ACL Lifetime Achievement Award

Old Linguists Never Die, They Only Get Obligatorily Deleted*

Eva Hajičová**
Charles University

1. Introduction

Martin Kay, in his speech delivered in 2005 on receipt of his ACL Lifetime Achievement Award, specified computational linguistics as follows: "Computational linguistics is trying to do what linguists do in a computational manner" (Kay 2005, page 429). I believe it is a legitimate question for a computational linguist to ask what linguists do. Coming from Prague, it is then quite a natural question for me to look back and to recollect what the "old" linguists (who never die but get obligatorily deleted on the visible surface) with the background of the world-famous Prague Linguistic School (PLS) contributed to linguistic studies and perhaps to suggest what aspects of their heritage are even today fruitful for computational linguistics.

First, to place the PLS in the course of the development of linguistic studies, it should be recalled that the Prague Linguistic Circle belongs to the first bodies that took part in the transition of the older diachronic paradigm of linguistics to a synchronic theory of language. Soon after its first session (taking place in 1926 in the office of the chairman of the Circle till his death in 1945), the Circle entered the international scene first of all with its systematically elaborated phonological theory. Starting with the Hague Linguistic Congress (see *Actes* 1928), Praguian phonology became the pilot discipline of structural linguistics. This approach was far from unified, but the strength of the Circle was in its spirit of dialogue, which kept the Circle receptive to new ideas, rather than in any set of postulates commonly professed. In my talk I will concentrate on three fundamental Prague School tenets, which I believe to have their validity also in the modern context of linguistic theory and computational linguistics. What I have in mind here is the Circle's structural and functional orientation, as well as the attention it has paid to the opposition of the center and the periphery of language structure, based on the concept of markedness.

2. The Structural Point of View of PLS

The PLS is generally (and truly) characterized by two attributes: structural and functional. Let us first turn to the **structural** point of view, namely, the School's endeavor to view language as a system of systems rather than to study individual phenomena as ad hoc, non-systematic issues. The Circle shared de Saussure's understanding of language as a system of (bilateral) signs, in which only oppositions rather than fixed

---

entities play a role. As mentioned above, this attitude was most apparently reflected in the study of phonology as a system displaying distinctive features and employing the notion of binary oppositions. Jakobson (1929) presented the phonological repertory (both in synchrony and in diachrony) as a system of oppositions (mainly binary, privative), based on acoustic distinctive features and understood as the clue to the sound and meaning relationship.

Along with phonemes and morphemes, the sentence was also recognized as one of the fundamental fields of systematic oppositions, that is, as an ingredient of *la langue*. Mathesius (1928, 1936) formulated a concept of functional syntax; a structural view of syntax, based on the dependency relation, was elaborated by Tesnière (1934), a French member of the Circle, who was a professor of the Ljubljana University; his monograph was published only posthumously (Tesnière 1959), but his papers were known in Prague, and his approach to syntax was applied to Czech by Šmilauer (1947), who combined dependency syntax with a constituent-based view of the relation between predicate and subject.

Dependency-based approaches, which understand the verb as the center of the sentence structure and describe this structure on the basis of binary relations between heads and their modifiers, have been for a long time a matter of Continental syntactic theories rather than of the mainstream syntactic approaches on the other side of the Atlantic. However, the notion of head can be found also in Bloomfield (1933) when referring to the names of the main constituents of the sentence, that is, NP (*noun phrase*, with N as its head) and VP (*verb phrase*, with V as its head). In the framework of the Chomskyan approach, originally based exclusively on the concept of immediate constituents, the notion of head becomes the basic notion of X-bar theory. Originally, four categories were singled out as possible heads of their respective maximal projections, namely, N, V, Adj, and P(rep); as remarked by James McCawley (personal communication, around 1990), such a theory may be interesting unless the specification of the set of basic categories grows beyond some reasonable limit. McCawley's critical remark reflected the gradual development of X-bar theory, which allowed practically any constituent (or, more generally speaking, any arbitrary symbol for a grammatical value) to act as the head, dependent on the needs of the analysis of this or that construction.[1]

The very name *head-driven phrase structure grammar*, an influential theory combining an immediate constituent approach with elements of a dependency-based approach, as proposed by Pollard and Sag (1994), explicitly points out that the theory takes account of the main element within a constituent. Although their approach is constituent based (working with a lexically based X-bar syntactic theory; [Pollard and Sag 1994, page 362]), the authors are aware that the notion of constituent structure is widespread but that it is not based on sufficiently convincing direct evidence. The authors refer to Hudson's (1984) approach and claim that it belongs to exceptions that do not overestimate the constituent structuring of sentence elements.

