LING 101: Introduction to Linguistics

Salvador Mascarenhas
Michael Goodale (TA)

21 Septembre 2023

Contents

1 Language: Psychological and Social Entity

Introduction

DEC is called like that because is would have bothered litteraries Prof: Master at NYU, Junior Research at Oxford, here since 2016. Using language as a window into human mind
Linguistics is a broad term for serious principled study of language. Many perspectives, from the cognitive point or study language by an external perspective (structuralism: Ferdinand de Saussure, or Leonard Bloemfield) looking out on structures. Also quite general here. Language looked at as a social entity. History of languages, typology of languages. Not only about cognitive studies, pretty broad look out.
Teaching assistant: Michael Goodale PhD student, LRSCP, computational models of Language Acquisitions. Statistical inference and model language after formal tools. Practical Skills really.
Assessments: Homework, graded on a qualitative schedule, due on lecture days. Final: Last Lecture, 30% of grade 10% of grade in TA participation. Website: Moodle hosted by Université de Paris, Syllabus Syllabus:

Schedule

1. Language as a psychological and social entity
2. Language (non-)variation: Universals, variation within parameters
3. **Morphology**, language typology
4. **Syntax I** constituent structure; selection and subcategorization
5. **Syntax II** subcategorization; X-overline theory
6. **Semantics I** first look at meaning *Studied by Salvador*
7. **Semantics II** philosophy of language and the case for methodological solipism.
8. **Phonology** phonetic macro classes;
9. Language and Reasoning
10. Language and Reasoning II
11. Language in the brain; deficit-lesion method; functional brain imagining; psycholinguistics; parsing, reading, lexical access.
12. Language and thoughts in minds vs. machines

1.1 Remarks and Observations about the Nature of Language

R.Descartes *"Discourse on the Method"*: Humans, everyone of them, can speak. Animals, though they have what is needed, can't express their thoughts. Insights:

1. It doesn't matter on general intelligence, social intelligence or any measure of your intellectual abilities. Happens despite any other difficulties.
2. *To Our Knowledge* Not any other animal can do what we can do. Article from the *NYT*, saying animals **can** speak, though it's emoverlineassing. A Chasm *appears* between humans and animals. Yet, it is continuous of what happens in the animal kingdom.

3. Animals are not incapable of language because they can reproduce human language, or use signs to communicate. Studies on Non-Human primate vocalizations, 3-4 words, all alarm calls → Language is independent on organs and communication systems. Yet is it a panic reaction or a real communication. Cannot conclude on A.Is then...

1.2 The Goal of Linguistics

A complete understanding of how sound (/sign/etc) relates to meaning:

- in terms of the speaker's knowledge: the state their mind is in by virtue of having acquired a natural language (competence). Distinguished from mastery of language/way it is produced. Describing skill, not usage.
- in terms of using that knowledge in linguistic tasks, like uncovering meaning from sound in real-time comprehension; executing motor commands necessary to externalize meaning in language production (performance)

Competence/performance is not really a sort of Chomsky. Chaz Firestone (Yale) published on competence/performance saying machines have been tested on performance and not competence. Tight connection between thinking and speaking. Behaviourism = school of thought that tried to figure out a way of studying humans that postulated and said absolutely about our mind. Not only linguistic behaviour: we shouldn't describe what happens in people's heads, just study what outputs comes from what inputs, without looking at the stimuli. Not what we will postulate. Freud sucks. Cannot look at functions: functions = algorithms, studied based on input/output pairs. Cannot do deduction nor induction but only abduction. Yet, we have no less reasons to believe it is right than to believe black holes exist.

1.3 Different levels of Study

Example: *Mushrooms are an edible fungus*

1. Sound Categories: Studying the sound signal based on the phonemes, represented in the mind.
2. Morphemes: first chucks of phonemes that has a meaning: Morphology. Here: *Mushroom* and *z* or *edi* and *able*. Sometimes they are not pronounced: need for a rigorous description. Sometimes, they're redundant, and appear with the same meaning in different places: compare theories. What is the probability of that happening ? And how about irregularities: Past = laugh-*ed* or g*a*ve ? Past in a concept that can manifest itself in different places: simple theory. FMRI theory can improve this theory. Morphemes don't always come in the same orthograph nor sound.
3. Words: Not much to do here.
4. Semantics: Organizing words into phrases. Here: *edible fungus* is a phrase, but *edible* is not. Must be done formally.

Three way of looking: Us, looking from a native's judgement - introspection **will** answer some questions. Exaggeration, yet: no written language, only looking at spokn language.

1.4 Language and Societies

1.4.1 Language and Classes

Language display depends only on human factors, political relationships, genetic factors: distinction between the animal and the meat (names coming from French, spoke by the upper class) in english. Happened in other languages. Different from hyperonyms like *poultry* for *chicken*. Context can explain linguistic aberrations.

1.4.2 Language and Dialect

Language and Dialect are political constructs and arbitrary decisions:

> A language is a dialect with an army and a navy. (M.Weinreich)

1.4.3 Infinity of Language

Sentences are of arbitrary length, and can always be augmented. Yet infinite-ish sentences are impossible to comprehend because Performance is finite, i.e. cognitive resources are finite. There is a *finite* system generating infinitely many linguistic representations: recursions are of the order.

1.4.4 Description

Not looking for rules prescripting language (fuck l'académie française), but only for rules describing them. No better way to speak, the way you ought to speak has nothing to do with linguistics and only with politics. Yet, language are principled, even those which are *proscribed*: adding *fucking* in the middle of a word: *Phila-fucking-delphia*. Rule here: *fucking* comes before the stressed syllable and the material right before needs to be heavy. Heavy comes from phonology, rigorous theory of the weight of syllables.

1.4.5 Phonological Differences

Languages have different constraints on the syllables composing their words: **pnick* works in French but not in Engl*sh.

1.4.6 Internal Structure of Sentences

Sentences are made of constituents that don't act up the same in every language: *des* is not used in Engl*sh ($\sim$ *of the*). They cannot be broken: *des burgers et des frites*. It is mysterious tho ? Maybe language has something else to do for us than communicate...
They cannot be considered alone: *Fat cats eat* and *Fat cats eat accumulates*. Supposition that two words next to each other are related in written language. Also, prosody is a big help in understanding.

1.5 TA 1 27/09

1.5.1 Animal Communication

Language is also communication, not only hearing (trees ?). Many (if not most) Animals Communicate, and almost all react to sound. (cf. NYT Article) : dolphins communicating by signs, bees dancing, monkeys having muscles/organs to 'talk', bird songs, ant pheromones, great apes...

Differences between human and animal communication ? Human language has: composition (recursion: meaning of sentence can be derived from its constituents), abstraction, no hypothetical/long term/prevention discussion, intentionality, arbitrary length of sentences, systematic neologisms/nonce words (when learning a word, it is usable immediately), non-instrumental.

Many experiments about teaching great apes language suck and were not really concluent. There is a poor, noisy, contradictive and unrepresenting stimulus that child have to make do with. Deaf child make their own language if they need one. For example of the poverty of the stimulus: 2 Layer Embedding of possessive happened 67 times in 120k sentences, while kids at 6 can do 4 level possessive embedding.

The words *stop, mat, tap, butter* all have *'t'* yet have different sounds: there is a sense where this is the same sound, but another one where they have different sounds.

2 (Non-)Variation and Languages

2.1 Different yet Similar

Langugages are made of signs, composed of a form and a meaning. They only look like they have multiple forms/meanings, on another level of analysis they only have one, e.g. past tense in english and the morpheme *ed*. Variants depend only and purely on properties of the root, and are entirely predictable. With *bat*, there is an ambiguity phenomenon, but there is also a phenomenon of polysemy, e.g. *book*, also, similar meanings often derive from a central point.

> A sign presents itself to the senses, and something distinct from itself to the mind - St-Augustine

There are two types of signs based on the link between form and meaning:

1. Those with a causal link: *smoke* implies there is a fire
2. Those with a arbitrary, conventional link: *black attire* implies mourning (*in some cultures but...*)

After Ferdinand de Saussure,

> Language is a systemn of conventional (arbitrary) signs.

Example: The word *man* in different languages: German Mann, Spanish hombre, Français homme, Hungarian ember, Turkish adam. With the sound produced by the rooster, the words differ in languages, but there is still a partial causal link, because you cannot mimic perfectly the sound of the animal, given the differences between vocal organs. Arbitrary doesn't mean random: it just doesn't matter what choice is made. It is no accident there is a resemblance between German and English words for *man*

2.2 Similarities between distant unrelated languages

2.2.1 Reciprocal pronouns

Pronouns marking reciprocity always have a mysterious constraint where they must be in the same, finite clause as their antecedents: *They thought I talked about each other* seems weird.
Generally, it seems that reciprocal pronouns must refer to a thing that lies in the same proposition, but why ?
First question is: Do we have a reason to say why they seem weird, just because they are longer ? No, *I thought that they talked about each other*, is equally long as *They thought that I talked about each other* and doens't sound half as bad.
Then, the sentence *They thought we talked about each other* shows it's not about third person.
Coming up with a precise answer implies looking for phrases where each bit of the sentence has been replaced, one at a time, to isolate the *issue*.
In *They thought that I talked about each of them*, *each of them* is not necessary reciprocal, as it only include (*each other*)'s meaning so it is compatible.

2.2.2 Sentivity to negative elements

Words occur sometimes with negative elements. Every natural language seems to contain at least one lexical items that is sensitive to whether the context in which they occur exhibits negative or positive polarity.
Exemple: *Jean a fait le moindre effort* doesn't seem natural, but *Jean* n' *a* pas *fait le moindre effort* does.
Facts are subtle: the mere presence of a negatvive item is enough to license a negative polarity, and sometimes negative polarity is infered without the obvious presence of a negative item. A sentence with an empty slot in place of the negative polarity item is the context that needs to be negative in order to license an NPI.

Positive polarity also exists: *John didn't see someone* is really weird, and it requires a really particular prosody and/or context to work. You have a meta interpretation of this. The word anyone, more than a negative item is also a free choice item.

2.2.3 Relations between Sentences

In all languages, declaratives and interrogatives are linked by *transformations* ; i.e. a reorganisation of the terms of the declarative to build the interrogative. It creates pairs of assignations, not necessarily questions and answers. Also some declaratives link to multiple questions.
There is a finite palette of strategies for these transformations: no known human language builds questions by mirroring completely the order of the words. Grammars do not count, e.g. there is no language with transformation swap words 1 and 2.

2.2.4 The Puzzle

Languages are systems of arbitrary signs. Any sign will do for internal calculation, and any convention widely known within a given community will do for communication. This observation does not predict the existence of strong systematic, pervasive similarities whose speech commuinties have had no contact.
How can the conventional character of language be reconciled with pervasive cross-linguistic similarities.

> We would learn so much if we could do horrible things to babies. We can do horrible things to animals though, not saying we should, but we can. - Salvador Mascarenhas

2.3 Universal Grammar

2.3.1 Chomsky's Hypothesis

Human faculty for language:

- enables humans to acquire and use langugae;
- delimits what linguistic structures humans are capable of acquiring and using
- delimits the kind of linguistic convetions that a community of humans can develop and successfully hand down to new generations

Thus we need to know all about gene reproduction, biological evolution and so on, to understand it fully
Universal Grammar in this sens is section of all humans' biological endowment and it is reflected in all natural languages.

2.3.2 Universal Grammar

This is not a language in itself and it doesn't imply the idea all languages come from a common source. This does not imply there is no other factors of relevance shaping actual natural languages than *a priori* constraints, there being no rhyme or reason to those constraints. Chomsky added it might be to optimize computation. This is *not* a collections of generalizations about trends in linguistic diversity, other projects (Dryer, 2005) have been doing it, especially on word order.
It has for purpose to classify the languages the human mind can learn and the *impossible* ones. It is a two steps projects: Finding facts, photographing the human languages, then deriving generalizations.

2.3.3 Language Acquisition

2.3.3.1 Critical Period

There is a critical period for human (and animal aswell) acquisition of language: post-puberty, acquisition is severely impaired, e.g. Jeanie, kid discovered in L.A. in 1970 at 13.5y. Any child can learn any language, it depends on features of the environment.
Acquisition doesn't seem to be full related to mathematical/intellectual/reasoning skills and so on... Learning a second/third language is totally different tho, it is correlated to musical and computational skills it seems, there even seems to be a purely linguistic talent. Almost nothing about adult phenological is due to purely biology and genes.

2.3.3.2 Problem of Induction

Grammars make predictions for infinitely many word sequences, yet the input if finite. Therefore, there is something not in the input is playing a key role in learning.

2.3.3.3 Absence of Negative Evidence

Experiments, can give negative information: they can show that under certain conditions, an outcome is *not* observed, but a child has extremely limited access to such negative data.

2.3.3.4 Conclusion

A child is far better at figuring their native language than a linguist. Linguistics is an empirical science not a fundamental one.
Linguistic data consist of judgements on utterances, grammatically judging strings of characters, truth-value judging sentences/scenarios. It can come from introspection, but introspection is limited: there is a limit to what one can infer, e.g. *I have more pictures of my kid on my phone than my dad ever saw me* isn't grammatical, but it is understandable. The unboundness of human mind might mean we can find sense in any sentence.
The better way is to speak with other linguists or conduct experiments on a large amount of naive participants. There is a risk of falling into delusion from thinking to much about the same sentence.

