JMORF — Morpho-Syntax

Introduction, Organization Morphology and Syntax

Francis Bond
Palacký University
https://fcbond.github.io/
bond@ieee.org

Lecture 1
Location: SV 2.39

Overview

- Syllabus; Administrivia
- Prescriptive/descriptive grammar; Competence/performance
- Some history
- Why study syntax?
- Morphology

Administrivia

Coordinator Francis Bond <bond@ieee.org>

All other details on the web page

100% Continuous Assessment

- Weekly Problems (50%)
- Mid-term (20%)
- Final (30%)

What do you learn?

On completion of this module, students should be able to:

- Understand the basics of morphology
- Recognize certain classes of syntactic phenomena
- Build analyses of those phenomena in a precise framework
- Apply the process of building a formalized analysis to test linguistic hypotheses
- Know a little about different approaches to the study of syntax
- Be able to present linguistics using LaTeX

Textbook and Readings

- Textbooks
  - Sag, Wasow and Bender 2003 *Syntactic Theory: A Formal Introduction* 2nd ed. CSLI (**required**)
  - Panocová, Renáta *Basic Concepts of Morphology I* Univerzita Pavla Jozefa Šafárika (**recommended**)
  - You should read all chapters assigned before class.
  - Ideas from the book will be pursued in parallel with the topics given above.

Student Responsibilities

By remaining in this class, the student agrees to:

1. Make a good-faith effort to learn and enjoy the material.

2. Read assigned texts and participate in class discussions and activities.

3. Submit assignments on time.

4. Attend class at all times, barring special circumstances (see below).

5. Get help early: approach us when you first have trouble understanding a concept or homework problem rather than complaining about a lack of understanding afterward.

6. Treat other students with respect in all class-related activities, including on-line discussions.

Attendance

1. You are expected to attend all classes.

2. Be on time - lateness is disruptive to your own and others' learning.

3. Valid reasons for missing class include the following:

   (a) A medical emergency (including mental health emergencies)
   (b) A family emergency (death, birth, natural disaster, ...)
   (c) An important event (tournament, job interview, ...)

4. There will be significant material covered in class that is not in your readings. You cannot expect to do well without coming to class.

5. If you miss a class, it is your responsibility to get the notes, any handouts you missed, schedule changes, etc. from a classmate.

Remediation and Academic Integrity

1. No late work will be accepted, except in the case of a documented excuse.

2. For planned, justified, absences on class days or days on which assignments are due, advance notice must be provided.

3. Cheating will not be tolerated. Violations, including plagiarism, will be seriously dealt with, and could result in **a failing grade for the entire course**.

4. For all other issues of academic integrity, refer to the University Honour Code

5. As always, use your common sense and conscience.

The winning strategy

- Read the books before class (and after again, if necessary)
- Work together: make study groups
- Homework: Discuss as much as you want, write up your own answers
- Exams: No discussion
- Ask questions ...early and often!

Resources

- Glossary at back of textbook
- Grammar summaries and Appendix A
- Answers to exercises at back of book
- Each other, grad-students, office hours, ...
- Online:
  - English Resource Grammar: delph-in.github.io/delphin-viz/demo/
  - Wikipedia has good summaries and many links

Layers of Linguistic Analysis

1. Phonetics & Phonology
2. **Morphology**
3. **Syntax** (Grammar)
4. Semantics
5. Pragmatics
6. Stylistics

Two Conceptions of Grammar

- PRESCRIPTIVE
  - Rules against certain usages. Few if any rules for what is allowed
  - Proscribed forms generally in use
  - Explicitly normative enterprise
- DESCRIPTIVE
  - Rules characterizing what people do say
  - Goal to characterize all and only what speakers find acceptable
  - Tries to be scientific

Uses of Grammar

- PRESCRIPTIVE
  - Identify speaker's socioeconomic class & education level
  - Identify level of formality of a particular usage
  - Standardize language for smoother communication
- DESCRIPTIVE
  - Understand how people produce & understand language
  - Identify similarities & differences across languages
  - Development of language technologies

Prescriptive grammar

- Examples of silly prescriptive rules?
- Examples of useful prescriptive rules?
- Some applications which might need to encode prescriptive rules?

Fill in the blanks:

he/his, *they/their*, or something else?

