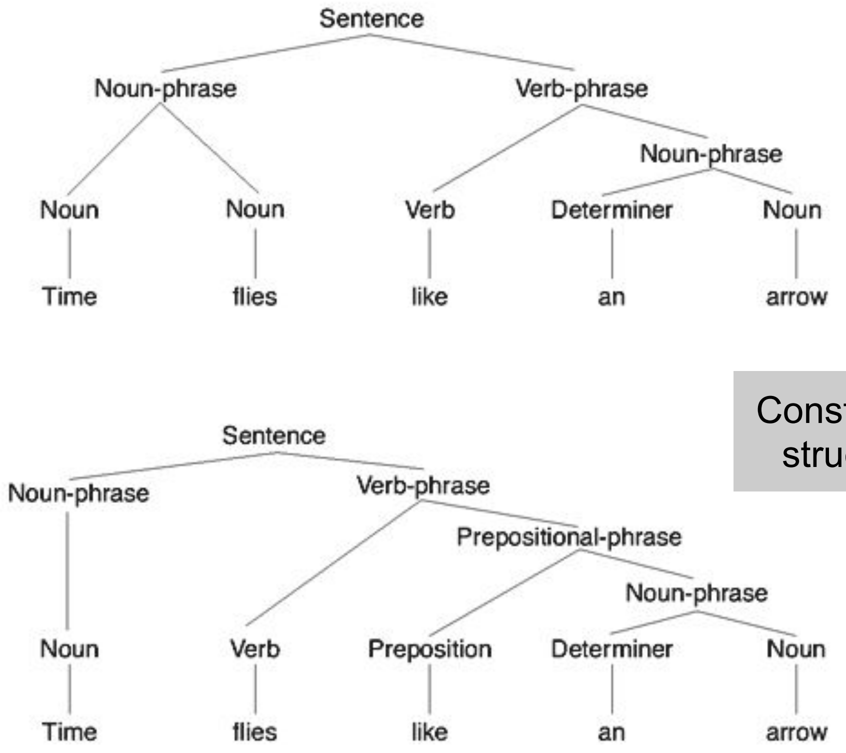
Lemmatisation = associate each form in a sentence with its lemma

Morphological analysis = associate each form in a sentence with its lemma AND morphological tags

