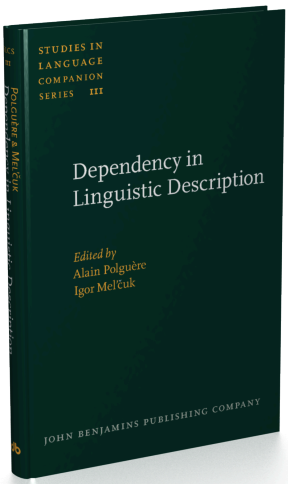
Studies in Language Companion Series
111

Foreword

Igor Mel'čuk
Alain Polguère

https://doi.org/10.1075/slcs.111.02mel



Pages xiii–xxii of
Dependency in Linguistic Description
Edited by Alain Polguère and Igor Mel'čuk
[**Studies in Language Companion Series**, 111]
2009. xxii, 281 pp.

Foreword

Igor Mel'čuk and Alain Polguère

Many forewords are written in order to observe a nice and useful tradition, a kind of politeness ritual. We have nothing against this approach, but we want this foreword to be more than that: it has to help the readers of the volume to make their way through its four papers. These papers are highly technical and not easy to assimilate, so we decided to try to "chew" them first in order to facilitate their digestion for the assiduous reader. This is said in order to send the right signal: our foreword—rather than just a friendly gesture—is intended to be a pedagogical introduction to this interesting, promising and in fact crucial domain of modern linguistics, which is dependency syntax. We will proceed in three steps:

1. general considerations about dependency syntax, as conceived of in this volume;
2. presentation of the papers by Mel'čuk, Kahane, Iordanskaja & Mel'čuk, and Milićević;
3. background on the making of the present volume and acknowledgments.

1 Dependency syntax: general considerations

We will start with two assumptions shared by many practitioners of the dependency approach to syntax, not necessarily those working in the Meaning-Text framework:

> A sentence has associated with it a formal object depicting its internal organization called the *syntactic structure*. (The sentence is taken in one particular sense; an ambiguous sentence can have different syntactic structures.)
>
> The syntactic structure of a sentence is a set of lexical units of this sentence linked together by *syntactic relations*.

Based on these assumptions, we will first formulate the properties of dependency syntactic structures on which most, if not all, researchers agree (1.1). Then we will proceed with properties that are more Meaning-Text related (1.2).

1.1 *Definitorial properties of dependency syntactic structures*

The syntactic structure of a sentence, presented in terms of dependencies between its words, has the following four definitorial properties, which we will explain in turn:

1. connectedness of the syntactic structure;
2. directedness of syntactic relations;
3. strict hierarchical organization of the syntactic structure;
4. "meaningfulness" of syntactic relations.

1.1.1 *Connectedness of the syntactic structure*
The syntactic structure forms a united whole, that is, a continuous system of syntactic relations. Consequently, any lexical unit L_1 being part of a sentence is syntactically related to at least one other lexical unit L_2; no lexical unit in the sentence is left out of the syntactic structure.

Note that a string made up of two syntactically connected lexical units L_1 and L_2 is a *minimal phrase*: L_1—L_2, where the dash indicates a syntactic connection; for example: *very surprising*, *eat bananas*, *for life*, *and John*, etc. The notion of minimal phrase is generalized to obtain the notion of phrase, which is essential to any further discussion of syntax. Anticipating on what is said in Section 1.1.3, we can characterize a phrase as a projection of a syntactic subtree.

Formal consequence of this property: the syntactic structure is a connected graph.

1.1.2 *Directedness of syntactic relations*
Syntactic relations are directed. This formal property reflects the asymmetric nature of phrases; namely, one component of a minimal phrase dominates the other. This is shown by the fact that the ability of a phrase to be added to a lexical unit inside a sentence is controlled by one of its components. As a result, any phrase behaves rather like its dominant component, or *head*. For instance, *very surprising* behaves like the adjective *surprising* and not like the intensifier adverb *very*; in a sentence, it can replace an adjective but not an intensifier adverb:

(1) a. *He had **strong** opinions. ~ He had **very surprising** opinions.*
b. *He had **incredibly** much money. ~ *He had **very surprising** much money.*

A syntactic relation must be directed in order to reflect this hierarchical organization of the phrase: $L_1 \rightarrow L_2$. The lexical unit L_1 is the (direct) syntactic *governor* of L_2, while L_2 is L_1's syntactic *dependent*. As a result, the head of a phrase is its element that is the direct or indirect governor of all its other elements.

