Syntactic Theory

Background and Transformational Grammar

Dr. Dan Flickinger & PD Dr. Valia Kordoni

Department of Computational Linguistics
Saarland University

October 28, 2011

Early work on grammar

- There is a long tradition of describing the structure of language
- In most cases, language was analyzed so classical texts could be read
- Grammar described archaic forms of language

Examples of early grammarians and linguistic work

- India: Pāṇini (estimated 4th century B.C.)
- China: Erya (author unknown) (3rd c. B.C.)
- Greece: Dionysius Thrax (2nd c. B.C.), Apollonius Dyscolus (2nd c. A.D.)
- Rome: Donatus (4th c. A.D.), Priscian (6th c. A.D.)
- France: Lancelot et al. (1660) *Grammaire générale et raisonnée* (Port Royal)

Pāṇini's grammar

- Sanskrit grammar, said to be short and complete
- Includes topics of syntax, morphology, phonology and pragmatics
- Especially known for the *Aṣṭādhyāyī*
  - describes algorithms that can be applied to lexical items to form words
  - systematic and highly technical
  - focus on brevity: difficult to read
- Pāṇini is said to have influenced the foundations of many aspects of modern linguistics
  - Structuralism (Ferdinand de Saussure and Leonard Bloomfield)
  - Generative grammar (Noam Chomsky)
  - Optimality theory

Diachronic linguistics

- Discovery of Sanskrit and its obvious resemblance to Latin and Greek led to development of comparative linguistics
- Originally focused on languages with written records
- Gradual shift of focus from prescriptive to descriptive grammars

Ferdinand de Saussure (1857–1913)

- Sanskrit scholar
- His course notes were published posthumously by his students in *cours de linguistique générale* (1916)
  - Turned attention from *diachronic* linguistics to *synchronic* linguistics
  - Formulated the arbitrariness of the sign
  - Introduces the terms *langage*, *langue*, and *parole*

de Saussure (cont.)

- *langage*, *langue*, and *parole*
  - **langage** is the faculty of speech: it is heterogeneous, consisting of physical, physiological, and psychological facts
  - A **langue** is a homogeneous system of symbols that may be mapped to meaning: a social product, exterior to individuals
  - **parole** is the act of using language; it is also here where psychology comes into play.
- Saussure's work is seen as the starting point of Structuralism, introducing "syntagmatic analysis": what elements can occur in which context, and what is their contribution to the meaning?

Towards modern syntax

- Structuralism – 1920s-30s: Bloomfield
- Distributionalism – 1950s: Hockett, Harris
- Categorial grammar – 1930s: Adjukiewicz
- Dependency grammar – 1930s: Tesnière

Generative grammar

Chomsky's *Syntactic Structures*

- Main task for linguist: separate grammatical from ungrammatical strings
- Two issues:
  - How to define grammatical strings?
    - Corpus-based or statistical methods fail because of the creative nature of language
    - Grammaticality cannot be determined by 'meaningfulness'
    - His proposed method: native speaker judgments
  - What kind of system can describe all grammatical strings of a language? It must
    - consist of a finite set of rules
    - be descriptively adequate
    - be explanatory

Level of formal grammar

- Will finite-state grammar suffice? Clearly not for English (Chomsky 1957):
  1. If S_1, then S_2
  2. Either S_3 or S_4
  3. If either S_3 or S_4, then S_2
  4. * If S_1, or S_2
- Phrase structure grammar?

Phrase structure grammar (PSG)

Chomsky on PSG:

- Not flawed in the same obvious way that a finite state grammar is
- There are probably languages that cannot be described by a PSG
  - Shown in the 1980s to be correct, for at least for Swiss German and Bambara
- English may be within the descriptive power of a PSG (context-free)
- But there may be other reasons beyond formal power to reject PSGs for representing natural languages ...

Adequacy of a linguistic theory

How to test whether a linguistic theory is adequate?

