Language & Computation

UNIB20005

Lecture 4.2

Introduction to Language (2):

Syntactic structure

Review (1)

- What is linguistics?
- What is (a) language?
- What do linguistics and computation have to do with one another?
- Some types of linguistic information/structure

Review (2)

- The ‘sign’ and sign systems
- Arbitrariness of the sign
- Structure in grammar

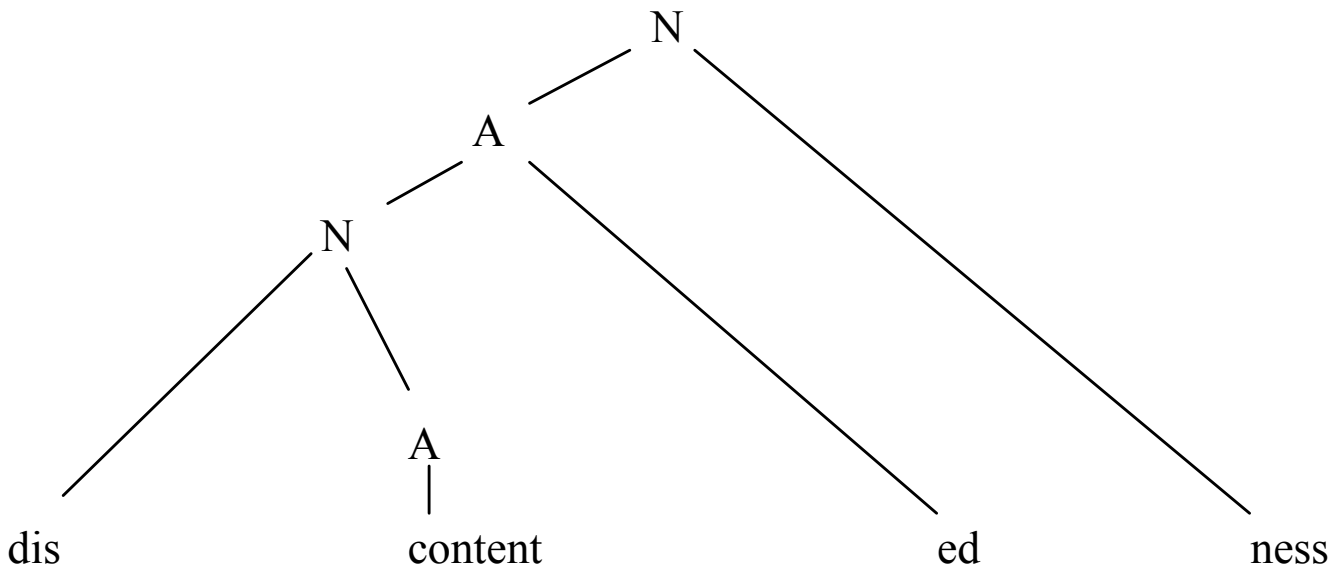
Review (3)

- The ‘word’
- Word class / lexical category
- Morphemes and morphology
- Morphological structure of words

An example of structural complexity of words

[[[dis- [content]] -ed] -ness]
N A N A

[[[OPP-TO [CONTENT]] CHARZD-BY] STATE OF BEING]


N
A
N
A
dis
content
ed
ness

Some languages have more morphemes per word than English allows:

Abanyawoihwarrgahmarneganjginjeng

‘I cooked the wrong meat for them again’ (Gun-djeihmi; Arnhem Land)

... & some have fewer (e.g. Korean, Mandarin)

Lexemes

Sets of words conceptually group together into **lexemes ...**

The **lexeme**:

- *cat, cat's, cats, catty, catlike* — 5 words, 3 lexemes
- *(I) work, (I) worked, (I'm) working, worker* — 4 words, 2 lexemes

A lexeme is a 'mental dictionary entry': independent of the particular entities it is applied to.