(1) *Everyone insisted that _____ record was unblemished.*
(2) *Everyone drives _____ own car to work.*
(3) *Everyone was happy because _____ passed the test.*
(4) *Everyone left the room, didn't _____?*
(5) *Everyone left early. _____ seemed happy to get home.*

Descriptive Grammar: an example

(6) *F_____ yourself!*
(7) *Go f_____ yourself!*
(8) *F_____ you!*
(9) **Go f_____ you!*

- Who taught you this?
- How did you learn it?

Kinds of Things We'll Worry About

- Where to use reflexives (e.g. *myself*) vs. ordinary pronouns (*I, me*)
- Agreement (e.g. *We sing* vs. **We sings*)
- Word order (e.g. **Sing we*)
- Case (e.g. **Us sing*)
- Coordinate conjunction (e.g. *We sing and dance*)
- How to form questions, imperatives, negatives, ...

... and much more

Competence vs. Performance

- The Distinction
  - Competence - knowledge of language
  - Performance - how the knowledge is used

- Examples

  (10) *That Sandy left bothered me.*

  (11) *That that Sandy left bothered me bothered Kim.*

  (12) *That that that Sandy left bothered me bothered Kim bothered Jo.*

  (13) *The horse raced past the barn fell.*

Competence v. Performance

(14) *You are what you eat*

(15) *You are what what you eat eats, too*

(16) *You are what what what you eat eats eats, too*

Acceptability vs. grammaticality

- A sentence is **acceptable** if native speakers say it sounds good.
- A sentence is **grammatical** (with respect to a particular grammar) if the grammar licenses it.
- Linguists are sometimes sloppy about the difference.
- Some people argue that it should be modeled probabilistically rather than as a binary distinction
  - It depends on individual speakers
  - But we often want to model groups of speakers
  - It is good to combine judgments with attested data
    but language is infinite, so we may not find the example we need attested

Some History

- Writings on grammar go back at least 3000 years
- Until 200 years ago, almost all of it was prescriptive
- Until 70 years ago, most linguistic work concerned sound systems (phonology), word structure (morphology), and the historical relationships among languages

The Generative Revolution

- Noam Chomsky's work in the 1950s radically changed linguistics, making syntax central.
- Chomsky has been a dominant figure in linguistics ever since.
- The theory we will develop (HPSG) is in the tradition started by Chomsky, but diverges from his work in many ways.

Main Tenets of Generative Grammar

- Grammars should be formulated precisely and explicitly.
- Languages are infinite, so grammars must be tested against invented data, not just attested examples.
- The theory of grammar is a theory of human linguistic abilities.

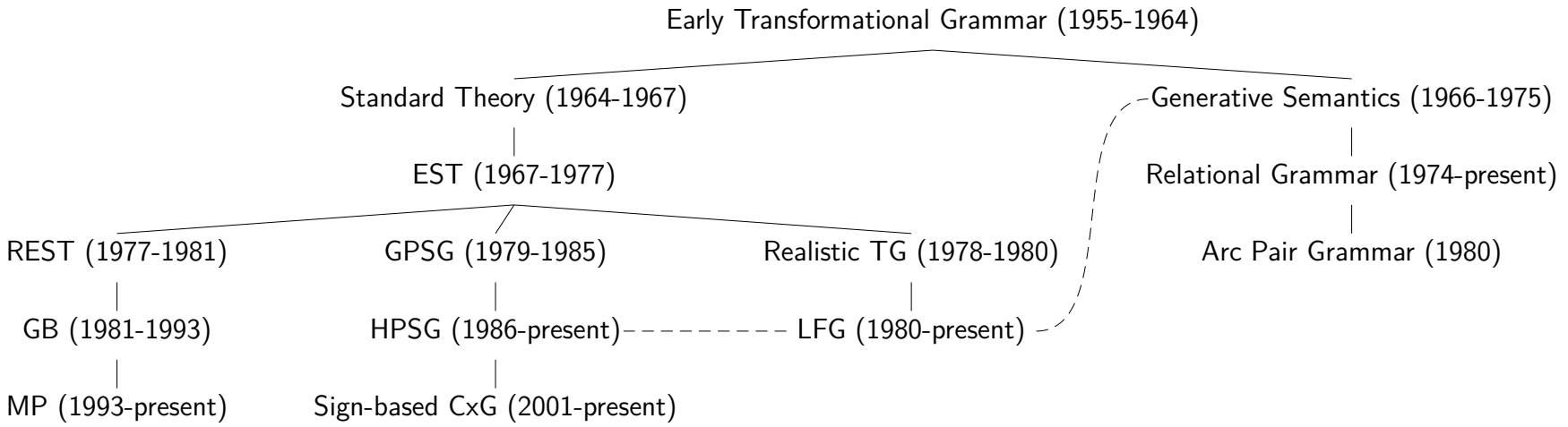
What does a theory do?