It is sometimes doubted if the direction of the dependency relation, namely, the determination of the governor and the dependent in each pair (syntagm) can be reliably stated. We believe that in the prototypical case, the main criterion for this distinction

---

1 To be fair, I should add that one of the rare attempts at a more explicit characterization of the notion of head can be found in Adger's monograph on minimalism (Adger 2003, page 75), which mentions two criteria: one based on the distribution of the whole constituent and the other taking into account which constituent determines the reference of the whole constituent.

can be based on the possibility that, in the endocentric constructions, the dependent can be absent (not just deleted on the surface). Thus, for example, in *Yesterday, my father worked for the whole day in the garden* it is possible to leave out the dependents *yesterday*, *my*, *for the whole day*, and *in the garden* without the sentence losing its grammaticality. However, there are exocentric pairs such as *(to) find something*, where neither of the two members of the pair can be deleted and for which, therefore, the mentioned method by itself cannot help to find out which element is the governor and which is the dependent. What helps here is the principle of analogy on the level of parts of speech: On the basis of the existence of verbs without objects it can be concluded that in such pairs the verb is the governor (in our particular example, *(to) find*). In the same vein, subject (Actor) can be understood as dependent on the verb, because there are verbs also without a subject: In *It is raining* (Latin *Pluit*), *It* is just a surface "filler," absent in the sentence structure proper. This view is also supported by the observation based on the annotators' agreement when assigning the dependency structures (trees) in the Prague Dependency Treebank: The annotators did not have too many troubles with the determination of the direction of dependency; if there was a disagreement, it concerned their assignment of the type (value) of the dependency relation (see Hajičová, Pajas, and Veselá 2002).

There is one linguistic phenomenon—present more or less in every language—all syntactic theories have to bother about, namely, the relation between syntactic structure and word order (the discontinuity of constituents, for which Gazdar [1981] introduced the term *unbounded dependencies*, used, for example, also by Pollard and Sag [1994, pages 157ff.]; see also the term *long-distance dependencies* used by some other authors) or, in terms of formal dependency descriptions, the non-projectivity of syntactic constructions. Informally speaking, the strongly restrictive condition of *projectivity* says that if a node *a* depends on *b* and there is a node *c* between *a* and *b* in the linear ordering, *c* is subordinated to *b* (where *subordinated* means an irreflexive transitive closure of *depends*). See Figure 1 for an example of non-projective parts of a tree; the vertical line leading from *a* intersects the dependency edge leading from *a* to *c*.[2]

The more restricted the formal syntactic description is, the more valuable are the observations based on it; in this sense, the condition of projectivity might well serve its purpose. However, there are seemingly many non-projective constructions in the surface shapes of the sentences. The task then is to attempt to classify the constructions in which the condition of projectivity is not met in the surface shape of the sentence and to attempt at a description not only meeting the condition as far as the core of the language system (see Section 4) is concerned, but also accounting by some simple, well-defined means for the cases of superficial non-projectivity (a preliminary formulation of movement rules specified as a transition from projective underlying trees to strings of morphemes in which the condition of projectivity cannot be applied can be found in Sgall [1997] and Hajičová and Sgall [2003]). This is, of course, a rather strong hypothesis that has to be verified and made more precise on the basis of systematic empirical research. It should be mentioned in this connection that it is in line with the Praguian approach that function words are distinguished from autosemantic words and that only the latter constitute nodes of their own in the underlying trees; from this it follows that in the numerous cases in which the "non-projectivity" of surface word order concerns

---

2 Projectivity as a property of word order important for a formal description of language was already stated by Hays (1960, 1964) and Lecerf (1960) in formal grammar; the condition of projectivity (in different forms that have been proved as equivalent) was defined by Marcus (1965) and used in many writings working with dependency descriptions.
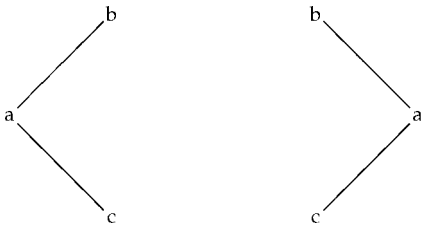
Figure 1
Examples of non-projective parts of a dependency tree.

auxiliary verbs or conjunctions, and so forth, the projectivity of underlying syntactic structure is not at stake.

The introduction of the notion of a head brings into the foreground the connection of grammar and lexicon; the necessity of such a relationship was already quite apparent in the earlier works of Fillmore (1966, 1968) introducing the so-called case grammar, which explicitly follows up Tesnière's notion of valency. The term "case" does not directly refer to case as a morphological category but to the meaning (function) of a (morphological) case: for example, Addressee is a prototypical meaning (function) of Dative, Agentive is a prototypical meaning of Nominative, and so on. The concept of valency is crucial in that it reflects the fundamental aspect of the presence of grammatical information in the lexicon: The valency frame is a part of the lexical entry, in which the obligatory and optional syntactic kinds of dependents of the given (head) word are registered. It is also well known that Fillmore's theory (as well as the thematic roles of Gruber [1967]) played a substantial role in the introduction of theta roles (and *theta grids*) in the Chomskyan model of *government and binding* (later called, more appropriately, the Principles and Parameters model).