- Example: *mangerons* = MANGER(v)+ind.fut.1pl

Syntactic ambiguity

Time flies like an arrow

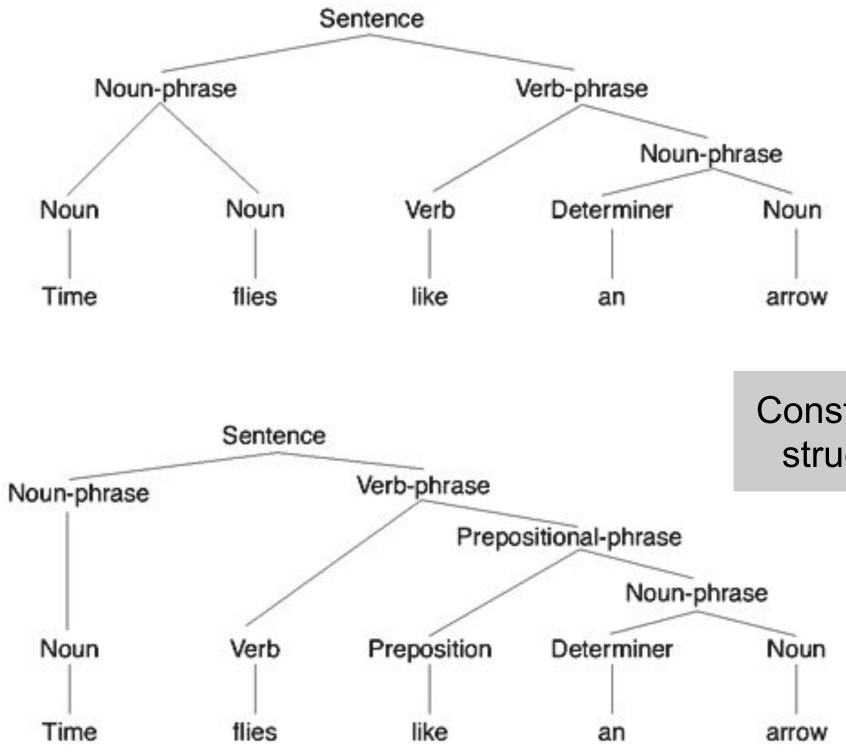


Constituency structures

Syntactic ambiguity

Time flies like an arrow

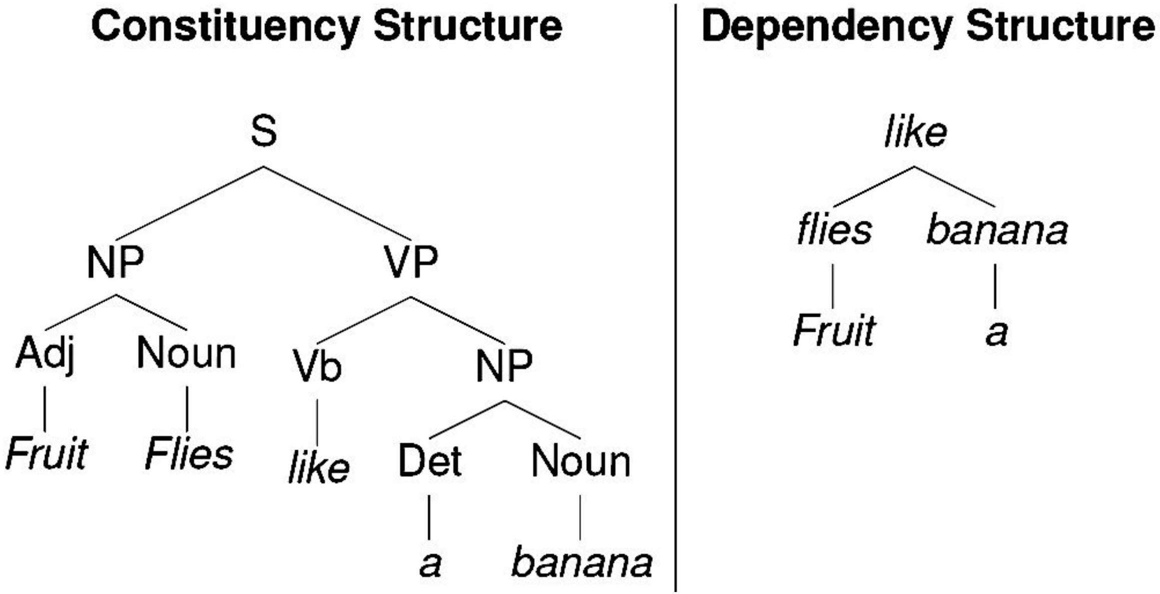
Cf. *Fruit flies like a banana*



Constituency structures

A bit of terminology

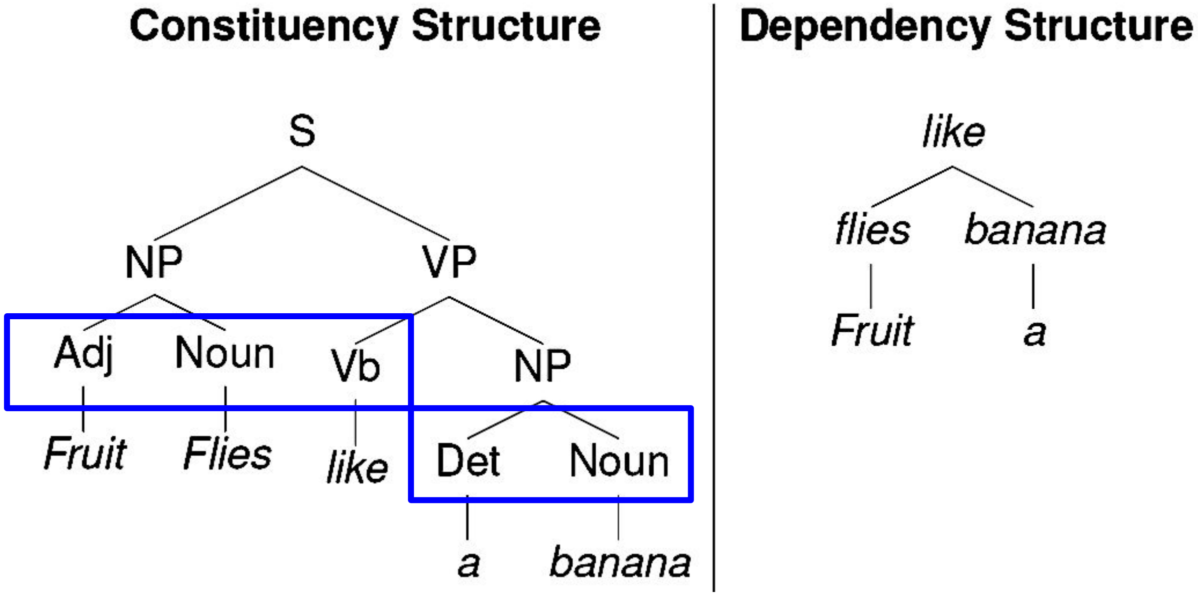
Fruit flies like a banana



A bit of terminology

Fruit flies like a banana

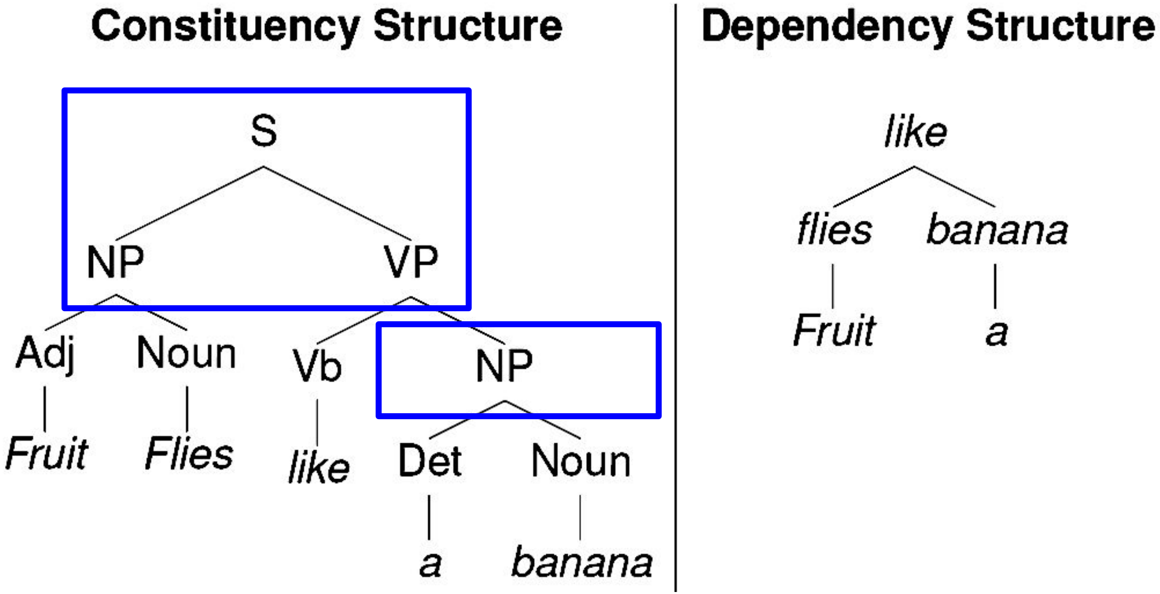
Parts-of-speech
(PoS)



A bit of terminology

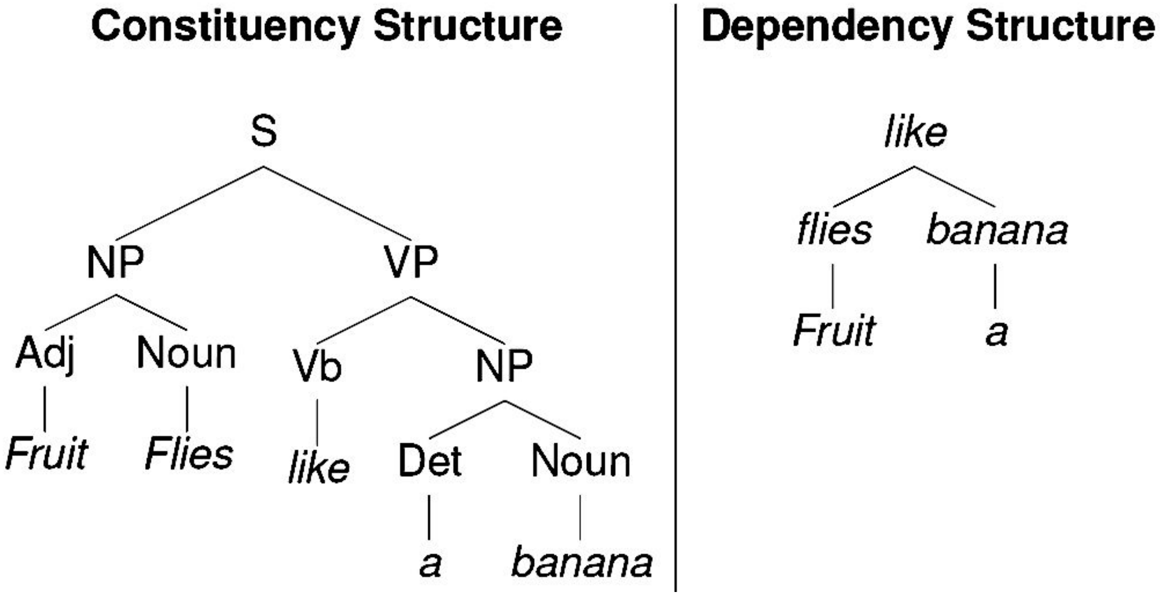
Fruit flies like a banana

Phrases (or constituents)



