Foundations of Linguistics

MSc Cognitive Systems
University of Potsdam

Tatjana Scheffler

tatjana.scheffler@uni-potsdam.de

Oct. 16, 2018
Course Details

- Large self-study component
- Approximately biweekly schedule:
  - Reading
  - Meet to discuss reading; previous and new problem sets
  - Do problem sets
- Exam: Oral exams
- Next Meeting: Oct 30
  - discuss Morphology reading and any question about problem sets
Foundations of Linguistics

Why do we care?
In modern linguistics, a grammar is viewed as a set of abstract devices, rule systems and principles that serve to characterize the well-formed sentences of a language.

(1) I ate lunch with him.  \[\Rightarrow\] well-formed, grammatical

(2) * Lunch with ate I him.  \[\Rightarrow\] ill-formed, ungrammatical

Cf. a formal language like html:

```html
<meta name="description" content="Die Webseite von Tatjana Scheffler."
```
Grammar 2/3

- Descriptive grammar vs. prescriptive grammar:
  (1) John doesn’t wanna eat.

- Grammaticality vs. processing difficulty:
  (2) The mouse the cat the kid likes caught escaped.
      The mouse escaped.
      The mouse the cat caught escaped.
      The mouse the cat the kid likes caught escaped.
Grammars of natural languages are psychologically real, they are in our minds, they are part of our cognitive systems.

Linguistic competence

≠

Linguistic performance
Ways of „doing Linguistics“

- **Generative grammar** (→ Noam Chomsky)
  - What do natural languages have in common („universals“)?
  - How can we model man’s knowledge of language and language processing?
  - Important method: Introspection, intuitive judgement on „grammaticality“

- **Structuralism** (→ Ferdinand de Saussure, Roman Jakobson)
  - Investigate the mechanisms of culturally-transmitted symbol systems
  - How can we describe a linguistic entity/phenomenon in relation to the overall system?
  - Important method: Qualitative analysis of language data

- **Corpus Linguistics** (→ Henry Kucera)
  - What patterns can be observed in language data?
  - How can we model „language use“ for one particular language?
  - Important method: Quantitative analysis of language data („corpora“)
Levels of Linguistic Analysis

- Pragmatics – Discourse/Context
- Semantics – Meaning
- Syntax – Grammar
- Morphology – Word Formation
- Phonology – Sound System
- Phonetics – Sounds
Levels of Linguistic Analysis

- Pragmatics – Discourse/Context
- Semantics – Meaning
- Syntax – Grammar
- Morphology – Word Formation
- Phonology – Sound System
- Phonetics – Sounds
Morphology
Syntax

Nach der zu langen Party sind wir sehr schnell zum See gefahren.
Semantics

“What’s your sister doing?”

“She’s painting a cow.”
Semantic knowledge

**SEMANTICS**
- **LEXICAL**
  - Simple units
  - Features
  - Semantic fields
  - Characterization in terms of Model Theory
  - etc.
- **COMPOSITIONAL**
  - Complex units
  - Procedure to derive the meaning of complex units from that of simple units.
Pragmatics

@James_Kpatrick: These two books contain the sum total of all human knowledge

Pragmatics

Grice’s ‘Logic and conversation’ (1975) is a defining moment in pragmatic theory. It strikes a balance between the two extremes described above, and it outlines a general theory of how to allow semantics and pragmatics to work together to produce linguistic meaning.

1.3 Some pragmatic phenomena

i. Quantifier domains

Why does everyone so rarely quantify over everyone?

The ‘everything bagel’

@patrickmarkryan: Come on, Everything Bagels, who you tryin’ to fool? You got like 6 seasonings on there. That’s a lot, but it ain’t everything.

@dwineman: Last time I had an everything bagel I got poppy seeds, Mira Sorvino, and Hegel’s Phenomenology of Spirit all over my shirt.