Fillmore himself explicitly mentioned that, when proposing his case grammar, he did not primarily consider which formal description his approach would fit into. However, he presents an example of how his approach can be formulated in terms of a phrase structure model: The sentence S can be decomposed into two parts, Modality and Proposition; the Proposition in turn can be articulated into the verb and a set of noun phrases, which are characterized by one of the case markers, that is, K_1NP, K_2NP, ..., K_nNP. Each of these noun phrases can then be decomposed into the noun phrase proper and the given case marker k_i (Agentive, Addressee, Objective, etc.). The analysis of Robinson (1969, 1970) devoted to the relation between Fillmore's approach and that of transformational grammar (of that time) throws an interesting light on the possiblity of a smooth transition from a phrase-based approach to a dependency-based one, which is more transparent and economical. In Fillmore's proposal, the case relations, that is, the relations of the noun phrase to the verb, are actually captured twice, once as the marker k_i and once as the characteristics of the given phrase (K_iNP). It is then possible to work with a pure dependency tree structure, where the root of the tree is the verb, and the nouns (or, as the case may be, other word categories) depend on the root as dependents with a certain type of relation.

It goes without saying that Fillmore's case grammar and its follow-up *frame nets* conception has influenced in a substantial way many of the contemporary approaches not only to treebank annotation and computational lexicology (cf., e.g., Fillmore et al. 2003) but also the work on underlying sentence structure in general.

Two "historically" motivated and seemingly contradictory observations are in place here: It can be documented by references to the development of linguistic theory in the

past 50 years that the deeper the analysis goes, the more the need of an introduction of the distinction between the notions of "head" and "modifier" (predicate, argument) is felt. Let us only recall here such approaches as:

(i)   the lexicosemantic analysis by Katz and Postal (1964), who work with the notions of head and modifier when specifying selection restrictions;

(ii)   the distinction between surface-oriented constituent structure and the (underlying) functional structure in lexical functional grammar by Bresnan (1978) and Kaplan and Bresnan (1982);

(iii)   the above-mentioned case grammar by Fillmore, motivated by the conviction that Chomskyan deep structure (with its specification of "deep" subject as a constituent of S regardless of the (different) semantic relations of the given NP to the verb) is not deep enough to capture the real underlying structure of the sentence; and

(iv)   the consecutive introduction of theta roles into the government and binding theory.[3]

On the other hand, dependency-based considerations have gradually and evasively penetrated to the "data"-oriented statistical methods and treebank annotations. As the freshest example, let us only refer to the recent EACL 2006 conference in Trento and the HLT-NAACL 2006 conference in New York with its CoNLL-X Shared Task on Multilingual Dependency Parsing (working with treebanks of 12 languages, of different sizes). In other words, the seemingly surface oriented analysis is prevailingly dependency based.[4]

A possible explanation for this apparent contradiction may be looked for in the economy and transparency of the dependency-based trees: In their applications, the data-oriented systems also aim at a representation of the meaning of the surface shapes of sentences (whatever one can understand by "meaning"), so that their attention is focused on a most transparent and economic way (avoiding "extra" nodes for phrases such as NPs, VPs, . . . , etc.) from the surface to the depth. Dependency analysis offers such a way.[5]

---

3 We have restricted our attention here only to systems staying in principle within the development of the Chomskyan paradigm or originating as a reaction to it. However, when discussing the relation between or combination of constituent-based and dependency-based grammars, special attention should be paid to the *lexicalized tree-adjoining grammars* (LTAG) continuing the original conception of *tree-adjoining grammar* (TAG) as proposed by A. K. Joshi (see, e.g., Joshi 1985), which has served as a basis for many studies on formal grammar as well as from the NLP domain. The similarity between LTAG and a dependency-based description in relation to the model using the so-called supertags (which encode syntactic information in terms of dependency) is analyzed by Joshi and Srinivas (1994).

4 It should be recalled in this connection that within machine translation dependency-based systems (sometimes in combination with phrase structure) were already at play in the early times of MT; see, for example, the works of B. Vauquois, one of the founders of computational linguistics (Vauquois 1975; Vauquois and Chappuy 1985) and the systems developed in Japan under the influence and guidance (direct or indirect) of M. Nagao (see the survey in Nagao [1989]).

5 In a similar vein, Steedman (2005) argues that the use of statistical language models is the only way to create a computer program that automatically analyzes sentences on the basis of broadly conceived grammars (with due regard to ambiguities) such as dependency-based grammars or grammars specifying heads (governors); according to Steedman, these grammars work well because they reflect a mixture of semantic information and information based on knowledge of the world.