2.3.3.5 Chomsky's Argument

There is such a gap between what a child is exposed to and the sophistication of the acquired grammar the child must have expectations as to what a language can be, analog to the concept of triangle: You've never seen a triangle but you know what it is. Mathis Hademeyer
Universal Grammar is a set of principles with a number of free parameters.
Principles and Parameters is a method now limited to small studies on variations, mostly in Italy (just you and me). By this we mean the fact that language, unrelated, share a structure. In fact most languages only occupy a minuscule section of the mathematical space of possibilities for a language. Principles and Parameters is a way to cash out on this idea: the problem is figuring what values has a child gotten for its parameters, and how he can deduce from some principles, the structure of language.
The null-subject parameter: In Catalan, *he speaks* is said $\varnothing$*parla*, there is no need for a pronoun, contrary to French or English. How do children hande this situation ? It seems that Italian children go through a phase where they omit subject when learning French.

2.4 TD 2: 04/10

2.4.1 Homework 1

2.4.1.1 Question 1

The word *nous* is used in formal context, so it is expected that people use formal grammar in the whole sentence, hence the *ne* of negation is expected.

2.4.1.2 Question 2

You can drop the copula in AAVE when you can contract the verb.

2.4.2 Exercise

2.4.2.1 Swedish Extrapolation

Separate the sentences in groups of morphemes, then extrapolate the meanings based on the other sentences.

2.4.2.2 Rule Extrapolation

Language is Chikisaw or something.

3 Morphology: The Structure of Words

3.1 Use of Morphology

Do you spell kick-ass or kickass ? Is it one or two words ? This is not very well understood.

3.2 The Study of Morphemes

Morphology is the study of words, and morphemes. Morphemes are the smallest linguistic unit that makes sense, e.g. *a* or *I*. The sound has no meaning, but the morpheme has. It is the smallest entity with both a form and a meaning.
Words are formed by processes, so they have a constituent structure, they can be constituted from multiple morphemes.
They can be derived from rules of morphology. Languages have a huge range of variations in what they can do.

3.2.1 Words

A Word is an indentifyable unit of phonology with a prosody (cf. phonology class).

Words have a structure: The *noun* (section of speech) *reuseables* is made of four morphemes: *re, use, able, s*. Here, the morphemes are, in order, *a prefix, a root, a suffix, a suffix*.
Words can be derived from other words also: *to invite - an invite*. They can then have an ambiguity in sense: is *reusable* a noun or an adjective.

3.2.2 Morpheme Types

Prefixes, suffixes, infixes (e.g. *abso*-fucking-*lutely*), circumfixes (e.g. em-*bold*-en - not a good example as historically it might have never been a circumfix) together are called affixes. Thos morphemes are called bound morphemes, meaning they cannot appear by themselves (e.g. *mang* root of *manger* in French)
When looking at the derivation, the thing that is not an affix is called the stem. The root is the smallest stem.
Things called clitics can also come into words, e.g. *l'*, it has more information than another morpheme: *l'aime* has two concepts in it.
Roots and Stems are called open class morphemes, new instances can easily emerge or be invented, e.g. *blick-ing* would be easy to understand. Affixes are closed class morphemes, meaning new instances develop slowly.

Inflectional morphemes only add grammatical information. They can have the same sound shape as some derivational morpheme. Derivational morphemes on the other hand, create new

Table 1: (Simplified) Derivational Rules in English

Affix	Rule	Output
-able	Verb + -able	= Adjective
re-	re- + Verb	= Verb
un_1-	un- + Adjective	= Adjective
un_2-	un- + Verb	= Verb

concepts out of existing ones.

There is a universal attested word-structure: [[[root] [derivational affixes]] [inflectional affixes]]

3.3 Forming Words

Particular grammars make certain derivational processes available, which we can describe by means of rules.

3.3.1 Derivational Processes

A rule specifies the category of the input and the category of the output, see below. Affixes typically add further restrictions on stems.

3.3.2 Constituent Structure

From a derivational analysis of a word in terms of the processes that generated it, we can extract its constituent structure. We can represent it in a sort of phylollogical tree.
For *un|enjoy|able*: *unenjoy* is not a word, as you cannot reverse enjoyment. Even though *un-see* exists, it is more of a creative product of language as you might not understand its meaning at first glance; it is understood because it is frequent. So we get a structure like: Node(un, Node(enjoy,able)). We might go deeper in analysis, but there is no real rule behind *en-joy*.
For *re|read|able*: We cannot base an argument on a stem that doesn't exist, because both hypothetical stems do. Yet, as *re-* does not work with an adjective, we can derive the structure to be: Node(Node(re,read), able)
For *un|wrap|able*: Here, both Node(Node(un, wrap), able) and Node(un, Node(wrap, able)) are valid processes of derivation involving different rules. The second is the negation of the adjective *wrapable*, which means something that cannot be wrapped. The first on the other hand, is something that can be *unwrap(ped)*, since we add *able* to a verb. This is an example of structural ambiguity, that is often disambiguated by prosody e.g. *fat cats eat...* Yet, this not only allows to express ambiguity, but also to explain it.

3.3.3 Compounding

This is a quite mysterious area of morphology.

Compounding is a process to build words from two or more stems. It constitutes in combining those stems with a novel meaning, with lower predictability than derivational processes, e.g. *bitter|sweet*. There is compounding between, adjectives and adjectives, adjectives and nouns, nouns and nouns, nouns and verbs, verbs and verbs, verbs and nouns... There is a very blurry line between what people think and historical reasons, are people thinking of *sleep* and *walk* in *sleepwalk* ?
It can be an ambiguous process: If a *houseboat* is a boat that is a house, what is a *housecat* ? English compounds are typically headed by their rightmost element, meaning the precedents are qualifying it, but sometimes compounds are headless: *kick-ass*, *ceasefire*...
Are there more rules ? Is an N-N compound necessarily a *for* relation, a *from* relation ? Is there simply just a hidden preposition in it ?

3.4 TA 3: 11/10

3.4.1 Homework 2

Nothing to add, see file.

3.4.2 Cross-Linguistic Variation[1]

Languages can be put on a spectrum from analytic to polysynthetic, based on how much they use morphemes and/or syntax. Analytic use the most syntax, polysynthetic don't care about it and only add morphemes.

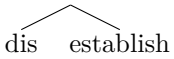
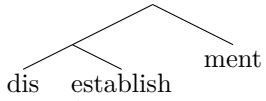
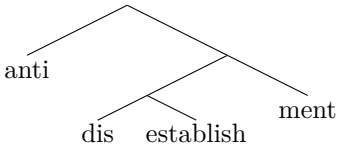
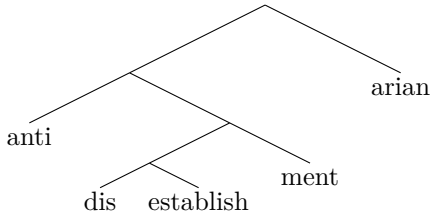
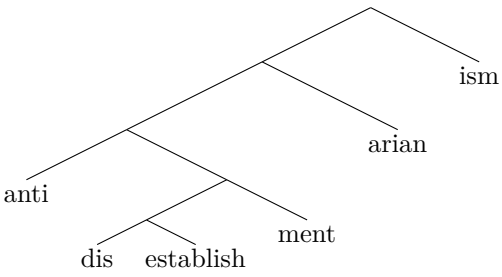
- Analytic (Isolating) languages where each morpheme is a word on its own (e.g. Chinese)
- Agglutinating languages where each grammatical bit of meaning is an affix to a stem (e.g. Japanese, Korean, Hungarian)
- Fusional languages where an affix or change in the stem can lead to multiple grammatical meaning variations (e.g. Latin, German, Arabic)
- Polysynthetic Languages where affixes have extremely rich content (e.g. many native North American languages)

Both agglutinating and polysynthetic languages pack a lot of information into a single word, the difference being that agglutinating use affixes for grammatical meaning only.

3.4.3 Morphological Derivation

3.4.3.1 Deriving *Antidesestablishmentarianism*

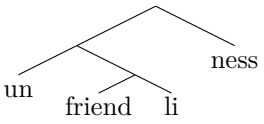
We can see this word from a sequence of trees:

	Tree:	Explanation
Step 1:		$dis : verb \mapsto verb$
Step 2:		$ment : verb \mapsto noun$
Step 3:		People are against *disestablishment*, not the people
Step 4:		$arians : action \mapsto peopleinit$
Step 5:		$ism : people/concept/action \mapsto idea$

[1] This is section of Lecture 3, see the handout

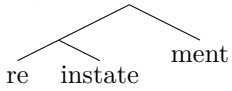
3.4.3.2 Deriving *unfriendliness*

We only give the tree:



3.4.3.3 Deriving *reinstatement*

You cannot attach *in* to a noun, it needs to be attached to a verb, but as it is a rare example of borrowing from latin, we cannot separate *instate*

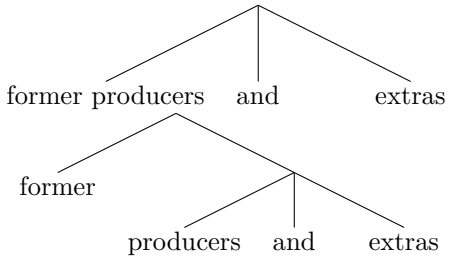


4 Syntax 1

4.1 Syntactic Competence

Syntactic competence is a wide concept, it includes :

- The knowledge of what sentence is grammatical and what isn't. This comes from the fact there is no syntax without semantics in linguistics, unlike in arithmetics. The idea of an autonomous syntax has disappeared since the 70s. It is however useful to talk about syntax in a somewhat disconnected way from semantics. For example, you cannot contract auxiliaries when they carry meaning: *I shoulda bought it* and **I shoulda more money*.
- An abitity to recognize a sentence as well-formed even if it does not make sense or consists of highly sequential transitions, meaning, even though some transitions have really low probability, e.g. *Colorless green ideas sleep furiously* and *Colorless Green* or *sleep furiously...*
- An ability to determine what the sentence means, whether it is ambiguous or not either structurally or lexically:

lexical e.g. *I went to the* bank	bank $= \begin{cases} \text{financial institution} \\ \text{river bank} \end{cases}$
structural, e.g. *former producers and extras*	

- Every language has a countably infinite set of sentences (proven by showing there is no longest sentence in a language). Native speakers can handle this set despite having finite minds. We can construct arbitrary long sentences based on the center embedding idea, i.e. *le jeu du Johnny Depp*.

4.2 Syntactic Analysis Building Blocks

Before writing rules for syntactic analysis, we need blocks to build those rules on.

4.2.1 Lexical Categories

The lexical categories (or parts of speech) such as *N(oun)*, *V(erb)*... This does not mean you can only tell what section of speech a word is if you know its meaning.

The thing about thing is that everything and anything is a thing - Salvador

Definition 4.2.1. *We define lexical categories given the sentential environments in can occur in and the affixes it can take, i.e. the grammatical distribution of the word.*

You shall know a word by the company it keeps - Ferguses

For example, a verb in English satisfies all of the following:

1. It can occur right after the auxiliary *will*
2. It can take the endings *-s* and *-ing*

4.2.2 Constituency

Definition 4.2.2 (A Heuristic)**.** *A sequence of words forms a constituent in a sentence if:*

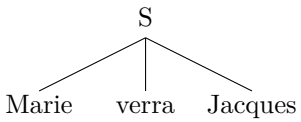
1. *It can be replaced with a minimal unit, ideally one word, presevervin grammaticality and conversely*
2. *Occurences of that minimal unit can be replaced by the sequence of words, preserving grammaticality*

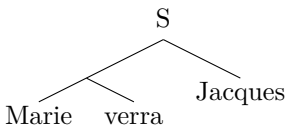
If those conditions hold, then the sequence has the same category as the minimal unit.

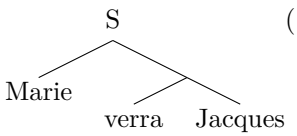
A constituent can also be found by Fronting and Pronominalization, Clefting, Eliding or Coordination

4.3 Rewriting Grammar

Marie verra Jacques leads to three possible structures in principle:

- 

- 

- 

(*verra Jacques* is a Verb Phrase = VP)

We might build a rewriting grammar for a fragment of French:

1. *Marie verra/entendra/... Jacques*
2. *Un type verra Jacques*
3. *Marie verra un type*
4. *Le type connu a vu Marie*
5. *Marie dormira*

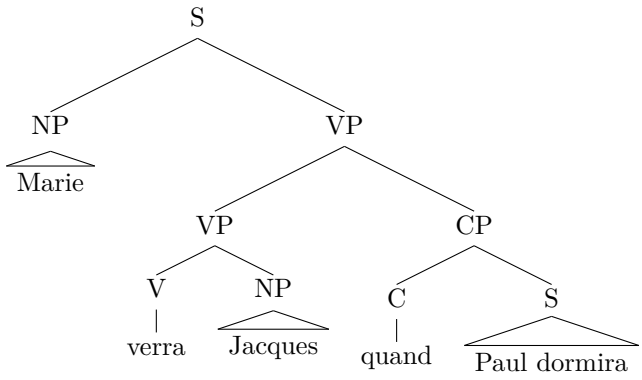
6. *Marie pense que Paul verra Jacques*

7. ...