@James_Kpatrick: These two books contain the sum total of all human knowledge

ii. Modality

In elementary school, students were often saying “Can I go to the bathroom” rather than “May I...”. I consider this an injustice; can has both ability and deontic (permission-oriented) readings, and the teachers knew which we meant!

iii. Scalar inferences

Why do some and most tend to exclude all? Why does three tend to mean exactly three? Why does few tend to exclude no?

iv. Focus effects

How does Well, Ellen didn’t read the book come to suggest (implicate) that Ellen did something else with the book? (Compare, ELLEN didn’t read the book; see Kratzer 1991; Büring 1999.)

Francis Ford Coppola’s The Conversation

The entire movie turns on whether the man being recorded said He’d kill us if he got the chance. or HE’d kill US if he got the chance.

Harry Caul is tormented by the question of which he hears on the recording.

3
Semantics vs Pragmatics

- **Semantics**: What does a sentence / an utterance mean „generally“, i.e. independent of context

- **Pragmatics**: What does a sentence / an utterance mean in a particular context
  - the situation of uttering
  - the preceding linguistic context
The meaning of a complex expression does not only depend on its lexical units but also on the way these are combined syntactically and phonologically.

(1) a. The panic among the visitors caused a stampede.
   b. A stampede caused the panic among the visitors.

(2) a. I only gave ANNA a book. \(\Rightarrow\) “only Anna”
   b. I only gave Anna a BOOK. \(\Rightarrow\) “only a book”
Linguistic Knowledge

- Semantic-pragmatic knowledge is productive. We understand utterances that we have never heard before.

(1) I saw a pink whale in the parking lot.

cf. Addition of two new numbers:

(2) 1437,952 + 21,84
Knowledge of the meaning of the lexical units and of the combination procedure is by and large unconscious (as opposed to that in arithmetic operations).

Beispiel 1: Partikel *ja*

(1) Der Tatort letzte Woche spielte ja wieder in Münster.
Why bother?
(1) a. Kim sent Pat Chris.
   b. Kim sent Pat to Chris.
   c. Kim was sent to Pat by Chris.
   d. Kim was sent Pat by Chris.
(2)  a. 田中 が  ライオン を 食べた。
    Tanaka ga  raion  wo  tabe-ta
    Tanaka NOM lion  ACC eat-PST
    ‘Tanaka ate the lion.’ [jpn]

b. 田中 を ライオン が 食べた。
    Tanaka wo  raion  ga  tabe-ta
    Tanaka ACC lion  NOM eat-PST
    ‘The lion ate Tanaka.’ [jpn]

c. 田中 が ライオン に 食べられた。
    Tanaka ga  raion  ni  tabe-rare-ta
    Tanaka NOM lion  DAT eat-PASS-PST
    ‘Tanaka was eaten by the lion.’ [jpn]

d. 田中 が ライオン に ケーキ を 食べられた。
    Tanaka ga  raion  ni  keeki  wo  tabe-rare-ta
    Tanaka NOM lion  DAT cake  ACC eat-PASS-PST
    ‘The lion ate the cake (to Tanaka’s detriment).’ [jpn]
(3)  a. Kim gave Sandy a book.
b. Kim gave a book to Sandy.
c. A book was given to Sandy by Kim.
d. This is the book that Kim gave to Sandy.
e. Which book do you think Kim gave to Sandy?
f. It’s a book that Kim gave to Sandy.
g. This book is difficult to imagine that Kim could give to Sandy.
\(\sim 7000\) languages

- All current NLP methods are worse for non-English.
- Not only due to sparser data.
- In some ways, many languages are genuinely ‘harder’ than English:
  - Free word order, scrambling.
  - Rich morphology, prefixing, etc.
  - Spelling systems.
  - Sounds.
  - …
Next time
For October 30

- read:
  - Bender (2013), ch. 2-4
  - Fromkin et al. (2003), ch. 3 (optional)

- answer the discussion questions

- Please have a look at the problem set in advance. It is due the week after.
Thank you

tatjana.scheffler@uni-potsdam.de