CAT = *cat, cats, cat's*

WORK= *work, worked, working*

Another distinction

Inflectional morphology

- Leaves the lexeme the same
- Does not change the part of speech
- Encodes grammatical meaning specific to that part of speech
  - *sweet—sweet-**er***
  - *director—director-**s***
  - *direct—direct-**ed***

Derivational morphology

- Changes the lexeme
- May change the part of speech (lexical category)
- Optional
  - *electric—electric-**ian***
  - *sweet—sweet-**en***
  - *director—**ex**-director*
  - *fortunate—**un**-fortunate*

Lexical category again

- A lexical category is: a set of words which share a common set of linguistic (esp. morphological and syntactic) properties.

Lexical category (word class) ambiguity

An example from Ch 5:

They refuse to permit us to obtain the refuse permit.

'Homonyms' = An ambiguous word (i.e. a word form corresponding to two separate lexemes), whose senses are far apart and not obviously related, other than through mere accident or coincidence e.g. *mug, bank* (cf. homophones: *threw/through; rode/rowed*; homographs: *wind/wind*).

- **We need to increase productivity.**
- **We need an increase in productivity.**
- **Why do you torment me?**
- **Why do you leave me in torment?**
- **We might transfer him to another club.**
- **He’s asked for a transfer.**
- **Are you some kind of pervert?**
- **Don’t try and pervert the course of justice.**

- **We need to inCREASE productivity.**
- **We need an INcrease in productivity.**
- **Why do you torMENT me?**
- **Why do you leave me in TORment?**
- **We might tranSFER him to another club.**
- **He's asked for a TRANsfer.**
- **Are you some kind of PERvert?**
- **Don't try and perVERT the course of justice.**

Lexical category ambiguities

- British left waffles on Falkland Islands
- Lung cancer in women mushrooms
- Clinton wins on budget, but more lies ahead
- Juvenile court to try shooting defendant
- Deer kill 17,000

Lexical category ambiguities

- British left waffles on Falkland Islands
  Is *left* a N or a V? Is *waffles* a N or a V?
- Lung cancer in women mushrooms
- Clinton wins on budget, but more lies ahead
- Juvenile court to try shooting defendant
- Deer kill 17,000

Example: the class ‘Noun’ in English

Distributional criteria:

- Occurs as ‘head’ of Noun Phrase (NP)
  - Cooccurs with **Determiners** (like *a, the, this, some*)
  - And with Numeral Determiners (*one, two ....*)
  - Cooccurs with **Adjectives** (like *red, special, clever*)
- Within its NP, occurs as subject, object etc. of verb
  - *The elephant grabbed the banana.*

Inflectional criteria:

- plural morpheme –s (*elephant – elephants*)
- Some affixes characteristically form Nouns (*-ness, -tion, -ity*)

Other:

- Is associated with other kinds of categories marked in other ways within the NP – e.g. Definiteness, indicated by *a/the*

And...

Meaning: Refers to things, people, abstract concepts (i.e. not typically events or states). Eg. ‘rock’, ‘mouse’, ‘truth’.

BUT generally meaning is a poor criterion for lexical category:

- *vaccination* is a noun—but an action
- *bright* is a property, and an adjective;
  glow is a property—and a verb
- Mismatches from language to language:
  English *I am hungry* (not *I hunger*)
  Greek verb *pino* (not adjective *pinasmenos*)

An exercise

- How do you detect structure if you don’t know the language?
- Or can you ....?

Mbabarum

1. **dogul mog njab**
2. **mogul dog njarə**
3. dog lob
4. mog lonə alŋgi
5. mog lonuŋ
6. dogul mog njarə alŋgi
7. mog lob anmın
8. dog lim lob
9. dog lonə
10. mogul limul dog njaruŋ
11. mogul dogul njab
12. dogul mog njab anmın
13. mog lim lob anmın

Mbabarum

- Mbabarum is a language of North Queensland, spoken inland from Cairns.
- It is highly unusual in that through historical change it has lost the first syllables of words, so words are largely monosyllabic –
- and ‘dog’ in the data really does mean ‘dog’, quite independently of English.

Solution

- The incorrect sentence is 11. There should not be two words with the -ul ending. The correct sentence would be:

mogul dog njab

Structure above the level of the word

- The huge wombat chased a tiny fruit fly.
- A huge fruit fly chased the tiny wombat.