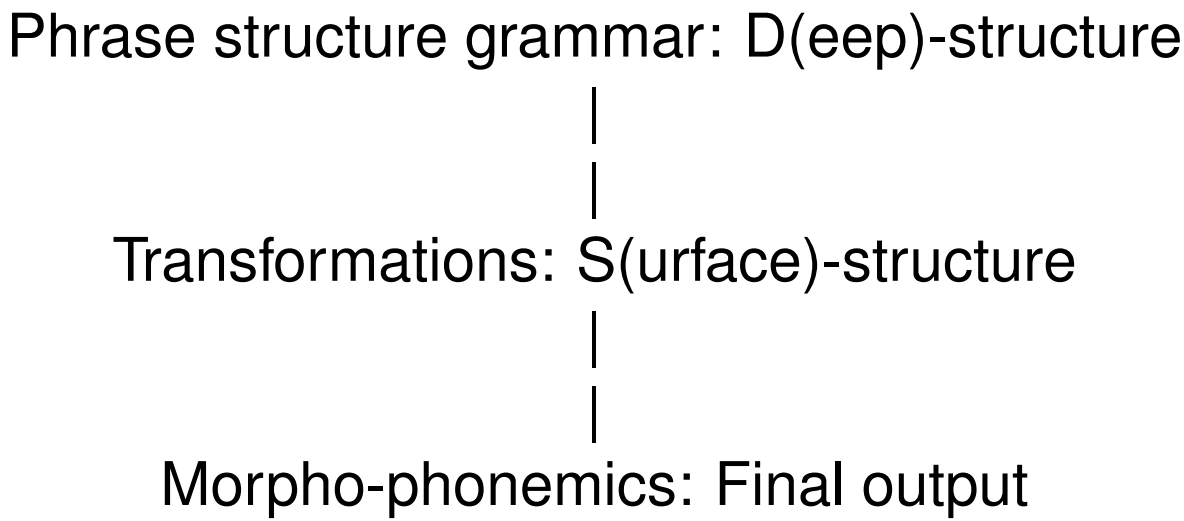
- Can it account for all of the data? (basic requirement)
- Can it account for the data in an elegant, straightforward way, or does it lead to extreme complexity? (cf. learnability)
- Can the same system be used to construct grammars for all languages? (cf. universal grammar)

Limitations of phrase structure grammar

A PSG may be able to generate all strings, but has difficulty with capturing regularities in relations between expressions

- Coordination
  1. The topic of the lecture is syntax.
  2. The topic of the book is syntax.
  3. The topic of the lecture and of the book is syntax.
- Passivization:
  1. Noam Chomsky wrote *Syntactic Structures*.
  2. *Syntactic Structures* was written (by Noam Chomsky).

Three levels of morpho-syntactic representation



Transformations

How to capture grammatical phenomena such as agreement, coordination, passivization?

- Main idea: separate syntactic structures into a deep (underlying) structure and a surface structure (roughly what is observed directly)
- The phrase structure grammar rules define D-structures
- Transformations apply to D-structures to derive S-structures
  – so an active sentence and its passive variant both have the same D-structure

Information in syntactic structures

In addition to defining how a sentence can be analyzed into its constituents (its component parts), we want to know how the parts relate to each other:

- Definitions of grammatical functions
- The lexicon
- Features on categories

Grammatical functions and grammatical categories

- Grammatical functions (subject, object, predicate) are defined in relation to D-structure
  - Subject-of-S [NP, S]
  - Object-of-V [NP, VP]
  - Predicate-of-S [VP,S]
- Syntactic properties are generally represented by (boolean) features:
  - N: [+N,-V]
  - V; [-N,+V]
  - A: [+N,+V]

Subcategorization and lexical insertion

- Lexical items come with a *subcategorization* frame
  - *love*: [V;NP]
  - *smile*: [V:–]
  - *rely*: [V:PP]
  - *think*: [V:S]
- Here the subject is admitted structurally: the subcategorization frame o nly defines the structure of the VP.
- **Lexical Insertion Rule** (Ouhalla 1994, p.50):
  Insert lexical item X under terminal node Y, where Y corresponds to the categorial features of X, and YP corresponds to the subcategorization properties of X.

Transformations: Passivization

Passivization: optional

Structural analysis:

$$NP - Aux - V - NP$$

the dog – past – chase – the cat

Structural change:

$$X_1 - X_2 - X_3 - X_4$$

the dog – past – chase – the cat

$$\downarrow$$

$$X_4 - X_2 + be + en - X_3 - by + X_1$$

the cat – was – chased – by the dog

(Chomsky (1957, p. 112))

PSG and Transformation: Tense

- Starting with PS rule: $S \rightarrow NP\ Aux\ VP$
- Consider the following examples:
  - The boy watched the movie.
  - The boy will watch the movie
  - The boy doesn't watch the movie
  - The boy didn't watch the movie, but his friend did
  - Watch the movie? She wondered whether the boy will
- Tense seems to be part of *Aux* rather than *VP*:
  $S \rightarrow NP\ Aux\ VP$
  $Aux \rightarrow Tense\ (Modal)\ (Neg)$

(based on Ohalla (1994))

PSG and Transformation: Tense (cont.)

