Computational Linguistics: History & Comparison of Formal Grammars

Raffaella Bernardi

KRDB, Free University of Bozen-Bolzano

P.zza Domenicani, Room: 2.28, e-mail: bernardi@inf.unibz.it

Contents

1. Formal Grammars

We have seen that

- Formal Grammars play a crucial role in the research on Computational Linguistics.
- We have looked at Context Free Grammars/Phrase Structure Grammars, Categorial Grammar and Categorial Type Logic

But through the years, computational linguists have developed other formal grammars too.

Today, we will look at the most renown ones and how the history of Formal Grammar developed.

Tomorrow, we will look at their generative capacity and their complexity.

Remark: The historical overview is by Geert Jan Kruijff. (University of Saarland, Saarbruecken), the general comparison is from Paola Monachesi (University of Utrecht) and Hans Uszkoreit (University of Saarland).

2. History of Formal Grammars

Important steps in the historical developments of Formal grammar started in the 1950's and can be divided into five phases:

1. Formalization: Away from descriptive linguistics and behavioralism (performance vs. competence) [1950's 1960's]
2. Inclusion of meaning: Compositionality [1970's]
3. Problems with word order: Need of stronger formalisms [1970's 1980's]
4. Grammar meets logic & computation [1990's]
5. Grammar meets statistic [1990's 2000's]

In these phases, theoretical linguists addressed similar issues, but worked them out differently depending on the perspective they took:

- constituency-based or
- dependency-based.

2.1. Constituency-based vs. Dependency-based

Constituency (cf. structural linguists like Bloomfield, Harris, Wells) is a **horizontal** organization principle: it groups together constituents into phrases (larger structures), until the entire sentence is accounted for.

- Terminal and non-terminal (phrasal) nodes.
- Immediate constituency: constituents need to be adjacent (CFPSG).
- But we have seen that meaningful units may not be adjacent –Discontinuous constituency or long-distance dependencies.
- This problem has been tackled by allowing flexible constituency: "phrasal re-bracketing"

Dependency is an asymmetrical relation between a head and a dependent, i.e. a **vertical** organization principle.

2.2. Constituency vs. Dependencies

Dependency and constituency describe different dimensions.

1. A phrase-structure tree is closely related to a derivation, whereas a dependency tree rather describes the product of a process of derivation.
2. Usually, given a phrase-structrue tree, we can get very close to a dependency tree by constructing the transitive collapse of headed structures over nonterminals.

Constituency and dependency are not adversaries, they are complementary notions.

Using them together we can overcome the problems that each notion has individually.

3.1. Chomsky's Syntactic Structure

The preface of "Syntactic Structures" emphasizes the **heuristic role of formalization** in clarifying linguistic analyses, supporting empirical testing and falsification:

> " ... The search for rigorous formulation in linguistics has a much more serious **motivation** than mere concern for logical niceties or the desire to purify well-established methods of linguistic analysis. Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the **process of discovery itself**. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, **gain a deeper understanding of the linguistic data**. More positively, a formalized theory may **automatically provide solutions** for many problems other than those for which it was explicitly designed. Obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects."

3.2. Generative grammar

A context-free component, generating "kernel sentences", and a transformation component (cf. Harris (1957)). Two used approaches have been

- (A) Generate a (finite) set of **elementary sentences**, and use transformations to broaden it to the class of representations of all sentences for a language.
- (B) Generate a (finite) set of representations of **all sentences** of a language, and then use transformations to arrive at surface forms.

Variant (B) lead to stratificational grammar,

- Stratificational grammar, cf. e.g. (Hays, 1964; Lamb, 1966).
- Chomsky's (1965) "Aspects of the Theory of Syntax" adopts (B), and would later develop into "Government & Binding theory" (Chomsky, 1981), cf. (Haegeman, 1991; Higginbotham, 1997)

3.3. Early non transformational approach

The landscape of formal grammar was not covered solely by generative (transformational) approaches.