Formal consequence of this property: the syntactic structure is a directed connected graph.

1.1.3 *Strict hierarchical organization of the syntactic structure*
Each lexical unit in the syntactic structure has one and only one syntactic governor, except for one unit that does not have a governor at all. The non-governed unit is the *top node* of the syntactic structure—i.e., the *head* of the sentence. The unicity of the governor for each lexical unit and the presence of a head in each sentence is what is meant by *strict hierarchical organization*.

Postulating the unicity of the governor for each lexical unit is justified by linguistic facts. Namely, in prototypical cases, it is the governor that controls the linear position of

the dependent, which is ordered with respect to it. Thus, in *very←surprising* it is *very* that is positioned before *surprising* rather than *surprising* after *very*. It is natural to suppose that linear positioning is carried out with respect to a single reference point (before it ~ after it); therefore, each lexical unit must have just one syntactic governor, which controls its linear positioning.[1] As a result, there is one and only one lexical unit in the sentence that must have no governor: this is the head of the sentence—i.e., the top node of its syntactic structure.

> Formal consequence of this property: the syntactic structure is an acyclic directed connected graph, i.e., a hierarchized tree or, for short, a tree.

1.1.4 *"Meaningfulness" of syntactic relations*

It is not sufficient to indicate an oriented syntactic relation between two lexical units L_1 and L_2 in order to fully specify the corresponding phrase. A structure such as $L_1 \rightarrow L_2$ can correspond to two or more contrasting phrases. For instance, the structure SEND→MARY is underspecified: it applies to both occurrences of *sent Mary* in sentences (2) below, whereas these occurrences correspond to two semantically contrasting phrases.

(2) a. *Mother **sent→Mary** to the doctor.*

b. *Mother **sent→Mary** 200$.*

This shows the necessity to distinguish syntactic relations themselves according to the exact nature of the phrase they participate in : $\mathbf{r_1}$ for (2a) and $\mathbf{r_2}$ for (2b). It is convenient to use, instead of simple abstract distinguishers, more telling names based on accepted labels for the dependent component of the phrase: **direct-objectival** for (2a) and **indirect-objectival** for (2b), that is, SEND—**direct-objectival**→MARY and SEND—**indirect-objectival**→MARY. Note that we use here adjectival labels applied to the dependency itself, rather than to its dependent element (**direct object**, etc.). This is related to the fact that our own approach focuses more on dependencies themselves (see 1.2 below).

A syntactic relation carries much more linguistic information than simply indicating the hierarchical organization of the phrase. It is a bridge between the meaning of the phrase and its actual surface form, including morphology and prosody. It is in this sense that we call it meaningful—although a syntactic relation does not normally correspond to a specific meaning. Generally speaking, a syntactic dependent of a given type can fulfill different, even contrasting, semantic roles. For instance, the syntactic subject of an active verb can correspond to Agent, Patient or Location (to name just a few):

(3) a. *The University hired 15 new professors last year.*

b. *The University suffered a setback last year.*

c. *The University hosted an important conference last year.*

Syntactic relations do correspond to semantic roles (and vice versa) but these correspondences are by no means direct or systematic.

---

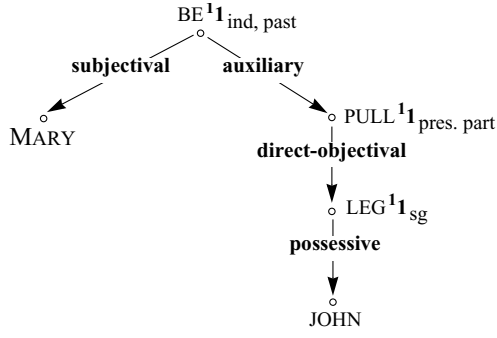
1. In point of fact, the situation is more complex: in some cases a dependent can control the position of the governor, as in PREP→N phrases—see Mel'čuk's paper, p. 26; or consider second-position clitics, which are positioned independently of their syntactic governor—see Milićević's paper.