A bit of terminology

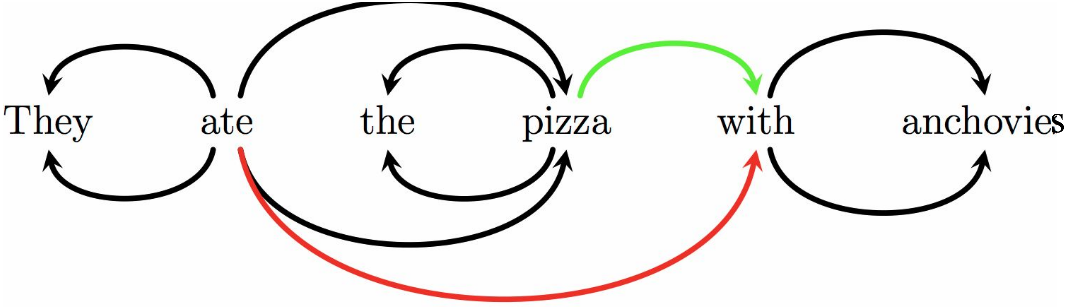
Fruit flies like a banana



Automatic syntactic analysis = parsing

- Constituency parsing
- Dependency parsing

Syntactic ambiguity: PP attachment



Garden-path sentences

The cotton clothing is usually made of grows in Mississippi

Until the police arrest the drug dealers control the street

Mary gave the child the dog bit a bandaid

The girl told the story cried

The dog that I had really loved bones

The old man the boat

The raft floated down the river sank

We painted the wall with cracks

Garden-path sentences

(The cotton (clothing is usually made of)) grows in Mississippi

(Until the police arrest) (the drug dealers control the street)

Mary gave (the child (the dog bit)) (a bandaid)

(The girl (told the story)) cried

(The dog that I had) really loved bones

(The old) man (the boat)

(The raft (floated down the river)) sank

We painted (the wall with cracks)

Semantic ambiguity: polysemy

Hyponymy: man (vs. animals) ⊃ man (vs. woman) ⊃ man (vs. boy)

Metaphor: mole (the animal) > mole (a spy)

Object/color: cherry (the fruit) > cherry (as a color, cf. *I like your cherry shirt*)

Object/Informational content: book (the object) // book (its content)

Object/Collective abstract: tramway (vehicle) // tramway (means of transportation)

Tree or plant/Material/fruit/vegetable it produces: cotton (plant) > cotton (material)

Animal/Its (edible) flesh: rabbit (animal) > rabbit (meat)

Semantic ambiguity

Named entities:

- Detection
- Linking

Not logged in Talk Contributions Create account Log in

Article Talk Read Edit View history Search Wikipedia

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Michael Jordan (disambiguation)

From Wikipedia, the free encyclopedia

Michael Jordan (born 1963) is an American basketball player.

Michael Jordan or **Mike Jordan** may also refer to:

Contents [hide]

People [edit]

Sports [edit]

- Michael Jordan (footballer) (born 1986), English goalkeeper (Arsenal, Chesterfield, Lewes)
- Mike Jordan (racing driver) (born 1958), English racing driver
- Mike Jordan (baseball, born 1863) (1863–1940), baseball player
- Michael Jordan (American football) (born 1992), American football cornerback
- Michael-Hakim Jordan (born 1977), American professional basketball player
- Michal Jordán (born 1990), Czech ice hockey player

Other people [edit]

- Michael B. Jordan (born 1987), American actor
- Michael Jordan (insolvency baron) (born 1931), English businessman
- Michael Jordan (Irish politician), Irish Farmers' Party TD from Wexford, 1927–1932
- Michael I. Jordan (born 1956), American researcher in machine learning and artificial intelligence
- Michael H. Jordan (1936–2010), American executive for CBS, PepsiCo, Westinghouse
- Michael Jordan (mycologist), English mycologist

Multiple ambiguity

- Most or all tasks in speech and language processing can be viewed as resolving **ambiguity** at one of the levels of signal or linguistic structure.
- The spoken sentence, *I made her duck,* has five different meanings.
  - (1) I cooked waterfowl for her.
  - (2) I cooked waterfowl belonging to her.
  - (3) I created the (plaster?) duck she owns.
  - (4) I caused her to quickly lower her head or body.
  - (5) I waved my magic wand and turned her into undifferentiated waterfowl.

Multiple ambiguity

- These different meanings are caused by multiple ambiguities.
  - PoS: *duck* can be a verb or a noun, while *her* can be a dative pronoun or a possessive pronoun -> part-of-speech tagging
  - Polysemy: the word *make* can mean *create* or *cook* -> word sense disambiguation
  - Syntactic ambiguity: the verb *make* is syntactically ambiguous in that it can be transitive (2), or it can be ditransitive (5). Moreover, *make* can take a direct object and a verb (4), meaning that the object (*her*) got caused to perform the verbal action (*duck*) -> parsing
  - In a spoken sentence, phonological ambiguity (homophones) is also present; the first word could have been *eye* or the second word *maid*.