- The structure of *The boy watched the movie* is *NP – tense – V – NP*
- The tense marker thus precedes the verb *watch* in the D-structure.
- How can we be sure the tense will be marked on the main verb in the 'spelling-out' phase?
  1. Apply a transformation moving *V* to *Aux*?

     S-structure: $[[_{NP}\textit{The boy}][_{Aux}\textit{watch}_i - \textit{ed}][_{VP}\__i\ \textit{the movie}]]$
  2. Apply a transformation moving tense to V?

     S-structure: $[[_{NP}\textit{The boy}][_{Aux}\__i][_{VP}\textit{watch} - \textit{ed}_i\ \textit{the movie}]]$

(based on Ohalla (1994))

Evidence for moving tense

- Adverbs can precede or follow a VP in English:
  1. The boy cleverly avoided Bill
  2. The boy avoided Bill cleverly
  3. The boy will cleverly avoid Bill
- If V moves to Aux, the verb precedes the VP on the surface
- Adverbs should be able to follow the verb, but
  - *The boy avoided cleverly Bill.
- The conjugated verb thus remains in situ, and tense must move to the VP, if there is no modal verb: ‘affix-hopping’

Transformational grammar: initial stages

- Standard Theory: interpretation from D-structure
- Extended Standard Theory: interpretation from D-structure, S-structure, and possibly the final derived structure
- Trace theory: when transformations move elements around, these elements leave a *trace*:
  $\rightarrow$ semantics can be interpreted from S-structure only

Assumptions in transformational syntax

- There is a difference between competence and performance, i.e. between what speakers know about the language and how they use it.
- Children can and do learn a complex system such as language because the basis is innate: we are born with Universal Grammar pre-installed
- Descriptive adequacy: describe the language (competence) as known by its speakers
- Explanatory adequacy: judge the plausibility of the analysis based on whether it is (easily) learnable given our Universal Grammar

Contributions to syntactic theory

- Syntax was positioned in the center of linguistic research
- The aims of syntactic theory go beyond description:
  - Attention to the (more) formal representation of generalizations
  - Attention to psychological aspects of grammar
- This led to more systematic research to develop relevant linguistic data
  - Native speaker judgments
  - Distinction between grammatical and ungrammatical utterances

Some remarks

- Transformational syntax states that
  - a grammar (PSG + transformations) must be able to generate all expressions that are part of the language
  - A speaker must have access to S-structure and D-structure (in Standard Theory) to interpret an expression
- Hence many take transformational grammar to be a language production model: This is not necessarily the case.
- The primary aim of the transformational approach is to explain how language works as a system that can be learned by children

Some more remarks

- Because a language production/interpretation model is not the aim of transformational grammar, this may not be the most suitable for studying generation
- Despite its considerable advances toward formal description, the details are often not explicit enough for direct encoding in computational models
  – e.g., how does Lexical Insertion work, exactly?
- This shortcoming also applies to some degree to later developments in the Chomsky tradition: $\bar{X}$-theory, Government and Binding, and Minimalism.

Conclusion

- Syntactic theory has deep roots
- Diachronic research led to descriptive linguistics
- Transformational grammar emerged from a need to improve on structural approaches
- Chomsky's Standard Theory provides the foundations for most current work in syntactic theory
- We will see a variety of ways in which research has diverged since then

What to retain from today

- Chomsky's ideas
  - What is the aim of syntactic research, and in particular of transformational grammars?
  - Chomsky's assumptions concerning grammaticality and the innateness of grammar
  -
- The basic architecture of the language model assumed in the transformational approach: D-structure, S-structure, and interpretation of these structures

What you don't have to retain

- Historical names and dates
- How to formalize transformations
- Details of the illustrative analyses for the transformations of Passivization and Affix-Hopping
- Motivation for these analyses over competing alternatives

Suggested further reading

- The first two chapters of Sag, Wasow and Bender (2003)
- The second chapter of
  Ouhalla, Jamal (1994) *Introducing Transformational Grammar*. New York: Oxford University Press.