Syntax deals with the way the ‘arrangement’ of words contributes to meaning.

Noun Phrases

- We have already seen that Nouns tend to occur in groups (phrases) with specific other kinds of lexemes:
  - Determiners (*a, the,* etc.)
  - Adjectives (*huge, tiny,* etc.)
- In English these go together in particular orders, and the phrases formed have particular functions in the sentence.

- All languages structure information in terms of
  - Kinds of things
  - Their properties
  - The actions & relationships they are involved in
- But they express these kind of relationships formally in a great variety of ways

- English: largely a **word-order** scheme
- That is, the words that constitute a description of an entity appear as a block together, and the relative order of these blocks with respect to words denoting actions and relationships (verbs) indicates the role of the entities within the event / situation.

But compare Warlpiri:

Wita	wajilipu-ngu	maliki	ngaya-ngku	wiri-ngki
small	chase-PAST	dog	cat-ERG	large-ERG

large cat chased small dog

Wita	wajilipu-ngu	maliki-rli	ngaya	wiri-ngki
small	chase-PAST	dog-ERG	cat	large-ERG

large dog chased small cat

Native speaker intuitions & grammaticality judgments

- After the show they all went back to Suzie's place.
- 'Twas brillig and the slithy toves did gyre and gimble in the wabe ...
- He roared with me the pail slip down his back.
- It was Jilly with the bag on her back that Fred was carrying her to safety.
- The worst part and clumsy looking for whoever heard light.

“constituents” & “constituency”

- A syntactic constituent is:

  a word or a group of words that functions as a single unit within a syntactic structure

Why do we need to recognise constituents?

The English question rule:

a. Lee can lift 100 kilos.
b. Can Lee lift 100 kilos?

a. Mathematicians are generally thought to be odd?
b. Are mathematicians generally thought to be odd?

a. They will want to reserve several rooms.
b. Will they want to reserve several rooms?

a. Mary has proved several theorems.
b. Has Mary proved several theorems?

English Question Rule #1

Lee can lift 100 kilos.

1 2 3 4 5

“To form a question from a declarative sentence, place word 2 at the beginning of the sentence.”

Can Lee lift 100 kilos?

2 1 3 4 5

But...

- Yesterday Lee could lift 100 kilos.
- Many mathematicians are thought to be odd.
- Those people will want to reserve two rooms.

- Lee yesterday could lift 100 kilos.
- *Mathematicians many are thought to be odd.
- *People those will want to reserve two rooms.

English Question Rule #2

"To form a question from a declarative sentence, place the first verb of the sentence at the beginning of the sentence."

Lee can lift 100 kilos.

Many mathematicians are thought to be odd.

.

.

.

BUT:

Yesterday Lee could lift 100 kilos.

*Could yesterday Lee lift 100 kilos?

And ...

- You know those women.
- Mary left early.
- They went to Sydney.

- *Know you those women?
- *Left Mary early?
- *Went they to Sydney?

- Do you know those women?
- Did Mary leave early?
- Did they go to Sydney?

- Auxiliary verbs but not main verbs are fronted when forming questions
- If a sentence contains no auxiliary verb, but has only a main verb, then the auxiliary verb *do* is introduced and used to form the question

Question Rule #3:

a. To form a question from a declarative sentence, place the first auxiliary verb of the sentence at the beginning of the sentence
b. If there is no auxiliary verb, place an appropriate form of the verb *do* at the beginning of the sentence (and make appropriate changes in the main verb)

But still...

- The people who are standing in the room will leave tonight.
- Many mathematicians who you will meet are thought to be odd.
- Anyone that can lift 100 kilos is eligible for our club.

- *Are the people who standing in the room will leave tonight?
- Will the people who are standing in the room leave tonight?

- The people who were saying that John is sick will leave soon.
- The people who were saying that Pat has told Mary to makeTerry quit trying to persuade David that many mathematicians are thought to be odd will leave soon.