Language sparsity

Corpora

Corpus = body of text stored in a machine-readable form

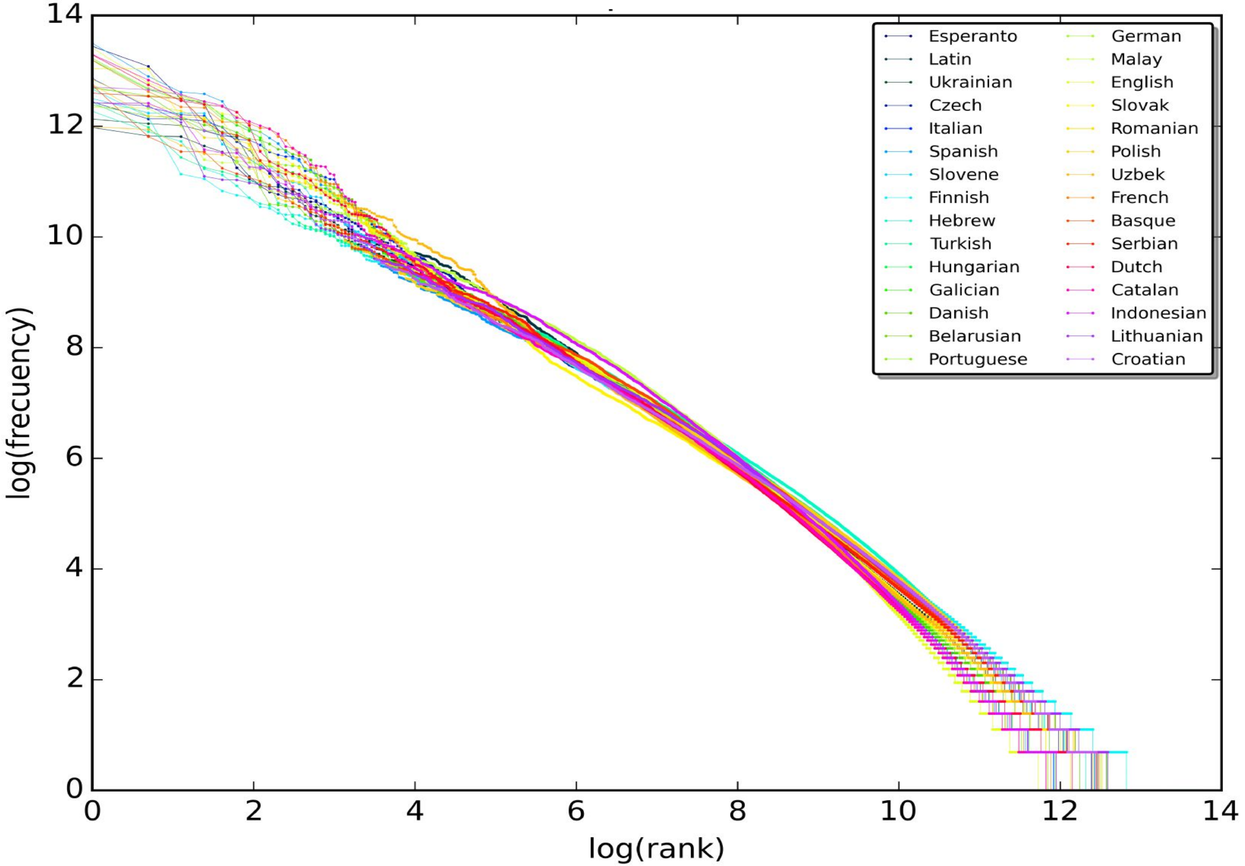
Corpora can be annotated, for serving as training, development or test data

- Morphosyntactically-annotated corpora
- Treebanks (syntactically-annotated)
- Semantically disambiguated corpora
- etc.

Zipf's law

A plot of the rank versus frequency for the first 10 million words in 30 Wikipedias

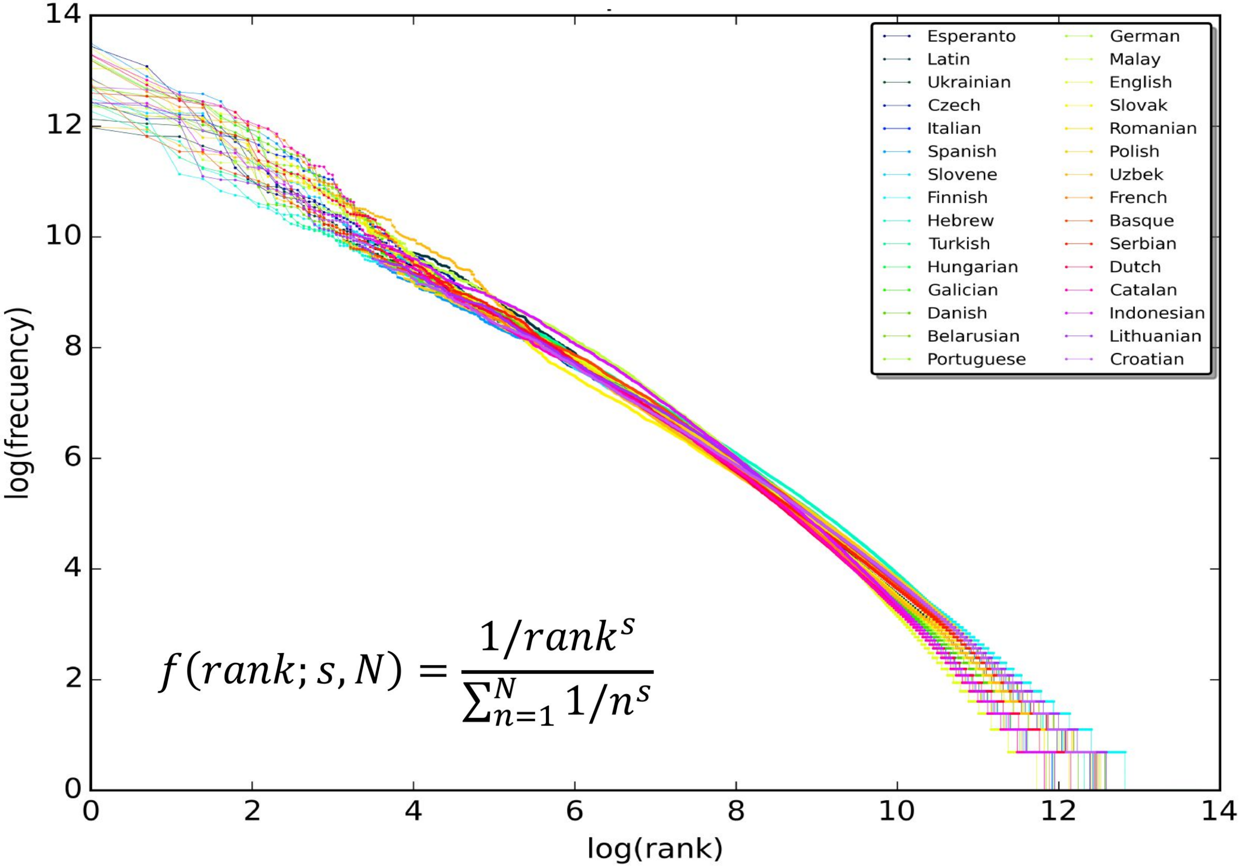
(source: Wikipedia; data: dumps from Oct 2015)



Zipf’s law

A plot of the rank versus frequency for the first 10 million words in 30 Wikipedias

(source: Wikipedia; data: dumps from Oct 2015)