Among urgent questions to be asked with regard to the approaches to sentence structure, there are then the following issues:

(i)   Is it more appropriate to analyze a sentence such as *In this garden, she was reading a book on the history of Spain yesterday* as having the complex verb form *was reading* as its head, with *she*, *(a) book*, and *garden* as its dependents, or to see the basic characteristics of its structure in distinguishing whether *in this garden* or *yesterday* is more immediately connected with its verb?

(ii)   Do we have clearer criteria for answering the former or the latter of these two questions?

3. Prague School Functionalism

The other attribute of Prague structuralism is **functional.** Mathesius (1928, 1936), inspired especially by the philosophy of language of Marty (1908), presented his theory of functional grammar, based on the concept of function as related to universal intentional acts and treated as a dichotomy of functional onomatology and functional syntax. Mathesius combined this universal dichotomy with the language-specific opposition of function and form. As Sgall (1987) pointed out, the core of the system of language was conceived of as consisting of levels, the units of which have their functions in that they represent or express units of the adjacent higher levels, up to the non-linguistic layer of cognitive content. The units of the system were understood as constituting hierarchies in which some of them function as certain parts of the others. Thus, for example, phonemes were defined and delimited one against the other on a functional basis (two different phonemes can distinguish two morphemes), and the established repertory of distinctive features gave a firm foundation to the description of the system of phonemes as a structured whole. Strings of phonemes (morphs, in more modern terminology) are understood as expressing morphemes, and sequences of morphemes as expressing sentence structure.

Another important aspect of the functional approach is to view language as a functioning system, adapted to its communicative role, diversified in more or less different social and local varieties, and to describe the sentence structure as adapted to its functioning in discourse.

This leads me to pay attention to the information structure (in our terms, topic–focus articulation) of the sentence. Let me first look again at the history of the issue. It was the Prague scholar Mathesius (1929, 1938) who introduced the study of information structure into structural linguistics, preferring the terms *Thema* and *Rhema* (used before in German linguistics by H. Ammann) to the older *psychologisches Subjekt* and *Prädikat* (used by G. von der Gabelentz, H. Paul, and others), and understanding the former (the topic) as one of the functions of the subject in English. He distinguished topic proper, comment (focus) proper, and the accompanying elements of either of these two parts. Later on, one of Mathesius' followers, Jan Firbas, extended the hierarchical understanding of the information structure of the sentence by postulating a scale of communicative dynamism. The Praguian concepts met a favorable response within continental linguistics (one should mention in this connection especially the works by British linguists M. A. K. Halliday and H. W. Kirkwood, several German linguists such as J. Esser, R. Bartsch, and J. Jacobs, the French linguist J.-M. Zemb, and others). However, only the syntactic or word order consequences (or, as the case may be, conditions)

of different sentence articulations into topic and focus were mostly taken into account, and its relevance for and effects on the coherence of discourse.

A new impetus into the study of information structure was given by Petr Sgall, who was the first to come up with examples testifying to the semantic effects of this issue and claiming that two utterance tokens differing in their topic–focus articulation are tokens of two different sentences, that is, that topic and focus belong to the language system rather than only to the use of language in communication (Sgall 1967, page 205ff.). As a matter of fact, the split of transformational grammar into the generative and the interpretative semantics wings coming out at the same time operated with arguments based on sentences that in Praguian terms differ only in their topic–focus structure (this fact, of course, not being recognized by the authors): See Lakoff's (1969) examples: *Many men read few books* against *Few books are read by many men*, *John talked about many problems to few girls* versus *John talked to few girls about many problems*. To be fair to the other side of the dispute, Chomsky (1965, page 224) noticed the semantic difference between the sentences *Everybody in the room knows at least two languages* and *At least two languages are known by everybody in the room* and was in doubt as to whether this difference should be ascribed to the difference between active and passive; he remarks that such a distinction might be described in terms of topic (as Lakoff [1969] notes, in this consideration, the influence of Halliday [1967–1968] played its role). In his reaction to the generative semanticists' criticism of the "shallowness" of his deep structure, Chomsky (1968) was even more inclined to use notions related to what in present-day terms would be called the information structure of the sentence; he claims that in the semantic interpretation of the sentence, one should take into account the distinction between what he calls presupposition and focus, and a related notion of the range of permissible focus. There are two interesting points in his approach: First, Chomsky connects these syntactic issues with the placement of the intonation center in the spoken form of the sentence, and second, he connects the possible operational criterion for the determination of the choice of focus from the range of permissible focus with the scope of negation. In particular, to decide what is the focus of the answer to *Was it an ex-convict with a red SHIRT that he was warned to look out for?*, one should consider possible different negative continuations such as *No, he was warned to look out for an AUTOMOBILE SALESMAN*, or . . . *for an ex-convict wearing DUNGARIES*, . . . *for an ex-convict with a CARNATION*, . . . *for an ex-convict with a red TIE*.

