Natural Language Processing: How do humans process language?

Philipp Gabler <pgabler@student.tugraz.at>

2020-05-07

Outline

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Motivation

What does NLP have to do with humans, at all?

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Fundamental questions of linguistics

- What do you know when you know a language?
- What do you know when you understand an utterance?

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Too much theory is bad? But why?

- "Every time I fire a linguist, the performance of the speech processing system goes up." (Frederick Jelinek)

- Does it mean we should refrain from linguistic inspiration?

    - (NLP already does that. Ask a linguist.)

- Cf. the good, bad, and ugly parts of artificial neural networks

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Linguists and Engineers tend to have different focus

- Computational: what is explained?

  - Description of linguistic performance vs. explanation of linguistic competence

- Algorithmic: how is it done?

  - Cognitive realism, computational complexity/efficiency

- Implementational: how is is realized?

  - Neurological plausibility

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

What this lecture is about

A very short introduction to:

- Grammar theory
    - What is language built of?
- Cognitive linguistics
    - How does language work in the mind?

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Get a better understanding of what should work in language processing

- After all, it's <u>natural language</u> processing

- Comparison gives confidence:
  - NLU system behaviour vs. L1 acquisition
  - Observation of similar effects/errors, e.g., garden path sentences
  - Human performance is the ultimate (utopic?) benchmark!
  - We're not inventing something <u>new</u>…

We don't yet know how human language really works

- Very conflicting hypotheses, most of which work only on a computational level

- New ideas:
    - Shallow processing
    - Distributed, implicit, usage-based knowledge
    - Computational construction grammar
    - Computational semantics (λ calculus)

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Be warned!

- This is will be an extremely rough, simplified, and incomplete overview

- It is biased in favour of Cognitive Linguistics (and a bit against Generative Grammar)

- Linguistic theory is not rigorously formal
  - "Theory" = "proposed descriptive model", not "axiomatic system"

- If you're interested: go to the linguistics department
  - Sprache und Kognition, Sprachen der Welt, …
  - Learn more languages (for grammar, not talking)

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Models of human language

Some examples from different areas of linguistics and cognitive science

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Cognitive & Linguistic Development

Cognitive abilities develop in similar ways

- Typical progress:
    - Statistical learning (expectation & surprise)
    - Inductive learning (categorization & abstraction)
    - Social learning (imitation, intention, theory of mind)
- Sensomotory system has an important influence in learning!
- Critical periods vs. extreme robustness

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Language learning tends to follow a U-shaped progress

- Phases:
    - Simplification: How do you do dese…work/tortillas/in English
    - Overgeneralization: Yesterday I didn't painting; it noises
    - Restructuring How do you…make this/like it; how…do cut it
- Cf. exploration vs. exploitation in reinforcement learning
- Computational and associative learning

Creolization processes

Figure: Hotel room signs in Tok Pisin (Papua New Guinea)

Philipp Gabler <pgabler@student.tugraz.at>
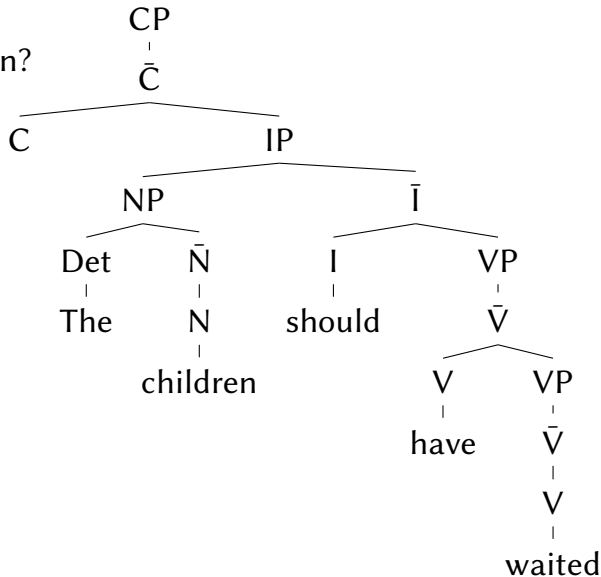2020-05-07

Is *langage*[1] special?