- The people who are standing in the room will leave tonight.
- Many mathematicians who you will meet are thought to be odd.
- Anyone that can lift 100 kilos is eligible for our club.

- Will the people who are standing in the room leave tonight?

- The people who were saying that Pat has told Mary to makeTerry quit trying to persuade David that many mathematicians are thought to be odd will leave soon.

English question rule #4

"To form a question from a declarative sentence, locate the first auxiliary verb that follows the subject of the sentence and place it immediately to the left of the subject."

The psychological reality of constituent structure

- “Click” experiments (Fodor, Bever & Garrett 1974)
- Subjects use major constituent boundaries in their perception of sentences
- Subjects wearing headphones heard a taped sentence in one ear and a click in the other; they were asked to write down where in the sentence they heard the click.
- That the girl was happy | was evident from the way she laughed.

  . . . .
- Subjects showed a tendency to ‘mishear’ the location of the click – as closer to the major break than it actually was
- Major constituents resist interuption
- (However, not uncontroversial)

Identifying constituents: tests for constituent structure

We have already seen one of the main tests for constituent structure: **substitution**.

The elephant looks unhappy.

It looks unhappy.

*The it looks unhappy.

*The it unhappy.

The very big elephant with the purple mottled trunk looks unhappy.

It looks unhappy.

Identifying constituents: tests for constituent structure

- **Movement tests:** Can the same string of words appear as a unitary sequence in different positions in the sentence?

- Classic problem

Sue ran up the hill.

Sue rang up her mother.

Up the hill, Sue ran.

*Up her mother, Sue rang.

- **Sentence fragment test**: Can the sequence be used alone?

Where did Sue run?
Up the hill.

Who did Sue ring?
Up her mother.

- **Coordination test**: Constituents of the same type can be coordinated

Sue ran up the hill and over the other side.

*Sue rang up her mother and up her aunty.

- **‘cleft construction’ test**

it is/was X that/who/which Y

It was up the hill that Sue ran.
*It was up her mother that Sue rang.

Sue ran [up the hill].
Sue rang up [her mother].

*Sue ran the hill up.
Sue rang her mother up.

Grammars

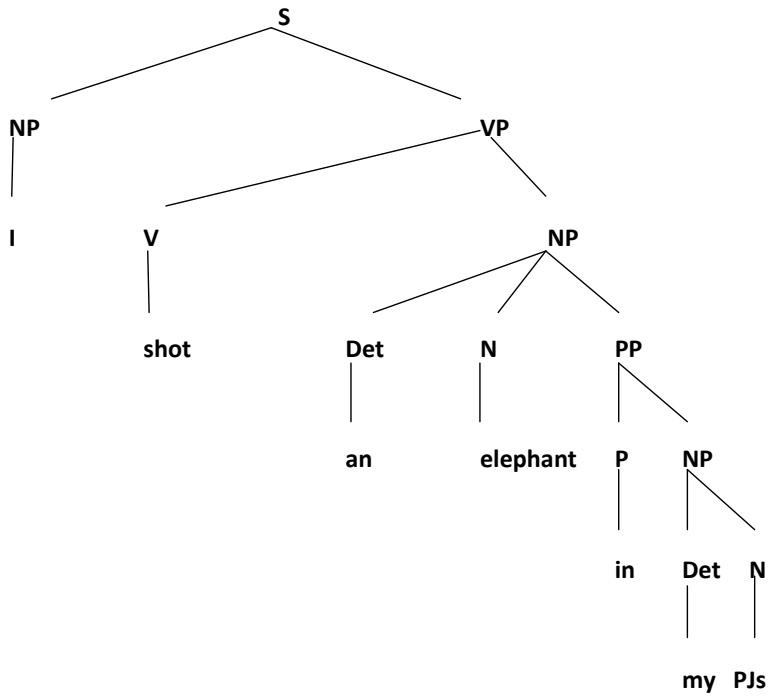
- One of the things that grammars do is to specify the constituent structure of grammatical sentences ...