- Monolingual
  - Model grammaticality/acceptability
  - Model relationships between sentences (internal structure)
- Multilingual
  - Model relationships between languages
  - Capture generalizations about possible languages

Some of Chomsky's Controversial Claims

- The superficial diversity of human languages masks their underlying similarity.
- All languages are fundamentally alike because linguistic knowledge is largely innate.
- The central problem for linguistics is explaining how children can learn language so quickly and easily.

Family Tree of Generative Syntactic Theories



- Many Other Theories
  - Dependency Grammar (links words not phrases)
  - Combinatory Categorical Grammar (allows multiple derivations)
  - Tree Adjoining Grammar (links subtrees)
  - Functional Grammar (considers function to be central)
    - Systemic Functional Grammar
    - Role and Reference Grammar
  - Biosemiotics (how living organisms produce, interpret, and exchange signs and meanings)

Which theory is best?

- Most theories are better at explaining some things than others
  - HPSG is good at modelling structure and (some) meaning
  - Minimalism is good at modelling similarities
  - Dependency grammar is easy to implement (and good for case marked languages)
  - Functional grammar is good at describing why we choose to use certain constructions
  - Conversation Analysis is good at modelling turn taking and social interaction

- I teach HPSG because I know it well
  - I know HPSG because I sat next to Tsunkeo Nakazawa at NTT
  - I think it is a good model of syntax and semantics
  - I don't think it is the only valid way of studying language

- The general approach to analysing language should be transferable to any theory

What is Morphology

- **Morphology** is the study of form and structure.

- In linguistics, it generally refers to the study of form and structure of words.

horses	horse-s
talked	talk-ed
unhappiness	un-happy-ness
went	go-ed
yes	yes
$talk_N$	$talk_V$
psa	pes+l

Words and morphemes

There are two main usages of the term **word**:

1. Surface form (spoken or written representation)

2. Abstract form (lemma or dictionary entry, e.g. bare infinitives in English; nominative singular in Latin)

The class of forms representing a word in different contexts is called a **lexeme**.

$$\textbf{\textit{sing}} = \{\textit{sing}, \textit{sings}, \textit{sang}, \textit{sung}, \textit{singing}\}$$

A definition of words?

- **Words** can be described as units of language (sequences of sounds or signs) that function as meaning bearers. But this is a fuzzy notion, e.g.:
  - *sang* expresses both "singing" and past tense.
  - Is *more or less* "roughly" one word, or are there three words?

A structuralist solution: **morphemes**.

Morphemes and Morphological analysis

- **Morphemes**
  - **Morphemes** are minimal meaning-bearing units.
    Example: *talked* contains two morphemes: *talk* and *-ed* (past).
  - Form–function pairs (sound/sign–meaning): basic units of morphology.
  - The realisations of morphemes are called **morphs**.
    Example: English plural morpheme **[number pl]** has allomorphs *-s*, *-es*, *-en*, $\emptyset$: *boy-s*, *box-es*, *ox-en*, *sheep*.
  - These different realisations of the same morpheme are called **allomorphs**.

- **Morphological analysis**
  - Segmentation of expressions into basic units (mostly starting from word-level).
  - Classification of these basic units according to function.

A language:

11–112 **phonemes**

↓

4,000–10,000 **morphemes**

↓

An infinite number of **sentences**

Types of morphemes

- **Free Morphemes**
  - **Free morphemes** can occur independently. Common in Chinese and English. Examples: *boy*, *sing*, 狗 *gǒu* "dog".

- **Bound Morphemes**
  - **Bound morphemes** must be attached to another morpheme and cannot be used independently.
    Example: **[number pl]** *-s* → *boys*.
  - Typical bound morphemes are: **affixes** (*boy+s*, *talk+ed*); **clitics** (French: *je ne sais pas*, *je* and *ne* cannot occur without a verb); **roots** (Spanish *habl-* needs an ending indicating person, number, mood, etc.).

Formatives and pseudo-morphemes

Morphemes are form–meaning pairs, but not all segmentable forms have an identifiable meaning.

- **Formatives**
  - Forms without identifiable meaning.
    Example: linking elements in German compounds: *Geburt+s+tag* (birthday), *Schwan+en+hals* (swan neck).