We then can get an idea on French Grammar:

Id	Node	$\longrightarrow$	Rephrased as
0	S(entence)	$\longrightarrow$	N(oun) P(hrase) + VP
1	NP	$\longrightarrow$	Proper Name
2	NP	$\longrightarrow$	Det + N + (Adj)
3	VP	$\longrightarrow$	V_{intr}
4	VP	$\longrightarrow$	V_{tr} + NP
Up to this point, we cannot generate an infinite number of sentences			
5	C(omplementizer) P(hrase)	$\longrightarrow$	*que* + S
6	VP	$\longrightarrow$	$V_{cl(ausal)}$ + CP
7	V_{tr}	$\longrightarrow$	voir, entendre,...
8	V_{intr}	$\longrightarrow$	dormir, briller,...
9	V_{cl}	$\longrightarrow$	penser, croire,...
There is now a loop in rules, so there is recursion and we can generate an infinity of sentence			
10	C_{temp}	$\longrightarrow$	avant que, quand
11	CP_{temp}	$\longrightarrow$	C_{temp} + S
12	S	$\longrightarrow$	NP + VP + (CP_{temp})
	VP		a) V_{intr} + (CP_{temp}) b)V_{tr} + NP + (CP_{temp})
14	S	$\longrightarrow$	S + CP_{temp}
15	VP	$\longrightarrow$	VP + CP_{temp}

We can then infer a tree from a sentence with the lexical categories:



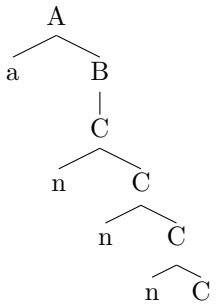
4.4 TA 4: 18/10

4.4.1 Homework 3

Nothing to say.

4.4.2 Study of an abstract language

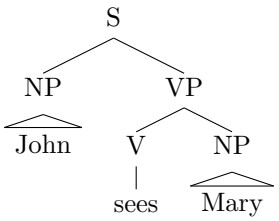
A→aB
B→ c
B→ z
B→ C
C→nC
C→ d



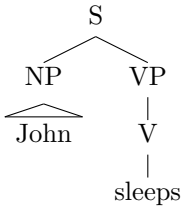
This gives us many strings possible: *ac*, *az*, *an***d*...
We can draw trees from those strings, e.g. *annnd*, see above.

4.4.3 Structure of sentences

4.4.3.1 John sees Mary

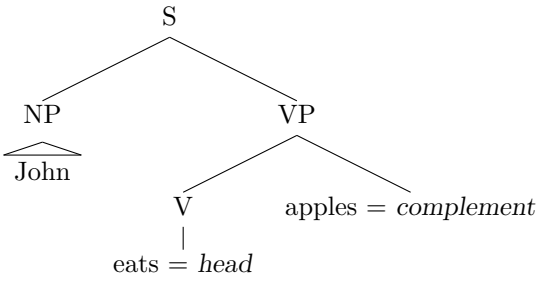


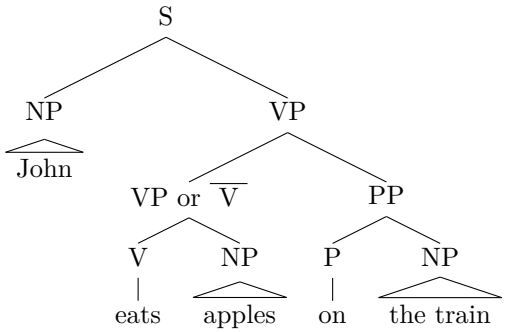
4.4.3.2 John Sleeps

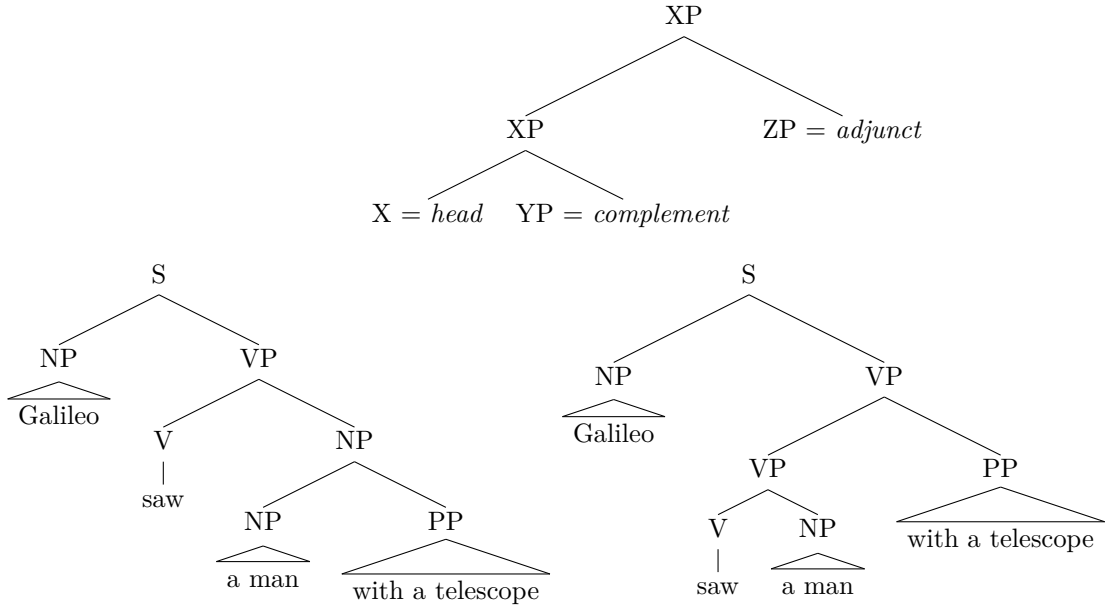


4.4.4 Heads, Complements, Adjuncts

Definition 4.4.1. *A head is a word. A* head *is connected to a phrase. Anything that is the sister of a* head *is a* complement*. When a* complement *is unnecessary, it is called an* adjunct*.*

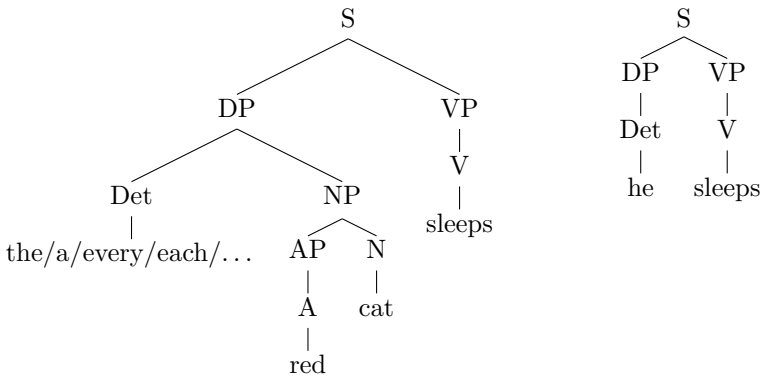




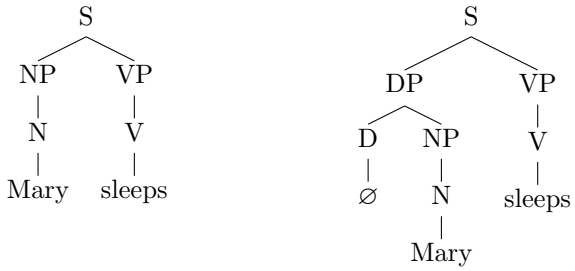


4.4.5 Structure of Noun Phrases

To analyze Noun Phrases with adjectives, we can base on the model from the first example below, introducting Determiner Phrases. We can do the same with verb phrases, seeing pronouns as determiners (this is called the DP hypothesis):



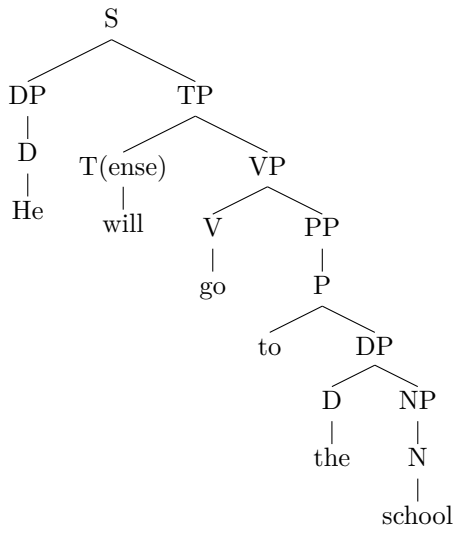
We may guess that in the sentence *Mary sleeps*, *Mary* is a determiner, yet we could say *every Mary sleeps*. There are two ways of seeing the sentence *Mary sleeps*: either *mary* is a noun phrase or *Mary* is a determiner phrase with an empty determiner.



The latter is better from the study of other languages such as Portuguese where the determiner is not empty: *a Maria dorm*.

4.4.6 Teaser for next class

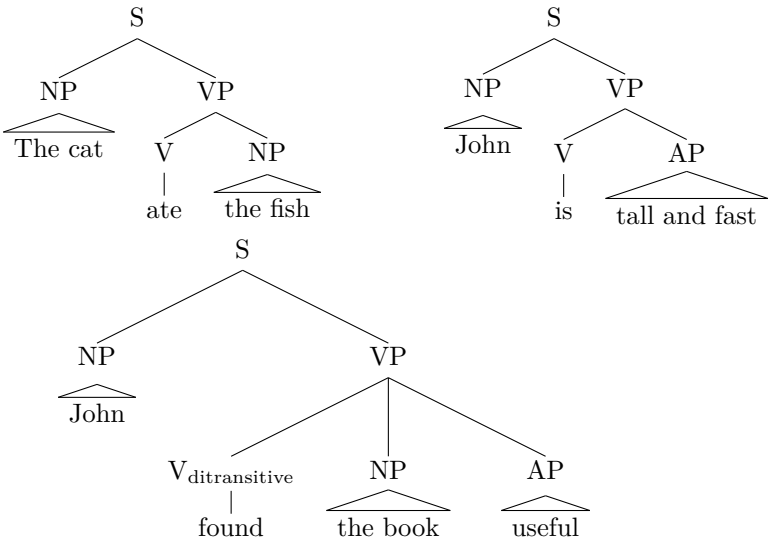
Consider *He will go to the school*:



We here only did binary branching, but there might be ternary branching sometimes.

4.5 TA 5: 25/10

4.5.1 Kinda like Homework



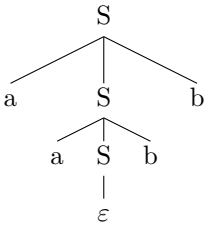
We could leave the VP in the last example with only two children by leaving the topological order. We would need to implement inside the heads a rule to tell the order of pronunciation. Another way to achieve this is by saying *the book useful* has secretely a more complex structure such as *the book to be useful*.

4.5.2 Grammar Generating a Language

Language: $L = \{a^n b^n \mid n \in \mathbb{N}\}$
Grammar:

$$S \to aSb + \varepsilon$$

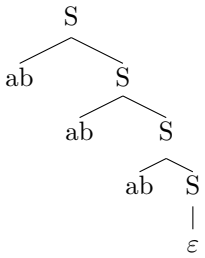
We get, for *aabb*:



Language: $L = (ab)^\star$.
Grammar:

$$S \to abS + \varepsilon$$

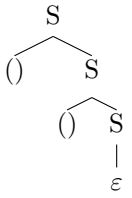
For *ababab*:



Language: Well-Parenthesized Expressions - Dyck's First Language
Grammar:

$$S \to \ \varepsilon$$
$$S \to ()S$$
$$S \to (S)$$

For *()()*:



4.5.3 On Disambiguation

John looks like Mary might mean *John looks in the same manner as Mary*. Syntactician do not care about priority rules, they don't want to disambue their grammar. They just add probability to their production rules, to generate parses probabilistically.

5 Syntax 2

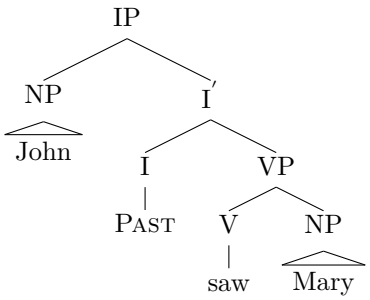
5.1 About Homework 4

In the sentence *The milk perished*, we cannot interpret *perished* as a verb directly, but as a verb in a verb phrase.
The sentence *Time flies like an arrow* can be read as *Time flies* (a particular kind of fly) *like an arrow* would then mean that a special type of fly likes a particular arrow. Same thing: *fruit flies like a banana*.
The charges against him may be read by the rule $NP \to NP\ PP$ or by looking at *against him* as an argument of the noun *charges*. By definition, grammatical operations can only target nodes of the syntax tree. Then, we cannot really have noun phrases with structure $NP \to Det\ N\ PP$. Yet, as all X-phrases have an X in them, except for S(entences), it seems unreasonable to regroup the noun and the PP inside a block inside a NP. We then might want to introduce intermediate generalizations with new parameters such as $\overline{X}$.
In *The accident deprived him of his mobility*, we cannot regroup *him of his mobility* under one block, and then we may look at *deprived* as having two complements, *him* and *of his mobility*.

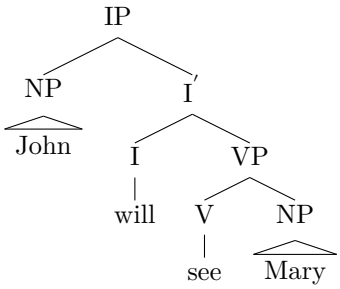
Passive sentences cause mayhem in syntax, so transformational grammar was created to derive transformation rules that would modify an hypothetical arrangement of the informations in the sentence into the arrangement that comes out of speaking.