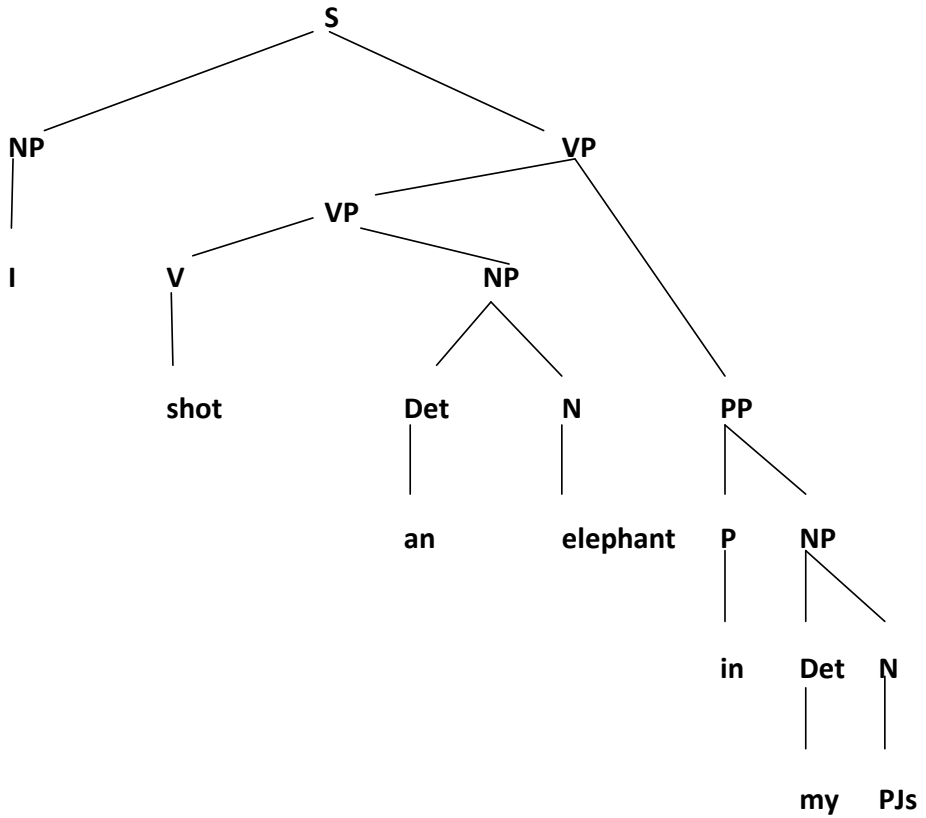
More Intuitions

- While hunting in Africa, I shot an elephant in my pajamas. How an elephant got into my pajamas I'll never know.

  [From the Groucho Marx movie *Animal Crackers*]

‘PP attachment ambiguities’





Recursion

S $\rightarrow$ NP VP

NP $\rightarrow$ Det N PP

PP $\rightarrow$ P NP

Recursion arises where a ‘phrase structure rule’ or a pair of phrase structure rules, allow for the same category to occur on the left of the arrow as on the right.

Languages (and therefore grammars) are recursive, creative and infinite....

Summary

Types of syntactic information encoded in grammars:

- The minimal lexical units of the sentence (words and morphemes)
- Their linear order
- Constituent structure
- Lexical and phrasal categories
- Grammatical function [maybe]

Additional reading

- Linguistic sections of Ch 8 of your textbook, esp. 8.1,8.2,8.3
- Any introductory linguistics textbook, chapter on syntax, e.g.:

Fromkin, V, R Rodman, N Hyams, P Collins, M Amberber and M Harvey. 2009. *An Introduction to Language* (sixth edition) Melbourne: Cengage Learning.

Finegan, E., Blair, D. and Collins, P. 1997. *Language. Its structure and use.* (Second edition.) Sydney: Harcourt Brace.

[This is the Oz edition: or, Finegan 1997 referenced in your text.]

O'Grady, W., M. Dobrovolsky and F. Katamba. 2004. *Contemporary Linguistics: An Introduction* . New York: Longman.

- See also the references to syntax texts at the end of Ch 8