One could argue that it is the presence of structures with quantification rather than the topic–focus articulation of the quoted examples that is responsible for the indicated semantic distinction. However, the Praguian writings from the sixties convincingly demonstrate that it is not difficult to find sentences without quantification that exhibit the same phenomenon (for reasons I will mention in a minute, in the examples, the capitals indicate the intonation center): *Russian is spoken in SIBERIA* versus *In Siberia, RUSSIAN is spoken*, or *John works on his dissertation on WEEKENDS* versus *On weekends, John works on his DISSERTATION*. In Russian linguistics, such examples have been discussed as *Kurit' ZDES'* versus *Zdes' KURIT'*. The sentences quoted also document that the difference cannot be ascribed to the active/passive distinction; neither can it be claimed that the word order always plays a decisive role: Consider Halliday's (1970) famous example from a London underground station: *Dogs must be CARRIED*. With the same word order, but with a change in the placement of the intonation center one gets a certainly unwanted interpretation: *DOGS must be carried* would imply that everybody stepping on the escalator has to carry a dog (in a similar vein as *Carry DOGS*.). A plausible explanation of the semantic difference covering all these cases is to describe them in terms of difference in their information structure.

This had not been recognized or at least commonly accepted for some time on an international scale until the appearance of Mats Rooth's Ph.D. dissertation in 1984.[6] Although restricted to prosodic focus (pointing out that the difference in truth conditions between such sentences as *Mary only introduced BILL to Sue* and *Mary only introduced Bill to SUE* is only in the location of focus, denoted here by capitals), Rooth's work was an impetus for an increasing interest in the related issues, first in the domain of formal semantics (here the influential role of Barbara H. Partee should be emphasized), but soon literally everywhere. Let us mention in this context that semantic considerations apparently stood behind the conception of combinatory categorial grammar first proposed by Steedman (1996, 2000); his introduction of *floating constituents*, the division line between which is given by the articulation of the sentence with regard to its information structure rather than fixed, determined once for all. Steedman, in contrast to many other researchers presently working in the domain of information structure, pays a due respect to the close relation between information structure, syntactic sentence structure, and prosody; in this respect, also the work on corpus annotation led by him is a pioneering enterprise (Calhoun et al. 2005).[7]

Due respect paid to the description of the information structure of the sentence is also crucial for the study of discourse structure and coherence. It might be interesting to note in this connection that the first systematic study indicating such a relation— although in terms influenced by the then prevailing psychological view of language— is Weil's (1844) study on the order of words. The author introduces the notions of progression of "ideas," distinguishing "progression" and *marche parallèle*: In the former sequence (segment), the given sentence B is connected to the preceding sentence A by starting with a reference to the idea that was at the end of A, whereas in the latter, the sentences "march in parallel," that is, they begin with a reference to the same idea. It is not difficult to see an analogy between this view and a more modern and explicit treatment of the so-called centering theory (Grosz, Joshi, and Weinstein 1995) and its shifts of centers.

I am dwelling at such length on the issues of information structure not just because it is my favorite child (and, indeed, it is), but because I am fully convinced about the importance of this issue for an adequate description of the sentence structure, both in formal description of language as well as in natural language processing. Let me illustrate by a personal recollection that I am not beating a straw man. Some time ago (if I am not mistaken, it was in 1989) I was invited for an IBM-organized MT conference in Garmisch-Partenkirchen to deliver a talk on the Praguian approach to MT. Naturally enough, I devoted most of my time to illustrate examples of translations from and to several European languages that topic–focus articulation as an important aspect the translation (be it human or automatic) has to take into account. The group of

---

6 From a different perspective, the term focus was used by Grosz (1977). The author adopted a
    psychological point of view of focus of attention and considered the sentence focus to be that item that is
    in such a focus, that is, in our terms the topic of the sentence (what the sentence is *about*). Grosz'
    approach has found its continuation in the centering theory mentioned below.
7 Let us note in this connection that the difficulties of a syntactic description based on phrase structure for
    an adequate capturing of the topic–focus articulation were pointed out already in Sgall, Hajičová, and
    Benešová (1973, page 163ff.) and in Hajičová and Sgall (1975) and illustrated on examples such as *This
    year we will spend two weeks on Mallorca* used in the context of *How will you spend your holidays this year?*, i.e.
    with the focus part of the sentence being *two weeks on Mallorca*. Working with phrase structure, it would
    be very difficult to characterize the two groups as a single phrase; in a similar vein, to determine the
    topic of the sentence as a single phrase is also difficult: if the question were *Where do you spend two holiday
    weeks this year?*, the focus of the answer would be on *Mallorca*, with the topic being *this year we will spend
    two weeks*.

people attending the meeting was extremely nice and friendly, and it was no wonder that the program chairs, Margaret King and Jonathan Slocum, could dare to make the concluding session a sort of fun ascribing to each of the papers some characteristic evaluative attribute: It was quite symptomatic of those times that the issues discussed in my paper were characterized as "least important for MT" (it may be of interest to recall that Mercer's paper on the IBM statistical approach to MT delivered there was evaluated as "crazy science fiction").