- Bar-Hillel focused primarily on categorial grammar (Bar-Hillel, 1953), elaborating Ajdukiewicz's (1935) syntactic calculus, though provided with his **algebraic linguistics** (Bar-Hillel, 1964) a notion that was intended to cover a broader range of approaches to formal description of grammar.
- Lambek (1958; 1961) similarly focused on categorial grammar, though of a more logical (**proof-theoretical**) kind than Bar-Hillel's.

4. Meaning entered the scene

Chomsky was, in general, **sceptical of efforts to formalize semantics**. Interpretative semantics or the autonomy of syntax: Syntax can be studied without reference to semantics (cf. also Jackendoff).

Criticism on both transformational and non-transformational approaches:

- Transformations do not correspond to syntactic relations, relying too much on linear order.
- Similarly, Curry (1961; 1963) criticized Lambek for the focus on order (directionality).

4.1. Different ongoing efforts

- Developing a notion of (meaningful) logical form, to which a syntactic structure could be mapped using transformations. Efforts either stayed close to a constituency-based notion of structure, like in generative semantics (Fodor, Katz), or were dependency-based (Sgall et al, particularly Panevová (1974; 1975); Fillmore (1968)). Cf. also work by Starosta, Bach, Karttunen.
- Montague's formalization of semantics – though Montague and the semanticists in linguistics were unaware of one another, cf. (Partee, 1997)

4.2. Montague and the development of formal semantics

The foundational work by Frege, Carnap, and Tarski had led to a rise in work on modal logic, tense logic, and the analysis of **philosophically interesting issues in natural language**. Philosophers like Kripke and Hintikka added model theory.

These developments went hand-in-hand with the **logical syntax** tradition (Peirce, Morris, Carnap), distinguishing syntax (well-formedness), from semantics (interpretation), and pragmatics (use).

Though the division was inspired by language, **few linguists attempted to apply the logician's tools in linguistics as such**.

This changed with **Montague**.

> "I reject the contention that an important theoretical difference exists between formal and natural languages." (Montague, 1974)(p.188)

A compositional approach, using a "rule-by-rule" translation (Bach) of a syntactic structure into a first-order, intensional logic. This differed substantially from transformational approaches (generative or interpretative semantics).

5. The trouble with Word Order

Traditional phrase-structure grammar (Bloomfield) is context-free (CFPSG) and therefore it's unable to account for **long-distance dependency** without an extra apparatus.

Chomsky (1957) therefore added transformations on top of a CFPSG.

But, there are both linguistic **problems with transformations** (no corresponding linguistic concept), and formal problems, (Peters & Ritchie (1971; 1973)).

Moreover, Chomsky's arguments against CFPSG (incapable of generalization, mathematical proof concerning string languages) were shown to be awed (e.g. by Gazdar, Pullum –as we have seen last time).

Finally, studies in **nonconfigurational** languages (e.g. Australian) starting in the 1970's gave rise to a more relational view on structure, in contrast to the configurationality of English.

These problems led to the development of new, non-transformational grammar frameworks like Relational Grammar and Arc Pair Grammar, Lexicalized Formal Grammar (LFG), Generalized Phrase Structure Grammars (GPSG). And further

frameworks influenced by the latter, as Head-Driven Phrase Structure Grammar (HPSG).

5.1. Heads entered the scene

Heads (asymmetric relations) start entering the scene, in various guises.

Studies in nonconfigurational languages revealed that **relations** rather than phrases are typologically significant for the expression of meaning (cf. also (Bresnan, 2001)); Relational Grammar, Arc Pair Grammar, LFG.

The developments in GPSG, LFG, and Arc Pair Grammar showed the **feasibility of a nontransformational perspective**, employing a relational perspective to obtain better generalizations.

5.2. Heads enter the scene: Categorial Grammar

Related work was going on in categorial grammar (functional rather phrasal structure): (Venneman, 1977) binding dependency and functional structure – again, combining vertical and horizontal organization.

