

LING 101 : Introduction to Linguistics

Salvador Mascarenhas

Michael Goodale (TA)

21 Septembre 2023

Table des matières

I	Cours 1 21/09	2
1	Remarks and Observations about the Nature of Language	2
2	The Goal of Linguistics	3
3	Different levels of Study	3
4	Language and Societies	3
4.1	Language and Classes	3
4.2	Language and Dialect	3
4.3	Infinity of Language	4
4.4	Description	4
4.5	Phonological Differences	4
4.6	Internal Structure of Sentences	4
II	TA 1 27/09	4
5	Animal Communication	4
III	Cours 2 28/09	4
6	Different yet Similar	5
7	Similarities between distant unrelated languages	5
7.1	Reciprocal pronouns	5
7.2	Sensitivity to negative elements	5
7.3	Relations between Sentences	6
7.4	The Puzzle	6
8	Universal Grammar	6
8.1	Chomsky's Hypothesis	6
8.2	Universal Grammar	6
8.3	Language Acquisition	7
8.3.1	Problem of Induction	7
8.3.2	Absence of Negative Evidence	7
8.3.3	Conclusion	7
8.3.4	Chomsky's Argument	7

Première partie

Cours 1 21/09

Introduction

DEC is called like that because it would have bothered literatures 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 Bloomfield) 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 : <https://moodle.u-paris.fr/course/view.php?id=7374> Syllabus : https://moodle.u-paris.fr/pluginfile.php/1002151/mod_resource/content/5/intro-ling-syllabus-2023.pdf

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-bar theory
6. **Semantics I** first look at meaning *Studied by Salvador*
7. **Semantics II** philosophy of language and the case for methodological solipsism.
8. **Phonology** phonetic macro classes ;
9. Language and Reasoning
10. Sign languages (guest lecture)
11. Language in the brain ; deficit-lesion method ; functional brain imaging ; psycholinguistics ; parsing, reading, lexical access.
12. Language and thoughts in minds vs. machines

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 embarrassing. 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...

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.

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 chunks 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 gave ? 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 spoken language.

4 Language and Societies

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.

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)

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.

4.4 Description

Not looking for rules prescribing 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.

4.5 Phonological Differences

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

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 (*~ 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.

Deuxième partie

TA 1 27/09

5 Animal Communication

Language is also communication, not only hearing (trees?). Many (if not most) Animals Communicate, and almost all react to sound. (cf. https://www.nytimes.com/2023/09/20/magazine/animal-communication.html?unlocked_article_code=_mKij4e1jtSDj61vUQZVjQCPQ678h069vto7sqwbRaA3kmyw5BAInBSESFgm2rBtU7GDoS_gyv_G6GUUnUjV5Wb8L_Cb4YjsG-BFKXy_y03FYnEC0JFaCdmGPS7pCbPH81qQcH514mixJE4IfNzBFm94Uh6YT-76HUyTwGvPuYKgrb0-F-xoAdiItiAZoUDJzWBY2GIujc08Hw7TiORkZXfc8MRihzb4S7i6_eZR1mWD4-yafAQQP4hFCSV-wmJKxSHpSSrMFpoK9n4sdL&smid=url-share) : 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.

Troisième partie

Cours 2 28/09

6 Different yet Similar

Languages 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 an arbitrary, conventional link : *black attire* implies mourning (*in some cultures but...*)

After Ferdinand de Saussure,

« Language is a system 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*

7 Similarities between distant unrelated languages

7.1 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 doesn'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 includes (*each other*)'s meaning so it is compatible.

7.2 Sentivity to negative elements

Words occur sometimes with negative elements. Every natural language seems to contain at least one lexical item 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 negative item is enough to license a negative polarity, and sometimes negative polarity is inferred 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.

7.3 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.

7.4 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 communities 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

8 Universal Grammar

8.1 Chomsky's Hypothesis

Human faculty for language :

- enables humans to acquire and use language;
- delimits what linguistic structures humans are capable of acquiring and using
- delimits the kind of linguistic conventions 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 part of all humans' biological endowment and it is reflected in all natural languages.

8.2 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.

8.3 Language Acquisition

8.3.0.1 Critical Period

There is a critical period for human (and animal as well) acquisition of language : post-puberty, acquisition is severely impaired, e.g. Genie, 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 fully related to mathematical/intellectual/reasoning skills and so on... Learning a second/third language is totally different though, it is correlated to musical and computational skills it seems, there even seems to be a purely linguistic talent. Almost nothing about adult phonological is due to purely biology and genes.

8.3.1 Problem of Induction

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

8.3.2 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.

8.3.3 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 too much about the same sentence.

8.3.4 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