Formal consequence of this property: the syntactic structure is a tree whose nodes are labeled with lexical units and whose arcs are labeled with the names of syntactic relations.

The above formal characterization of the syntactic structure can be illustrated by a simple example. The dependency syntactic structure of sentence (4) appears in (5). More precisely, the tree diagram in (5) represents the surface-syntactic structure of sentence (4); we will see in the next subsection that, in the Meaning-Text approach, this level of representation is distinguished from a deep-syntactic level.

(4) *Mary was pulling John's leg* ‹Mary was teasing John›.

(5)



Comments

1. The nodes of a syntactic structure are labeled with the lexical units of the language under analysis; their names are printed in small capitals and supplied with sense-distinguishing lexicographic numbers. In the above example, these numbers are borrowed from *Longman Dictionary of Contemporary English ONLINE*.
2. The names of syntactic relations, which label the branches of the tree, are meant to reflect the specific nature of the corresponding construction. These names are of two types:
   - Some are built on the dependent member. Thus, the dependent of the **subjectival** relation is the subject, and that of the **possessive** relation, a possessive complement.
   - Some other names are built on the governor member. For instance, the governor of the **auxiliary** relation is an auxiliary verb, and that of the **prepositional** relation is a preposition.
3. The suffix of the so-called Saxon Genitive (= *-'s*) is a marker of a syntactic relation and therefore it does not appear in the syntactic structure (as all other agreement and government markers). The dependent member of the **possessive** relation is also characterized by its obligatory anteposition and its incompatibility with determiners.

1.2 *Syntactic dependencies, with more Meaning-Text flavor*

The specific contribution of the Meaning-Text approach to syntactic dependency can be summed up in four points: meaning-to-text perspective, emphasis on the description of

dependencies themselves, deep *vs.* surface distinction, and ban on linear order in syntactic structures.

1.2.1 *Meaning-to-text perspective*

In our approach, syntactic structures are considered within a meaning-to-text perspective. More precisely, a syntactic structure is conceived of as a convenient intermediate structure between a source, which is a semantic non-hierarchized network, and a target—a linearly ordered morphological string. It is this perspective that allows the researcher to lay bare the language rules that relate the semantic, syntactic and morphological structures and actually make up language as a device for expressing thoughts. This approach also seems to be more fruitful from a pedagogical point of view: it makes sense to teach people to speak a language, that is, to teach language structures in a text-production—rather than text-interpretation—setting. The meaning-to-text orientation does not of course preclude the elaboration of analytical procedures as well. On the contrary, the formal proximity of syntactic dependency structures (graphs consisting of connected lexical units) to semantic networks (graphs consisting of connected lexical meanings) makes the analysis even easier.

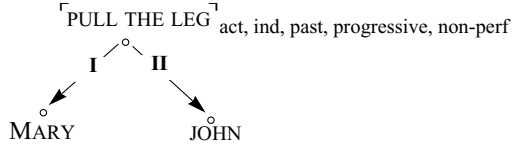
1.2.2 *Emphasis on the description of dependencies*

We put main emphasis on the description of syntactic dependency relations themselves, rather that on the sentence elements connected by them. Dependency relations are considered as information-carrying entities: each syntactic relation (**subjectival**, **direct-objectival**, **auxiliary**, **prepositional**, **conjunctional**, etc.) is treated as a linguistic unit in its own right. Means and techniques are developed in order to establish the inventories of syntactic relations for particular languages. (For French valence-controlled dependencies, see Iordanskaja & Mel'čuk's paper in this volume.)

1.2.3 *Deep* vs. *surface distinction*

Two levels of syntactic dependency are distinguished: the deep-syntactic structure, closer to meaning, and the surface-syntactic structure, closer to the "physical" form of the sentence. For instance, sentence (4) above is associated to the following deep-syntactic structure:

(6)



This structure obviously contrasts with the surface-syntactic structure in (5) in that it presents explicitly only the hierarchization of the three full lexical units whose meanings are expressed in sentence (4).[2] Unlike the surface-structure, it does not directly reflect all that is needed in order to properly linearize and morphologize all lexical units actually appearing in the sentence.