In general though, categorial grammar tried to deal with flexible word order by introducing means of composition that were more powerful than application:

- Bach's wrap operations (1984).
- Ades and Steedman's combinatorial rules (1982) (also Jacobson, Szabolcsi, and later Hoffman (1995), Steedman (1996; 2000)).
- Moortgat's generalized connectives (cf. (1988), also work by Oehrle, Morrill, Van Benthem).
- Only towards the end of the 1980's, early 1990's is dependency again explicitly introduced into categorial grammar: (Steedman, 1985; Hepple, 1990; Pickering, 1991; Moortgat and Morrill, 1991; Barry and Pickering, 1992; Moortgat and Oehrle, 1994).

5.3. Combining Constituency and Dependencies

In 1975, Joshi et ali introduced a grammatical formalism called Tree-Adjoining Grammars (TAGs), which are tree-generating systems. The application of TAGs to natural language is known as LTAGs.

- New way of thinking of domain of dependencies
- Localization of dependencies : elementary structures of a formalisms over which dependencies such as agreement, subcategorization and filler-gap relation can be specified.

5.4. Disadvantages of Transformational Grammars

Chomsky's Extended Standard Theory has been criticized from various perspectives.

- **Semantics**: Generative Semantics vs. interpretative semantics. Connections with the formal models developed from the work of Montague and with the proponents of Categorial Grammar. This line influences the development of Phrase Structure grammars such as GPSG and HPSG.
- **Psycholinguistics**: experiments did not show that the transformational model is a plausible one. It doesn't seem to be able to represent the competence of speakers. Within an unconstrained grammar, the analysis of a sentence becomes an indecidable problem. Development of more constrained theories of grammar such as LFG.
- **Representation**: Inadequacy of tree representation for non configurational languages: use of graphs or of features.
- **Implementation**: transformational grammars cannot be easily integrated in computational systems for the analysis of natural language.

6. Grammars meet Logic & ...

Logics to specify a grammar framework as a mathematical system:

- Feature logics: HPSG, cf. (King, 1989; Pollard and Sag, 1993; Richter et al., 1999)
- Categorial Type Logics (Kurtonina, 1995; Moortgat, 1997)

Logics to interpret linguistically realized meaning:

- Montague semantics: used in early LFG, GPSG, Montague Grammar, Categorial Type Logic, TAG (Synchronous LTAG)
- Modal logic: used in dependency grammar frameworks, e.g. (Broeker, 1997; Kruijff , 2001).
- Linear logic: used in contemporary LFG, (Crouch and van Genabith, 1998).

7. .. Computation

Computation of linguistic structures

- Unification (constraint-based reasoning): LFG, HPSG, categorial grammar (UCG), dependency grammar (UDG, DUG, TDG)
- "Parsing as deduction": CTL
- Optimality theory: robust constraint-solving, e.g. LFG

7.1. Unification

The development of Unification Grammars has strongly been influenced by the:

- use of tools developed in Logics and in AI;
- the progress made in the area of Natural Language Processing;
- Development of Logic Programming: Prolog.

    1. Declarative character: grammar is not a set of rules, but a set of constraints that a sequence needs to satisfy in order for it to be a grammatical phrase.
    2. Constraints do not need to be ordered.

Transformational grammars are inadequate if faced with implementation problems. Derivations proceed from deep structures while automatic sentence analysis requires the inverse process.

7.2. Importance of Unification Grammars

Unification grammars or constraint based grammars represent the new syntactic models of the 80's.

Four models which are representative of this trend are: Lexical Functional Grammar (LFG) (Bresnan 1982), Generalized Phrase Structure Grammar (GPSG) (Gazdar et al. 1985), Head-driven Phrase Structure Grammar (HPSG) (Pollard Sag 1987, 1994), Tree Adjoining Grammar (TAG) (Joshi et al. 1991).

They are grammar models which try to find an explicit division of labour among the **lexicon, syntax and semantics**.

They are **based on logical models** which are well studied and for which programming techniques have been developed.

They represent an adequate compromise between linguistic expressivity and possibility of implementation.

7.3. Common aspects

These four models have some common properties:

1. surface based;
2. use of complex features to develop syntactic descriptions;
3. definition of general principles of grammaticality;
4. integration of lexicon, syntax and semantics.

Rewrite rules are not interepreted in a procedural way, but as a description of well formed syntactic structures.