5.2 Inflection Phrases

Modern syntactic theories postulate the existence of Inflection Phrases (IP) instead of S phrase. This addresses the emoverlinerassment of having a phrasal constituent without a head. The inflection (tense, mood, person/number), seen on the verb, is the head category of a sentence. Then, using $\bar{\text{I}}$ we can retrieve the information on the tense/modality by separating the verb from the inflection:



In sentences with modals or auxiliary verbs, we can see phonologically overt material occupying the I node, e.g.:



Most importantly, certain facts about the placement of adverbs help us figure out where precisely in the tree the main verb is to be found. The position is dependent on the language.
Inflection phrases can explain the difference in positions of the word *often* and *souvent* in English and French.

5.3 Displacement

From Inflection Phrases, we can derive the structure of sentences from others by using transformational rules.
Many pairs of sentences in English are intuitively "variations" on each other. They are related by displacement, in that one appears derived from the other by moving certain constituents around. In the first formal theories structures, the pairs (Affirmative, Interrogative) were taken to have a common D(eep)-structure but different S(urface)-structure. S-structures are the results of applying certain transformations to D-structures.

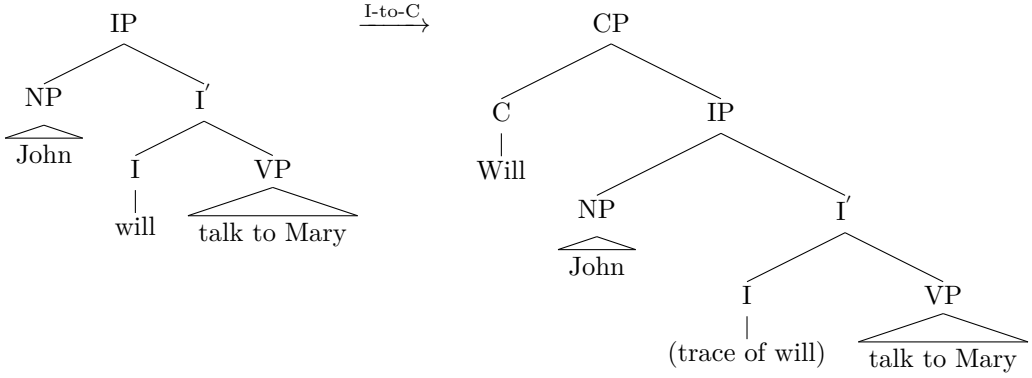
Remark 5.3.0.1. *Contemporary syntactic theories in the Chomskyan tradition instatiate this intuition differently, taking a more derivational approach.*

From this postulate, we can derive from languages transformational rules. For exemple, the I-to-C transformational rule is used to build a yes-no question from a D-structure with phonologically overt material in I by:

1. Projecting a CP above IP

2. Moving I to C

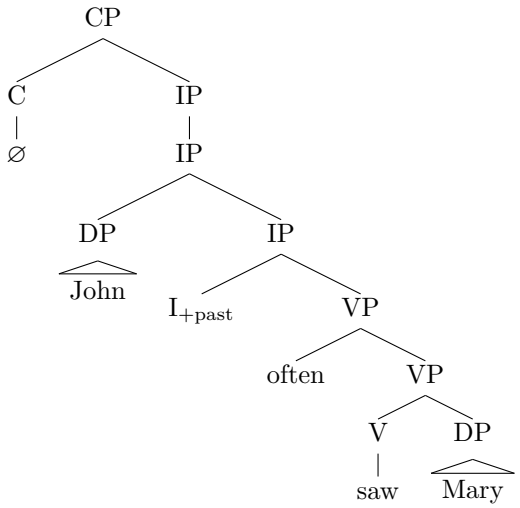
For example, we get:



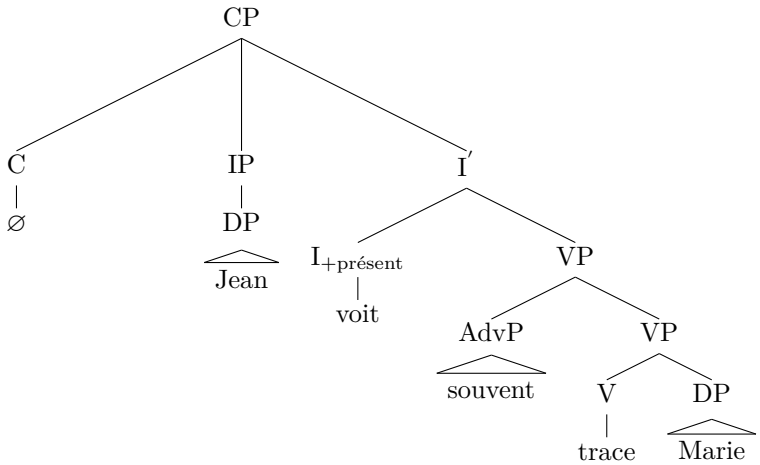
5.4 TA 6: 8/11

5.4.1 On the words *Often*, *Souvent*, and Displacement

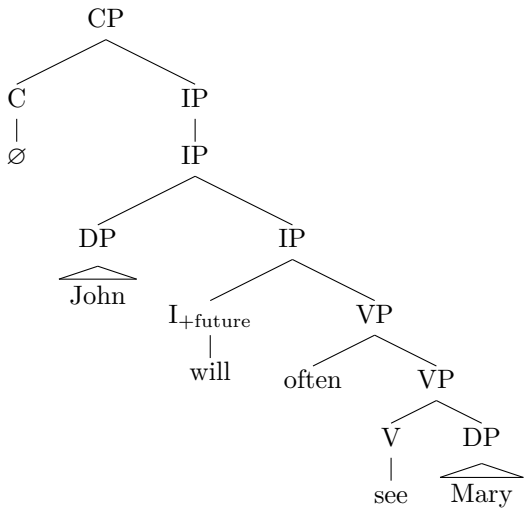
In english, we can do:



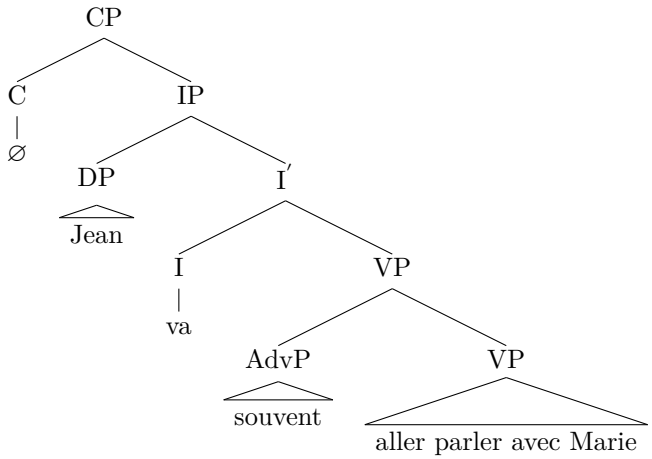
But in French, a similar structure feels unnatural because of a crossing: *Jean voit souvent Marie*. We thus propose to have the verb in the inflection phrase and use a trace in the Verb Phrase:



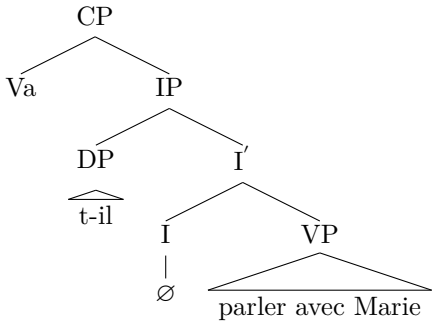
Then, in future tense, in English:

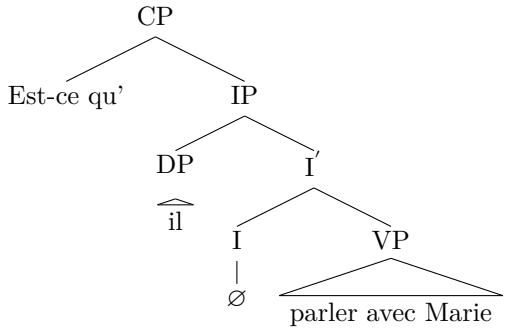


And in French:

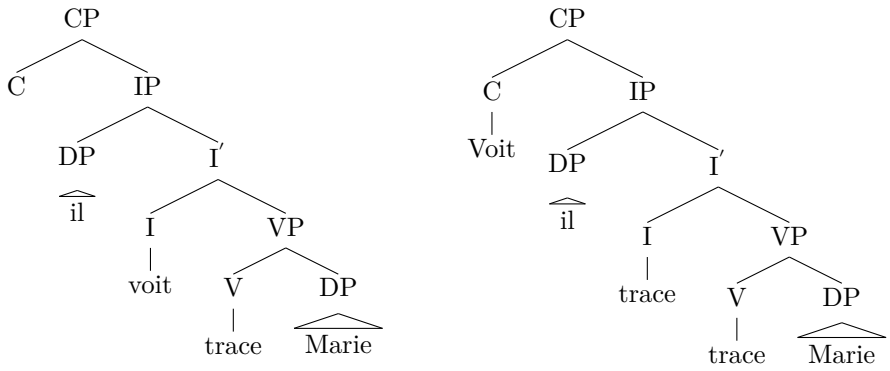


And for questions ? In English, we move the verb before the subject, but in French, there are multiples solutions:



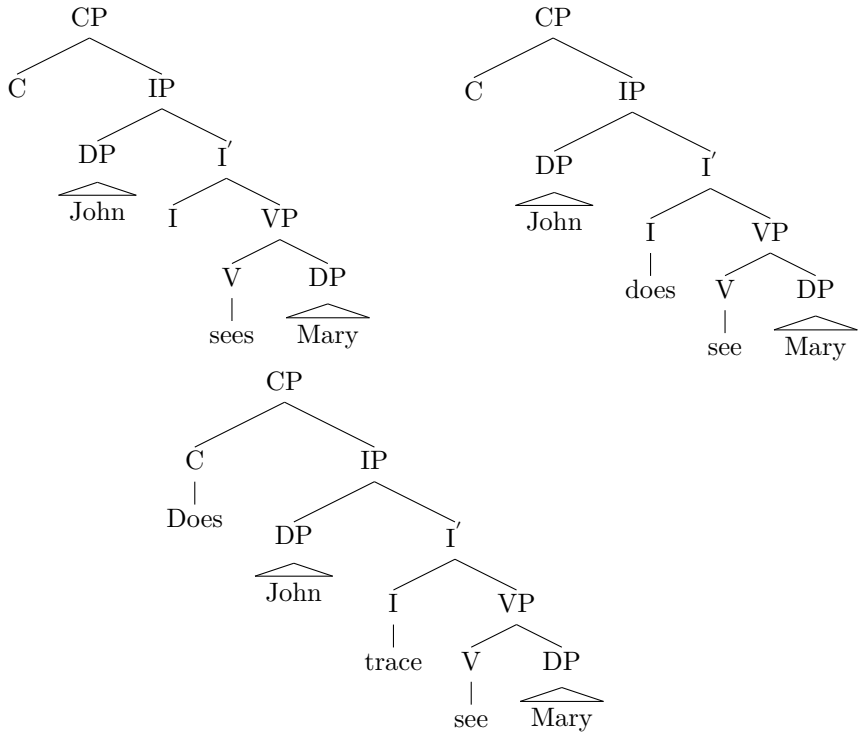


And in the present tense:



The verb here moves twice, from the V to the I then to the C.

We can do the same in English:



6 Semantics I

6.1 Syntax and Grammaticality

Syntax as we studied it in this course has been about how words are put together to form grammatical sentences, with only limited consideration of meaning. For typical artificial languages, we define semantics over an already defined syntax, and we do not need that every symbol has a sense. Yet, for natural languages, we need to ask for that, and there needs to be a connection between semantics and syntax. In linguistics, there is no such thing as a syntax without a semantics.
We need to have things that can be interpreted in particular models. Yet, even tho some words don't have sense or are not words in english (e.g. Jabberwocky) we can figure grammaticality without sense. Same, we can figure out sense without grammaticality, but this is not of interest.

6.2 Semantics

6.2.1 The Study of Meaning

The meaning of linguistic expressions is what we care about, not really their form. Semantics studies:

1. The meaning of linguistic expressions
2. The way in which form and meaning are connected
3. The different kinds of meanings that natural languages use

Two important uses of language are exchanging information in communication and internal calculation.
Propositions are the kinds of meanings that contain information, which can in turn be used in communication and in internal calculation. Indeed, it is very unclear how much (if any) information interrogative sentences contain; they structure, give a goal.
The paradigmatic linguistic exponents (syntactic categories) of propositions are declarative sentences.

6.2.2 Understanding Sentences

What is "understanding a declarative sentence" ?

Consider *The planet Earth is roughly spherical*, *France is a country in East Asia*. One property that propositions have is that they can be true or false. We get a first stab at an answer: *To understand a sentence is to know whether it is true or false.*

But we clearly don't need to know whether a sentence is true or false to understand its meaning: *Salvador likes almonds.* (Here we suppose that the speaker is cooperative, thus Salvador is the one we think about, the teacher.)

> I am the most salliant Salvador in your lives - Salvador.