2. Though lexemes PULL and LEG appear in the sentence, their meanings are not expressed because they are no more than formal constituents of the lexical unit (the idiom) ⌜PULL THE LEG⌝.

We cannot justify here the deep *vs.* surface distinction at the syntactic level. Suffice it to indicate four aspects in which it is particularly useful: 1) handling mismatches between semantic and syntactic structures; 2) processing idioms, which are simplexes at the deep-syntactic level but complex structures at the surface level, see (4)-(6) above; 3) providing means for a systematic account of collocations and facilitating the choice of collocates at the surface-syntactic level (see Polguère 2000, for lexicalization in text generation); 4) ensuring an efficient description of syntactic paraphrasing (Žolkovskij & Mel'čuk 1967; Mel'čuk 1988; Milićević 2007).

1.2.4 *Ban on linear order in syntactic structures*

We proscribe linear order in syntactic structures because ordering is the main and universal means of **expression** of those structures; therefore, it is not part of them.[3] A dependency syntactic structure of a sentence must contain all the information necessary to properly compute all possible word orders in the sentence. This task is taken, basically, by syntactic dependency relations. For each individual relation, syntactic rules indicate the ordering of its dependent element with respect to the governor.

It is impossible to conclude this section without mentioning the fact that dependency approach is in fierce competition with phrase-structure approach. The relation between the two viewpoints is touched upon in Mel'čuk's and Kahane's papers in this volume. Let us simply mention here that although phrase-structure is rejected as a means of representing the syntactic organization of the sentence, phrases themselves are indispensable, even in the strictest dependency approach. They appear at the deep-morphological level of the sentence representation and are treated as genuine linguistic units with particular linear order and prosody.

2 Presentation of the papers

The present volume is not a well-organized manual or systematically arranged anthology. It is a collection of four papers, each dealing with a specific aspect of dependency syntax, and arranged in a natural order:

- the first paper, by Mel'čuk, is a general theoretical discussion of the notion of dependency as applied to language;
- it is followed by Kahane's paper, which broadens the theoretical perspective by presenting an example of a formal dependency grammar and an in-depth comparison with phrase structure grammar;
- the next paper, by Iordanskaja & Mel'čuk, presents the application of theoretical principles for establishing Surface-Syntactic relations in French—more specifically, valence-controlled (that is, actantial) ones;
- finally, Milićević considers a challenging word order problem for dependency approach: second position clitics in Serbian.

In a collection of this type, repetitions are unavoidable; we eliminated the most obvious and irritating ones, but by no means all of them—first, this would have resulted in

3. Of course, drawing a dependency tree on a page we have to somehow order its nodes; however, this order is chosen only for the readers' convenience.

too deep a recasting of the volume and second, isn't *Repetitio mater studiorum*? We also standardized the terminology and notations in all four papers, as well as the presentation of the references—which are kept separate for each paper.

Now we offer short abstracts of the four papers.

2.1 *Mel'čuk: a sketch of dependency theory*

Based on such basic notions as wordform, clause, semantic predicate, inflectional category, etc., the paper starts with a demonstration of the existence of three types of dependency relations between two wordforms in sentences of natural languages:

- semantic dependency between two wordform meanings: predicate→argument dependency;
- syntactic dependency between two wordforms as such: governor→dependent dependency, which controls the passive valence of the phrase (its potential governors), as well as the mutual linear positioning of its wordforms;
- morphological dependency between two wordforms, of which one controls the inflectional values (grammemes) of the other: controller→target dependency.

Fourteen cases of possible combinations of different types of dependency between two wordforms in a sentence are considered: for instance, Fr. *Lida semble heureuse* 'Lida seems happy', where *Lida* depends on *heureuse* semantically, while there is no direct syntactic dependency between these two wordforms, and where morphologically *heureuse* depends on *Lida* for its singular and feminine gender. This discussion contributes to systematically distinguishing between different kinds of dependencies; a failure to do so has resulted on many an occasion in confusion and serious mistakes.