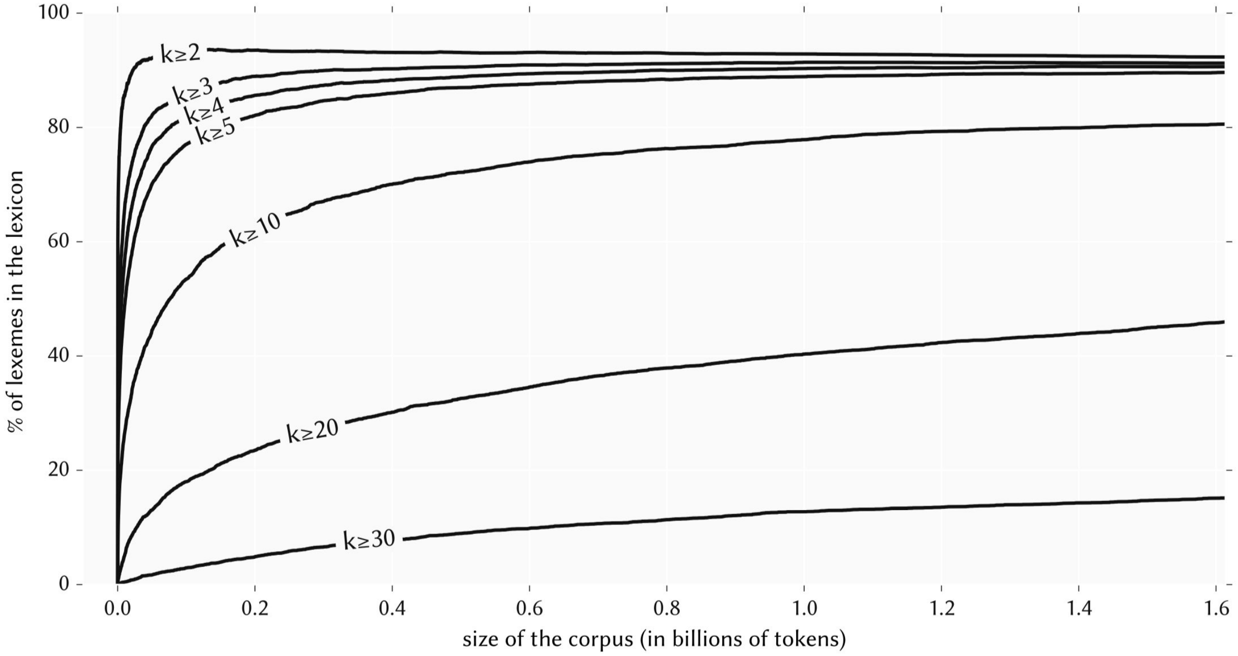


Zipf's law

In language data, many phenomena follow a zipfian distribution

Heaps's law

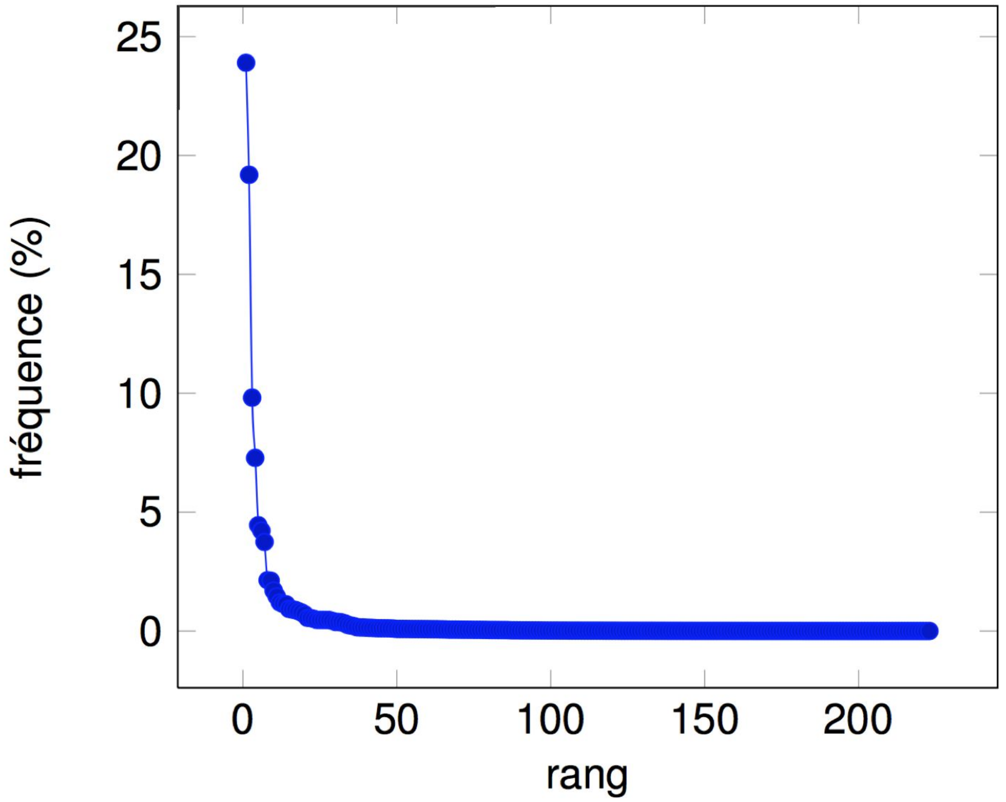
Example: proportion of verbal lemmas (known to a pre-defined lexicon) attested in at least *k* inflected forms as a function of vocabulary size in a large web-based corpus of French (FrWaC) for various values of *k*



Zipf’s law

In language data, many phenomena follow a zipfian distribution

Example: frequency of syntactic constructions in an automatically parsed 500M-word corpus



Logic of the course

The personal assistant

Amazon Alexa,
Google Home,
Baidu Raven, etc

Such systems can

- Identify the talker
- Recognize the words
- Understand the query
- Respond orally

Such systems can

- **<u>Identify the talker</u>**
- Recognize the words
- Understand the query
- Respond orally

performance: better than human (for long utterances)

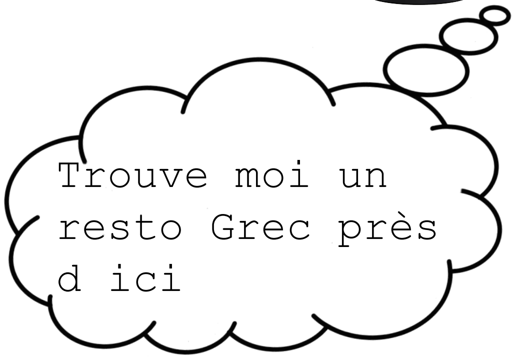
Such systems can

- Identify the talker
- **<u>Recognize the words (speech to text)</u>**
- Understand the query
- Respond orally

performance: roughly like humans

Such systems can

- Identify the talker
- **<u>Recognize the words (speech to text)</u>**
- Understand the query
- Respond orally



performance: roughly like humans (not for casual or noisy)