Sometimes, there is a way to verify the truth value of a sentence *The world will come to an end in 2025*, but sometimes there might not be *Humans will never inhabit Jupiter*, and sometimes it is impossible to be certain *Socrates never existed*, we only have clues it did. We get a second stab at an answer: *To understand a sentence is to know under which conditions it would be true, and under which conditions it would be false.*

This is clearly not enough, how to understand swear words in this paradigm. How to understand questions and imperatives ? In the early days, we tried to reduce interrogatives to declaratives, but it doesn't work, since it implies wrapping a declarative aroung an interrogative.

6.3 Models

Theories formulated in artificial languages are interpreted with respect to standard models: mathematical structures that represent what the artificial language is meant to be about, allowing us to be completely rigorous about the conditions under which statements in the artificial language are true or false.

6.3.1 Models and Interpretations

We have something of such:

- Statement (with sugar): $1 + 2 = 4$
- Standard Model: $\mathcal{M} = \langle \mathbb{N}; s(\cdot), \cdot + \cdot; \leq; 0 \rangle$

- Truth Conditions of The Statement: To apply the successor function $s(\cdot)$ of $\mathcal{M}$ three times is the same as to apply it four times. (This is False under a successor function that works as intended and as can be defined using a decent axiomatization of arithmetic.)

Remember that the interpretation of a formal system (whatever complex), is always given in a natural language. This is only an attemps to reduce the complexity of the interpretation in natural language of a system. The idea is to provide an idea of the number of properties that models need to have to get the truth value.

6.3.2 First Order Logic

FoL is a particularly powerful artificial language that constitutes the logical foundation for most of modern science.

$\exists x$	There is an individual, call it x.
xKf	x stands in the K relation with f
$\forall y$	For any individual, call it y
$yKf \rightarrow y = x$	y stands in the K relation with f imposes y to be equal to x
$\wedge$	conjunction, meaning and
$\exists x.xKf \wedge (yKf \rightarrow y = x)$	There is exactly one individual standing in the K relation with something called f.
$B(x)$	x has the property B
$\exists x.xKf \wedge (yKf \rightarrow y = x) \wedge B(x)$	There is exactly one individual standing in the K relation with something called f. Let's call that individual x.

This table gives a way to model in FoL the sentence *The King of France is Bald*.

This is insufficient to understand the sentence, since we don't know what *King*, *France* nor *Bald* means. We have only explained the word *the*, by asserting the existence and the unicity. $B(x)$ here can be viewed as an interpreation of *is* and xKf as an interpretation of *of*. Yet, here, *of* seems to only be verbal tissue. Here we have understood a section of functional words, but not lingual functions of words, and this is insufficient to end the analysis: *The day was lovely*, is not fully understood, there might have been many loverly days. We then need to use some kind of contextual restriction.
Also remember that, following Russel, *The King of France is not Bald* doesn't negate the whole formula, but only $B(x)$. And, moreover, saying something like *John knows the Earth is flat*, also implies the speaker thinks the Earth is flat.

FoL models are structure which tell us precisely, which individuals there are, what properties they have, what relations they participate in and with whom.
A FoL formula will be true in somme FoL models, and false in others, e.g.:

- In $\mathcal{M} = \langle \{a, b, f, g\}; K = \{\langle a, f\rangle, \langle b, g\rangle\}; B = \{a, b\}\rangle$, the sentence is true.
- In $\mathcal{M}' = \langle \{a, b, f\}; K = \{\langle a, f\rangle, \langle b, f\rangle\}; B = \{a\}\rangle$, the sentence is false since there are two x such that xKf.
- In $\mathcal{M}'' = \langle \{a, b, f\}; K = \{\langle a, f\rangle\}; B = \varnothing\rangle$, the sentence is false since there are no elements in B.

We can take the class of models where the sentence is true to fully specify the truth conditions of the sentence. The task for truth-conditional semantics is to provide an interpretation function that associates each sentence with its class.

6.4 Compositionality

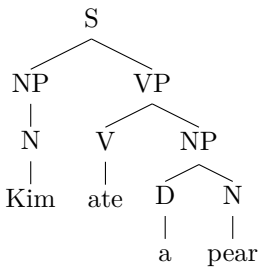
6.4.1 Frege's Principle

The meaning of a sentence is determined by the meaning of its parts and the way in which those parts are assembled.

This comes from the fact that we interpret sentence we've never heard just as easily as sentences we've already heard. This implies a huge section of compositionality.

6.4.2 Combining Meanings

6.4.2.1 Infering from Structure



We first infer $[\![pear]\!]$ then $[\![apear]\!]$, then $[\![ateapear]\!]$ and so on...

We use the phrase structure rule: $S \to NP\ VP$ where NP is the subject of the predicate VP. In many sentences of natural languages, the predicate attributes some property to the individual contributed by the subject.
What is the meaning of a name or of a definite noun phrase ? These are all individuals, real or fictional. The noun phrase *Salvador* refers to Salvador. Salvador is the referent of *Salvador*.
Predicates correspond to properties: *is a carpenter* corresponds to the property of being a carpenter. Properties can apply to (be true of) a number of individuals. At the very least, properties must be able to tell us who has them, and who lacks them. So, at a first approximation, we might identify properties with the sets of individuals that they are true of.

Now, since Subject NPs refer to individuals and Predicate VPs denote sets of individuals, we can relate the two so as to capture the truth conditions of the sentence: we say $[\![Maryisacarpenter]\!]$ is true if and only if $[\![Mary]\!]$ is a member of the set $[\![isacarpenter]\!]$. These are the truth conditions of this sentence.

Proposition 6.4.1 (Semantic Rule I)**.** *If S is a sentence...*

6.4.2.2 Intersective Adjectives

Yet, for more complicated predicates (e.g. *is a smart student*), how do we do ? Intuitively, this VP is composed of two properties (*begin smart* and *being a student*). We may want to take the intersection of those two sets then apply Semantic Rule I, this is Semantic Rule II.

Definition 6.4.1. *Entailment: A sentence X entails another sentence Y (noted $X \models Y$) just in case, whenever X is true, Y is true. For example:*

- *John is a blond carpenter* $\vDash$ *John is blond*
- *John loves cats* $\nvDash$ *John loves dogs*
- *John is a blond carpenter* $\vDash$ *John is a carpenter*

6.4.2.3 Non-Intersective Adjectives

What about: *Pat is a fake carpenter* ? Does it entail those sentences ?

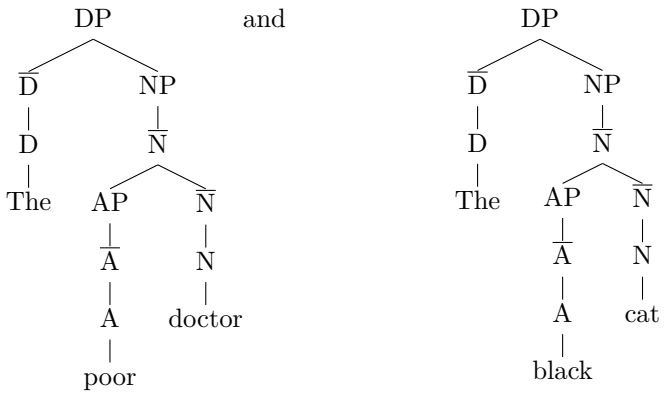
- *Pat is fake*
- *Pat is a carpenter*

This sentence shows our rule for intersection does not work for all adjectives. We restrain rule II to intersective adjectives (such as *prime* or *even*), and call non-intersective adjectives those that fail the entailment test. Worse: *That gun is fake* or *This is a toy gun*; can we then call it a *gun* ? We might think that fake means not, but Salvador is not a fake carpenter, and is not a carpenter. Then, is a *fake blah* something that is not a *blah*, but seems to be a *blah* ? Fakeness seems to require intentionality. Then, *fake* is some kind of functional element that applies to nous, and generates a new set of properties, maybe the set of things that intend to look like the argument intersected with the complement.
And even then, what can we say about *tall* ? It cannot simply represent a set. *Tall* is a subsective adjective since it returns a set beneath a context. Moreover, there is often typicality: when shown a picture of a robin, people will more easily say "This is a bird" than when shown a more atypical bird such as a penguin.
More over, even for non-intersective adjectives, sometimes the argument is not made of morphemes: *Pat is an excellent carpenter* and *That carpenter is excellent*. Then, we may say that there is a placeholder of some sort filling the argument, which is resolved by contextuality. Same thing: *Last night there was an excellent politician and a pretty awful carpenter*, what is the carpenter *awful* at ? Eventually, there is some intersection between excellent and carpenter.

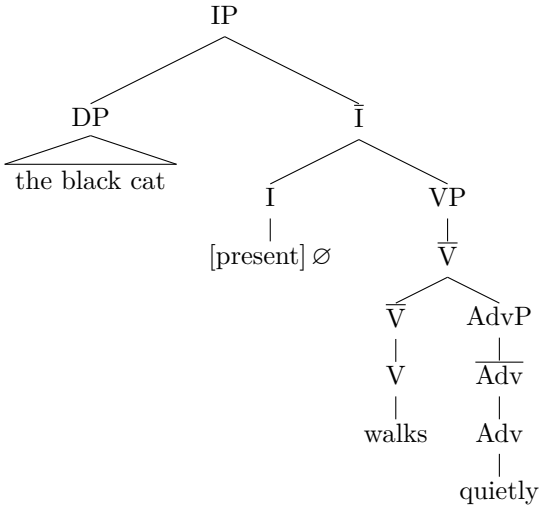
6.5 TA 7: 15/11

6.5.1 One Last Big Syntax Tree

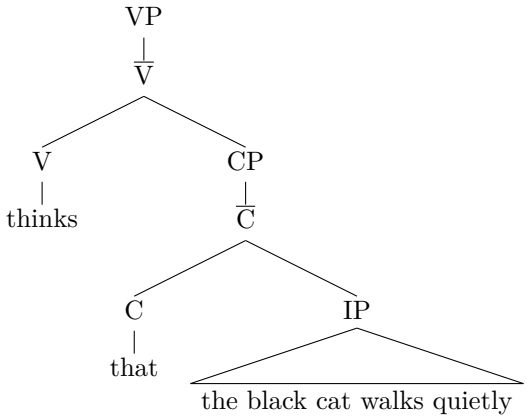
We draw the tree for *The poor doctor thinks that the black cat walks quietly*. We start by mapping the words to their parts of speech. We assume the DP hypothesis and will use $\overline{\text{X}}$-trees. For example, *The poor doctor* and *The black cat* become:



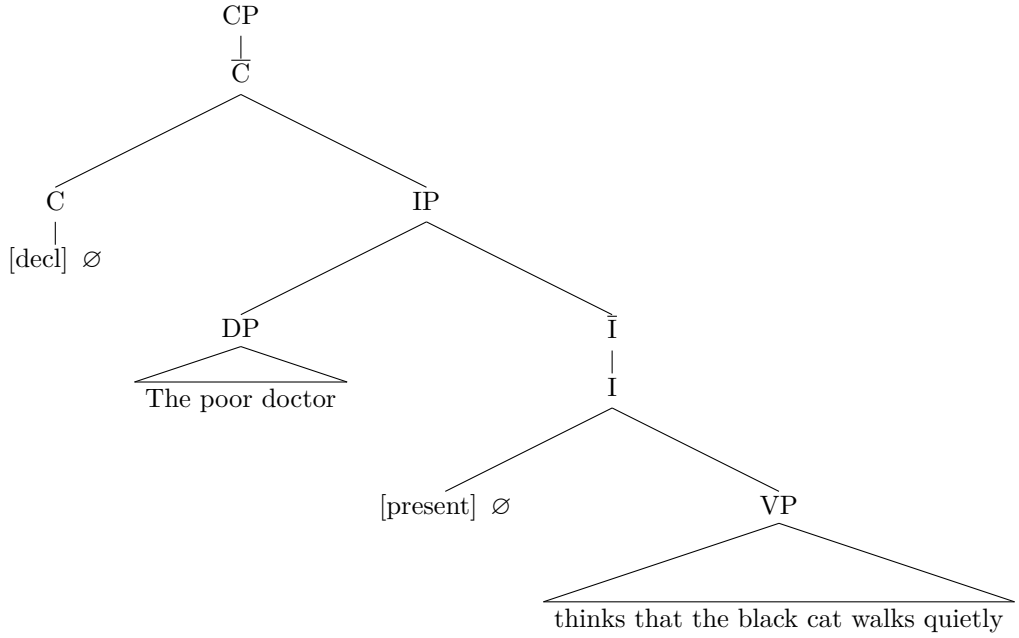
Then, we get for *the black cat walks quietly*:



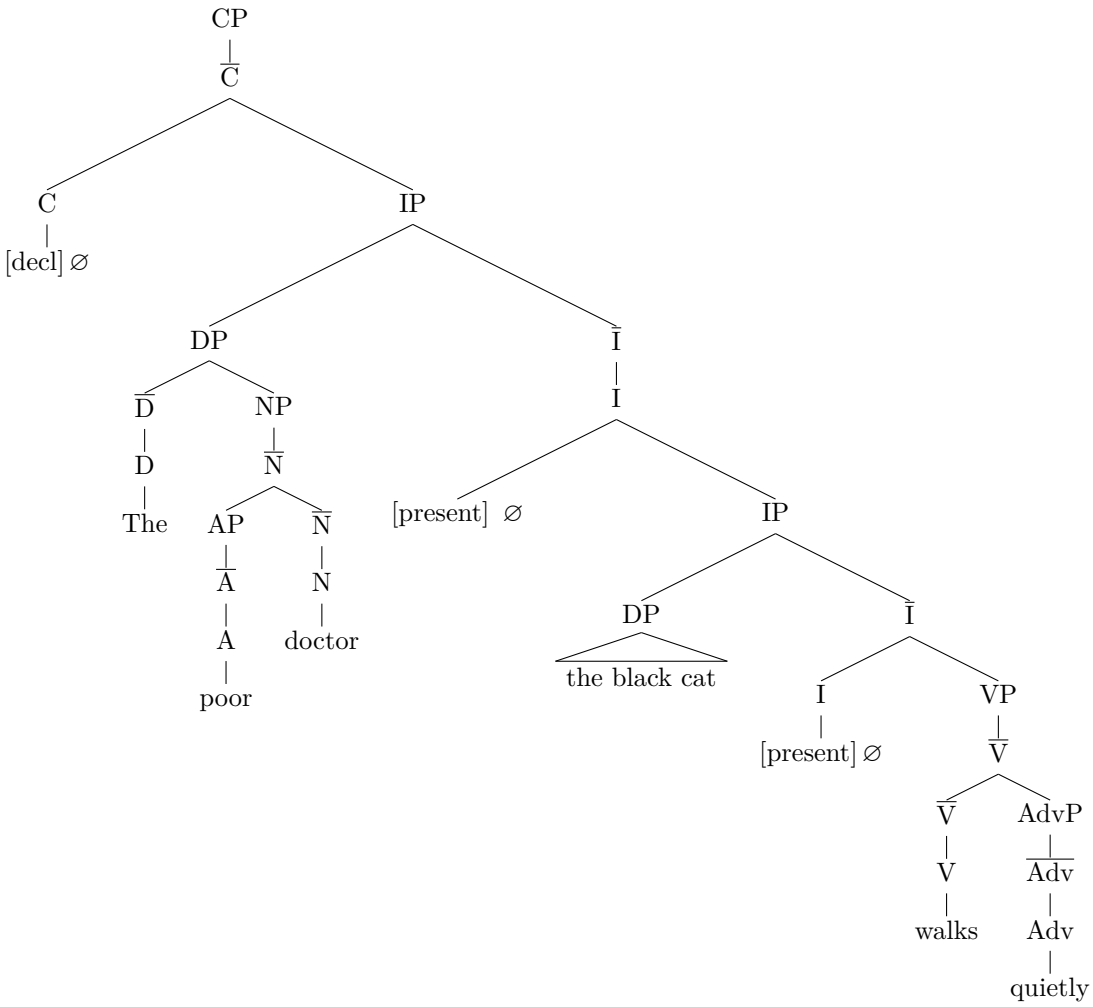
This shows why we need to use the $\bar{\cdot}$ structure, as it allows us to differentiate complements (that attach to the lowest $\overline{X}$, next to the head X), adjuncts (that attach to the other $\overline{X}$) and an eventual specifier that attaches directly to the XP. Then:



Then, the global structure becomes:



Finally:



6.5.2 Entailment

For sentences, we describe their entailment:

- *John is a man* $\vDash$ *John is a human*: This gives an idea on the way we understand words.
- *John blicketed Mary* $\equiv$ *Mary was blicketed by John*: This is purely based on syntax.
- *Everybody left* $\vDash$ *John left*: This is only true based on the contextual restriction of *Everybody*: it is true if and only if *John* is a member of the *Everybody* considered.
- *John read and Mary sang* $\vDash$ *John read* and *John read and Mary sang* $\vDash$ *Mary sang*: This comes from the definition of the logical and.
- *John read* $\vDash$ *John read or Mary sang*: This is always true.
- *All lemons are yellow and not all lemons are yellow* $\vDash$ *Charles the Fifth is a horse*: A false statement implies any statement.
- *John killed Mary* $\vDash$ *John caused Mary to die*: The reciprocal seems to not be always true. It depends on the way we interpret *killed*. Yet, it seems hard to find a counter-example: maybe thinking about a long enough causal chain: *Franz Ferdinand* caused all the deaths in WW1, but he didn't kill some random soldier. A way to visualise the difference would be to look at the scenari that come to mind when hearing these sentences.
- *John did some of his homework* $\nvDash$ *John did not do all of his homework*: *Some* implies a section of his homework, which can be the whole homework. Thus, the first sentence does not

entail the second. We then talk about *implicatures*, which are away to think about pragmatic reasoning. This comes from the fact that people mostly want to be as precise as possible. They might disappear under *if-then* statements.

- *The King of France is not bald* $\nvDash$ *The King of France exists*: Again, this is not entailment, but the fact that the first sentence does not make sense if the second if false makes a *presupposition* mandatory. Another example is *I stopped smoking* which implies that *I used to smoke*.

7 Semantics II

7.1 Limits of our Definition

7.1.1 Problems

Something is said to be denotational when it directly points out to something from the real world.

In our definition, if two sentences are true in the same conditions, then their meaning are identical. Yet, talking about *Kim is a bachelor* and *Kim is an unmarried man*. Then, what about the Pope? What about a 5 year old? But then again, consider:

- i) *Two plus two equals four* and ii)*Forty-five times six equals two-hundred and seventy*. Those sentences are true under every circumstances, yet they make sense very differently. Then those two sentences mean exactly the same thing ?
- iii) *I want to take the ENS Semantics course in the Spring* and iv) *I don't not want to take the ENS Semantics course in the Spring*. Fuck excluded middle.

Someone whose mind represents i) doesn't nessecarily assent to ii). So they must, in some sense mean different things, if meanings are supposed to exist in our minds. Rather than mapping sentences to stuff and facts out there in the world, rather than talking about truth conditions, perhaps we should map linguistic elements to the mental representations that give rise to them and that they evoke. If you're a brain in a vat, is there anything about the semantics of your language that changes ? Can we have both projects ? Can we combine the mentalistic project of talking about the structure of mental representations and the naturalistic project that involves connecting language to the external world ?

7.1.2 Solutions ?

Putnam argues that "wide beliefs", that is beliefs that are about something outside the mind, cannot be given a purely internalistics semantics. Say I display some behvaior that every human with basic common sense can see was directly caused by some belief or other of mine. Any representational analysis of this belief now needs to connect it to things and facts out there in the world, otherwise our theory of representations does absolutely nothing in way of explaining our behavior, since the behaviour is precisely all about our interactions with the outside world.

Fodor bites the methodologically solipsistic bullet: perhaps you can't have your cake and eat it, perhaps you can't have a formal semantics that is both psychological and connected to the external world. But that's not such a big deal, precisely because for almost all cases, and notably many of the most mysterious cases, common sense will tell what in your beliefs connects to what in the world, we can just accept this big hole in our understanding of how mental representations connect to the world, and just move on, focusing our scientific enterprise on the matters we do know how to build properly general theories of: what must representations look like, what structural relations do they have, what are their truth conditions, how are they used in internal calculation and in communication.

The same problems happen in all science, but there are more traces of arguments about that in cogsciences.

7.1.3 On the trouble with adaptationist explanations or What Fodor Got Right[2]

Say we have two phenotypical traits T_1 and T_2, whose precise genetic underpinnings we don't yet understand and we can see pretty clearly using basic common sense that T_1 is the trait responsible for an increase in fitness and that T_2 is completely orthogonal:

$$T_1 \xrightarrow{\text{fit for selection}} \text{selected}$$
$$T_2 \xrightarrow{\text{not fit for selection}} \text{not selected}$$

But if the traits are coextensional, no amount of observational data can resolve this question[3]. We need a causal theory of how traits connect to fitness. There is no such general theory of intensional causation, and no one knows if one can even in principle be given.

$$\begin{array}{ll} T_1 & \xrightarrow{P(F|T_1)>P(F|\neg T_1)} \text{selected for} \\ \updownarrow & \\ T_2 & \xrightarrow{P(F|T_2)>P(F|\neg T_2)} \text{not selected for} \end{array}$$

And for the most theoretically opaque cases, your grandparents could tell you easily which trait is causally connected to fitness. Fodor's conclusion: there's a big mind shaped hole in the theory of natural selection.

$$\begin{array}{ll} T_1 & \xrightarrow{T_1 \rightsquigarrow \text{fitness}} \text{selected for} \\ \updownarrow & \\ T_2 & \xrightarrow{T_2 \not\rightsquigarrow \text{fitness}} \text{not selected for} \end{array}$$

Much like the problem with representations, here we have a hole in the theory for which we don't seem to know how to give a general account. Yet, we can give special accounts[4]. We can conclude that the validity of adaptationist accounts is predicated on the validity of the special causal theories presumed, and we can demand that those causal theories be rigorous. Whether this is too tall an order is an important open question.

At least, we can study the structure of meanings qua elements of mental representations determined by linguistic properties and more generally based on facts observed in the environment.

7.2 Assertion and Presupposition

7.2.1 The notion of Assertion

The assertive content of a sentence is everything in its meaning that is not preserved in the negation or the interrogation

7.2.2 The notion of Presupposition

Take sentences *Kim has stopped smoking* and *Kim used to smoke*. The first one entails the second, and so do the negated, interrogative and modal version of it, we say that it presupposes the second. More formally, S_1 presupposes S_2 if $S_1 \vDash S_2$ and so do the negative, interrogative and modal versions of S_1.

As an example, *Pat got married again* presupposes that *Pat was married*, *John knows that Mary is a blonde carpenter* presupposes *Mary is a blonde carpenter*.

Presupposed content is assumed to be section of the conversational common ground, and is hard or impossible to negate. The assertion is the section of the content that is not actually presupposed. *The King of France is not bald* presupposes that *There is a King of France*. It can be challenged in certain special ways:

- *No, the King of France is* NOT *bald, there* IS *no King of France*

[2] What Darwin Got Wrong
[3] Bareinboim et al. 2022
[4] André et al. 2023, on the design features of moral cognition, with precise models of cooperation and interaction

- *Hey, wait a minute ! There is no King of France !*
- But it seems weird to just say: *That's false.*

Here, the negation is sort of 'meta', since you can in fact utter the words.

7.2.3 Literal Meaning and Implicatures

The asserted and presupposed contents of a sentence form its literal meaning, but naturally occurring conversational is often not literal at all:

- Phone Call: *Is Bill there ?*
- At dinner: *Can you pass the Salt ?*
- Sarcasm: *Not at all, I absolutely love being stepped on*
- Conversation: *What's the temperature outside ? Oh, you won't need a coat.*

Besides having literal meaning, all sentences also acquire a non-literal meaning contained in the context. The propositions that sentences may non-literally suggest, as in the previous slide, are called implications.

- Mary: Is Pierre a good cook ?
- John: Hum, he's French.
- Implicature: He is a good cook.

John's reply suggests that Pierre is indeed a good cook, given the stereotype. But how are implicatures differents ? Implicatures can be canceled without the speaker inccurring a contradiction. The same is not true of entailments: *He's French, but he's actually not a French citizen.*

7.2.4 Cooperation

Proposition 7.2.1 (Cooperative Principle)**.** *Participants in a conservation assume that speakers are cooperative: everyone is doing their best to further the goals of conservation. Conservations follow certain maxims that stem from the assumption of cooperativeness.*

- *Relevance: Every statement made in a conversation must address the issues that conversation is about*
- *Quality: Every time a speaker utters a sentence they say something they believe to be true.*
- *Quantity: Every statement needs to be as informative as the situation requires it to be.*

These maxims are explicit rules that all speakers follow because they are cooperative. Sometimes, literal content of a statement seems to go against one or more of these maxims. They can be superficially flouted. Hearers will draw all manner of conclusions about what speaker meant by their utterance (implicatures of the speaker's utterances), while following the assumption that the speaker is cooperative.

- In the example, John seems to go against the maxim of relevance, but the hearer assumes that John is being cooperative, and thus that there must be some way to interpret John's statement as adressing the question. When an answer really goes agaisnt the maxim of relevance, we call them an infelicitous answer, e.g. *What time is it ? FISH !*.
- The maxim of quality makes sure that cooperative speakers aren't absolute skeptics, information exchange is impossible unless we curtail skepticism. Sarcasm is an example of flouting this maxim.
- The maxim of quantity implies that speakers will typically be assumed to be imparting as much information as they can.

7.2.5 Implicatures and Operations

Either:

- Disjunction is conjunctive in NL
- Disjunction is ambiguous in NL
- Disjunction is semantics {disjunction, conjunction} and the actual typing comes from another property of speech.

For example: *John or Mary came to the party. In fact they both did* and *John or Mary, but not both, came to the party. In fact, they both did.* Since the second is contradictory but not the first, they can't mean the same thing. Take *If John or Mary came to the party, then Bill was pleased* and *If John or Mary, but not both, came to the party, then Bill was pleased.* Those sentences have different meanings under the assumption *John and Mary came to the party* so they can't be equivalent.

For the anecdote: the US Supreme Court ruled that "A jury should not be permitted to engage in conjecture whether an unresponsive answer, true, and complete on its face, was intended to mislead or divert the examiner", and thus established the literal truth rule.