The question should then be discussed in which way is it possible to describe the interplay of the dependency (or constituency)-based patterning of the sentence and the topic–focus articulation (or information strcuture) of the sentence as two basic aspects of syntax (now cf. Hajičová and Sgall, in press). Is it true that a dependency-based view of the underlying structure as the core of the language system (in which there are no nodes corresponding to function words and the left-to-right order of lexical items meets the condition of projectivity; cf. Section 2 above) might be useful in this respect?

I am happy to see that much has changed in this particular domain of studies since those times; I cannot say I welcome all the changes, but it is encouraging to see that the two Praguian tenets I have discussed so far—namely, the dependency approach and the due regard to the information structure—have found an undisputable appraisal within our field.

4. The Core of Language and the Periphery

The third Praguian notion I would like to mention is the distinction made between the **center** (core) of language and the **periphery.** This distinction is closely connected with the notion of markedness; markedness, characterizing the intrinsic asymmetry of binary (and other) oppositions (not only in phonology, but also in morphology, in semiotics, and in many other domains), was first systematically presented by R. Jakobson. It was properly understood and used as an organizing principle of sign systems, also in connection with language universals and language acquisition. As Battistella (1995) notes, this notion belongs to those aspects of the Prague linguistic theory that in some form have been taken over by Chomsky, who applied it, albeit in a different shape, in his Principles and Parameters theory, as proposed in the early 1980s.

Although the relationships between the two oppositions of marked versus unmarked phenomena and the core versus the periphery of the system of language are far from straightforward (see Sgall 2002, 2004), it can be claimed that because language is more stable in its core, regularities in language should be searched for first in this core; only then is it possible to penetrate into the subtleties and irregularities of the periphery. The relatively simple pattern of the core of language (in Sgall's view, not far from the transparent pattern of propositional calculus) makes it possible for children to learn the regularities of their mother tongue on the basis of shared human mental capacities, instantiated also by systems such as those of elementary arithmetic or Aristotelian logic. The freedom of linguistic behavior, limited only by the speakers' desire to be understood by their audience, offers space for the flexibility of the large and complex periphery (i.e., not only of individual exceptions, but also by most different sets of marked phenomena determined by contextual conditions and lists).

The question to be asked then is which of the two possible approaches to how to project this view to a formal description of language is to be preferred: to attempt to describe all phenomena "at once," that is, to consider language as a whole and to describe all phenomena "at a single layer," or to proceed from the core of the system to its periphery.

5. Corpus Annotation as a Test of Linguistic Theories

At the beginning of my talk, I promised to suggest which aspects of Praguian heritage (and in a more general view, of linguistics as such) I believe to have been fruitful for computational linguistics. When talking about the three particular aspects I have chosen, I have tried to make some suggestions as to urgent questions to be asked. Let me finish my talk by an illustration taken from the presently flourishing field of language resources, corpus annotation, and evaluation.

It has been already commonly accepted in computational and corpus linguistics that grammatical (or lexical semantic, etc.) annotation does not "spoil" a corpus, because the annotation is done "in addition" to the raw corpus. Thus, on the contrary, annotation may and should bring an additional value to the corpus. However, there are some necessary conditions for an annotation to fulfil this aim: Its scenario should be carefully (i.e., systematically and consistently) designed, and it should be based on a sound linguistic theory. This view is corroborated by the existence of annotated corpora of various languages (even if their creation is mostly done manually but supported by annotator-friendly software tools or semiautomatic procedures): the Penn Treebank, its successors as PropBank or the Penn Discourse Treebank, Tiger, the Prague Dependency Treebank, and several others. These conditions being met, corpus annotation serves, among other things, as an invaluable test for the linguistic theories standing behind the annotation schemes, and as such represents an irreplaceable resource of linguistic information for the construction and enrichment of grammars, both formal and theoretical.

This claim can be documented by the case of the multilayered annotation of the Prague Dependency Treebank (PDT; see, e.g., Hajič 1998), which is based on the framework of the Functional Generative Description (see, e.g., Sgall, Hajičová, and Panevová 1986). It is important to note that the PDT annotation concerns not only the surface and morphemic shape of sentences, but also (and first of all) the underlying sentence structure (tectogrammatical layer), which elucidates phenomena hidden on the surface although unavoidable for the representation of the meaning and functioning of the sentence, for modeling its comprehension and studying its semantico-pragmatic interpretation, for the work on lexical semantics, and for dictionary buildup and many other aims.