- Is language based on common cognitive machanisms?
    - Categorization, association, memory, hierarchy…
- Or is there a specialized, innate language mechanism?
    - Mental grammar, language acquisition device, Universal Grammar

---

[1]This is not a typo, but French.
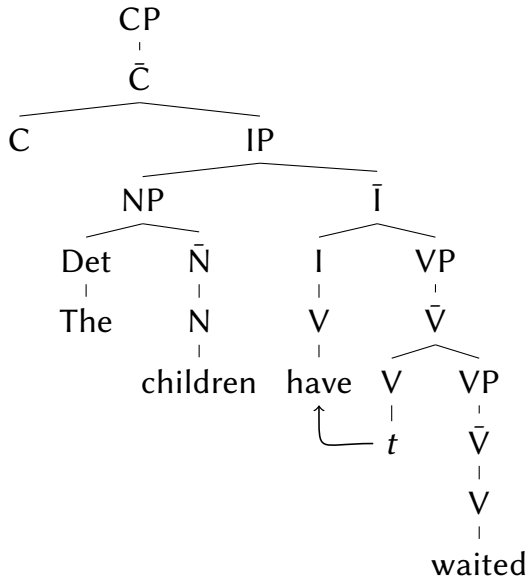
Generative Grammar = trees + transformations

- Grammatical construal in terms of rules
    - from <u>deep structure</u> to <u>surface structure</u>
- Exlaining all languages in terms of <u>principles and parameters</u>
    - Solution to fast, one-shot L1 acquisition

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Triangles in the brain?


```
                    CP
                    |
                    C̄
           ┌────────┴────────┐
           C                 IP
                    ┌────────┴────────┐
                   NP                 Ī
              ┌─────┴─────┐      ┌─────┴─────┐
             Det          N̄      I           VP
              |           |      |           |
             The          N    should        V̄
                          |         ┌─────────┴─────────┐
                       children     V                   VP
                                    |                   |
                                   have                 V̄
                                                        |
                                                        V
                                                        |
                                                      waited
```

CP
|
C̄

C   IP

NP   Ī

Det   N̄   I   VP

The   N   V   V̄

children have   V   VP

t   V̄

V

waited

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Criticism of this kind of analysis

- Explicitely not empirical (at least by Chomsky)

    - Against "behaviourism", focus on competence
    - Tends to categorize everything in terms of recursive symbolic structures
    - Good for English – what about Chinese? Pirahã? *Conversational* English?

- Computationally complex, cognitively... difficult to explain

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Pushing the Boundaries of Generative Grammar

Language processing is basically an inverse problem:

- Colorless green ideas sleep furiously

- The Sally hugged him the Thomas

- Time flies like an arrow

- The apartment that the maid who the service had sent over was decorated

- Keine Kopfverletzung ist zu harmlos um sie nicht zu ignorieren

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Language is conveying mental state through symbols

- Grammar is only an "artifact" to structure the transportation of mental state
    - Or: only an instrument for performative utterance
- Semantics from a cognitive perspective: meaning is…
    - perspectivic (relative to utterance context)
    - dynamic (system changes with environment)
    - encyclopedic (association with experiences & culture)
    - determined by usage (a system derived from concrete experience)

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Some cognitive approaches to semantics and grammar

- How is meaning represented?

    - Prototypes, radial networks, schemata, …
    - Metaphor

- How is meaning expressed through form?

    - Construction grammar, grammatical construal, usage-based grammar…
    - Information structure

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Information Structure (aka Information Packaging)

Conveying more information beyond denotation

- Intonation can focus different parts of an utterance

    - John only introduced Bill to <u>Sue</u>
    - John only introduced <u>Bill</u> to Sue
    - John only introduced <u>Bill</u> to <u>Sue</u>
    - John only <u>introduced</u> Bill to Sue

- Differences in meaning independent of linguistic form!

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Information Structure (aka Information Packaging)

Constructions that relate meaning in conversation[2]

- Different pragmatic practices are associated with:

    - As for John, he lost his wallet
    - What happened was that John lost his wallet
    - What John did was lose his wallet
    - It was John who lost his wallet
    - What John lost was his wallet

---

[2]See Martin Hilpert's lectures: https://www.youtube.com/watch?v=PJecXZp_SYw

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Constructions everywhere

- Constructions are patterns whose form or meaning is not strictly predictable from their components:

  - He has whiffled my borogroves completely vorpal again
  - *The knife chopped the carrots into the salad

- Embedded items are <u>coerced</u>:

  - There was cat all over the road
  - She smiled herself an upgrade

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Not just arbitrary idioms and poetry!

- We understand things in terms of metaphor, and use it all the time[3]

- Abstract term = container

  - An argument has a hole, has less substance, does not have content
  - To find something in an argument

- Argument = journey

  - The content of the argument proceeds, path to the core of the argument, the direction has no substance

---

[3]See *Metaphors we live by* by John Lakoff

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Applications

What does theory have to do with NLP, at all?

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07

Thank You!

Next: ???

Philipp Gabler <pgabler@student.tugraz.at>
2020-05-07