Such systems can

- Identify the talker
- Recognize the words
- **<u>Understand the query</u>**
- Respond orally

```
QUERY("restaurant",
type="grec",
geo-distance=2 km)
```

performance:
only for simple
requests

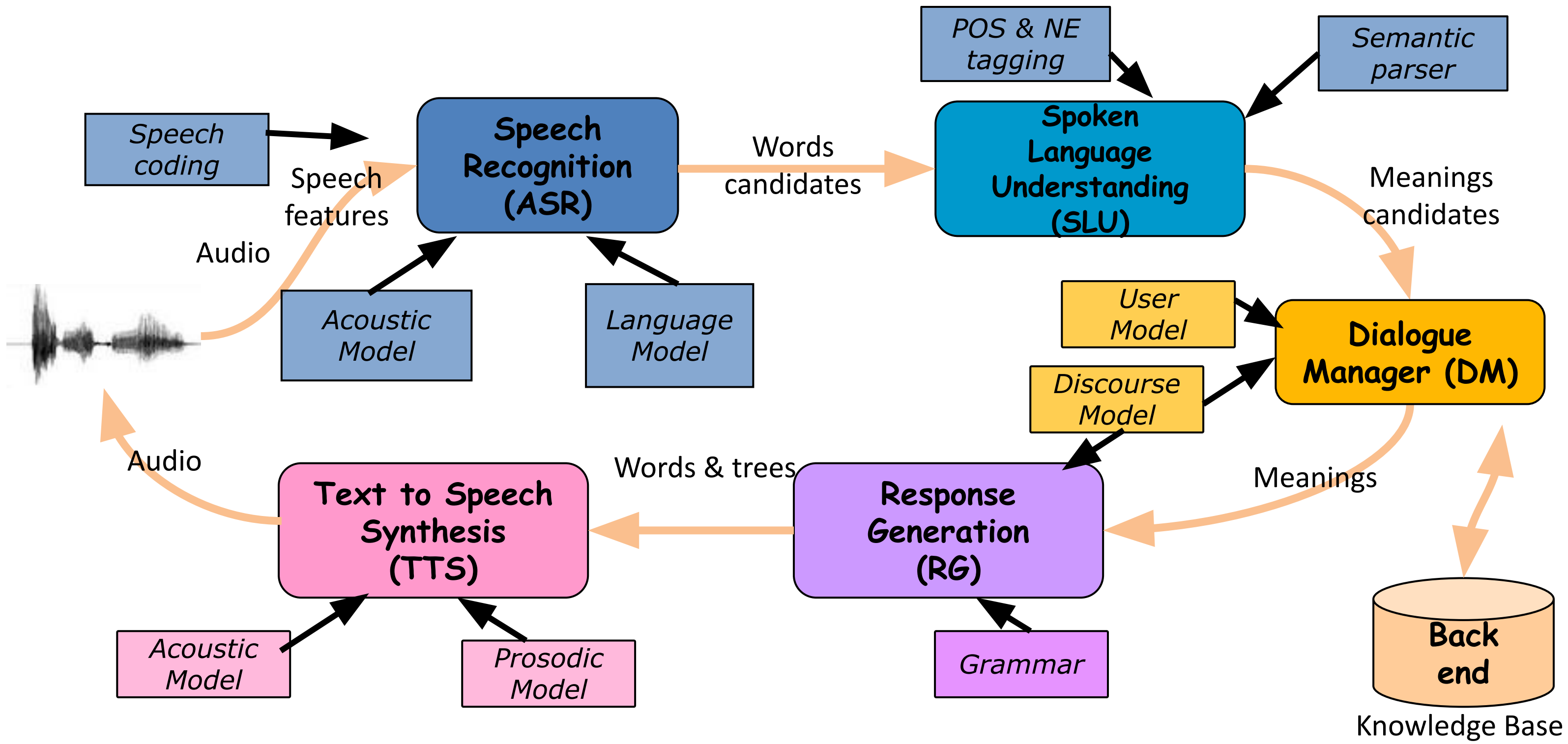
Such systems can

- Identify the talker
- Recognize the words
- Understand the query
- **<u>Respond orally</u>**

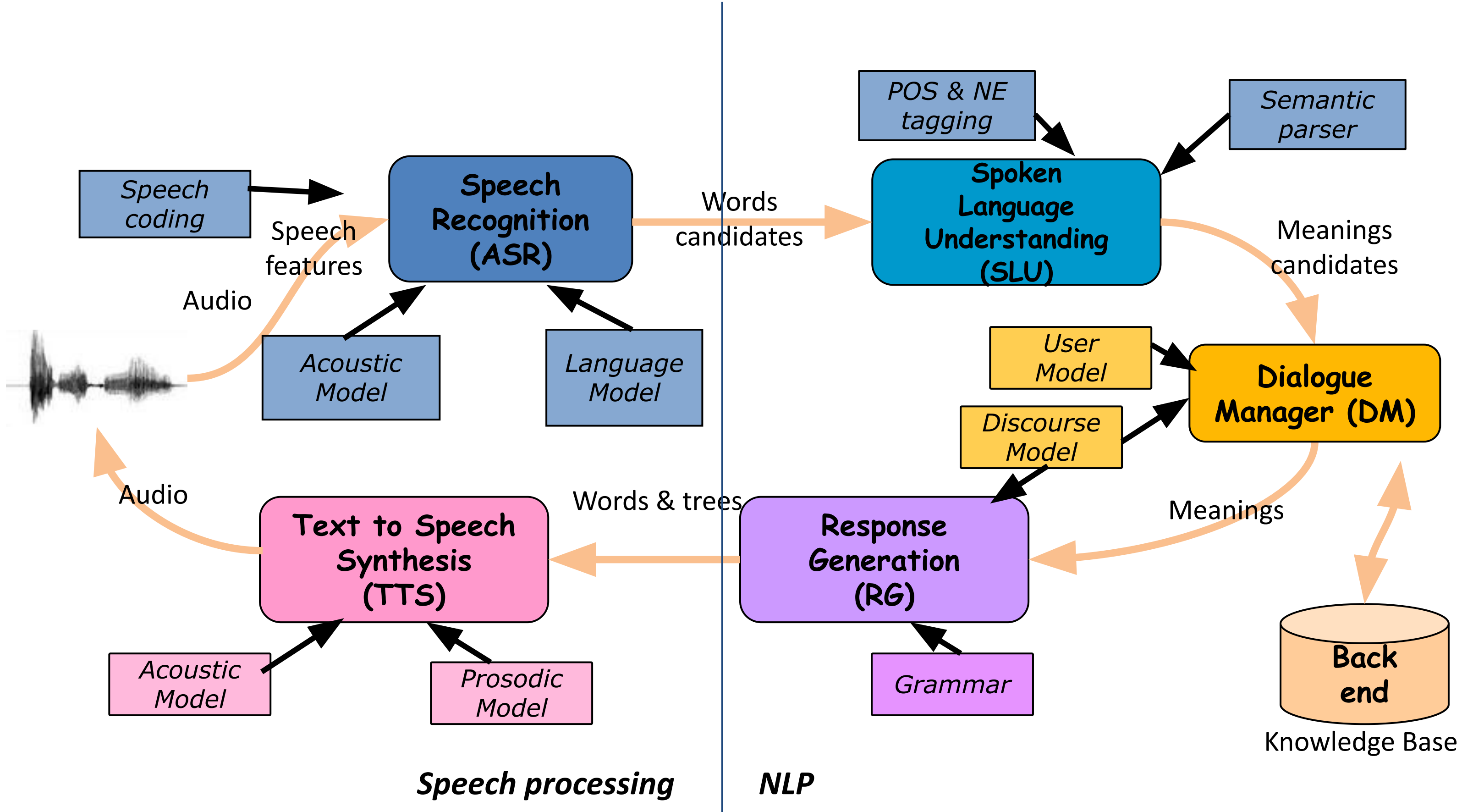
“Il y a un restaurant Grec ouvert au coin de la rue; je le réserve?”