Concentrating on syntactic dependency, the author moves to propose three groups of criteria for establishing a syntactic dependency between two wordforms $\mathbf{w_1}$ and $\mathbf{w_2}$ in a sentence:

- Criteria A establish the presence of a syntactic link between $\mathbf{w_1}$ and $\mathbf{w_2}$, based on determining their mutual linear arrangement and possible prosodic unity;
- Criteria B establish the direction of the syntactic link between $\mathbf{w_1}$ and $\mathbf{w_2}$, based on the passive syntactic valence of the phrase $\mathbf{w_1}$—$\mathbf{w_2}$, its external morphological links, and its semantic content;
- Criteria C establish the specific type of the surface-syntactic relation **r** holding between $\mathbf{w_1}$ and $\mathbf{w_2}$, based on the absence of semantic contrast, syntactic substitutability of the dependent subtree, and repeatability of **r**.

The introduction of these criteria allows for a deeper characterization of syntactic dependency. Several particular problems related to syntactic dependency are analyzed (the top node in a syntactic structure, verbless sentences, zero verb forms, ellipses, etc.); an illustrative list of Surface-Syntactic relations for English is proposed: 52 surface-syntactic relations, of which 50 subordinate and 2 coordinate ones. The paper ends with a general overview of some residual problems.

2.2 *Kahane: phrases in Head-Driven Phrase Structure Grammar*

Kahane's paper, unlike the other three papers in this volume (which are situated within the Meaning-Text approach), crosses theory boundaries. By examining the role played by phrases in the process of sentence building, it bridges the gap between dependency and phrase structure approach to syntax. More specifically, the paper offers a dependency interpretation of extraction phenomena as modeled within the Head-driven Phrase Structure approach, or HPSG (Pollard & Sag 1994). The focus on extraction is justified by the fact that so-called Complementizers and Relativizers have received great attention in phrase structure approaches, with analyses that are primarily based on phrases as specific syntactic entities. The paper offers an alternative—more precisely, lexicalist—modelling of extraction, where HPSG formalism is used to implement a purely dependency interpretation of this phenomenon. An interesting point developed by Kahane is the treatment of relative pronouns, for which double dependency is proposed: a relative pronoun is considered, as a pronoun, to be a dependent of the Main Verb of the relative clause, and at the same time, as a transferer, to be its governor.

The main thesis defended by Kahane is that the modeling of extraction belongs to the syntax-semantic interface, while phrases are entities of syntax proper (Gerdes & Kahane 2007). The paper claims that a lexical-based approach to extraction will therefore be sounder and more economical than a phrase-based one, and it supports this claim by offering precise, well-formalized analyses of specific syntactic structures. The use of HPSG formalism presents the advantage of a rigorous mathematical modeling of extraction, while demonstrating that this formalism can very well support a dependency approach to syntax.

2.3 *Iordanskaja and Mel'čuk: valence-controlled surface-syntactic relations between a verb and its dependents in French*

This paper is an elaborate exercise in the methodology of establishing surface-syntactic relations [= SSyntRels] for French—those that are controlled by the active syntactic valence of a verbal governor. The central idea is to fit the techniques for establishing SSyntRels in a language into the accepted theoretical and typological frame used for establishing the inventories of other linguistic units (e.g., phonemes and grammemes).

Sixteen relevant linguistic properties of possible actantial syntactic dependents of French verbs are put forward: being present in any full-fledged clause, being able to depend on the Main Verb only, etc. Based on the similarity of the dependents with respect to these properties, the classification of the verbal syntactic constructions of French is proposed. Three criteria for the specific types of SSyntRels are defined (these are Criteria C in Mel'čuk's paper): absence of semantic contrast (known in linguistics as Minimal pair test), mutual substitutability (Substitution test), and repeatability (Cooccurrence test). The commonality of the syntactic properties of dependents, combined with the application of these criteria has resulted in a list of sixteen SSyntRels. For each SSyntRel described, the authors supply: 1) properties of this SSyntRel; 2) formal types of its dependents (prepositionless noun, a DE-infinitive, a subordinate clause, etc.); 3) linguistic comments and a justification (if need be) in the form of a comparison with other SSyntRels.

2.4 *Milićević: Serbian second-position clitics in a dependency framework*

This paper describes a known word order problem of Serbian syntax: linear placement of clitics. The problem, previously studied in phrase-structure approaches, is tackled here in terms of dependency representation.