7.3 TA 9: 22/11

7.3.1 Free Variables and Unpronounced Argument

7.3.1.1 Free Variables

Talented Lawyer seems to mean that the man is talented at being a lawyer, but again, in the example about Dancing with the Stars, it might mean the lawyer is talented at dancing. We want to say that we make an assumption about the argument of *talented*, while it actually is a free variable.
In *John ate to a local restaurant*, *local* has a free variable, as it might mean 'of the locality' as well as 'serving a cuisine from the region'.
We can see a free variable as in the phrase: $ate(j, x)$ where x is not pre-defined and j is a constant.
We can often bind free variables with a prepositional phrase: *local (to-x)* or *beautiful (at-x)*.

7.3.1.2 Unpronounced Argument

However, in *John ate*, while we can have an argument to *ate*, the sentence is sufficient to itself. It thus seems that there is only an un pronounced argument, from the logic phrase $\exists x, ate(j, x)$.
In French, *Jean a vu* when answering a affirmation *Il s'est passé ça*.

8 Phonology

8.1 Preliminaries

The vast majority of human languages primarily use sound as a mean of externalization. Phonetics asks questions about the properties of the sounds of languages: How are they characterized acoustically ? How are they produced ? How are they processed ? Phonology (within cogscience) asks questions about the mental representations of the sounds of language: What are they like ? What computations do they participate in ? What sound-related processess do we find in natural language ?

8.1.1 Phonological Competence

There are many things speakers know about the phonology of their native language. This knowledge isn't conscious, but some of it is accessible to introspection with the right training.
There are individual sounds that are ok, but sequences as a whole that are not: e.g. *bnick* is made

of sounds 'b', 'n', 'ik' that are ok while *bn* is not. *Abnormal* seems a counter-example but *bn* is cut between two syllabs. And in russian, *rtut* is fine, but not in english. This is all about the way sounds are organised.
Sometimes, the problem with sounds comes from the internal word position, e.g. the first sound of *hang* cannot be at the end of a word, and the last one cannot be at the beginning of a word.

8.2 Phonetics

8.3 Definitions

Phonetics studies the physical aspects of how speech sounds are produced and perceived:

- What muscles and movements are involved in articulating speech ? What do speakers do to produce different sounds ?
- How does the auditory system perceive different sounds ? Is it just the auditory system ? (McGurk effect)
- What are the acoustic properties of speech sound ? What differentiates *sh* from *s* in perception ?

There is a big chasm between orthography and the actual sound of words:

	letters	sounds
phlegm	6	4
blame	5	4
ophthalmology	13	10
awe	3	1
music	5	6

We need a better way than orthography to talk about the sounds of human speech.

8.3.1 Speech Sounds

The sounds of speech can be studied from two complementary perspectives, articulation and spectrography ?

Producing speech sounds is done by moving air through the vocal tract. The air vibrates, and airwaves travel from vocal tract to the listener's eardrums. Sound quality changes when the shape of the vocal tract is modified by moving the jaw and the soft tissue organs inside. Speech sounds can be classified into different groups, depending how they are articulated. The main dimensions in which phoneticians classify sounds are:

- voicing: do the vocal folds vibrate ? The vocal folds can be pulled together and made to vibrate. If the vocal folds are vibrating while the speech sound is being produced, then the sound is voiced.
- manner of articulation: how is the airflow articulated ? The air only moves through the vocal tract freely when vowels are produced.
- place of articulation: ? Sounds often pattern differently depending on where in the vocal tract is obstructed. The terms for differnet places of articulation are based on vocal tract anatomy.

There are major classes of sounds:

- Non-continuants (n, m, p, t...) vs continuants (f, θ, l, r...): Complete vs. partial obstruction of the oral cavity.
- Obstruents (p, t, k, f...) vs sonorants (m, n, a, e...): Severe vs mild obstruction to the airflow
- Consonants vs vowels: Obstruction vs no obstruction to the air flow.

8.3.2 Phones and Phonemes

We have been studying speech sounds and their intrinsic properties (phones), from the perspective of articulation and acoustics. What abous speech sounds in the minds of speakers as section of their linguistic competence ? The abstract mental representations of the speech sounds are called phonemes. A phoneme can be realized as two different phones: In english, /p/ can be realized as [p] and [p^h]. We say that [p] and [p^h] are two allophones of the phoneme /p/. We don't say that those phones correspond to different phonemes for multiple reasons:

- Untrained Speakers have a very hard time distinguishing two allophones of one and the same phoneme, but no such difficulty distinguishing allophones of different phonemes.
- The two allophones will behave in completely predictable ways across the phonology.
- Mental representations can be less redundant and eliminate predictable aspects of phonetics.

How can we tell when two sounds in a language correspond to different phonemes or are allophones of one and the same phoneme ? Well, current methods do not allow us to answer this question directly: we cannot directly inspec mental representations and phonological competence. Phonologists address it by asking more sophisticated and indirect questions.

We can construct minimal pairs if the differ in just one sound or one feature of a sound. When possible, it's best to use words of the same morphosyntactic category, as well as to find pairs where the segment in question appears in different positions within the words. A feature is distinctive in a certain context if two sounds differing in just that feature in that context can distinguish words. Then, since [p^h] appears only in the beginning of a word and between two sonorants if the following vowel is stressed, and [p] occurs everywhere else, they are two allophones of the same phoneme.

Some allophonic variation is motivated by coarticulation, i.e. the situation wherein the phonetic realization of a phoneme is influenced by a preceding or following speech sound. It is extremely common in human languages for a nasal to acquire the same place of articulation features. Do all phonological processes involve considerations about what is easier or harder to pronounce ? No, there are examples of arbitrariness in phonological processes as well: *Laden* in both German and Dutch.

8.3.3 Loan Words

Loan words are words that were imported into a language from another, such as *sito* meaning *store* in Yoruba, but here, the scheme consonant-vowel was kept.

Languages have overarching constraints on phonology, including sequences of sounds and syllabic structure. This is called phonotactics.

8.4 TA 10: 06/12

8.4.1 On Previous Homework

8.4.1.1 Difficult

Two main free variables: *to X* and *for X*

8.4.1.2 Legal

Main free variables: *in X* where *X* is a jurisdiction, same with *according to X*, e.g. Korean law forbids a Korean resident to smoke weed in another country.

8.4.1.3 Of the utmost importance

When someone says something is important, there is a sense of *to achieve X*. The level of degree that stands in the word *importance* that is filled in by *utmost*.

8.4.2 On the distinction between allophone and phoneme

The phoneme is the abstract phonology that can be represented by an allophone, a sound. The allophones are characterized by the sounds composing them, the latter being sorted as in the following tabular, based on their place and manner of pronunciation:

	Nasal	Plosive/Stop	Fricative	Sibilants	Tap	Trill
Bilabial	m	p | b			bfra	
Labiodental			f | v			
Dental			θ | viii			
Alveolar						
Palatal						
Velar	ŋ	k				r
Uvular		q | g	rh			
Glottal						

A minimal pair is a pair of words that separate two allophones of a same phoneme, e.g. *Either* and *Ether* or *Thigh* and *Thy* on the difference between [z] and [s].
A complementary distribution is a set of words such that two allophone cannot appear in the same situation.

8.4.3 Optimality Theory

This states that all the languages have the same set of rules that encode the way that ways are formed, e.g. Consonant - Vowel alternance, No Consonant at the end, No Cluster of consonants and so on, but the order of priority changes in the set of rules, so that the less valued can be more easily broken. This rule is especially useful in borrow-words.
In Japanese, CVC is the most important, then Agree, then Dependency and so on, while in Hawaïan, No Coda, then Dependency then Agree, and so on...
In principle, we should consider the infinity of strings to see which represents the real word.

9 Language and Reasoning

9.1 Reasoning

Basing on simple problems: *A bat and a ball cost $1.10 together. The bat costs $1.00 more than the ball. How much does the ball cost ?*, we want to see how semantics connect to reasoning. The process that connects us to the wrong answer $10c$ might have to be with the fact that we're looking for the answer in the data that we've gotten. It might also have to do with availability heuristics.

9.1.1 Systematicity

Human reasoning fails in strikingly systematic ways, we do something when we try to answer questions like these, but it looks like what we do is the wrong thing. Human reasoning works on mental representations that are often the product of interpretive processes, linguistic processes feed into reasoning processes. We will look at number of interactions between interpretive processes and reasoning. We will study theory of the operations of the faculty for reasoning.
For example, in the sentence: *John speaks German and Mary speaks French or else Bill speaks Italian*, we get trouble on entailment:

- Classical Entailment: If *Bill doesn't speak Italian* then *John speaks German*
- Naive "Entailment", 20/20 acceptance: If *John speaks German* then *Mary speaks French*

Those two different notions of *What follows* are linked to two different representations of deduction:

Truth Conditions	Inquisitive Semantics
$(j \wedge m) \vee b$	$\{j \wedge m, b\}$
(worlds where John speaks German intersect worlds were Mary speaks French) union (worlds where Bill speaks Italian)	Two Alternatives: 1) J. speaks German and M. speaks French. 2) Bill speaks Italian

The same type of situations appear when talking about the number of days it takes for lilypads to cover a lake or about the kings and queens of Europe that Mary met.
When looking at *every A or every B*, a reasonable reasoning suggests that either it's *every A and no B* or the opposite. The same thing happens when talking about *most of the homework*. Yet, those pragmatic reasoning seem to be systematic across every human being, as anyone could think how the hearer would respond to this signal.

9.1.2 Formalisation

- Compelling Fallacies are invalid inference patterns that we often accept
- Repugnant Validities are valid inference patterns that we often reject
- *Failures* of reasoning ? It's possible some of the problems we've seen so far are in fact the result of sound reasoning acting on non-obvious but perfectly reasonable and predictable interpretations of the premises. If we don't countenance this possibility in a systematic and sophisticated way, we run the risk of misdiagnosing interesting interpretive processes as failures of reasoning.

Indeed, we resist disjunction introduction : $\varphi \vee \psi$ suggests strongly one doesn't know which of φ and ψ is in fact true. This is because if the speaker did know, they would have given us the true one, which would be more informative that the disjunction. Then, take Moore's Paradox : *It's raining but I don't believe it.* And then, *If it's raining but I don't believe it, then there is something I don't know.*

9.2 Theories on Reasoning

9.2.1 Mental Logic and Mental Models

These share the idea that intuitively humans do something that is related to formal logic.

- In Mental Logic : Humans do some form of proof theory, they have deductive rules they use to transform sentential representations into other sentential representations.
- In Mental Models : Humans do some form of model theory, and build mental models of the informations they are given, update these representations sequentially and inspect the resulting model for constitutive parts that are of interest.

9.2.2 Heuristics and Biases

We do not reason with logic or mathematics, instead we use heuristics that give quick decisive answers to our questions. Our reasoning is rendered fallible by the imperfection of these heuristics as well as a collection of innate biases :

- Judgement by Substitution : Faced with a hard question, we substitute for it an easier question whose answers we can map into answers to the hard question. We work on answering the easier question, and transpose the answer we find.
- Base-rate neglect : We often ignore general information about a representative sample and go purely on information about a particular individual.
- Anchoring effet : We tend to incorporate information we recognize as altogether irrelevant into our decision making.

9.2.3 Other Theories

There also are more theories such as a Probabilistic Approach and one called Argumentative Theory. The latter states that reasoning appeared not to help make better decisions but to help justify decisions and persuade others to adopt our views. Argumentation then is an essential tool allowing for coordination of complex social activities, political structures, and so on. Reasoning helps ground argumentation.

Many experiments can be run to see the way that people reason from visible looks.

9.3 Manipulations on a Problem

9.3.1 Contextuality

When changing the context of a problem, the answers given by participants change. For example, partipants had better answers to an ecologically valid problem such as the age and alcohol version of the Wason Selection Task than in the pure context of letters and numbers. This was used to validate the environmental influence on reasoning. Multiple reasons have been advanced :

- J'ai pas eu le temps de lire...

And even in a particular environment, just changing the formulation of the problem changes the results : a disjunction case gives perfect results while introducing negations into the antecedent, the consequent of a conditional or both has a negligible effect on the distribution of people's choice. That is, they seemingly ignore negation. In fact, people, having no idea what is going on except for the fact it has something to do with vowels and even numbers, always answer the same thing : people are stumped by conditional, thus they go back to something they know.

9.3.2 Illusory Content

Sometimes (cf. beginning of lecture sentences), disjunction-like elements induce illusory inferences. Indeed, sentences that are expressed with a lot of conjunctions and disjunctions can often be simplified significantly, yet, giving the difficulty of parsing, there are problems. The erotetic principle says :

- Our natural capacity for reasoning proceeds by treating successive premises as questions and maximally strong answers to them. (Problem of Failure)
- Systematically asking a certain type of question as we interpret each new premise allows us to reason in a classically valid way. (Problem of Success)

There is commitment on interpretation, since disjunctions raise alternatives and put pressure towards choosing an alternative : disjunctions are like questions in this regard.

9.4 TA 11: 06/12

9.4.1 On Last Homework

- *pass* should be pronounced : [pʰɑːs] in british english or [pʰæs].
- For two words without a minimal pair, you can find with [ʒ] and [t]

9.4.2 IIFD

We study this phenomenon : Given the proposition $a \vee (b \wedge c)$ and b, people conclude c.

There are seemingly two systems inside the head of peoples :

1. One that is based on heuristics : b looks like $b \wedge c$ so people talk conclude c.
2. One that can actually do reasoning.