We have tried to demonstrate on certain selected grammatical and discourse phenomena (in Hajičová, Sgall 2006) that the process of the annotation during the last decade and its results have allowed for an enrichment of this framework in several regards. In particular, our examples were taken from the domain of the condition of projectivity, classification of dependency relations, topic–focus articulation (the bipartition of the sentence into topic and focus and the cannonical underlying word order in the focus of the sentence), and some aspects of discourse structure.

6. Final Remarks

I have always been an optimist, and therefore let me go back to the Indiana University Linguistic Club logo from 1984 quoted in the title of my talk: I strongly believe that old linguists never die, they only get obligatorily deleted. Deletions concern the surface rather than the underlying structure so that we may hope that while the old linguists' bodies may lie a-moldering in their graves, the best of their ideas will be marching on.

Acknowledgments

References

Actes du Premier Congrès international des linguistes à la Haye. 1928. A. W. Sijthoff, Leiden.

Adger, David. 2003. *Core Syntax. A Minimalist Approach*. Oxford University Press, Oxford.

Battistella, Edwin. 1995. Jakobson and Chomsky on markedness. In E. Hajičová, M. Červenka, O. Leška, and P. Sgall, editors, *Prague Linguistic Circle Papers 1*. John Benjamins, Amsterdam, pages 55–72.

Bloomfield, Leonard. 1933. *Language*. Holt, Rinehart and Winston, New York.

Bresnan, Joan. 1978. A realistic transformational grammar. In M. Halle et al., editor, *Linguistic Theory and Psychological Reality*. MIT Press, Cambridge, MA, pages 1–59.

Calhoun, Sasha, Malvina Nissim, Mark Steedman, and Jason Brenier. 2005. A framework for annotating information structure in discourse. In A. Meyers, editor, *Pie in the Sky. Proceedings of the ACL Workshop 2005*. Ann Arbor, MI, pages 45–52.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. MIT, Cambridge, MA.

Chomsky, Noam. 1968. Deep structure, surface structure and semantic interpretation. In D. D. Steinberg and L. A. Jakobovits, editors, *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*. Cambridge University Press, Cambridge, pages 183–216.

Fillmore, Charles J. 1966. Toward a modern theory of case. In D. A. Reibel and S. A. Schane, editors, *Modern Studies in English*.

Prentice-Hall, Englewood Cliffs, NJ, pages 361–375.

Fillmore, Charles J. 1968. The case for case. In E. Bach and R. Harms, editors, *Universals in Linguistic Theory*. Holt, Rinehart, and Winston, New York, pages 1–90.

Gazdar, Gerald. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry*, 12:155–184.

Grosz, Barbara J. 1977. *The Representation and Use of Focus in Dialogue Understanding*. Ph.D. thesis, University of California, Berkeley, CA.

Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

Gruber, Jeffrey S. 1967. Functions of the lexicon in formal descriptive grammar. Technical Report (TM)-3770/00, Systems Development Corporation, Santa Monica.

Hajič, Jan. 1998. Building a syntactically annotated corpus: The Prague Dependency Treebank. In E. Hajičová, editor, *Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová*. Karolinum, Prague, pages 106–132.

Hajičová, Eva, Petr Pajas, and Kateřina Veselá. 2002. Corpus annotation on the tectogrammatical layer: Summarizing of the first stages of evaluation. *The Prague Bulletin of Mathematical Linguistics*, 77:5–18.

Hajičová, Eva and Petr Sgall. 1975. Topic and focus in transformational grammar. *Papers in Linguistics*, 8(1–2):13–58.

Hajičová, Eva and Petr Sgall. 2003. Dependency syntax in functional generative description. In Vilmos Agel et al., editors, *Dependenz und Valenz, Vol. 1*. Walter de Gruyter, Berlin, pages 570–592.

Hajičová, Eva and Petr Sgall. 2006. Corpus annotation as a test of a linguistic theory. In *Proceedings of LREC 2006*, Genoa.

Hajičová, Eva and Petr Sgall. Forthcoming. The fundamental significance of information structure. In C. Caffi and H. Haberland et al., editors, *Future Prospects of Pragmatics*.

Halliday, Michael A. K. 1967–1968. Notes on transitivity and theme in English. *Journal of Linguistics*, 3:37–81, 199–244; 4:179–215.

Halliday, Michael A. K. 1970. *A Course in Spoken English: Intonation*. Oxford University Press, Oxford.

Hays, David G. 1960. Grouping and dependency theories. In *Proceedings of the*

National Symposium on Machine Translation, pages 258–266, Englewood Cliffs, NJ.

Hays, David G. 1964. Dependency theory: A formalism and some observations. *Language*, 40:511–525.

Hudson, Richard. 1984. *Word Grammar*. Blackwell, Oxford.

Jakobson, Roman. 1929. Remarques sur l'évolution phonologique du russe comparée á celle des autres langues slaves. *Travaux du Cercle Linguistique de Prague*, volume 2.