A crucial distinction is stated between two major types of clitics: verb-hosted clitics, like those found in French, Spanish, etc., and second-position clitics, characteristic of Serbian. The latter include unstressed pronouns, auxiliaries and particles (boldfaced in Serb. *Gde **li sam ih** sreo?* lit. ‘Where on-Earth am them having-met?’ = ‘Where on Earth have I met them?’

Two defining properties of second-position clitics are stated: 1) they cluster, i.e., are brought together in a rigidly specified linear sequence, which is linearly positioned as a whole, and 2) the clitic cluster must stand in the clause, roughly speaking, after the first appropriate constituent. A Meaning-Text type constituent is a string of wordforms representing the continuous projection of a dependency subtree and treated by word-order rules as a whole; such constituents appear in the Deep-Morphological Structure. Milićević establishes the syntactic and prosodic properties of constituents that allow/disallow a constituent to host the clitic cluster. According to these features, constituents are identified as non-hosts or potential hosts, the latter being subdivided into skippable *vs.* non-skippable hosts and insertable *vs.* non-insertable hosts.

Linear positioning of a clitic cluster is carried out as follows: first, establishing the linear order of all the constituents of the clause—except for the clitic cluster; second, processing of all the constituents to determine which ones can or must host the cluster.

Two special topics are introduced and discussed: 1) skipping of constituents when looking for a landing site for the clitic cluster (in fact, the cluster can end up after second, third, etc. constituent); 2) insertion of the cluster inside the hosting constituent.

3 Making of the volume

This volume is the fruit of an informal symposium held in a bucolic atmosphere of the French region of Auvergne, in the village of Saint Just, in 1999. A group of colleagues and friends, all working in Meaning-Text dependency approach, gathered there in order to exchange ideas about the use of the dependency formalism in theoretical as well as computational perspectives. The first sketches of the four papers were presented and discussed during the *Saint Just Symposium*. It took, however, many years to finalize the papers and prepare the volume for publication. Nevertheless, the results of the research reported here are not dated because, in spite of so much effort and many interesting achievements, dependency approach to syntax still does not receive enough attention.

This book is intended to partially fill the gap and thus contribute to linguistic dependency coming to the fore in linguistics.

In conclusion, we will pay a pleasant debt of gratitude and acknowledge the financial assistance of the Fonds québécois de la recherche sur la société et la culture (FQRSC), which supports the work of the Observatoire de linguistique Sens-Texte (OLST) at the Université de Montréal, and the Blaise Pascal research grant given to Igor Mel'čuk, through the Fondation de l'École Normale Supérieure (Paris) and the Région Île-de-France. Next, we would like to thank Danielle "Dan" Collignon for her contribution to the preparation of the manuscript. Last, but not least, we are eternally grateful to Mr. René Cussac, mayor of Saint Just (Cantal, France), who provided us with everything we needed (and more) for our little scientific gathering.

References

Gerdes, Kim & Kahane, Sylvain 2007. "Phrasing It Differently". In *Selected Lexical and Grammatical Issues in the Meaning-Text Theory*, L. Wanner (ed.), 297-328. Amsterdam/Philadelphia: Benjamins.

Mel'čuk, Igor 1988. "Paraphrase et lexique dans la théorie linguistique Sens-Texte: Vingt ans après". *Cahiers de lexicologie* 52(1): 5-50 & 53(2): 5-53.

Milićević, Jasmina 2007. *La paraphrase: modélisation de la paraphrase langagière*. Bern: Peter Lang.

Polguère, Alain 2000. "A 'Natural' Lexicalization Model for Language Generation". *Proceedings of the Fourth Symposium on Natural Language Processing (SNLP'2000)*, 37-50. Chiangmai (Thailand).

Pollard, Carl & Sag, Ivan 1994. *Head-Driven Phrase Structure Grammar*. CSLI Series, Chicago: University of Chicago Press.

Žolkovskij, Aleksandr & Mel´čuk, Igor 1967. "O semantičeskom sinteze" [On Semantic Synthesis (of Texts)]. *Problemy Kibernetiki* 19: 177-238.