7.4. Approaches to rewrite rules

Lexical rules replace certain transformations The various models have different approaches with respect to rewrite rules:

- in LFG there is a parallel level of syntactic representation which is called functional structure;
- in HPSG rewrite rules are replaced by typed feature structures which undergo general principles with respect to the distribution of features;
- in TAG rewrite rules are replaced by elementary trees which are directly associated with the lexical items and combined among each other by specific operations among which is unification

8. Criticisms to Formal Grammars

General criticism:

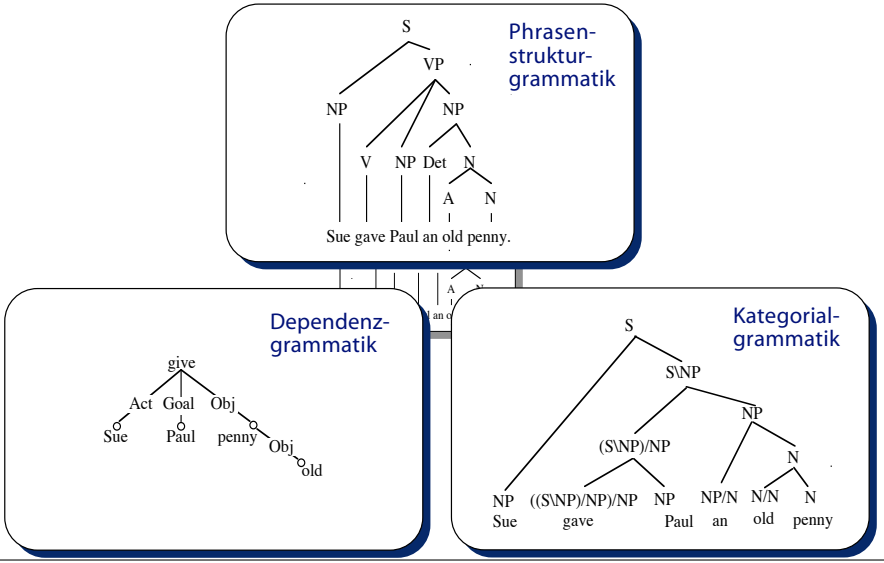
- formal grammars have the advantage of being suitable for implementation, but are not interesting from a cognitive point of view.
- They make a sharp division between grammatical and ungrammatical sentences while this division is normally less sharp.
- It is possible to establish degrees of violation of grammatical principles, which constitute the basis for psycholinguistic hypotheses (Fodor 1983).

In the various models certain principles are formulated to define which features are appropriate, how they cooccur and how they propagate.

The formulation of principles make these models more plausible from a cognitive point of view.

9. DG & CFPSG & CG

THREE TRADITIONS

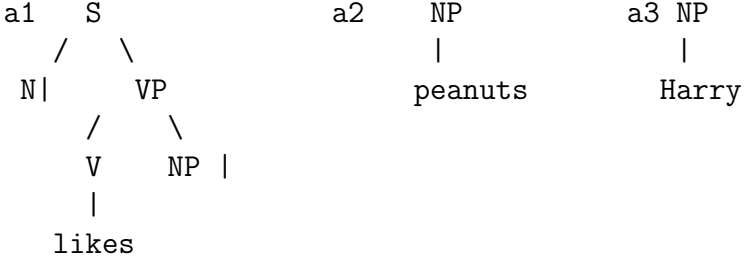


10. TAG & CFG

CFG:

```
S -->  NP VP      NP --> Harry       ADV --> passionately
VP --> V NP       NP --> peanuts
VP --> VP ADV     V --> likes
```

TAG:



Then TAG uses the operation of substitution and adjunction to assemble trees.

11. Grammars meeting Statistics

Read:

> Lillian Lee "I'am sorry Dave, I am afraid I can't do that: Linguistics, Statistics and NLP circa 2001*"

or ask Pasquale!

12. Conclusion

Tomorrow: Room E 4.12, 14:00-16:00.

We will look at the generative power of these Grammars and their complexity. Moreover, we will get an historical overview of Computational Linguistics by looking at its past, present and feature.

Remark Bring the sample exam with mistakes marked!!!