Joshi, Aravind. 1985. Tree-adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions? In D. Dowty, editor, *Natural Language Processing*. Cambridge University Press, Cambridge, pages 206–250.

Joshi, Aravind and Bangalore Srinivas. 1994. Disambiguation of super parts of speech (or supertags): Almost parsing. In *Proceedings of the 15th International Conference on Computational Linguistics* (COLING 1994), pages 154–160, Kyoto, Japan.

Kaplan, Ronald and Joan Bresnan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In Joan Bresnan, editor, *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA, pages 173–281.

Katz, Jerrold J. and Paul M. Postal. 1964. *An Integrated Theory of Linguistic Descriptions*. MIT Press, Cambridge, MA.

Kay, Martin. 2005. A life of language. *Computational Linguistics*, 31(4):425–438.

Lakoff, George. 1969. On generative semantics. In D. D. Steinberg and L. A. Jakobovits, editors, *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Pyschology*. Cambridge University Press, Cambridge, pages 232–296.

Lecerf, Yves. 1960. Programme des conflits, modèle des conflits. *Traduction Automatique*, 1(4):11–18; 1(5):17–36.

Marcus, Solomon. 1965. Sur la notion de projectivité. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 11:181–192.

Marty, Anton. 1908. *Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie 1*. Halle/S.

Mathesius, Vilém. 1928. On linguistic characterology. *Actes*, pages 56–63.

Mathesius, Vilém. 1929. Zur satzperspektive im modernen Englisch. *Archiv für das Studium der neueren Sprachen und Literaturen*, 155:202–210.

Mathesius, Vilém. 1936. On some problems of the systematic analysis of grammar. *Travaux du Cercle Linguistique de Prague*, volume 6, pages 95–107.

Nagao, Makoto. 1989. *Machine Translation: How Far Can It Go?* Oxford University Press, Oxford.

Pollard, Carl and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. University of Chicago Press, Chicago and London.

Robinson, Jane J. 1969. Case, category and configuration. *Journal of Linguistics*, 6:57–80.

Robinson, Jane J. 1970. Dependency structures and transformational rules. *Language*, 46:259–285.

Sgall, Petr. 1967. Functional sentence perspective in a generative description. *Prague Studies in Mathematical Linguistics*, 2:203–225.

Sgall, Petr. 1987. Prague functionalism and topic vs. focus. In René Dirven and Vilém Fried, editors, *Functionalism in Linguistics*. John Benjamins Publishing Company, Amsterdam/Philadelphia, pages 169–189.

Sgall, Petr. 1997. On the usefulness of movement rules. In B. Caron, editor, *Actes du 16e Congrès International des Linguistes (Paris 20-25 juillet 1997)*. Elsevier Science, Oxford.

Sgall, Petr. 2002. Freedom of language: Its nature, its sources and its consequences. *Prague Linguistic Circle Papers*, 4:309–329.

Sgall, Petr. 2004. Types of languages and the simple pattern of the core of language. In P. Sterkenburg, editor, *Linguistics Today—Facing a Greater Challenge (Plenary lectures from CIL 17)*. John Benjamins, Amsterdam/Philadelphia, pages 243–265.

Sgall, Petr, Eva Hajičová, and Eva Benešová. 1973. *Topic, Focus and Generative Semantics*. Scriptor, Kronberg/Taunus.

Sgall, Petr, Eva Hajičová, and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Reidel, Dordrecht; Academia, Prague.

Šmilauer, Vladimír. 1947. *Novočeská skladba [The syntax of Modern Czech]*. Mikuta, Prague.

Steedman, Mark. 1996. *Surface Structure and Interpetation*. The MIT Press, Cambridge, MA.

Steedman, Mark. 2000. Information structure and the syntax–phonology interface. *Linguistic Inquiry*, 31:649–689.

Steedman, Mark. 2005. Grammar acquisition by child and machine. Invited Talk at the 17th European Summer School of Language, Logic and Information, Edinburgh.

Tesnière, Lucien. 1934. Comment construire une syntaxe. *Bulletin de la Faculté des lettres de Strasbourg*, 12(7):219–229.

Tesnière, Lucien. 1959. *Eléments de Syntaxe Structurale*. Klinksieck, Paris.

Vauquois, Bernard. 1975. La traduction automatique à grenoble. *Documents de linguistique quantitative*, 24.

Vauquois, Bernard and Sylviane Chappuy. 1985. Static grammars: A formalism for the description of linguistic models. In *Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation*, pages 298–322, Colgate University, Hamilton, New York.

Weil, Henri. 1844. *De l'ordre des mots dans les langues anciennes comparées aux langues modernes* (*The Order of Words in the Ancient Languages Compared with That of Modern Languages*), Paris; Amsterdam [1978].