- **Pseudo-morphemes / cranberry morphemes**
  - Special cases of formatives: segmentable parts of a complex word without independent meaning.
    Examples: *cran+berry*, *rasp+berry*; *re+ceive*, *con+ceive*.

What is morphology? (follow up)

Morphology can refer to three different things:

1. Description of the behaviour of morphemes and how they are combined.

2. Derivational, inflectional and compositional processes of word formation occurring in a specific language.
   Example: "German has a richer morphology than English."

3. Description of such word-formation processes (i.e., the theory/grammar of a language's morphology).

Inflectional Morphology

- **Inflection** is required by syntactic criteria (e.g., an English verb must have tense).

- It marks grammatical (**morphosyntactic**) distinctions:

  Conjugation (verbal): person, number, gender; tense, aspect, mood; agreement
  Declension (nominal): case, number, gender, degree, definiteness

- Meaning of the general concept is generally not changed, though *when*, *who/what*, and sometimes *where/how/whether* may be specified by inflectional morphemes.

- Some people consider bound and free inflectional morphemes:

  – *go [TENSE past] : went*
  – *go [TENSE future] : will go*

Inflection — paradigm

"A set of forms having the same root/stem, one of which must be selected in a certain syntactic environment" (based on Crystal 1997:277; Payne 1997:26).

For instance, German conjugation:

present			**past**		
	singular	**plural**		**singular**	**plural**
1.	*dehn-e*	*dehn-te*	1	*dehn-en*	*dehn-te-n*
2.	*dehn-st*	*dehn-te-st*	2.	*dehn-t*	*dehn-te-t*
3.	*dehn-t*	*dehn-te*	3.	*dehn-en*	*dehn-te-n*

Paradigm — Czech noun first declension *pes* "dog"

case	singular	plural
NOM	*pes*	*psi*
GEN	*psa*	*psů*
DAT	*psoive, psu*	*psům*
ACC	*psa*	*psy*
VOC	*pse*	*psi*
LOC	*psovi, psu*	*psech*
INST	*psem*	*psy*

- **syncretism**: the same form expresses different feature combinations.
  psa is both **genitive** and **accusative**

- **exponence**: relation between form and function is $m:n$.
  - **multi-exponence (cumulation)**: one form expresses several functions.
    -ů expresses both **genitive** and **plural**.

Derivational Morphology

- Builds complex words by combining bound and free morphemes.
- **Derivational** operations are optional (not required by syntactic criteria).
- They may change:
  1. **semantics**, e.g. *clear* $\rightarrow$ *un+clear* = *unclear*
  2. **syntactic category**, e.g. $derive_{\mathsf{V}} + ation_{\mathsf{N}} + al_{\mathsf{Adj}} = derivational$
  3. **valency of a verb**, e.g. Havasupai *qaw* 'it breaks' $\rightarrow$ *t+qaw* 'he breaks it'
  4. and more

Compounding

- Builds complex words by juxtaposition of free morphemes.
- Examples: *[sale]+s+[man]*, *[dish]+[washer]*
- Productive compounding results in an infinite lexicon.

$$\left\{\begin{array}{l} English \\ German \\ Czech \end{array}\right\} \left\{\begin{array}{l} phonetics \\ morphology \\ syntax \end{array}\right\} \left\{\begin{array}{l} teacher \\ researcher \\ student \end{array}\right\}$$

There are many other derivational patterns

- **Reduplication**: *teeny-weeny*
- **Clipping**: *ad[vertisment]*, *[in]flu[enza]*
- **Acronym**: *laser* "Light Amplified by Stimulated Emission of Radiation", *hpsg* "Head-Driven Phrase Structure Grammar", **UPOL**: "U[niverzita] P[alackého] [v] Ol[omouci]", " "
- **Blending**: *brunch* "breakfast + lunch", *srandista* "sranda (fun, joke) + humorista (humorist)" → "joker, funny person"
- **Expletive infixing**: *Abso-fucking-lutely* (but not *Absolute-fucking-ly* or *Abso-fucking-lute*!)

Conclusions

- Next week we will look into some simple theories of

Acknowledgments and References

- Course design and slides borrow heavily from Emily Bender's course: *Linguistics 566: Introduction to Syntax for Computational Linguistics*
  http://courses.washington.edu/ling566

- Morphology slides borrow from Antske Fokkens

- Thanks to Na-Rae Han for inspiration for the student policies (from *LING 2050 Special Topics in Linguistics: Corpus linguistics*, U Penn; adapted).