performance: close to natural speech (not emotional speech)

Inside the box: language processing

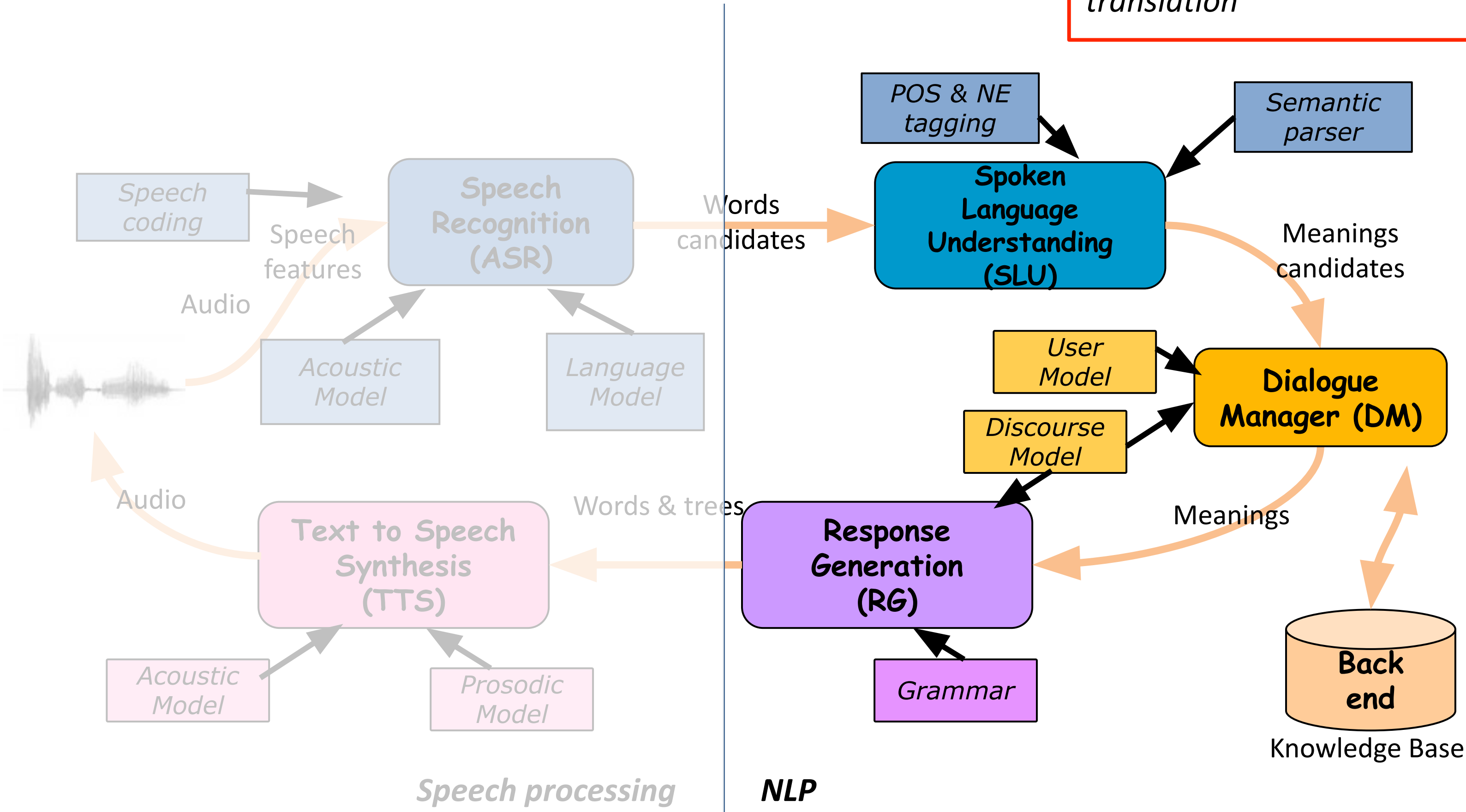


Inside the box: language processing

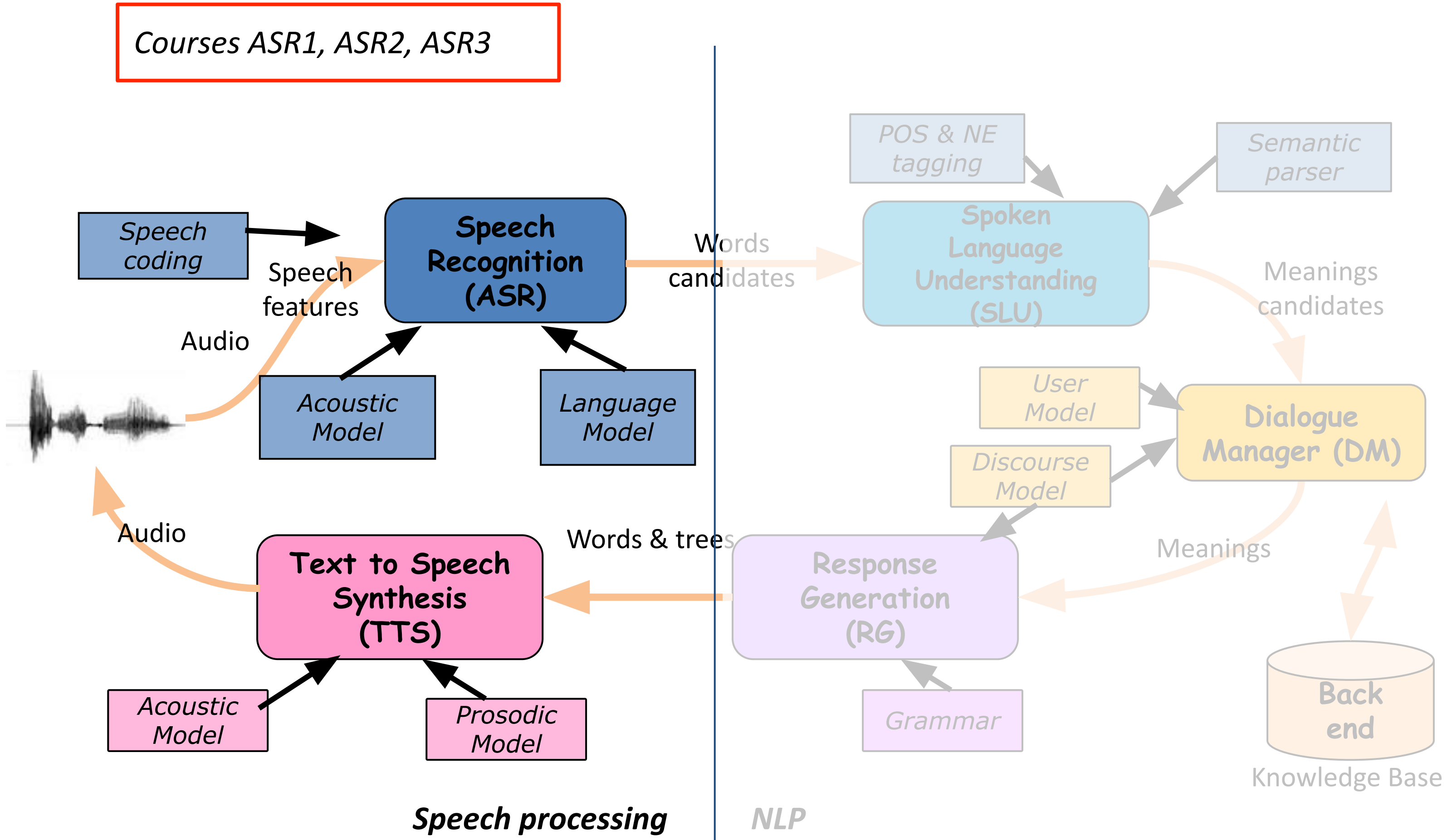


Inside the box: language processing

Courses NLP1, NLP2, NLP3, translation

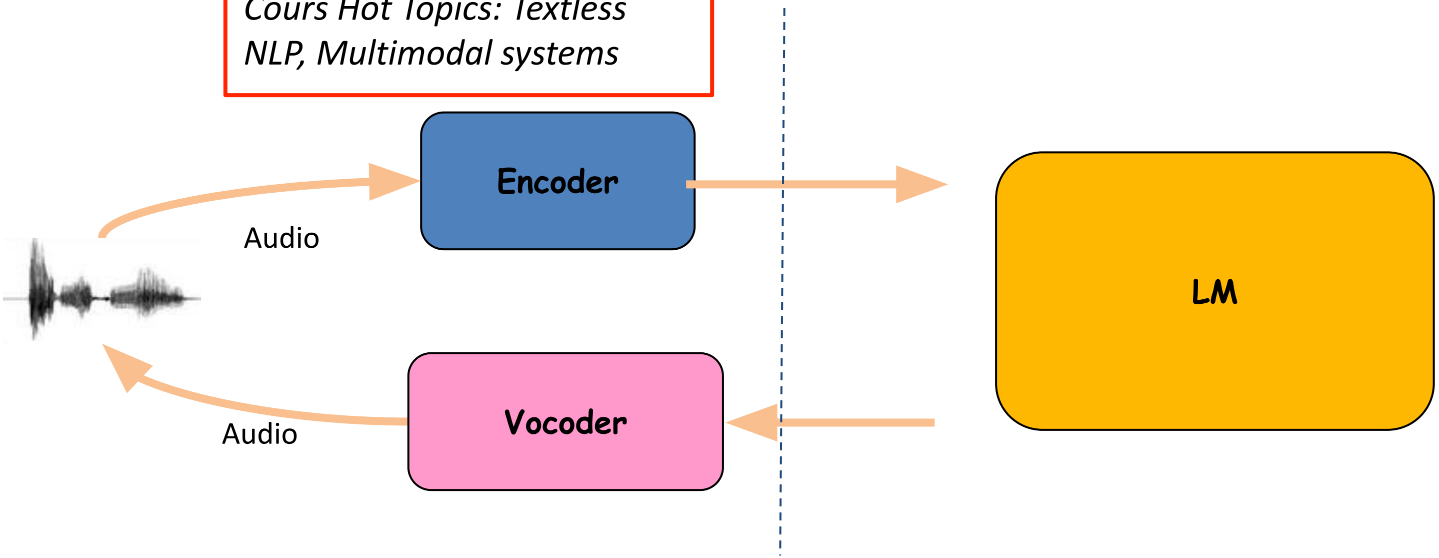


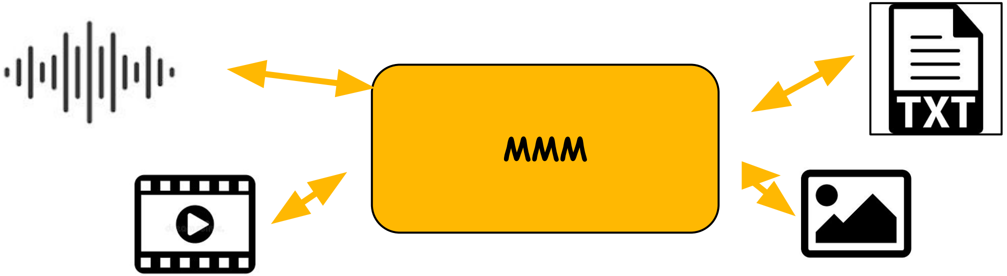
Inside the box: language processing



Inside the box: language processing

Cours Hot Topics: Textless NLP, Multimodal systems





Course logistics and evaluation

Course logistics

The course involves classes and a project. It spans over **10 weeks:**

- 9 weeks with an approx. 2-hour class+30-40min of quiz+debrief+Q&A (more on quizzes on the next slides)
- A 10th week for project defenses

Up-to-date information can always be found on the course's GitHub page:
https://github.com/edupoux/MVA_2023_SL

Teachers can be contacted by e-mail at:
mva.speech.language@gmail.com

Course details (see GitHub page for precise dates):

#1: Intro (Sagot & Dupoux)

#2: NLP1 (Sagot & Algayres) +Q

#3: NLP2 (Sagot & Algayres) +Q

#4: NLP3 (Guest: Schwenk)

#5: NLP4 (Sagot & Algayres) +Q

#6: ASR1 (Dupoux & Zeghidour) +Q

#7: ASR2 (Dupoux & Zeghidour) +Q

#8: ASR3 (Dupoux & Zeghidour) +Q

#9: Hot topics (Sagot & Dupoux)

Evaluation

- 6 on-line Quizzes (courses #2, #3, #5-#8)
  - 30% of the final grade (on your 5 best scores out of 6 quizzes)
  - is open right after the class, for 15 min
  - Followed by 15min of discussion on the quiz's questions and answers

<u>No compensation (pas de rattrapage)</u>

- 1 Project
  - 70% of the final grade
  - Based on a recent paper+code
  - Objective: replicate the paper and extend it with an additional experiment
  - Group: 2 to 4 students
  - One-page outline (due on week #3's Thursday evening at midnight — see GitHub page for confirmation)
    Strictly enforced deadline (1/24th point /20 substracted every late hour)
  - Oral presentation (10min presentation + 5min questions) + 4 pages of summary (week #10)

Questions?