Multiple theories arise :

- One where Pragmatic reasoning is encouraged : Through studying the alternatives, people strengthen the statement in : $(a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge c)$. This second sentence looks like the `xor` of the *natural language* sentence, and might come from a pragmatic interpretation of the natural language *or*. This is sometimes tested by tring to overflow the brains of people by asking them to remember a grid. It has been found that people doing this were more precise.
- One where Pattern Matching happens
- One where Erothetic judgement comes in

10 Follow-up : Language and Reasoning

10.1 Erotetic Theory

A standard employment of *or* is in the specification of possibilities : This comes from the principle that questions are modeled as sets of propositions.
Disjunctions are approaches to free choice, counterfactuals, exceptional scope-taking. Inquisitive Semanctics say disjunctions are the building blocks of questions.
Many natural languages (Polish, Korean, Japanese, Slavic...) have the same morphemes for the interrogative complementizer and disjunction operator.

Then, illusory inferences in the erotetic theory come from incomplete answers.

10.1.1 Mental Models

We formalize Erotetic Theory through mental models.

10.1.2 Success in the Erotetic Theory

Sound reasoning is possible. So there must be some strategy using our natural faculties guarantees it. This strategy has to be innate or learnable but it shouldn't be costly :

Theorem 10.1.1 (Soundness Theorem)**.** *The ETR derivation strategy where an update with* Δ *is immediately preceded by a sequence of inquire steps for each atoms that occurs somewhere in* Δ *is sound for classical propositional models.*

10.2 Testing this Theory

We can induce the representations of a disjunction with visual animations. This has been done to demonstrate that infants can construct complex representations of alternative possibilities such as induced by a disjunction.
From other experiments, we get that children under 4 have impoverished symbolic system and cannot represent alternative possibilities. Children acquire more complex representational arsenal thanks to linguistic bootsrapping wia the language of modality which appears around the same age.

10.3 TA 12 20/12

10.3.1 Erotetic Theory Formalisation

We don't use $P \wedge R$ to symbolise P *and* R, but prefer $P \sqcup R$.
For disjunctions, we use sets containing all the possible propositions : $\{P \sqcup R, J\}$.
We represent our situation as a tuple, with our propositions on the first dimension, and the set og things we know to be true in the second dimensions : $\langle \{P \sqcup R, J\}, \varnothing \rangle$.
We symbolise updates such as this : $\langle \{P \sqcup R, J\}, \varnothing \rangle [j]^{up}$ which means that we have learned for

sure that j is true.
Then, we can chain updates : $\langle \{P \sqcup R, J\}, \varnothing \rangle [j]^{up} [\{m, r\}]^{up}$ is equivalent to $\langle j, \{\{j\}\} \rangle [\{m, r\}]^{up}$ which is equivalent to $\langle \{j \sqcup m, j \sqcup r\}, \{j\} \rangle$

10.3.2 Disjunction Reduction

People are better at resolving disjunctions when we update with the negative : we talk about $\neg P$ and $P \vee Q$ which implies Q :

$$\langle 0, \varnothing \rangle [\neg P]$$
$$\langle \neg P, \{\{\neg P\}\} \rangle [\{P, Q\}]$$
$$\langle \neg P \sqcup Q, \{\{\neg P\}\} \rangle [Q]^{MR}$$

is easier than

$$\langle 0, \varnothing \rangle [\{P, Q\}]^{up}$$
$$\langle \{P, Q\}, \varnothing \rangle [\neg P]$$
$$\langle \{\neg P \sqcup P, \neg P \sqcup Q\}, Q \rangle$$

Here, MR means *Moleculary Reduction.*

10.3.3 Inquiry Operation

The inquiry operation works as such :

$$\langle \{\neg a, \neg b, \neg c\}, Q \rangle [a]^{inq}$$
$$\langle \{\neg a, \neg b \sqcup a, \neg b \sqcup \neg a, \neg c \sqcup a, \neg c \sqcup \neg a\} \rangle$$

We can then go and explicit every formula based on any possible atoms.
This is of use when resolving c and *a and b or c and d*

$$\langle \{a \sqcup b, c \sqcup d\}, \varnothing \rangle [c]^{up}$$
$$\langle \{c \sqcup d\}, \{\{c\}\} \rangle$$
$$d$$

Remind this is not supposed to be classically true, but it models what people do. Then using inquiries :

$$\langle \{a \sqcup b, c \sqcup d\}, \varnothing \rangle [c]^{inq}$$
$$\langle \{a \sqcup b \sqcup c, a \sqcup b \sqcup \neg c, c \sqcup d\}, \varnothing \rangle [c]^{up}$$
$$\langle \{a \sqcup b \sqcup c, c \sqcup d\}, \{\{c\}\} \rangle$$

11 Neurolinguistics and Psycholinguistics

11.1 Localizing Language

The brain is composed of two hemispheres. For humans, as well as for all modern vertebrates, sides swapped. Handedness makes a difference: right ear advantage for speech processing in right-handed people (and so for left)
Split-brain surgery is a common treatment for seizures. When different images are presented to split vision field, patients can describe and point out to images on the right but cannot describe images presented on the left.
There is a possibility to localize everything in the brain by stimulizing electrically the brain.

- Tests with right-handed people demonstrate about 90% left-hemisphere language localization and 10% mixed distribution.
- With left-handed people, 75% LH language, 15% mixed distribution and 10% RH language.
- The correlation between handedness and language localization is robust but very imperfect.

Handedness is not trivial to ascertain, varying by task.

How can we localize language in the brain ?

11.1.1 Broca's Aphasia

The lesion-deficit paradigm :

1. Find Someone with significant lesion in area X
2. Characterize their deficit in linguistic tasks Y
3. Conclude that X plays a big role in Y

This is how it went for Pierre Broca with his patient Tan. The autopsy reveale a lesion in what is known as Broca's area (BA44 & BA45) of left-hemisphere.
In similar cases, there are different manifestations depending on the person. Mainly, people often keep their lexicon words, still have vocabulary, but lose in capacity using function words. This resonates with the case of Jeanie (cf. 2.3.3.1). We also find :

- Halting, non-fluent production, with verbs left out more than nouns, a reduced mean length of utterance and a frequent omission of function words and functional morphemes.
- Comprehension impairment especially when it crucially depends on syntactic distinctions
- Non linguistic problems (e.g. asymmetric paralysis most often of face and right hand)

In a 1960 documentary, we see a person having issues understanding a passive way sentence. We then see that syntax does have an influence on comprehension : *The leopard was killed by the lion* is tougher to understand than *The lion killed the leopard*. This comes from the fact that things that come later in a sentence are often newer.
The role of Broca's area in language is not completely clear

- Damage must extend to surrounding areas and underlying whit matter.
- Broca's aphasics tend to display language impairments not entirely describable in terms of syntax of phonology.
- Broca's area is by no means the only area of the brain where damage will reliably produce systematic language impairment.
- Yet, there is a strong correlation with syntax, processing and speech production.

11.1.2 Wernicke's Aphasia

Similarly to Broca's Aphasia, Wernicke's sees patients after a trauma produce linguistics modification :

- Speech is fluent and well articulated, though repetition is impaired
- Patients produce frequent semantic and phonemic errors, including non-words
- Speech appears "empty" or confused
- Associated comprehension deficit, even with word meanings.

11.2 Brain Imaging

11.2.1 The System

- Functional Neuroimaging consists to measuring blood flow in the brain. Working harder needs more oxygen and thus more blood. This gives a high spatial resolution. However, blood moves slowly compared to electrical brain activity. There is a 2-6 second delay between electrical activity and the result appearing.
- EEG measures electry potentials directly from the scalp, but the electric signal is distorted by the cranium. This gives poor spatial resolution.

- MEG measures magnetic fields associated with electrical signals. Since fields aren't distorted by the skull, space resolution is better.

MEG allows to sea, about 350ms after reading a word, a peak of activity in the left temporal lobe. Changes depending on the linguistic properties of the stimulus appear, and the activity is quicker when the word is more frequent. The hypothesis associated is that activity is associated with lexical access.

11.2.2 Morphological Decomposition

How to interpret the difference between *date + [PAST] = dated* and *give + [PAST] = gave*. The first one is clearly bimorphemic, but is the second ?
The first answer that seems to be is NO. Indeed, this comes from priming, the ability to recognize words faster if one has seen the word recently. Since *dated* is recognized faster after seeing *dated*, and *give* is not recognized faster after seeing *gave*. Since *nine* is not recognized faster after seeing *asinine*, nothing here has to do with phonetics nor orthographics.
Yet, phonologically and orthograpically similar words inhibit priming effects. The hypothesis is that *gave* IS bimorphemic.
If we can measure priming before interference has taken place, then *give/gave* and *date/dated* should have the same effect. We find that M350 happens in both *give/gave* and *date/dated* but that there is no response in both *give/gave* and *boil/broil*. Moreover, we see that priming does improve the result in the first two examples but not on *boil/broil*. Yet, *teach/taught* is not improved since it is already faster to recognize than the 3 other pairs.
Morphological Decomposition is always needed, even when morphology is not to be seen in phonetics (e.g. *gave*).

11.3 Psycholinguistics

This field talks about how we deploy our knowledge of language :

- Grammar : This is a formal specification of structures allowed in a language along with their interpretation. A central element of theory of competence.
- Parsing : This is the process of determining the syntactic structure of an input string. A parser takes as input a string of words and produces as output a parse tree: a labeled bracketing of the input sentence.

11.3.1 Parsing

AAAAAAAAAAAAAAAAAAAH

11.3.2 Ambiguity

There is ambiguity at different levels of representation :

- Lexical - the same phonological string may correspond to more than one word.
- Lexical Semantic - a word may have more than one sense
- Structural - a string may have more than one legal parse

Also, we need to distinguish :

- Global Ambiguity - no information in the sentence resolves the ambiguity
- Local Ambiguity - at one or more points during parsing, there is insufficient information to determine the correct structure. But the ambiguity disappears once the entire string is parsed.

The Garden Path phenomenon tells us that the parsing of a sentence is entirely serial : Frazier and Rayner theorised that

- Only one syntactic representation is built at any given time
- Everything is based on heuristics : Late clorsure where awe attach new items to the phrase or clause currently being processed and Minimal Attachment where we attach new items to the existing structure in a manner that requires establishing as few new nodes as possible.

12 TA 13: 10/01

12.1 Handedness

For RH people, the language processing areas are mostly in the left hemisphere while it's more diverse for the other people.
In fact, when cutting corpus clausum (link between hemispheres), people cannot describe what they see with their left eye.

12.2 Aphasias

Distinction between aphasias :

Broca	Wernicke
Aware Non-Fluent, no function words No syntax Struggle with Passives and Syntactic Movement	Non-Aware Fluent, Long sentences, Nice Prosody No meaning

Mostly, it is impossible to recover from such traumas.

12.3 Priming

Priming studies seem a bit weird a times, but a lot of them are replicable.
Studying *dated* it seems that it's *give* + [PAST], but studying *gave*, it seems not. In the end, there is no way to be sure that one or the other theory is false.
Then, priming effects suggest that the decompositional view is seemingly better. But sometimes, the atomic view seems better. Yet, looking at *boil* and *broil*, it seems that we check if the words are indeed the same. In the same way, *give* and *gave* similarity may inhibit the priming effect. Looking for a special signal (M350) in a EEG seems to validate this effect. On a neurological level, this hypothesis seems to be valid, but in the behaviour it is not clear.
Yet in priming studies, subjects are facing written words, but how much more is this looking at effects on reading than language ?

12.4 Parsing

There are two main types of parser : top down and bottom up, c.f. Compilation. Honestly I don't want to write down the parsing algorithms he shows.
From the meanings of *buffalo* (being : a bison, a city, to bully), we can build a sentence with any number of *buffalo* following each other. Homophony makes stuff tough sometimes, and it might change the way that parsing works.

13 Machine Learning and Linguistics

13.1 Machine Learning 101

The point of Machine Learning is approximating a function. Loss functions are measures of the distance from an approximation $\hat{f}$ to the real function f. The idea is to minimize the loss. When doing supervised learning (given a series of points and their images), the method of the least

squares is a solution. Yet, for many linguistics tasks, it is tougher. One may use a cloze task where we replace *key* words and have the LLM guess and retribute it with 1 for a correct guess and a 0 in the other case. One can also use ext word prediction. In linguistics model, we prefer to output a probability distribution. To turn NWP into something continuous, we use cross-entropy loss, doing dot product between a vector and the model result.
The goal of machine learning is to find a specific hypothesis f from a set $\mathbb{F}$ of potential hypothesis. $\mathbb{F}$ can be anything computable, for example, it can be the powerset of all valid rules we can format. Yet, we cannot build a loss function on pure CFGs, we need to make CFGs probabilistic to do so. A classic kind of model was n-grams. The idea is to assign probabilities for sequences of n words in a look-up table. 4-grams seem realistic. Yet, it is possible to find bigrams that could be valid, but are not possible by themselves, n-grams are fine, but they couldn't generate properly the language. Neural Networks are stacked linear regressions. Early Neural Networks were one layer, so they cannot learn XOR. Yet, if you have more than one non-linear layer, you can learn exclusive or. This is because you can rearrange the space in the first layer and transform it into a linear problem for the last layer. In principle, recurrent neural network are Turing-Complete, so they can represent natural language.
To train a neural network on linguistics issues, we use gradient descent.

13.2 Machine Learning and Linguistics

Many times, LLM have right answers, but their heuristics don't apply to all sentences. Then, while on entailed sentences, the accuracy is great, it is really low (15%-ish) on non-entailed sentences. Thus, most of the time, LLMs are right for the wrong reasons.